



**UNIVERSIDAD NACIONAL DE LA PLATA**

**FACULTAD DE CIENCIAS EXACTAS**

**DEPARTAMENTO DE CIENCIAS BIOLÓGICAS**

---

**Trabajo de Tesis Doctoral:**

***Búsqueda de nuevos fármacos  
antichagásicos inhibidores de la Tripanotión  
sintetasa***

---

**Tesista: Alice, Juan Ignacio**

**Director/a: Talevi, Alan**

**Codirector/a: Bellera, Carolina Leticia**

**Año: 2022**

Universidad Nacional de La Plata  
Facultad de Ciencias Exactas  
Departamento de Ciencias Biológicas  
Laboratorio de Investigación y Desarrollo de Bioactivos | LIDeB

Trabajo de Tesis Doctoral

# Búsqueda de nuevos fármacos antichagásicos inhibidores de la *Tripanotión sintetasa*



**Tesista: Alice, Juan Ignacio**

**Director: Talevi, Alan**

**Codirectora: Bellera, Carolina Leticia**

**Año: 2022**

Departamento de  
Ciencias Biológicas  
FACULTAD DE  
CIENCIAS EXACTAS



UNIVERSIDAD  
NACIONAL  
DE LA PLATA



# Índice

|   |           |
|---|-----------|
| Agradecimientos.....  | 7         |
| Acrónimos.....  | 8         |
| Objetivos e hipótesis de trabajo.....   | 10        |
| Objetivo General.....   | 10        |
| Objetivos Específicos.....  | 10        |
| Hipótesis de trabajo.....   | 10        |
| <b>Capítulo 1   Introducción.....</b>   | <b>11</b> |
| 1.1 Enfermedades Tropicales Desatendidas.....                                     | 12        |
| 1.2 Enfermedad de Chagas.....   | 12        |
| 1.2.1 Aspectos generales.....   | 12        |
| 1.2.2 Trypanosoma cruzi.....  | 13        |
| 1.2.3 Ciclo de vida de T. cruzi.....  | 15        |
| 1.2.4 Vías de transmisión.....  | 16        |
| 1.2.5 Epidemiología y distribución de la enfermedad.....                          | 17        |
| 1.2.6 Medicamentos disponibles.....   | 18        |
| 1.2.7 Perspectivas hacia nuevas farmacoterapias para la Enfermedad de Chagas..... | 20        |
| 1.3 Procesos redox dependientes de Tripanotión.....                               | 22        |
| 1.4 Tripanotión sintetasa.....  | 24        |
| 1.4.1 La enzima.....  | 24        |
| 1.4.2 TryS como blanco molecular.....   | 25        |
| 1.5 Referencias Bibliográficas.....   | 30        |
| <b>Capítulo 2   Descubrimiento de Fármacos.....</b>                               | <b>36</b> |
| 2.1 Descubrimiento de fármacos: Generalidades.....                                | 37        |
| 2.2 Cribado Virtual.....  | 40        |
| 2.2.1 Métodos basados en el blanco molecular.....                                 | 42        |
| 2.2.2 Métodos basados en ligando.....   | 44        |
| 2.2.3 Métodos combinados.....   | 48        |
| 2.3 Reposicionamiento de fármacos.....  | 48        |
| 2.3.2 Métodos sistemáticos de reposicionamiento de fármaco.....                   | 49        |
| 2.3.3 Reposicionamiento de fármacos para ETD.....                                 | 52        |
| 2.3.4 Casos prometedores y exitosos en tripanosomátidos.....                      | 53        |
| 2.4 Estrategia utilizada en el trabajo de tesis.....                              | 53        |

|   |           |
|---|-----------|
| 2.5 Referencias Bibliográficas .....  | 55        |
| <b>Capítulo 3   Metodología .....</b>   | <b>60</b> |
| 3.1 Introducción .....  | 61        |
| Primera campaña de cribado virtual y evaluación experimental .....            | 62        |
| 3.2 Construcción de modelos QSAR – Funciones Discriminantes.....              | 62        |
| 3.2.1 Recopilación y curado del set de datos.....                             | 62        |
| 3.2.2 Partición del set de datos.....   | 63        |
| 3.2.3 Selección de los conjuntos de entrenamiento y de prueba .....           | 66        |
| 3.2.4 Cálculo de descriptores.....  | 67        |
| 3.2.5 Construcción de los modelos.....  | 67        |
| 3.2.6 Aleatorización de subespacios.....                                      | 67        |
| 3.2.7 Selección de descriptores moleculares y generación de modelos.....      | 68        |
| 3.3 Validación de los modelos .....   | 70        |
| 3.3.1 Validación Interna.....   | 70        |
| 3.3.2 Validación externa.....   | 71        |
| 3.4 Comportamiento y desempeño de los modelos .....                           | 72        |
| 3.4.1 Análisis de las curvas ROC.....   | 73        |
| 3.4.2 Campaña retrospectiva de cribado virtual .....                          | 74        |
| 3.5 Combinación de Modelos.....   | 74        |
| 3.6 Selección del valor de corte .....  | 75        |
| 3.7 Cribado Virtual.....  | 76        |
| 3.8 Dominio de aplicación de los modelos.....                                 | 76        |
| 3.9 Sección Experimental .....  | 77        |
| 3.9.1 Evaluaciones biológicas.....  | 77        |
| 3.9.2 Evaluaciones bioquímicas.....   | 79        |
| Segunda campaña de cribado virtual y evaluación experimental .....            | 80        |
| 3.10 Construcción de modelos QSAR – Funciones Discriminantes.....             | 80        |
| 3.10.1 Recopilación y curado del set de datos.....                            | 80        |
| 3.10.2 Partición del set de datos .....                                       | 80        |
| 3.10.3 Selección de los conjuntos de entrenamiento y de prueba .....          | 81        |
| 3.10.4 Digitalización de las estructuras y cálculo de descriptores .....      | 81        |
| 3.10.5 Construcción de los modelos.....                                       | 81        |
| 3.10.6 Aleatorización de subespacios .....                                    | 81        |
| 3.10.7 Selección de descriptores moleculares y generación de los modelos..... | 81        |
| 3.11 Validación de los modelos.....   | 81        |
| 3.12 Comportamiento y desempeño de los modelos .....                          | 82        |
| 3.12.1 Factor de enriquecimiento.....   | 82        |
| 3.12.2 Métrica BEDROC .....   | 82        |

|   |            |
|---|------------|
| 3.12.3 Campaña retrospectiva de cribado virtual .....                   | 83         |
| 3.13 Combinación de modelos.....  | 83         |
| 3.14 Selección del valor de corte .....                                 | 84         |
| 3.15 Cribado virtual .....  | 84         |
| 3.16 Dominio de Aplicación de los modelos.....                          | 85         |
| 3.17 Sección experimental .....   | 85         |
| 3.17.1 Evaluaciones biológicas .....                                    | 85         |
| 3.17.2 Evaluaciones bioquímicas .....                                   | 86         |
| 3.18 Referencias Bibliográficas .....                                   | 87         |
| <b>Capítulo 4   Resultados.....</b>                                     | <b>92</b>  |
| 4.1 Resultados de la primera campaña: Estrategia Computacional.....     | 93         |
| 4.1.1 Compilado del set de datos .....                                  | 93         |
| 4.1.2 Partición del set de datos .....                                  | 94         |
| 4.1.3 Generación y validación de los modelos .....                      | 96         |
| 4.1.4 Evaluación del desempeño de los modelos.....                      | 98         |
| 4.1.5 Combinación de modelos .....                                      | 99         |
| 4.1.6 Selección del valor de corte para cribado virtual.....            | 101        |
| 4.1.7 Cribado virtual y selección de compuestos .....                   | 102        |
| 4.2 Primera Campaña de Evaluación Experimental.....                     | 105        |
| 4.2.1 Evaluación biológica frente a <i>T. cruzi</i> .....               | 105        |
| 4.2.2 Evaluación biológica frente a <i>T. brucei</i> .....              | 108        |
| 4.2.3 Evaluación bioquímica frente a <i>Tripanotión sintetasa</i> ..... | 111        |
| 4.3 Resultados de la segunda campaña: Estrategia computacional .....    | 115        |
| 4.3.1 Compilado de un nuevo set de datos .....                          | 115        |
| 4.3.2 Partición set de datos.....                                       | 115        |
| 4.3.3 Generación y validación de modelos .....                          | 117        |
| 4.3.4 Evaluación del desempeño de los modelos.....                      | 120        |
| 4.3.5 Combinación de modelos .....                                      | 121        |
| 4.3.6 Selección del valor de corte para cribado virtual.....            | 125        |
| 4.3.7 Cribado virtual y selección de compuestos .....                   | 127        |
| 4.4 Segunda campaña de evaluación experimental.....                     | 129        |
| 4.4.1. Evaluación biológica frente a <i>T. cruzi</i> .....              | 129        |
| 4.4.2 Evaluación biológica frente a <i>T. brucei</i> .....              | 132        |
| 4.4.3 Evaluación bioquímica frente a <i>TryS</i> .....                  | 134        |
| 4.5 Referencias Bibliográficas .....                                    | 138        |
| <b>Capítulo 5   Drogabilidad .....</b>                                  | <b>141</b> |
| 5.1 Introducción .....  | 142        |
| 5.1.1 TryS: una enzima para estudiar.....                               | 145        |

|   |            |
|---|------------|
| 5.2 Metodología.....  | 146        |
| 5.2.1 Conformación del Set de datos.....  | 146        |
| 5.2.2 Partición del set de datos.....   | 147        |
| 5.2.3 Conformación de los conjuntos de entrenamiento y validación.....                          | 148        |
| 5.2.4 Obtención de descriptores.....  | 148        |
| 5.2.5 Generación de modelos.....  | 151        |
| 5.2.6. Validación de los modelos obtenidos.....   | 151        |
| 5.2.7 Dominio de aplicación.....  | 153        |
| 5.3 Resultados.....   | 153        |
| 5.3.1 Conformación y partición del set de datos.....  | 153        |
| 5.3.2 Generación y validación de los modelos.....   | 154        |
| 5.3.3 Evaluación y resultados frente a la enzima TryS.....                                      | 158        |
| 5.3.4. Análisis visual complementario frente a LmTryS.....                                      | 158        |
| 5.4 Discusión de resultados.....  | 161        |
| 5.5 Referencias Bibliográficas.....   | 162        |
| <b>Capítulo 6   Conclusiones.....</b>   | <b>166</b> |
| 6.1 Conclusiones de la primera campaña.....   | 167        |
| 6.2 Conclusiones de la segunda campaña.....   | 168        |
| 6.3 Un balance global de las dos campañas.....  | 170        |
| 6.4 Conclusiones sobre drogabilidad.....  | 171        |
| 6.5 Perspectivas.....   | 172        |
| 6.6 Detalle de publicaciones, congresos y becas.....  | 172        |
| 6.7 Referencias bibliográficas.....   | 174        |
| <b>Anexo 1   Descriptores.....</b>  | <b>176</b> |
| <b>Anexo 2   Drogabilidad (PDB id).....</b>   | <b>184</b> |
| .....   | 184        |
| <b>Anexo 3   .....</b>  | <b>187</b> |
| <b>Estructuras de compuestos activos e inactivos del set de datos, segunda campaña de CV ..</b> | <b>187</b> |

# Agradecimientos

A mi director, **Dr. Alan Talevi**, por aceptarme en su grupo permitiéndome dar mis primeros pasos en la ciencia, por sus consejos científicos y enseñanzas compartidas.

A mi codirectora, **Dra. Carolina Bellera**, por la motivación para iniciar el recorrido docente y científico, por sus consejos diarios y la paciencia infinita.

A la **Facultad de Ciencias Exactas, UNLP y CONICET**, por brindar los medios que hicieron posible la realización de la tesis.

A la **Dra. María Laura Sbaraglini** por guiarme en el mundo de los ensayos biológicos y por las charlas en el tren.

A la **Dra. Catalina Alba Soto**, por su confianza y predisposición para colaborar con mi formación académica en su laboratorio y los consejos sobre la vida científica y académica.

Al **Dr. Marcelo Comini y su equipo de colaboradores**, por su tiempo dedicado para formarme y enseñarme con mucha paciencia cada una de las técnicas utilizadas, pero sobre todo por hacerme parte del Instituto durante mis estadías en Uruguay.

A la **Dra. Carolina Carrillo y Dani Ruiz**, por su ayuda con ensayos biológicos.

A la **Dra. Ma. Esperanza Ruiz**, por su colaboración, asesoramiento y su enseñanza en el tratamiento estadístico de los datos.

Al **Farm. Francisco "pachi" Morales**, por su infinita paciencia y colaboración cuando dejaban de funcionar los *scripts*.

A mis compañeros y compañeras del LIDeB, porque cada día se vive mejor si es compartido.

A mis compañeros de las distintas cátedras, por su confianza y enseñanzas de la vida.

A mis **amigos de Bragado y La Plata**, por escucharme y sostenerme ante las caídas, pero también por los momentos de alegría compartidos.

A mis **hermanos y sobrinos** por estar siempre presentes.

A **mi madre y mi padre**, por apoyarme en todos los caminos iniciados, pero sobre todo porque sin ellos ninguno de mis caminos podría haber comenzado. *Gracias!!!*

# Acrónimos

|                |   |
|----------------|---|
| <b>aa</b>      | aminoácido/s  |
| <b>ACM/MCA</b> | Análisis de Control Metabólico ( <i>Metabolic Control Analysis</i> )    |
| <b>ADL</b>     | Análisis Lineal Discriminante   |
| <b>ADME</b>    | Administración, Distribución, Metabolismo y Excreción                   |
| <b>ADN</b>     | Ácido Desoxirribonucleico   |
| <b>ARNi</b>    | Ácido Ribonucleico de interferencia                                     |
| <b>ATP</b>     | Adenosín Trifosfato   |
| <b>AUCROC</b>  | Área Bajo la Curva ROC ( <i>Area Under the Curve ROC</i> )              |
| <b>BZN</b>     | Benznidazol   |
| <b>CDC</b>     | Directorio de Cavidades Drogables ( <i>Druggable Cavity Directory</i> ) |
| <b>CV</b>      | Cribado Virtual   |
| <b>CVBBM</b>   | Cribado Virtual Basado en el Blanco Molecular                           |
| <b>CVBL</b>    | Cribado Virtual Basado en el Ligando                                    |
| <b>Cz</b>      | Cruzipaína  |
| <b>DA</b>      | Dominio de aplicación   |
| <b>DHFR</b>    | Dihidrofolato Reductasa   |
| <b>DMSO</b>    | Dimetilsulfóxido  |
| <b>DTU</b>     | Unidad de Tipificación Discreta ( <i>Discrete Typing Unit</i> )         |
| <b>EF</b>      | Factor de enriquecimiento ( <i>Enrichment factor</i> )                  |
| <b>E-5</b>     | Ensamble de los mejores 5 modelos individuales                          |
| <b>E-10</b>    | Ensamble de los mejores 10 modelos individuales                         |
| <b>E-15</b>    | Ensamble de los mejores 15 modelos individuales                         |
| <b>E-20</b>    | Ensamble de los mejores 20 modelos individuales                         |
| <b>E-90</b>    | Ensamble de los mejores 90 modelos individuales                         |
| <b>ETD</b>     | Enfermedades Tropicales Desatendidas                                    |
| <b>GDPDH</b>   | Gliceraldehído-3-Fosfato Deshidrogenasa                                 |
| <b>G6PDH</b>   | Glucosa-6-Fosfato Deshidrogenasa  |
| <b>GSH</b>     | Glutati3n   |
| <b>Gsp</b>     | Glutationilspemidina o mono glutationilspemidina                        |
| <b>GspS</b>    | Glutationilspemidina Sintetasa  |
| <b>HGPRT</b>   | Hipoxantina-guanina fosforribosiltransferasa                            |

|                         |   |
|-------------------------|---|
| <b>Hk</b>               | Hexoquinasa   |
| <b>HMG-CoA</b>          | 3-hidroxi-3-metil-glutaril-Coenzima A   |
| <b>HTS</b>              | Cribado de Alto Rendimiento ( <i>High throughput Screening</i> )  |
| <b>YECS</b>             | gamma-Glutamilcisteina Sintetasa  |
| <b>Ig</b>               | Inmunoglobulinas  |
| <b>IUPAC</b>            | Unión Internacional de Química Pura y Aplicada ( <i>International Union of Pure and Applied Chemistry</i> ) |
| <b>LBD</b>              | Descubrimiento Basado en Literatura ( <i>literature-based discovery</i> )                                   |
| <b>LmTryS</b>           | Tripanotión Sintetasa <i>de la especie Leishmania major</i>   |
| <b>NFX</b>              | Nifurtimox  |
| <b>OMS/WHO</b>          | Organización Mundial de la Salud ( <i>World Health Organization</i> )                                       |
| <b>OPS</b>              | Organización Panamericana de la Salud   |
| <b>PDB</b>              | <i>Protein Data Bank</i>  |
| <b>PPP</b>              | Vía de las Pentosas Fosfatos ( <i>Pentose Phosphate Pathway</i> )   |
| <b>PPV</b>              | <i>Positive Predictive value</i>  |
| <b>QSAR/QSPR</b>        | Relación Cuantitativa Estructura-Actividad. ( <i>Quantitative Structure-Activity Relationship</i> )         |
| <b>%BC</b>              | Porcentaje de Buenas Clasificaciones  |
| <b>RMN</b>              | Resonancia Magnética Nuclear  |
| <b>ROC</b>              | Característica Operativa del Receptor ( <i>Receiver Operating Characteristics</i> )                         |
| <b>6PGL</b>             | 6-Fosfogluconolactona   |
| <b>6PGDH</b>            | 6-Fosfogluconato Deshidrogenasa   |
| <b>SQO</b>              | Escualeno Epoxidasa   |
| <b>TbTryS</b>           | Tripanotión Sintetasa <i>de la especie Trypanosoma brucei</i>   |
| <b>TcTryS</b>           | Tripanotión Sintetasa <i>de la especie Trypanosoma cruzi</i>  |
| <b><i>T. brucei</i></b> | <i>Trypanosoma brucei</i>   |
| <b><i>T. cruzi</i></b>  | <i>Trypanosoma cruzi</i>  |
| <b>TGI</b>              | Tracto Gastro Intestinal  |
| <b>TryS</b>             | Tripanotión Sintetasa ( <i>Trypanothione Synthetase</i> )   |
| <b>TR</b>               | Tripanotión Reductasa ( <i>Trypanothione Reductase</i> )  |
| <b>T[SH]2</b>           | Tripanotión ( <i>Trypanothione</i> )  |

# Objetivos e hipótesis de trabajo

## Objetivo General

- *Identificar, mediante búsqueda racional asistida por computadora, nuevos compuestos químicos capaces de inhibir/modular la actividad de la enzima Tripanotión Sintetasa; los cuales puedan ser candidatos para utilizarse en la farmacoterapia de enfermedades originadas por tripanosomátidos.*

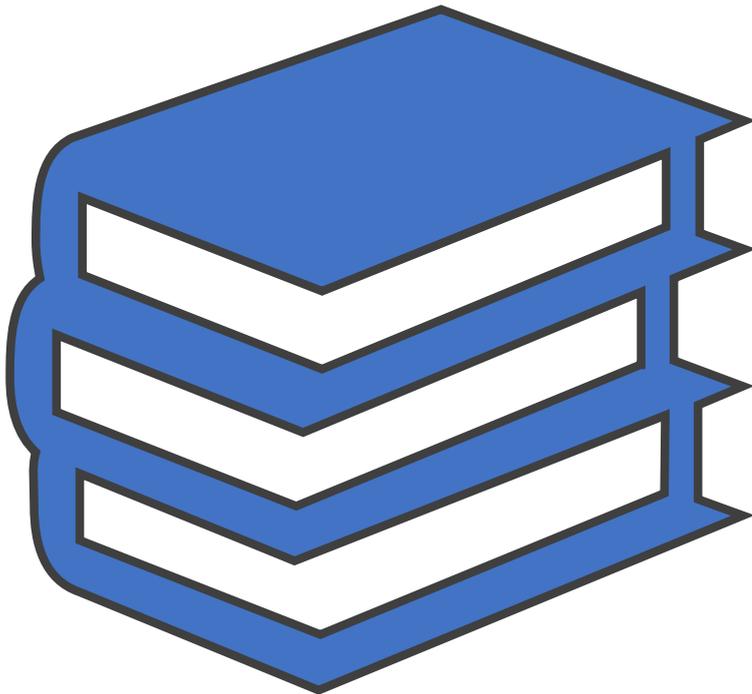
## Objetivos Específicos

- *Compilar, a partir de literatura, un set de datos de compuestos con actividad reportada frente a la enzima Tripanotión Sintetasa.*
- *Desarrollar y validar modelos computacionales basados en el ligando capaces de identificar compuestos químicos, tipo-fármaco.*
- *Aplicar los modelos generados en campañas del tamizado virtual de diversas bibliotecas químicas. En particular, la identificación de drogas aprobadas por diferentes agencias regulatorias, permitiendo detectar la actividad de interés en fármacos ya utilizados en clínica para otras indicaciones terapéuticas (estrategia conocida como knowledge-based drug repurposing, es decir, búsqueda de segundos usos médicos basada en el conocimiento).*
- *Adquirir los candidatos seleccionados a partir de la metodología in silico.*
- *Evaluar experimentalmente, en ensayos in vitro, los compuestos adquiridos.*

## Hipótesis de trabajo

*Mediante modelado in silico basado en el ligando se podrían identificar nuevos compuestos con actividad tripanocida, inhibidores de la enzima Tripanotión Sintetasa, entre los cuales existen posibles candidatos a fármacos, es decir, compuestos con las características farmacocinéticas/farmacodinámicas adecuadas para convertirse en potenciales agentes para el tratamiento de la enfermedad de Chagas.*

# Capítulo 1 | Introducción



### 1.1 Enfermedades Tropicales Desatendidas

Las **Enfermedades Tropicales Destendidas (ETD)** son un grupo heterogéneo de patologías infecciosas crónicas que prevalecen en regiones de climas tropicales y subtropicales, azotando principalmente a las poblaciones social y económicamente vulnerables. Las regiones más afectadas son África, América Latina y parte de Asia. La escasez de agua potable, el bajo ingreso per cápita, la falta de educación y la escasez de sistemas de salud adecuadamente preparados son algunas de las características más relevantes que comparten dichas poblaciones. Adicionalmente estas enfermedades son de carácter transmisible, por lo tanto, se transmiten de manera directa o indirecta (esto es, mediante vectores); en muchos casos también existe la transmisión zoonótica, es decir, de animales a personas. Si bien estas enfermedades suelen presentar baja tasa de mortalidad, poseen un alto índice de morbilidad, lo que produce un descenso en la calidad de vida de las personas enfermas, medido habitualmente en años de vida ajustados por discapacidad (DALYs, por su denominación en inglés *disability-adjusted life years*) (Andrade et al., 2019; Aya Pastrana et al., 2020, 2020; C. L. Bellera et al., 2018).

El término ETD fue propuesto por Hotez y colaboradores en el año 2003 con el fin de visibilizar la escasa inversión que se realizaba en patologías como tuberculosis o malaria, en relación con otras enfermedades que tenían la atención prioritaria de los sectores gubernamentales y privados (Smith & Taylor, 2016).

Actualmente la **Organización Mundial de la Salud (OMS/WHO)** reconoce 18 ETD presentes en más de 140 países y que afectan a más de 1.000 millones de personas en todo el mundo. En este punto la propia OMS alerta sobre el potencial riesgo a nivel mundial que significa la falta de atención y control de estas enfermedades, ya que en la era de la globalización el continuo flujo de mercancías y personas ha derivado en contagios en regiones no endémicas ([www.who.int](http://www.who.int), última consulta agosto-2020). A pesar de estas advertencias, no se cuenta aún con un programa centralizado de estadísticas sobre ETD. Los datos reportados provienen de encuestas desarrolladas mediante diferentes metodologías, y por lo tanto las métricas no son comparables entre sí. Esto dificulta la evaluación de la inversión en investigación y desarrollo en este área (Moran, 2009).

En la presente tesis aportamos nuestros esfuerzos para emprender un camino de búsqueda hacia nuevos tratamientos farmacológicos orientados a ETD causadas por tripanosomátidos; nos enfocamos principalmente en la Enfermedad de Chagas, aunque algunos resultados podrían ser extrapolados a la Enfermedad del Sueño y la Leishmaniasis.

### 1.2 Enfermedad de Chagas

#### 1.2.1 Aspectos generales

La Tripanosomiasis Americana, más conocida como Enfermedad de Chagas es una parasitosis endémica de América Latina causada por el protozoo *Trypanosoma cruzi* (*T. cruzi*).

Actualmente distintas organizaciones estiman que hay entre 8 y 10 millones de personas infectadas en todo el mundo. Si bien la enfermedad tiene una fuerte prevalencia en la zona de América Latina, el creciente flujo migratorio ha derivado en contagios de personas en regiones no endémicas como Europa o Asia ([www.mundosano.org](http://www.mundosano.org), última consulta agosto-2020).

La principal vía de transmisión del parásito, en las regiones endémicas, es mediante las heces de la vinchuca (*Triatoma infestans*) un insecto hematófago de la familia de los *triatomas*. Este insecto defeca mientras se alimenta, y en consecuencia las heces depositadas sobre la piel suelen ser arrastradas hasta la herida (producto de la picadura) durante el proceso de rascado, lo que permite que los parásitos ingresen al organismo y accedan al torrente sanguíneo. La transmisión alimentaria también se produce mediante las heces del vector, pero de manera indirecta, ya que en este caso el insecto defeca sobre alimentos o bebidas, los cuales al ser ingeridos resultan en la infección de la persona. Además, existen otras vías de transmisión como la transmisión vertical (madre-hijo), por trasplante de órganos, transfusiones de sangre, o accidentes laborales. Estas vías son las que predominan en las regiones no endémicas, donde la transmisión vectorial es escasa ([www.InfoChagas.org](http://www.InfoChagas.org), última consulta septiembre-2020) (Bern, 2015).

La enfermedad de Chagas presenta dos fases bien definidas una aguda y una crónica. La fase aguda suele cursarse con sintomatología inespecífica ya que se presenta con fiebre, malestar generalizado, y en algunos casos puede exhibirse un nódulo cutáneo o chagoma en el lugar de la inoculación o sobre el párpado superior. Sin bien durante esta fase es evidente los niveles de parasitemia, como así también la linfocitosis atípica, la falta de acceso a salud que presentan estas regiones dificulta el diagnóstico y tratamiento temprano de la patología. Sorteada esta primera fase la persona infectada se recupera del malestar, y aunque la parasitemia suele reducirse a niveles indetectables entre las 4 y 8 semanas posteriores al contagio, la infección no desaparece (Bellera et al., 2015; Bern, 2015; Sales Junior et al., 2017).

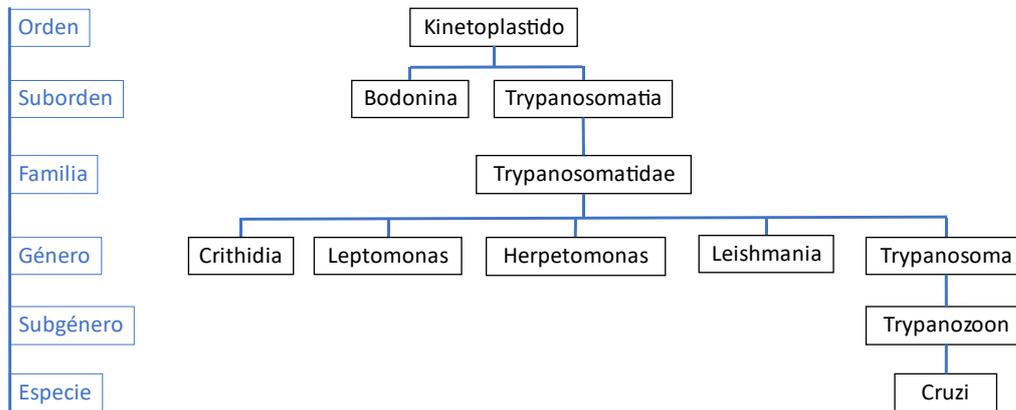
La fase crónica de la parasitosis podría separarse en dos etapas, la primera de ellas asintomática y sin complicaciones para el paciente, por lo que la persona infectada puede vivir años o incluso décadas sin presentar manifestaciones clínicas. La segunda etapa suele presentarse en el 30 a 40% de las personas infectadas e incluye manifestaciones clínicas graves, siendo la más frecuente las cardiopatías chagásicas que derivan en insuficiencia cardíaca (Bern, 2015; Sales Junior et al., 2017).

El complejo ciclo de vida del parásito, en conjunto con las condiciones socioeconómicas en la que se encuentran las poblaciones más afectadas, y la falta de inversión en soluciones terapéuticas por parte de la industria farmacéutica han derivado en que, a más de 100 años de ser descripta, hoy solo se disponga de dos fármacos para el tratamiento de la enfermedad de Chagas (Duschak, 2016).

### 1.2.2 *Trypanosoma cruzi*

*T. cruzi* es un parásito unicelular que afecta principalmente a los vertebrados. Pertenece al reino protista y dentro de este al orden de los kinetoplástidos, los cuales se caracterizan por ser organismos sencillos que presentan un núcleo, uno o dos flagelos y una

única mitocondria localizada en la base del flagelo. Dentro de la mitocondria se localiza el kinetoplasto, el cual consiste en múltiples copias de ADN circular del genoma mitocondrial. Dentro de este orden se encuentra la familia de los tripanosomátidos, a la cual pertenecen los géneros *Leishmania* y *Trypanosoma*. En la **figura 1.1** se muestra el árbol taxonómico de *T. cruzi* (Kaufer et al., 2017; Rassi et al., 2012).



**Figura 1.1.** Árbol taxonómico de *Trypanosoma cruzi*.

Dadas las diversas manifestaciones clínicas tanto en fase aguda como en fase crónica, sumadas a la complejidad epidemiológica y la distribución geográfica de la enfermedad, varios autores han sugerido la necesidad de clasificar a *T. cruzi* según el linaje genético de las diferentes cepas. A partir de 1999 comienzan a denominarse Unidad de Tipificación Discreta (UDT, o en inglés **DTU**) y se le asigna un número a cada grupo de cepas que comparten el mismo linaje genético (Guhl, 2013; Zingales, 2011).

Hasta la actualidad se han identificado siete DTU diferentes, de las cuales seis se denominan con la nomenclatura TcI-TcVI, mientras que la séptima, identificada en 2010, se denomina TcBat debido a que la misma fue aislada de murciélagos y no presenta características compatibles con ninguna de las DTU descritas hasta el momento (Calvopina et al., 2020; Lima et al., 2015).

Las DTU se diferencian tanto en el linaje genético como en su distribución geográfica. La más abundante y con mayor distribución sobre el continente americano es la TcI, asociada tanto a ciclos domésticos como selváticos. TcI afecta principalmente la región de América Central, siendo raros los reportes de esta DTU en la zona de América del Sur. Dicha región está principalmente colonizada por TcII la cual, en general, presenta un ciclo doméstico (Calvopina et al., 2020; Magalhães et al., 2015). Las DTU restantes se encuentran marginadas a pequeñas regiones del continente, como es el caso de TcBat que, hasta la fecha, solo fue reportada en murciélagos del Amazonas de Brasil y Panamá (Guhl, 2013; Lima et al., 2015).

Las infecciones provocadas por TcI en general progresan a complicaciones cardíacas y reactivación de la enfermedad, mientras que TcII, TcV y TcVI están asociadas, principalmente, a los trastornos gástricos como megaesófago o megacolon. TcBat aún no ha sido reportada como infectante de humanos (Calvopina et al., 2020; Lima et al., 2015; Magalhães et al., 2015).

### 1.2.3 Ciclo de vida de *T. cruzi*

El parásito presenta un complejo ciclo de vida diferenciado en cuatro estadios, los cuales están influenciados por condiciones de pH, temperatura, y medio en el que se encuentra. Estos estadios son **epimastigote**, **tripomastigote metacíclico**, **amastigote**, y **tripomastigote sanguíneo** (Brener, 1973; Guarner, 2019).

El ciclo de vida del parásito alterna entre el hospedador vertebrado e invertebrado. Comienza cuando el insecto vector ingiere sangre infectada desde un hospedador vertebrado. Los parásitos que sobrevivan a la exposición de las secreciones gástricas del estómago del vector seguirán su recorrido dentro del tracto gastrointestinal (**TGI**); al llegar a la zona media del intestino del insecto se diferenciarán a **epimastigote** replicativos. En este estadio los parásitos migran hacia la porción distal del intestino donde, adhiriéndose a las células endoteliales, mudan de epimastigotes a **tripomastigote metacíclico**, proceso conocido como meta-ciclogénesis. Esta forma no replicativa es excretada durante la defecación del vector. Cuando las heces del vector entran en contacto con la mucosa oral, ocular o nasal o con heridas de hospedadores vertebrados los parásitos pueden invadir diferentes células o tejidos. Una vez dentro de las células de los mamíferos los tripomastigote metacíclico se diferencian a **amastigote** que se replican mediante fisión binaria. Posteriormente y por diversos factores los parásitos mudan a **tripomastigote** (tripanogénesis), los cuales se liberan al torrente sanguíneo o mediante lisis celular. En este estadio pueden invadir nuevos tejidos, produciendo nuevamente el ciclo de amastigote. Cuando el mamífero infectado es picado por una nueva vinchuca no infectada se reinicia el ciclo. **Figura 1.2** (Mosquillo et al., 2020; M. L. Sbaraglini et al., 2016; Tanowitz et al., 2009). Actualmente algunos autores han sugerido que los epimastigote recientemente diferenciados también podría tener capacidad infectiva frente a células mamíferas (Kessler et al., 2017).

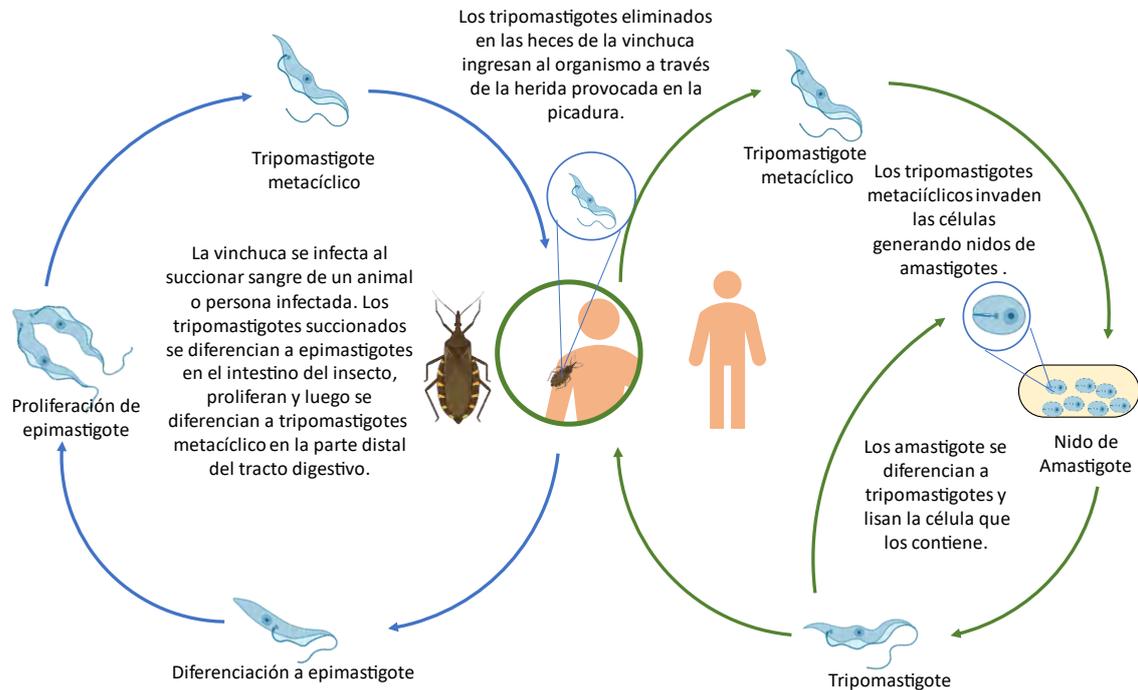


Figura 1.2. Esquema representativo del ciclo de vida de *T. cruzi*.

### 1.2.4 Vías de transmisión

Los primeros indicios de transmisión de *T. cruzi* a humanos hacían referencia principalmente a la transmisión vectorial, es decir el contagio a través de la picadura de una vinchuca infectada. Sin embargo, el avance sobre el conocimiento de la enfermedad, la biología del parásito y la del vector, hizo posible establecer otros modos de contagio que en la actualidad se clasifican en dos grandes grupos: vías de transmisión primarias y vías de transmisión secundarias. Se denominan vías primarias a aquellas que ocurren con mayor frecuencia, entre las que predominan la transmisión vectorial, la transmisión mediada por transfusiones sanguíneas, transmisión por alimento o bebidas contaminadas, y transmisión congénita o transplacentaria. Las vías secundarias, por otra parte, comprenden el trasplante de órganos, accidentes de laboratorio, ingesta de carnes crudas o manipulación de animales infectados, y suelen ocurrir con menor frecuencia (Coura, 2014; Tanowitz et al., 2009). Actualmente se ha demostrado la transmisión sexual en ratones, pero aún no hay reportes de infecciones mediante esta vía en humanos (Coura, 2014).

A pesar de los programas de control de vectores en la región del cono sur de América la transmisión vectorial sigue considerándose prioritaria, esto podría ser consecuencia de la gran cantidad de mamíferos, aves e insectos de los que se pueden alimentar los *triatomas*. Esta diversidad de fuentes de alimentos del vector, y el entorno agreste en el que desarrollan estas poblaciones, donde las falencias edilicias son habituales (casa de paja y barro) en conjunto con el acceso de los animales domésticos al interior de las viviendas, generan un entorno propicio para el contagio vectorial (de Fuentes-Vicente et al., 2018; Leite-Dias et al., 2016).

En el interior de los domicilios las vinchucas suelen resguardarse en lugares húmedos y oscuros como detrás de cuadros y muebles, o en grietas de la edificación. Al caer la noche los insectos salen en busca de alimento, picando a los habitantes del lugar y depositando sus heces mientras se alimentan. Dada la pequeña longitud del TGI del insecto, los tripomastigotes presentes en las heces pueden acceder directamente a la herida o ser arrastrados hacia la misma durante el proceso de rascado (de Fuentes-Vicente et al., 2018; Leite-Dias et al., 2016).

Las vías de transmisión que involucran las transfusiones de sangre y trasplante de órganos se relacionan principalmente con las migraciones de personas infectadas a regiones no endémicas, donde los bancos de órganos y sangre no realizaban, por desconocimiento de la patología, análisis para detección de *T. cruzi*, y en consecuencia los pacientes trasplantados o transfundidos recibían sangre u órganos de personas infectadas (Bern, 2015; Guarner, 2019).

Por otra parte, las infecciones mediante alimentos y bebidas contaminadas, la ingesta de carne mal cocida o la manipulación de animales infectados suelen vincularse con entornos de bajos recursos, como son las zonas endémicas; mientras que los accidentes laborales se registran en general en los profesionales de la salud (bioquímicos o técnicos en hemoterapia principalmente) como así también en investigadores (Guarner, 2019).

### 1.2.5 Epidemiología y distribución de la enfermedad

A principios de 1900 Carlos Chagas describió la enfermedad en la región de Minas Gerais - Brasil; y hacia 1934-1935 Salvador Mazza confirmó la presencia de pacientes infectados con *T. cruzi* en la provincia argentina de Santiago del Estero. Dado el bajo flujo de información con el que se contaba en ese entonces, se consideró por mucho tiempo que la enfermedad de Chagas estaba confinada a América del Sur. Pero pasado apenas medio siglo de su descubrimiento ya se había demostrado que la misma se encontraba presente en más de 20 países de Latinoamérica.

La distribución epidemiológica de la enfermedad, en un primer momento era netamente rural, lo cual comenzó a cambiar rápidamente hacia fines del siglo XX. Este cambio epidemiológico se produce en un contexto complejo ya que el mismo involucra al menos dos variables principales, por un lado, los factores socioeconómicos, políticos y ambientales que han generado migraciones, voluntarias o forzadas, de poblaciones rurales hacia regiones de mayor urbanización; y el segundo factor está vinculado con el ciclo de vida doméstico que presenta el parásito. Este nuevo perfil epidemiológico ya no es solo un tema latinoamericano y rural, sino que plantea una problemática a nivel mundial (*CHAGAS Una problemática vigente una deuda pendiente*, s. f.; Bern, 2015).

Actualmente la OMS ha estimado que existen más de 6 millones de casos activos de tripanosomiasis americana en el mundo. Sin embargo, algunos autores sugieren que estas cifras podrían estar subestimadas debido a las migraciones de las últimas décadas desde regiones endémicas hacia países no endémicos que no contaban con la instrucción e insumos de detección para la enfermedad, siendo posible que muchos contagios producidos mediante transfusiones de sangre o trasplantes de órganos no se hayan reportado como casos activos en su debido momento (Meymandi et al., 2017; [www.who.int](http://www.who.int), última consulta septiembre - 2020). Adicionalmente diferentes organizaciones vinculadas a la salud, entre las que se

destacan la OMS y la Organización Panamericana de la Salud (**OPS**), han alertado que más de 70 millones de personas en todo el mundo se encuentran en riesgo de contraer esta patología (*CHAGAS Una problemática vigente una deuda pendiente*, s. f.; [www.who.int](http://www.who.int), última consulta septiembre - 2020).

De los cuatro países con mayor cantidad de casos activos de *T. cruzi*, Argentina ocupa el primer puesto con más de 1.500.000 infectados, esto representa alrededor del 4% de la población local. En segundo lugar, se encuentra Brasil con más de 1.150.000 casos (~1%) luego lo sigue México con 870.000 (~1%), y en el último lugar se encuentra Bolivia con algo más de 600.000 casos. Sin embargo, este último es el país con mayor incidencia de caso, ya que este número representa más del 6% de la población boliviana (Antinori et al., 2017).

### 1.2.6 Medicamentos disponibles

Actualmente no existen vacunas ni farmacología profiláctica para la enfermedad. El tratamiento farmacológico se sustenta únicamente en dos drogas que datan de la década de los 60/70, **Nifurtimox (NFX)** y **Benznidazol (BZN)** (figura 1.3). Es decir que a la fecha acumulamos más de medio siglo sin innovación terapéutica para esta parasitosis. Si bien estos fármacos son bien tolerados en niños y presentan elevada tasa de curación durante la fase aguda, no puede dejar de mencionarse la baja tasa de curación en la fase crónica de la infección (10-20%), la alta incidencia de efectos adversos en pacientes adultos (>40%), y el prolongado tiempo de tratamiento (60-120 días según el fármaco) como principales limitaciones de las opciones terapéuticas disponibles. Adicionalmente se ha sugerido que el éxito de la farmacoterapia está estrechamente vinculado a la DTU de *T. cruzi* con la que se haya infectado el paciente, y que en las infecciones mixtas disminuye la tasa de curación (Kryshchshyn et al., 2014; Sbaraglini M. et al., 2016).

En 1967 el laboratorio Bayer introdujo al mercado el primer fármaco tripanocida, **NFX** (N-(3-metil-1,1-dioxo-1,4-tiazinan-4-il)-1-(5-nitro-2-furil) metanimina). El mecanismo tripanocida de este agente químico aún no es plenamente conocido. Inicialmente se creía que su efecto estaba ligado a la formación de radicales libres y peróxidos que inducían el estrés oxidativo en el parásito; diversos autores han sugerido recientemente que la muerte celular es producto de reacciones mediadas por nitro-reductasas tripanosomales (Chao et al., 2020; Ribeiro et al., 2020)

Este medicamento se administra por vía oral y, si bien la posología se ajusta según la región y criterios médicos, el esquema inicial es un tratamiento de 90 a 120 días con dosis que varían entre 8 y 10 mg/kg, repartidos generalmente en tres tomas diarias (Antinori et al., 2017).

Actualmente el uso del fármaco ha sido discontinuado en buena parte de la región de América del Sur (Argentina, Brasil, Chile y Uruguay); esta medida está fundada en parte en la falla clínica (la tasa de curación en la fase crónica ronda apenas un 10 %) y, sobre todo, en la falta de seguridad del tratamiento. Si bien los niños toleran bien el NFX (solo 4% de reacciones adversas), en pacientes adultos los reportes de reacciones adversas se observan en más del 40%, siendo algunos de estos eventos moderados o graves (Crespillo-Andújar et al., 2018; Ribeiro et al., 2020). A pesar de no ser el fármaco de primera elección es importante contar

con datos clínicos y farmacológico relevantes, ya que es la única alternativa a emplear cuando fracasa el tratamiento con BNZ. En este aspecto el estudio CHICO<sup>1</sup>, finalizado recientemente, intenta clarificar fenómenos de seguridad, eficacia y parámetros farmacocinéticos en niños y adolescentes, mediante un estudio aleatorizado y cuádruple ciego de una formulación pediátrica de NFX. Se estima que los resultados estarán completos para agosto 2021 (Clinicaltrials.gov: NCT02625974, última consulta febrero-2021).

**BZN** (N-bencil-2-nitroimidazol-1-acetamida) es un compuesto químico del grupo de los nitroimidazoles. Ejerce su efecto modificando las macromoléculas mediante uniones covalentes de los compuestos generados durante las reacciones de nitrorreducción (Rajão et al., 2014; Sbaraglini M. et al., 2016). También se ha reportado inducción del estrés oxidativo en el parásito, lo que contribuiría al efecto tripanocida (Rajão et al., 2014).

El fármaco se administra por vía oral, en dosis de 5 a 10 mg/Kg/día y por un periodo de 30 a 60 días. Dadas las limitaciones comerciales de este medicamento (falta de diversificación de dosis o presentaciones), el esquema posológico puede ser modificado según la disponibilidad de la droga en el mercado (Morillo et al., 2015).

La eficacia y seguridad resultaba un aspecto poco claro del BZN, ya que hasta la publicación del estudio BENEFIT<sup>2</sup> se contaba con poca información sobre estudios observacionales donde se expusiera con claridad los resultados obtenidos en cuanto a tasa de curación, o frecuencia de efectos adversos (Pinheiro et al., 2017). Si bien algunos estudios reportaban tasas de curaciones superiores al 70% durante la fase aguda de la enfermedad, no se registraban los mismos resultados para fases indeterminada o crónica de la infección, ya que los datos reportados indican que la tasa de curación varía entre el 8 y el 40% (Sales Junior et al., 2017).

BENEFIT fue un estudio aleatorizado, y doble ciego en el que se incluyeron pacientes con cardiopatías derivadas de la enfermedad de Chagas, separados en dos grupos (tratados y placebo). El grupo tratado reportó alta incidencia de efectos adversos, entre los que se destacaban malestares gastrointestinales, rash cutáneo y alteraciones del sistema nervioso, por lo que muchos pacientes abandonaron el tratamiento de manera temprana. Además, no se ha demostrado que los pacientes tratados tengan mejor pronóstico respecto a las alteraciones cardíacas. Si bien es cierto que BZN presenta buenos resultado en las fases iniciales de la infección, no tendría efecto curativo en pacientes que han evolucionado a fase crónica sintomática (Morillo et al., 2015).

---

<sup>1</sup> Prospective Study of a Pediatric Nifurtimox Formulation for Chagas` Disease

<sup>2</sup> Benznidazole Evaluation for Interrupting Trypanosomiasis

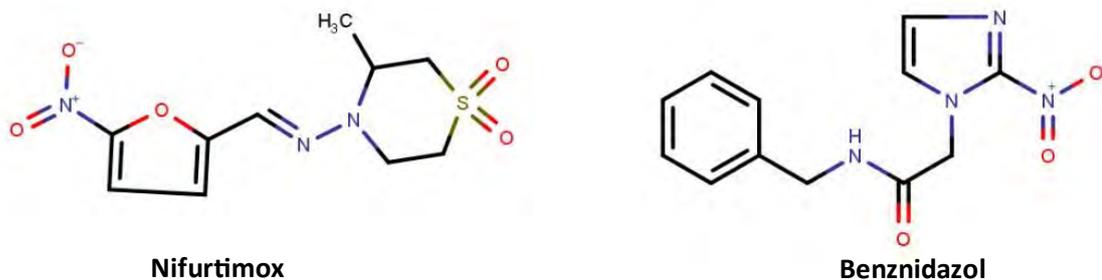


Figura 1.3 Estructura química de los fármacos utilizados en el tratamiento de la Enfermedad de Chagas.

Ninguno de estos fármacos ha sido autorizado para su comercialización en la Unión Europea; sin embargo, se pueden adquirir a través de la OMS (Antinori et al., 2017). BZN es el fármaco de primera línea para tratar la enfermedad de Chagas dado que presenta mejor perfil de seguridad, menor tiempo de tratamiento, y en principio se sospecha que posee una mejor eficacia que NFX. Además, estas características han generado mayor experiencia en el uso y manejo de BNZ frente a NFX. A pesar de ello, este último se utiliza en pacientes que no toleran adecuadamente el BZN (Crespillo-Andújar et al., 2018).

Dado el contexto socioeconómico y cultural en el que se encuentran inmersas las poblaciones más vulnerables, es evidente que la farmacoterapia no es una meta meramente clínica en la que solo priman las consignas de eficacia y seguridad. Muy por el contrario, el desarrollo de un nuevo fármaco requiere de un abordaje más complejo donde factores económicos y sociales deben ser contemplados. Es decir, que el medicamento debe ser accesible desde el punto de vista económico, pero también debe resultar de fácil conservación y almacenamiento, requiriendo, además, que el tiempo de tratamiento sea lo más corto posible y que garantice la cura de la infección en cualquiera de sus estadios (Bellera et al., 2019; Sbaraglini M.L. et al., 2016).

### 1.2.7 Perspectivas hacia nuevas farmacoterapias para la Enfermedad de Chagas

El descubrimiento de nuevas farmacoterapias es una tarea compleja, ya que debe contemplar la mejora en los perfiles de eficacia, seguridad y/o farmacocinéticos con respecto a los tratamientos disponibles hasta el momento. En este sentido, Planer y col. publicaron en 2014 un estudio enfocado en la búsqueda de nuevos agentes tripanocidas. En el trabajo se evaluaron 2000 drogas que reportaban datos de actividad biológica y se realizaron análisis de sinergismo de estas drogas con BNZ y, en algunos casos, entre sí; sin embargo, salvo contadas excepciones, ninguna de ellas ha superado los parámetros de eficacia para justificar evaluaciones clínicas (Planer et al., 2014).

La mayoría de los compuestos que han sido evaluado en fases más avanzadas corresponden al grupo de los antifúngicos azólicos, los cuales interfieren en la vía de síntesis del ergosterol. Si bien el grupo farmacológico es amplio, hasta el momento solo dos fármacos (**Posaconazol** y **Ravuconazol**) han sido evaluados en estudios clínicos avanzados de fase III; mientras que otras drogas del grupo tales como **Itraconazol**, **Ketoconazol**, **Fexinidazol** y

**Fluconazol** se encuentran aún en fases tempranas (Sales Junior et al., 2017; Sbaraglini M. et al., 2016).

Algunos de estos fármacos han sido evaluados en modelos celulares y animales, como el caso de Ketoconazol, o Itraconazol. Ketoconazol fue administrado en combinación con BNZ, y ha demostrado ser eficaz en modelos murinos de la enfermedad de Chagas, sin embargo, aún no se han reportados datos de eficacia en pacientes (Morillo et al., 2017). Itraconazol fue evaluado de manera independiente y combinado con BNZ y con Amiodarona (otro fármaco con promisorio actividad tripanocida) en ensayos independientes; sin embargo los resultados son poco claros, debido a las divergencias en los porcentajes de negativización (Cunha et al., 2019; Sass et al., 2019). Diversos estudios preclínicos con Itraconazol mostraron tasas de curación parasitaria entre 8 y 53% (Sass et al., 2019), sin embargo, hasta el momento no se registran estudios clínicos para esta droga frente a la enfermedad de Chagas (ClinicalTrials.gov ultima consulta febrero 2021).

Para Fexinidazol (probado recientemente como terapia oral para la tripanosomiasis africana) se inició en 2017 un estudio multicéntrico de fase II en España. Se estima que los resultados podrían estar disponibles a partir de 2020; de lograr superar exitosamente las fases II y III de los ensayos clínicos este tratamiento podría convertirse en la primera nueva terapia para la enfermedad de Chagas en más de medio siglo (www.dndi.org. última consulta septiembre - 2020; ClinicalTrials.gov: NCT03587766 ultima consulta febrero - 2021). Cabe mencionar que en la plataforma ClinicalTrials.gov se registra otro estudio sobre Chagas para esta misma droga, pero el mismo no ha sido actualizado desde julio 2015, aun cuando el estudio debió concluir en septiembre de ese mismo año (ClinicalTrials.gov: NCT02498782, última consulta febrero-2021).

Respecto a Posaconazol, se realizó un estudio clínico para evaluar su eficacia; el mismo incluyó 120 participantes con infección crónica y se llevó a cabo en 19 centros de salud de seis países. A cada paciente se le administró BNZ, Posaconazol, combinación de ambas drogas, o placebo vía oral por un lapso de 60 días; similar a un tratamiento estándar con BNZ. Desafortunadamente los resultados no mostraron mejoras en la eficacia del tratamiento para Posaconazol o su combinación con BNZ respecto al tratamiento estándar de BNZ (Morillo et al., 2017). Estos resultados son consistentes con datos publicados con anterioridad por el estudio clínico CHAGAZOL, en el cual participaron 79 pacientes divididos en tres grupos, BZN, Posaconazol baja dosis y Posaconazol alta dosis. Los resultados de este estudio mostraron una mayor incidencia de fracaso en los tratamientos realizados con Posaconazol respecto a BNZ (Molina et al., 2014).

Ravuconazol fue evaluado en modelos murinos y posteriormente en perros. Diniz y Col realizaron un estudio para evaluar la eficacia de esta droga en perros infectados con *T. cruzi* resistente a los tratamientos convencionales (BNZ y NFX), Si bien en la fase temprana del estudio se observó reducción de la parasitemia en perros, los niveles de inmunoglobulinas (Ig) IgG1 y IgG2 volvieron a elevarse luego de finalizado el tratamiento, lo que sugiere la presencia de parásitos en diferentes tejidos celulares (Diniz et al., 2010). En el año 2018 Torrico y col. realizaron un estudio clínico aleatorizado y doble ciego con el fin de evaluar la eficacia y seguridad del compuesto E-1224, un profármaco de Ravuconazol. El estudio incluyó a 231 pacientes con confirmación de infección por *T. cruzi*. Los pacientes se dividieron en cinco

grupos: Placebo, BNZ, y 3 tratamientos con E-1224; alta dosis (4000 mg en 8 semanas), baja dosis (2000 mg en 8 semanas) y dosis media (2500 mg en 4 semanas y luego 4 semanas de placebo). Si bien la seguridad fue aceptable, los resultados obtenidos son similares que para el resto de los azoles, donde se observa una disminución de la parasitemia inicial pero que no se mantiene a lo largo del tiempo. Por lo tanto, los autores no han profundizado en estudios de monoterapia para este compuesto (Torrico et al., 2018).

La Amiodarona es un antiarrítmico clase 3. Diversas observaciones clínicas han sugerido que pacientes con cardiopatías chagásicas tratados con este fármaco presentaban mejores resultados de eficacia en comparación con los pacientes que manifestaban las mismas cardiopatías, pero tratados con otros antiarrítmicos (Sbaraglini M. et al., 2016). Estas observaciones dieron origen a diversos estudios con el fin de evaluar el efecto de Amiodarona sobre la viabilidad de *T. cruzi*, y varios autores han demostrado que el fármaco interrumpe el ciclo de vida del parásito reduciendo el recuento de amastigote e inhibiendo luego la diferenciación a tripomastigote. Algunas conclusiones a la que han arribado son que la droga genera desregulación del calcio intracelular, altera la morfología de la célula y de estructuras internas, e interfiere en la vía de síntesis del ergosterol (Lourenço et al., 2018; Sbaraglini M. et al., 2016). Adicionalmente se realizaron estudios frente a modelos murinos observándose disminución de la parasitemia y mayor tiempo de sobrevida animal (Lourenço et al., 2018). Actualmente están concluyendo dos estudios clínicos para evaluar el efecto de Amiodarona en pacientes con cardiopatías chagásicas, y se estimaba que los primeros resultados estarían disponibles para fines de 2020 (ClinicalTrials.gov: NCT03193749 y NCT01722942 última consulta febrero-2021).

En esta sección presentamos algunos de los fármacos que se encuentran en estudios de fases clínicas, ya que si bien muchas otras drogas han sido reportadas como tripanocidas, varias de ellas han sido descartadas debido a fallas de seguridad o eficacia a nivel preclínico, o por presentar un perfil ADME inadecuado.

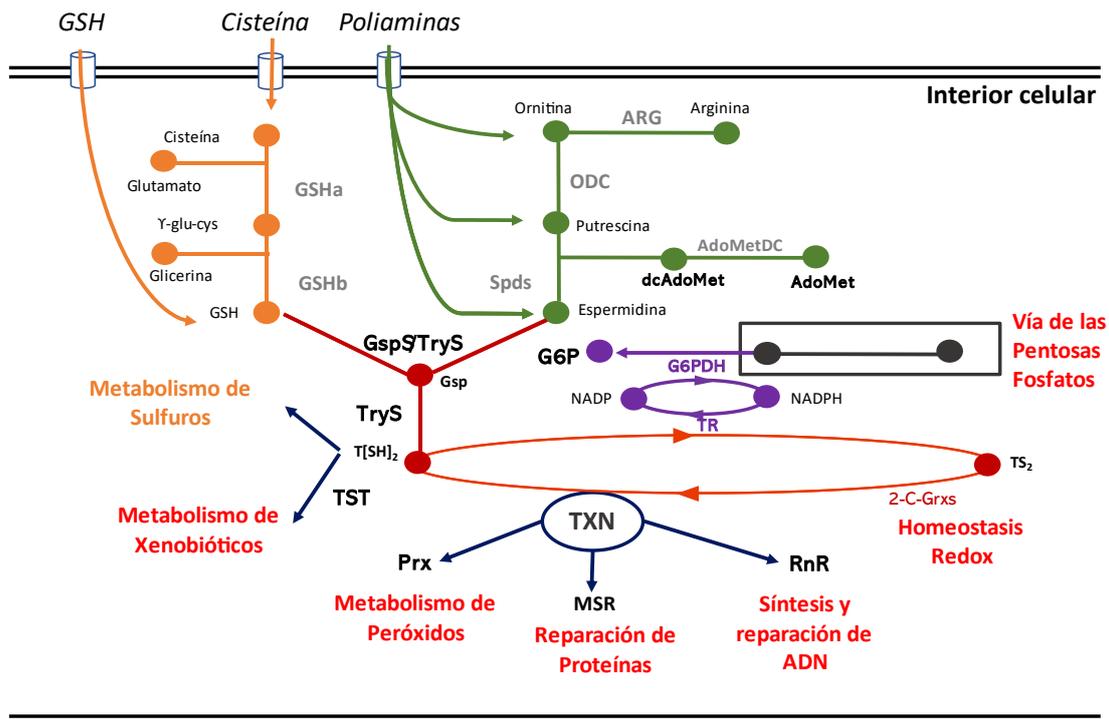
### 1.3 Procesos redox dependientes de Tripanotión

Todas las células, requieren para su normal desarrollo determinadas condiciones de estabilidad química, física y biológica. Las concentraciones iónicas, el rango de pH, las condiciones físicas y biológicas a la que se exponen las células condicionan sus funciones fisiológicas. En consecuencia, estos factores son determinantes en la viabilidad celular.

Dentro de los factores que regulan el funcionamiento celular cabe destacar el rol clave del potencial redox. Bien se sabe que la homeostasis redox es un equilibrio dinámico entre la formación y eliminación de las especies oxidantes derivadas, principalmente, de Oxígeno y de Nitrógeno y que, de no ser rigurosamente controladas, las concentraciones elevadas de estas especies generan daños celulares a nivel de ácidos nucleicos, proteínas y lípidos de membranas, lo que promueve la muerte o degeneración celular. Es por ello que todas las células presentan un exhaustivo control del metabolismo redox (Basria S.M.N. Mydin & I. Okekpa, 2019).

Como puede observarse en la **figura 1.4** la homeostasis redox de los tripanosomátidos esta exhaustivamente regulada por un complejo sistema de enzimas y segundos mensajeros que protegen al parásito del estrés oxidativo, sobre todo en etapas de infección donde a los subproductos del propio metabolismo se suman oxidantes proveniente de las defensas del huésped (Comini & Flohé, 2013). Sin embargo, a pesar de este sofisticado mecanismo, la homeostasis redox queda sujeta casi de manera exclusiva a cuatro tioles de bajo peso molecular (**Glutatión (GSH)**, **Glutathionespermidina (Gsp)**, **Tripanotión (T[SH]<sub>2</sub>)**, u **Ovotiol A**), los cuales son sintetizados según la especie y los requerimientos de l estadio del ciclo de vida de cada tripanosomátido (Comini et al., 2009; Krauth-Siegel & Comini, 2008).

Se conoce que la reactividad de los tioles aumenta cuando su pKa es próximo al pH del medio, y dado que T[SH]<sub>2</sub> presenta un pka =7,4, el cual coincide con el pH fisiológico de las células mamíferas, varios autores sugieren que el T[SH]<sub>2</sub> es el tiol con mayor peso en la defensa y mantenimiento del estado redox de los parásitos (Krauth-Siegel & Comini, 2008; Piacenza et al., 2013).



**Figura 1.4** La figura muestra la importancia del Tripanotión frente al metabolismo redox, y por lo tanto el rol clave de la enzima **TryS** en el parásito. Adaptada de Comini y Flohé 2013.

Las nuevas herramientas genéticas han clarificado el rol que diversas enzimas desempeñan frente a la homeostasis redox. El hallazgo más trascendente es la confirmación de la ausencia, en tripanosomátidos, de reductasas esenciales para eucariotas y procariontes y la importancia del T[SH]<sub>2</sub> para un adecuado metabolismo redox (ver **figura 1.4**). En los últimos años el estudio detallado de las vías de detoxificación redox en tripanosomátidos ha revelado una amplia ramificación de la vía dependiente del T[SH]<sub>2</sub> donde se observaron enzimas capaces de reducir ribonucleótidos, y metabolismo específico para grupos de azufre-hierro o

reducciones de sulfóxidos de metionina. Estos conocimientos permiten, en su conjunto, un enfoque racional hacia la búsqueda nuevos tratamientos farmacológicos (Comini & Flohé, 2013). La búsqueda se dirige sobre moléculas que puedan intervenir en el metabolismo del T[SH]<sub>2</sub>, y en consecuencia alterar las vías secundarias a este tiol.

### 1.4 Tripanotión sintetasa

#### 1.4.1 La enzima

Como se mencionó anteriormente los parásitos de los géneros *Trypanosoma* y *Leishmania* son altamente dependientes del metabolismo oxidativo regulado por el T[SH]<sub>2</sub>, un tiol de bajo peso molecular compuesto por dos moléculas de GSH interconectadas por una molécula de Espermidina; este compuesto es sintetizado por la enzima **Tripanotión Sintetasa (TryS)** (Benítez et al., 2016; Medeiros et al., 2020).

La enzima **TryS** es una flavoproteína monomérica con una masa molecular que oscila entre 72 y 74 KDa aproximadamente según la especie de tripanosomátidos/*Leishmania* (Fyfe et al., 2008; Oza et al., 2002), y que posee una longitud aproximada de 650 aminoácidos (aa), (648 aa reportados para *T. cruzi* y 652 aa para *Leishmania major*). Esta proteína es codificada a partir de un gen de copia única (Torrie et al., 2009).

La enzima, bifuncional y dependiente de ATP, cataliza de manera gradual la síntesis de tripanotión por el extremo C-terminal, el cual comienza con la síntesis del Gsp y luego incorporando una segunda molécula de GSH finaliza la reacción sintética con la producción de T[SH]<sub>2</sub>. En el otro extremo, el dominio N-terminal presenta actividad amidasa con capacidad de hidrolizar el producto final en los sustratos iniciales (GSH y Espermidina) (Comini et al., 2009; Krauth-Siegel & Comini, 2008; Sousa et al., 2014). Adicionalmente se ha demostrado que en los patógenos *T. brucei*, *T. cruzi* y algunas cepas de *Leishmania* la síntesis de T[SH]<sub>2</sub> esta medida expresamente por la **TryS**, por lo que la enzima resulta esencial para Tripanosomas y *Leishmania* (Comini et al., 2009). En la **figura 1.5** se esquematiza la síntesis de T[SH]<sub>2</sub>.

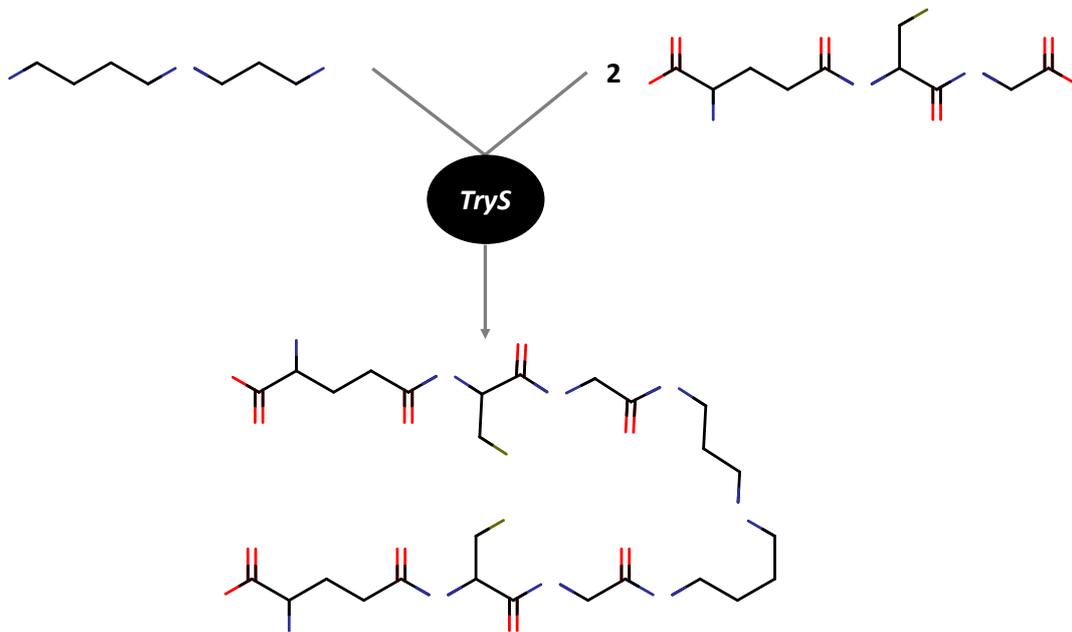


Figura 1.5. Síntesis de Tripanotión.

Se ha demostrado con claridad que la actividad sintética de la enzima es crucial en el ciclo de vida de los tripanosomátidos, y que la inhibición de esta tiene un fuerte impacto negativo tanto en la replicación como en la viabilidad (Mesías et al., 2019). Sin embargo, aún queda por comprender la actividad amidasa de la *TryS*. Esta actividad hidrolítica podría resultar de gran utilidad para la obtención y regulación de los niveles intracelulares de poliaminas, sobre todo en *T. cruzi*, que resulta auxótrofo para estos compuestos y, por lo tanto, su única posibilidad es captarlos del medio extracelular mediante transportadores específicos. En consecuencia, la hidrólisis del tripanotión podría ser crucial para la supervivencia de este parásito cuando se encuentre en presencia de medios extracelulares con bajas concentraciones de poliaminas (Oza et al., 2002).

Diferente es la obtención de poliaminas (Espermidina, Putrescina, Cadaverina, etc.) en *T. brucei* ya que carece de transportadores y por lo tanto es enteramente dependiente de la de las síntesis de *novo*, mientras que *Leishmania spp.* es capaz de obtener Espermidina a partir de Ornitina o Arginina tomada del medio (Comini & Flohé, 2013; Manta et al., 2013). En consecuencia, en estos tripanosomátidos la actividad amidasa sería menos relevante que para *T. cruzi*.

Estudios realizados con mutantes nulos para la enzima Ornitina Decarboxilasa han sido útiles para inferir que la homeostasis de poliaminas es un punto crítico para el desempeño de varias funciones celulares, ya que la ausencia de esta enzima en *L. donovani* y *T. brucei* deriva en la muerte o inviabilidad de los parásitos (Jiang et al., 2007; Talevi et al., 2019).

### 1.4.2 *TryS* como blanco molecular

El conocimiento de la biología y bioquímica del parásito ha dado lugar al abordaje de tratamientos farmacológicos enfocados en blancos moleculares (*targets*) específicos. Estos blancos son elegidos como tales evaluando si los mismos cumplen un conjunto de

recomendaciones que se basan en los conceptos de **esencialidad**, **ser modulable farmacológicamente** (*Druggable*, adaptado como drogable al español) y **vulnerabilidad**. Estos términos hacen referencia a que el blanco elegido debe ser esencial para la viabilidad del parásito y factible de ser modulado mediante interacción con pequeñas moléculas tipo-fármaco (Hajduk et al., 2005), mientras que el tercer término, vulnerabilidad, hace referencia a la fracción de copias del blanco molecular que debe ser modulado o alterados para obtener la respuesta farmacológica deseada. Los blancos altamente vulnerables son aquellos que con una pequeña fracción de ocupación del número de copias se observa la respuesta farmacológica deseada (Tonge, 2018). Es importante destacar que un blanco de baja vulnerabilidad requiere una elevada fracción de ocupación, lo que se traduce en altas concentraciones de fármacos cerca del objetivo terapéutico, y en consecuencia de la administración de una dosis elevada. Adicionalmente para las enfermedades infecciosas (bacterianas o parasitarias) se sugiere que el blanco molecular esté presente en el microorganismo, pero no en huésped o, al menos, de estar presente la divergencia sea tal que la modulación sobre el blanco molecular del microorganismo no repercuta significativamente sobre el blanco del huésped, minimizando así los efectos no deseados (Duschak, 2016; Terán et al., 2016).

Dado que la elección de un blanco molecular debería, idealmente, estar atravesada transversalmente por los conceptos de esencialidad, drogabilidad y vulnerabilidad, y que se ha demostrado que para *T. brucei* y *Leishmania* la enzima resulta esencial, dada la secuencia y funciones altamente compartidas entre las distintas especies, podría suponerse el mismo escenario para *T. cruzi* (Talevi et al., 2019). Otro dato relevante es la ausencia de la enzima **TryS** en los mamíferos, organismos hospedadores del parásito, por lo que esta proteína resulta un blanco farmacológico atractivo para hacer frente a las enfermedades de origen tripanosomátidos, ya que la inhibición o modulación de las funciones de la enzima no denotarían un efecto directo sobre el metabolismo del huésped (Medeiros et al., 2020; Talevi et al., 2019).

En cuanto al concepto de vulnerabilidad se puede inferir que la **TryS** resulta un blanco bastante más vulnerable que la enzima **Tripanotión Reductasa (TR)**. Este dato se desprende, de manera indirecta, de varios trabajos en los que la inhibición de la enzima **TR** de *T. brucei* no produce un aumento significativo de la actividad sintetasas, mientras que la inactivación de la **TryS** genera una duplicación en la actividad de la enzima **TR** (Krauth-Siegel & Comini, 2008). Se ha sugerido que los parásitos podrían sobrevivir con tan solo el 10% de la actividad reductasa, pero que su viabilidad se ve más comprometida frente a la intervención sobre **TryS** (Krauth-Siegel & Comini, 2008).

Por otro lado, se ha postulado que la mayoría de las vías metabólicas suelen presentar dos o tres proteínas (enzimas/transportadores) principales, y en consecuencia la intervención farmacológica sobre una de ellas presentará un mayor impacto con respecto a la modulación de una proteína “secundaria” (Saavedra et al., 2019)

El **Análisis de Control Metabólico (ACM, o MCA en inglés)** en conjunto con el modelado cinético permiten estimar qué proteínas presentan mayor control o regulación de una vía metabólica (Marín-Hernández et al., 2020), por lo que resultan de gran utilidad para priorizar el estudio de blancos terapéuticos, sobre todo en aquellos casos donde la mayoría de las dianas propuestas resultan esenciales (Michels et al., 2020; Saavedra et al., 2019). Es

importante recordar que el termino esencial hace referencia al efecto sobre el fenotipo de la supresión (química o genética) más que a su actividad (Michels et al., 2020; Saavedra et al., 2019).

El ACM es una herramienta teórico-experimental que estudia la contribución de cada proteína dentro de una ruta o vía metabólica completa. Este análisis requiere de la evaluación completa del sistema: se debe individualizar cada paso de reacción con el fin de asignarle a cada paso su valor de **coeficiente de flujo (C<sub>J*i*</sub>)**, donde **J** representa el flujo de la vía, y ***a*** es la actividad de la enzima/transportador en la vía ***i***. C<sub>J*i*</sub> puede tomar valores positivos, cuando el paso enzimático contribuye al flujo metabólico, o valores negativos cuando el paso resulta ser una vía de escape para la ruta principal (Zabdi González-Chávez et al., 2019; Michels et al., 2020). Es importante comprender que C<sub>J*i*</sub> es estimado a partir de datos experimentales, y para ello se requieren ensayos *in vitro* que permitan conocer con la mayor precisión posible la actividad y las condiciones de reacción óptimas (pH, concentración de sustratos, temperatura, iones, etc.) para cada una de las enzimas involucradas, como así también el detalle de la vía metabólica completa dentro del organismo en estudio (Michels et al., 2020; Saavedra et al., 2019).

En este sentido, un ACM sobre *T. cruzi* proporciona criterios adicionales para priorizar la selección y estudio de blancos terapéuticos, dado que la mayoría de las estructuras propuestas como dianas farmacológicas, hasta el momento, han resultado esenciales para el parásito (González-Chávez et al., 2015; Saavedra et al., 2019). Adicionalmente se ha demostrado que en la vía de síntesis, utilización y regeneración de T[SH]<sub>2</sub> todas las enzimas involucradas resultan esenciales. Este hallazgo fue realizado mediante manipulación genética de las especies *T. brucei* y *Leishmania*, debido a que existe una gran disponibilidad de datos en librerías de ARNi para ambas especies, sin embargo el ACM y la tecnología CRISPR serían las herramientas más adecuadas para evaluar el rol clave de cada una de las enzimas involucradas en el metabolismo del T[SH]<sub>2</sub> para *T. cruzi*, dado que aún no es posible aplicar técnicas de manipulación genética basadas en ARNi (González-Chávez et al., 2015; González-Chávez et al., 2019; Michels et al., 2020).

La vía de regulación redox de tripanosomátidos ha sido objeto de estudio para varios autores en los últimos años debido a la evidencia que denota como esenciales a todas las enzimas y transportadores involucrados en dicha vía (Saavedra et al., 2019); sin embargo la profundización de los estudios sobre esta ruta metabólica han sugerido que solo las enzimas **TryS** y **gamma-Glutamilcisteína Sintetasa (YECS)** presentan un control metabólico elevado, y en menor medida el transportador de Espermidina (González-Chávez et al., 2015; González-Chávez et al., 2019; Olin-Sandoval et al., 2012; Saavedra et al., 2019). En este sentido los autores reportan valores de C<sub>J*i*</sub> que varían entre 0,49 y 0,58 para **TryS** y entre 0,58 y 0,7 para **YECS**, mientras que dichos coeficientes presentaron valores insignificantes (<0,15) para el resto de las enzimas involucradas, con excepción del transportador de Espermidina para el cual se estimó un C<sub>J*i*</sub> de 0,24 aproximadamente (Olin-Sandoval, 2012; Saavedra et al., 2019).

Con los valores de C<sub>J*i*</sub> reportados se ha sugerido que se debe bloquear más del 90% de la enzima TR para lograr una disminución significativa de la regeneración de T[SH]<sub>2</sub>, mientras que la inhibición del 58% la enzima **YECS**, el 63% de la enzima **TryS** o el 73% del transportador

de Espermidina se traduce en un descenso de más del 50% en la síntesis de *nov* de T[SH]<sub>2</sub> (González-Chávez et al., 2019; Olin-Sandoval et al., 2012).

Es importante mencionar que independientemente de las discrepancias que pudiera haber entre los diferentes ensayos o metodologías de trabajo desarrolladas por los diversos grupos, la conclusión a la que arriban es prácticamente unánime, proponiendo en todo momento que el control metabólico de la vía redox en los tripanosomátidos está regulada principalmente por las enzimas  $\gamma$ ECS y *TryS*, y en menor medida por el transportador Espermidina o por la Espermidina sintetasa, dependiendo de la especie. En consecuencia, estas proteínas deberían captar la atención principal como blancos terapéuticos para la búsqueda de nuevos fármacos tripanocidas (González-Chávez et al., 2019; Olin-Sandoval et al., 2012).

Con el análisis realizado hasta aquí, y teniendo en cuenta que ya se han reportado inhibidores de *TryS* para las diferentes especies de tripanosomátidos (Benítez et al., 2016; Medeiros et al., 2020); postulamos que esta proteína se perfila como un buen blanco terapéutico para el tratamiento farmacológico de la enfermedad de Chagas, como así también otras enfermedades originadas por la familia de los *trypanosomatidae*. En la **figura 1.6** se presentan algunos de los inhibidores reportados.

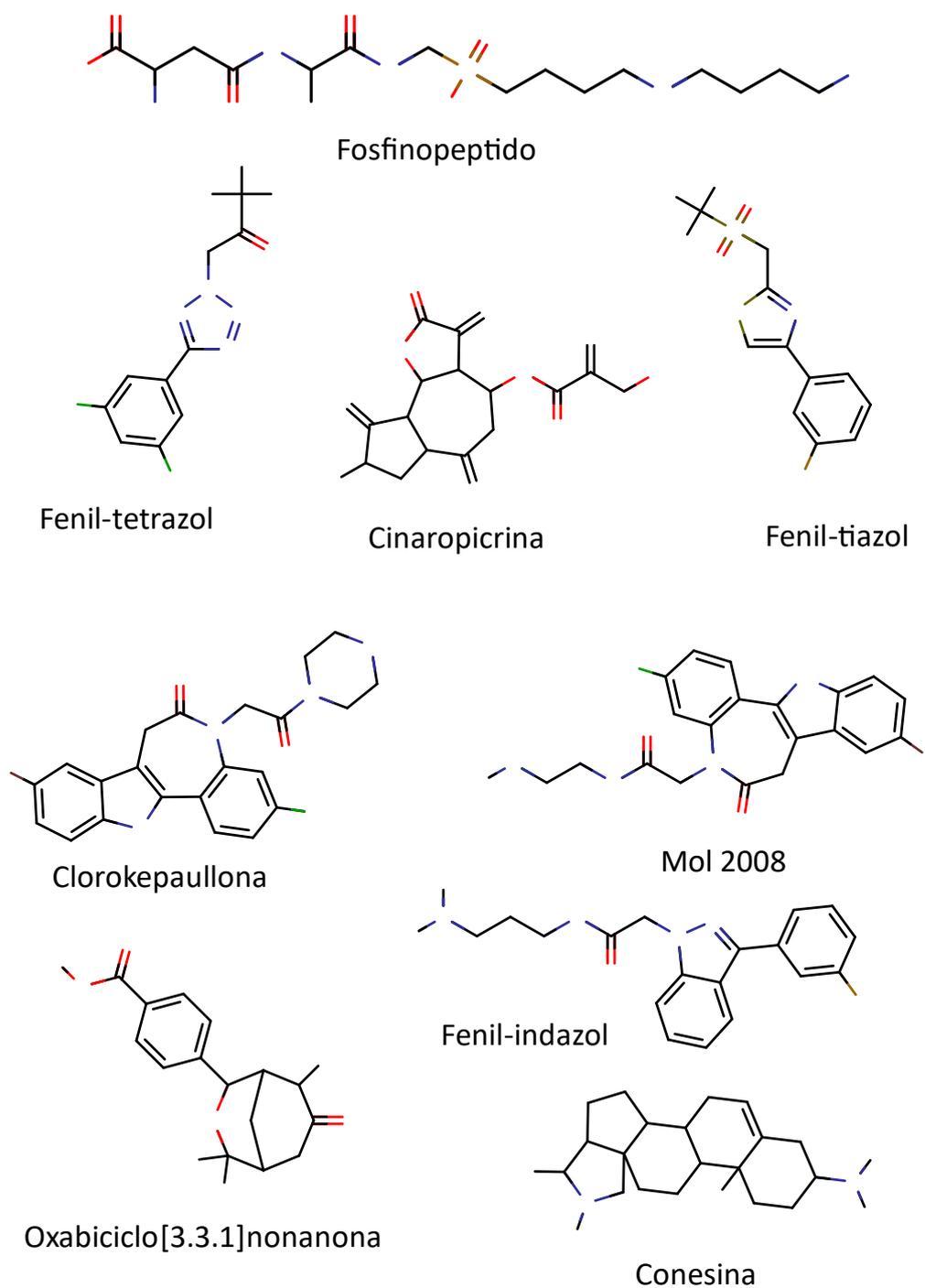


Figura 1.6. Algunos inhibidores de *TryS* reportados hasta el momento (Medeiros et al., 2020).

## 1.5 Referencias Bibliográficas

- Andrade, C. H., Neves, B. J., Melo-Filho, C. C., Rodrigues, J., Silva, D. C., Braga, R. C., & Cravo, P. V. L. (2019). In Silico Chemogenomics Drug Repositioning Strategies for Neglected Tropical Diseases. *Current Medicinal Chemistry*, 26(23), 4355-4379. <https://doi.org/10.2174/0929867325666180309114824>
- Antinori, S., Galimberti, L., Bianco, R., Grande, R., Galli, M., & Corbellino, M. (2017). Chagas disease in Europe: A review for the internist in the globalized world. *European Journal of Internal Medicine*, 43, 6-15. <https://doi.org/10.1016/j.ejim.2017.05.001>
- Aya Pastrana, N., Somerville, C., & Suggs, L. S. (2020). The gender responsiveness of social marketing interventions focused on neglected tropical diseases. *Global Health Action*, 13(1), 1711335. <https://doi.org/10.1080/16549716.2019.1711335>
- Basria S.M.N. Mydin, R., & I. Okekpa, S. (2019). Reactive Oxygen Species, Cellular Redox Homeostasis and Cancer. En F. Lasakosvitsch & S. Dos Anjos Garnes (Eds.), *Homeostasis—An Integrated Vision*. IntechOpen. <https://doi.org/10.5772/intechopen.76096>
- Bellera, C. L., Sbaraglini, M. L., & Talevi, A. (2018). Modern Approaches for the Discovery of Anti-Infectious Drugs for the Treatment of Neglected Diseases. *Current Topics in Medicinal Chemistry*, 18(5), 369-381. <https://doi.org/10.2174/1568026618666180509151146>
- Bellera, C., Sbaraglini, M., Balcazar, D., Fraccaroli, L., Vanrell, M., Casassa, A., Labriola, C., Romano, P., Carrillo, C., & Talevi, A. (2015). High-throughput Drug Repositioning for the Discovery of New Treatments for Chagas Disease. *Mini-Reviews in Medicinal Chemistry*, 15(3), 182-193. <https://doi.org/10.2174/138955751503150312120208>
- Bellera, C. L., Sbaraglini, M. L., Alberca, L. N., Alice, J. I., & Talevi, A. (2019). In Silico Modeling of FDA-Approved Drugs for Discovery of Therapies Against Neglected Diseases: A Drug Repurposing Approach. En *In Silico Drug Design* (pp. 625-648). Elsevier. <https://doi.org/10.1016/B978-0-12-816125-8.00021-3>
- Benítez, D., Medeiros, A., Fiestas, L., Panozzo-Zenere, E. A., Maiwald, F., Prousis, K. C., Roussaki, M., Calogeropoulou, T., Detsi, A., Jaeger, T., Šarlauskas, J., Peterlin Mašič, L., Kunick, C., Labadie, G. R., Flohé, L., & Comini, M. A. (2016). Identification of Novel Chemical Scaffolds Inhibiting Trypanothione Synthetase from Pathogenic Trypanosomatids. *PLOS Neglected Tropical Diseases*, 10(4), e0004617. <https://doi.org/10.1371/journal.pntd.0004617>
- Bern, C. (2015). Chagas' Disease. *New England Journal of Medicine*, 373(5), 456-466. <https://doi.org/10.1056/NEJMra1410150>
- Brener, Z. (1973). Biology of Trypanosoma Cruzi. *Annual Review of Microbiology*, 27(1), 347-382. <https://doi.org/10.1146/annurev.mi.27.100173.002023>
- Calvopina, M., Segovia, G., Cevallos, W., Vicuña, Y., Costales, J. A., & Guevara, A. (2020). Fatal acute Chagas disease by Trypanosoma cruzi DTU TcI, Ecuador. *BMC Infectious Diseases*, 20(1), 143. <https://doi.org/10.1186/s12879-020-4851-0>
- CHAGAS Una problemática vigente una deuda pendiente. (s. f.).
- Chao, C., Leone, J. L., & Vigliano, C. A. (2020). Chagas disease: Historic perspective. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease*, 1866(5), 165689. <https://doi.org/10.1016/j.bbadis.2020.165689>

- Comini, M. A., Dirdjaja, N., Kaschel, M., & Krauth-Siegel, R. L. (2009). Preparative enzymatic synthesis of trypanothione and trypanothione analogues. *International Journal for Parasitology*, *39*(10), 1059-1062. <https://doi.org/10.1016/j.ijpara.2009.05.002>
- Comini, M. A., & Flohé, L. (2013). Trypanothione-Based Redox Metabolism of Trypanosomatids. En T. Jäger, O. Koch, & L. Flohé (Eds.), *Trypanosomatid Diseases* (pp. 167-199). Wiley-VCH Verlag GmbH & Co. KGaA. <https://doi.org/10.1002/9783527670383.ch9>
- Coura, J. R. (2014). The main sceneries of Chagas disease transmission. The vectors, blood and oral transmissions—A comprehensive review. *Memórias Do Instituto Oswaldo Cruz*, *110*(3), 277-282. <https://doi.org/10.1590/0074-0276140362>
- Crespillo-Andújar, C., Chamorro-Tojeiro, S., Norman, F., Monge-Maillo, B., López-Vélez, R., & Pérez-Molina, J. A. (2018). Toxicity of nifurtimox as second-line treatment after benznidazole intolerance in patients with chronic Chagas disease: When available options fail. *Clinical Microbiology and Infection*, *24*(12), 1344.e1-1344.e4. <https://doi.org/10.1016/j.cmi.2018.06.006>
- Cunha, E. L. A., Torchelsen, F. K. V. da S., Cunha, L. M., de Oliveira, M. T., Fonseca, K. da S., Vieira, P. M. A., Carneiro, C. M., & de Lana, M. (2019). Benznidazole, itraconazole and their combination in the treatment of acute experimental chagas disease in dogs. *Experimental Parasitology*, *204*, 107711. <https://doi.org/10.1016/j.exppara.2019.05.005>
- de Fuentes-Vicente, J. A., Gutiérrez-Cabrera, A. E., Flores-Villegas, A. L., Lowenberger, C., Benelli, G., Salazar-Schettino, P. M., & Córdoba-Aguilar, A. (2018). What makes an effective Chagas disease vector? Factors underlying *Trypanosoma cruzi*-triatomine interactions. *Acta Tropica*, *183*, 23-31. <https://doi.org/10.1016/j.actatropica.2018.04.008>
- Diniz, L. de F., Caldas, I. S., Guedes, P. M. da M., Crepalde, G., de Lana, M., Carneiro, C. M., Talvani, A., Urbina, J. A., & Bahia, M. T. (2010). Effects of Ravuconazole Treatment on Parasite Load and Immune Response in Dogs Experimentally Infected with *Trypanosoma cruzi*. *Antimicrobial Agents and Chemotherapy*, *54*(7), 2979-2986. <https://doi.org/10.1128/AAC.01742-09>
- Fyfe, P. K., Oza, S. L., Fairlamb, A. H., & Hunter, W. N. (2008). *Leishmania* Trypanothione Synthetase-Amidase Structure Reveals a Basis for Regulation of Conflicting Synthetic and Hydrolytic Activities. *Journal of Biological Chemistry*, *283*(25), 17672-17680. <https://doi.org/10.1074/jbc.M801850200>
- Duschak, V. (2016). Targets and Patented Drugs for Chemotherapy of Chagas Disease in the Last 15 Years-Period. *Recent Patents on Anti-Infective Drug Discovery*, *11*(2), 74-173. <https://doi.org/10.2174/1574891X11666161024165304>
- González-Chávez, Z., Olin-Sandoval, V., Rodríguez-Zavala, J. S., Moreno-Sánchez, R., & Saavedra, E. (2015). Metabolic control analysis of the *Trypanosoma cruzi* peroxide detoxification pathway identifies tryparedoxin as a suitable drug target. *Biochimica et Biophysica Acta - General Subjects*, *1850*(2), 263-273. Scopus. <https://doi.org/10.1016/j.bbagen.2014.10.029>
- González-Chávez, Z., Vázquez, C., Mejía-Tlachi, M., Márquez-Dueñas, C., Manning-Cela, R., Encalada, R., Rodríguez-Enríquez, S., Michels, P. A. M., Moreno-Sánchez, R., & Saavedra, E. (2019). Gamma-glutamylcysteine synthetase and tryparedoxin 1 exert high control on the antioxidant system in *Trypanosoma cruzi* contributing to drug resistance and infectivity. *Redox Biology*, *26*, 101231. <https://doi.org/10.1016/j.redox.2019.101231>
- Guarner, J. (2019). Chagas disease as example of a reemerging parasite. *Seminars in Diagnostic Pathology*, *36*(3), 164-169. <https://doi.org/10.1053/j.semdp.2019.04.008>

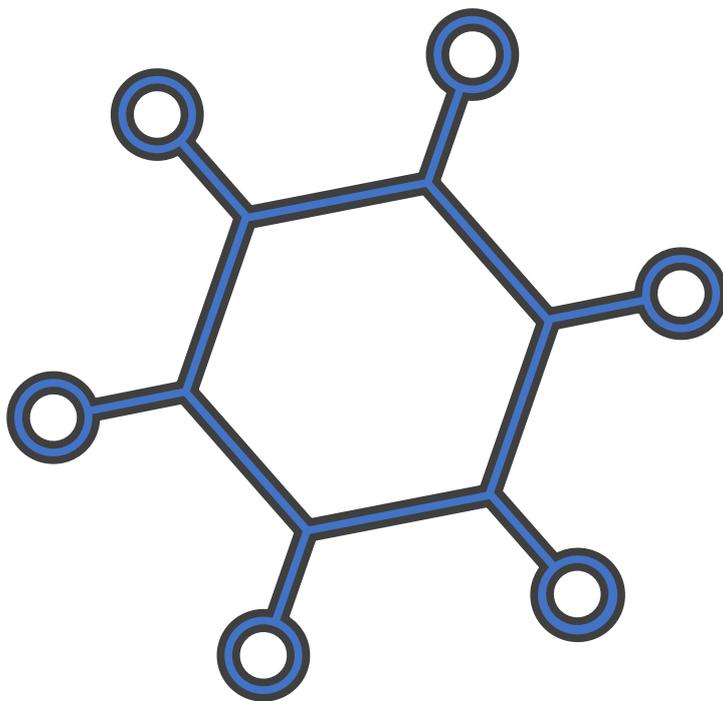
- Guhl, F. (2013). EPIDEMIOLOGÍA MOLECULAR DE TRYPANOSOMA CRUZI. *Rev Esp Salud Pública*, 9.
- Hajduk, P. J., Huth, J. R., & Tse, C. (2005). Predicting protein druggability. *Drug Discovery Today*, 10(23-24), 1675-1682. [https://doi.org/10.1016/S1359-6446\(05\)03624-X](https://doi.org/10.1016/S1359-6446(05)03624-X)
- Jiang, Q. C., Xu, C. L., Wang, H. Y., Wang, J. G., & Yang, Y. F. (2007). Estimation of the shifting distance of the eutectic point in hypereutectic Al-Si alloys by the lever rule. *Scripta Materialia*, 56(5), 329-332. <https://doi.org/10.1016/j.scriptamat.2006.11.023>
- Kaufer, A., Ellis, J., Stark, D., & Barratt, J. (2017). The evolution of trypanosomatid taxonomy. *Parasites & Vectors*, 10(1), 287. <https://doi.org/10.1186/s13071-017-2204-7>
- Kessler, R. L., Contreras, V. T., Marlière, N. P., Aparecida Guarneri, A., Villamizar Silva, L. H., Mazarotto, G. A. C. A., Batista, M., Soccol, V. T., Krieger, M. A., & Probst, C. M. (2017). Recently differentiated epimastigotes from *Trypanosoma cruzi* are infective to the mammalian host: *Trypanosoma cruzi* rdEpi are infective to mammal. *Molecular Microbiology*, 104(5), 712-736. <https://doi.org/10.1111/mmi.13653>
- Krauth-Siegel, R. L., & Comini, M. A. (2008). Redox control in trypanosomatids, parasitic protozoa with trypanothione-based thiol metabolism. *Biochimica et Biophysica Acta (BBA) - General Subjects*, 1780(11), 1236-1248. <https://doi.org/10.1016/j.bbagen.2008.03.006>
- Kryshchysyn, A., Kaminsky, D., Grellier, P., & Lesyk, R. (2014). Trends in research of antitrypanosomal agents among synthetic heterocycles. *European Journal of Medicinal Chemistry*, 85, 51-64. <https://doi.org/10.1016/j.ejmech.2014.07.092>
- Leite-Dias, J. V., Queiroz, D. R. M., Martins, H. R., Gorla, D. E., Pires, H. H. R., & Diotaiuti, L. (2016). Spatial distribution of triatomines in domiciles of an urban area of the Brazilian Southeast Region. *Memórias do Instituto Oswaldo Cruz*, 111(1), 43-50. <https://doi.org/10.1590/0074-02760150352>
- Lima, L., Espinosa-Álvarez, O., Ortiz, P. A., Trejo-Varón, J. A., Carranza, J. C., Pinto, C. M., Serrano, M. G., Buck, G. A., Camargo, E. P., & Teixeira, M. M. G. (2015). Genetic diversity of *Trypanosoma cruzi* in bats, and multilocus phylogenetic and phylogeographical analyses supporting Tcbat as an independent DTU (discrete typing unit). *Acta Tropica*, 151, 166-177. <https://doi.org/10.1016/j.actatropica.2015.07.015>
- Lourenço, A. M., Faccini, C. C., Costa, C. A. de J., Mendes, G. B., Fragata Filho, A. A., Lourenço, A. M., Faccini, C. C., Costa, C. A. de J., Mendes, G. B., & Fragata Filho, A. A. (2018). Evaluation of in vitro anti-*Trypanosoma cruzi* activity of medications benznidazole, amiodarone hydrochloride, and their combination. *Revista Da Sociedade Brasileira de Medicina Tropical*, 51(1), 52-56. <https://doi.org/10.1590/0037-8682-0285-2017>
- Magalhães, L. M. D., Viana, A., Chiari, E., Galvão, L. M. C., Gollob, K. J., & Dutra, W. O. (2015). Differential Activation of Human Monocytes and Lymphocytes by Distinct Strains of *Trypanosoma cruzi*. *PLOS Neglected Tropical Diseases*, 9(7), e0003816. <https://doi.org/10.1371/journal.pntd.0003816>
- Manta, B., Comini, M., Medeiros, A., Hugo, M., Trujillo, M., & Radi, R. (2013). Trypanothione: A unique bis-glutathionyl derivative in trypanosomatids. *Biochimica et Biophysica Acta (BBA) - General Subjects*, 1830(5), 3199-3216. <https://doi.org/10.1016/j.bbagen.2013.01.013>
- Marín-Hernández, Á., Gallardo-Pérez, J. C., Reyes-García, M. A., Sosa-Garrocho, M., Macías-Silva, M., Rodríguez-Enríquez, S., Moreno-Sánchez, R., & Saavedra, E. (2020). Kinetic modeling of glucose central metabolism in hepatocytes and hepatoma cells. *Biochimica et*

- Biophysica Acta (BBA) - General Subjects*, 1864(11), 129687. <https://doi.org/10.1016/j.bbagen.2020.129687>
- Medeiros, A., Benítez, D., Korn, R. S., Ferreira, V. C., Barrera, E., Carrión, F., Pritsch, O., Pantano, S., Kunick, C., de Oliveira, C. I., Orban, O. C. F., & Comini, M. A. (2020). Mechanistic and biological characterisation of novel  $N^5$ -substituted paullones targeting the biosynthesis of trypanothione in *Leishmania*. *Journal of Enzyme Inhibition and Medicinal Chemistry*, 35(1), 1345-1358. <https://doi.org/10.1080/14756366.2020.1780227>
- Mesías, A. C., Sasoni, N., Arias, D. G., Pérez Brandán, C., Orban, O. C. F., Kunick, C., Robello, C., Comini, M. A., Garg, N. J., & Zago, M. P. (2019). Trypanothione synthetase confers growth, survival advantage and resistance to anti-protozoal drugs in *Trypanosoma cruzi*. *Free Radical Biology and Medicine*, 130, 23-34. <https://doi.org/10.1016/j.freeradbiomed.2018.10.436>
- Meymandi, S. K., Forsyth, C. J., Soverow, J., Hernandez, S., Sanchez, D., Montgomery, S. P., & Traina, M. (2017). Prevalence of Chagas Disease in the Latin American-born Population of Los Angeles. *Clinical Infectious Diseases*, 64(9), 1182-1188. <https://doi.org/10.1093/cid/cix064>
- Michels, P. A. M., Ginger, M. L., & Zilberstein, D. (Eds.). (2020). *Trypanosomatids: Methods and Protocols* (Vol. 2116). Springer US. <https://doi.org/10.1007/978-1-0716-0294-2>
- Molina, I., Gómez i Prat, J., Salvador, F., Treviño, B., Sulleiro, E., Serre, N., Pou, D., Roure, S., Cabezos, J., Valerio, L., Blanco-Grau, A., Sánchez-Montalvá, A., Vidal, X., & Pahissa, A. (2014). Randomized Trial of Posaconazole and Benznidazole for Chronic Chagas' Disease. *New England Journal of Medicine*, 370(20), 1899-1908. <https://doi.org/10.1056/NEJMoa1313122>
- Moran. (2009). *Neglected Diseases Research and development: How Much Are We Really Spending?* plos Medice.
- Morillo, C. A., Marin-Neto, J. A., Avezum, A., Sosa-Estani, S., Rassi, A., Rosas, F., Villena, E., Quiroz, R., Bonilla, R., Britto, C., Guhl, F., Velazquez, E., Bonilla, L., Meeks, B., Rao-Melacini, P., Pogue, J., Mattos, A., Lazdins, J., Rassi, A., ... Yusuf, S. (2015). Randomized Trial of Benznidazole for Chronic Chagas' Cardiomyopathy. *New England Journal of Medicine*, 373(14), 1295-1306. <https://doi.org/10.1056/NEJMoa1507574>
- Morillo, C. A., Waskin, H., Sosa-Estani, S., del Carmen Bangher, M., Cuneo, C., Milesi, R., Mallagray, M., Apt, W., Beloscar, J., Gascon, J., Molina, I., Echeverria, L. E., Colombo, H., Perez-Molina, J. A., Wyss, F., Meeks, B., Bonilla, L. R., Gao, P., Wei, B., ... Reyes-López, P. A. (2017). Benznidazole and Posaconazole in Eliminating Parasites in Asymptomatic T. Cruzi Carriers. *Journal of the American College of Cardiology*, 69(8), 939-947. <https://doi.org/10.1016/j.jacc.2016.12.023>
- Mosquillo, M. F., Smircich, P., Ciganda, M., Lima, A., Gambino, D., Garat, B., & Pérez-Díaz, L. (2020). Comparative high-throughput analysis of the *Trypanosoma cruzi* response to organometallic compounds. *Metallomics*, 12(5), 813-828. <https://doi.org/10.1039/D0MT00030B>
- Olin-Sandoval, V. (2012). Drug target validation of the trypanothione pathway enzymes through metabolic modelling. *FEBS Journal*, 23.
- Oza, S. L., Tetaud, E., Ariyanayagam, M. R., Warnon, S. S., & Fairlamb, A. H. (2002). A Single Enzyme Catalyses Formation of Trypanothione from Glutathione and Spermidine in

- Trypanosoma cruzi*. *Journal of Biological Chemistry*, 277(39), 35853-35861. <https://doi.org/10.1074/jbc.M204403200>
- Piacenza, L., Peluffo, G., Alvarez, M. N., Martínez, A., & Radi, R. (2013). *Trypanosoma cruzi* Antioxidant Enzymes As Virulence Factors in Chagas Disease. *Antioxidants & Redox Signaling*, 19(7), 723-734. <https://doi.org/10.1089/ars.2012.4618>
- Pinheiro, E., Brum-Soares, L., Reis, R., & Cubides, J.-C. (2017). Chagas disease: Review of needs, neglect, and obstacles to treatment access in Latin America. *Revista Da Sociedade Brasileira de Medicina Tropical*, 50(3), 296-300. <https://doi.org/10.1590/0037-8682-0433-2016>
- Planer, J. D., Hulverson, M. A., Arif, J. A., Ranade, R. M., Don, R., & Buckner, F. S. (2014). Synergy Testing of FDA-Approved Drugs Identifies Potent Drug Combinations against *Trypanosoma cruzi*. *PLoS Neglected Tropical Diseases*, 8(7), e2977. <https://doi.org/10.1371/journal.pntd.0002977>
- Rajão, M. A., Furtado, C., Alves, C. L., Passos-Silva, D. G., Moura, M. B. de, Chamber-Reis, B. L., Kunrath-Lima, M., Zuma, A. A., Vieira-da-Rocha, J. P., Garcia, J. B. F., Mendes, I. C., Pena, S. D. J., Macedo, A. M., Franco, G. R., Souza-Pinto, N. C. de, Medeiros, M. H. G. de, Cruz, A. K., Motta, M. C. M., Teixeira, S. M. R., & Machado, C. R. (2014). Unveiling Benznidazole's mechanism of action through overexpression of DNA repair proteins in *Trypanosoma cruzi*. *Environmental and Molecular Mutagenesis*, 55(4), 309-321. <https://doi.org/10.1002/em.21839>
- Rassi, A., Rassi, A., & Marcondes de Rezende, J. (2012). American Trypanosomiasis (Chagas Disease). *Infectious Disease Clinics of North America*, 26(2), 275-291. <https://doi.org/10.1016/j.idc.2012.03.002>
- Ribeiro, V., Dias, N., Paiva, T., Hagström-Bex, L., Nitz, N., Pratesi, R., & Hecht, M. (2020). Current trends in the pharmacological management of Chagas disease. *International Journal for Parasitology: Drugs and Drug Resistance*, 12, 7-17. <https://doi.org/10.1016/j.ijpddr.2019.11.004>
- Saavedra, E., González-Chávez, Z., Moreno-Sánchez, R., & Michels, P. A. M. (2019). Drug Target Selection for *Trypanosoma cruzi* Metabolism by Metabolic Control Analysis and Kinetic Modeling. *Current Medicinal Chemistry*, 26(36), 6652-6671. <https://doi.org/10.2174/0929867325666180917104242>
- Sales Junior, P. A., Molina, I., Fonseca Murta, S. M., Sánchez-Montalvá, A., Salvador, F., Corrêa-Oliveira, R., & Carneiro, C. M. (2017). Experimental and Clinical Treatment of Chagas Disease: A Review. *The American Journal of Tropical Medicine and Hygiene*, 97(5), 1289-1303. <https://doi.org/10.4269/ajtmh.16-0761>
- Sass, G., Madigan, R. T., Joubert, L.-M., Bozzi, A., Sayed, N., Wu, J. C., & Stevens, D. A. (2019). A Combination of Itraconazole and Amiodarone Is Highly Effective against *Trypanosoma cruzi* Infection of Human Stem Cell-Derived Cardiomyocytes. *The American Journal of Tropical Medicine and Hygiene*, 101(2), 383-391. <https://doi.org/10.4269/ajtmh.19-0023>
- Sbaraglini, M., Cristina Vanrell, M., Leticia Bellera, C., Benaim, G., Carrillo, C., Talevi, A., & Silvia Romano, P. (2016). Neglected Tropical Protozoan Diseases: Drug Repositioning as a Rational Option. *Current Topics in Medicinal Chemistry*, 16(19), 2201-2222. <https://doi.org/10.2174/1568026616666160216154309>
- Sbaraglini, M. L., Bellera, C. L., Fraccaroli, L., Larocca, L., Carrillo, C., Talevi, A., & Alba Soto, C. D. (2016). Novel cruzipain inhibitors for the chemotherapy of chronic Chagas disease.

- International Journal of Antimicrobial Agents*, 48(1), 91-95.  
<https://doi.org/10.1016/j.ijantimicag.2016.02.018>
- Smith, J., & Taylor, E. M. (2016). What Is Next for NTDs in the Era of the Sustainable Development Goals? *PLOS Neglected Tropical Diseases*, 10(7), e0004719.  
<https://doi.org/10.1371/journal.pntd.0004719>
- Sousa, A. F., Gomes-Alves, A. G., Benítez, D., Comini, M. A., Flohé, L., Jaeger, T., Passos, J., Stuhlmann, F., Tomás, A. M., & Castro, H. (2014). Genetic and chemical analyses reveal that trypanothione synthetase but not glutathionylspermidine synthetase is essential for *Leishmania infantum*. *Free Radical Biology and Medicine*, 73, 229-238.  
<https://doi.org/10.1016/j.freeradbiomed.2014.05.007>
- Talevi, A., Carrillo, C., & Comini, M. (2019). The Thiol-polyamine Metabolism of *Trypanosoma cruzi*: Molecular Targets and Drug Repurposing Strategies. *Current Medicinal Chemistry*, 26(36), 6614-6635. <https://doi.org/10.2174/0929867325666180926151059>
- Tanowitz, H. B., Machado, F. S., Jelicks, L. A., Shirani, J., Campos de Carvalho, A. C., Spray, D. C., Factor, S. M., Kirchoff, L. V., & Weiss, L. M. (2009). Perspectives on *Trypanosoma cruzi*-Induced Heart Disease (Chagas Disease). *Progress in Cardiovascular Diseases*, 51(6), 524-539. <https://doi.org/10.1016/j.pcad.2009.02.001>
- Terán, D., Hocková, D., Česnek, M., Zíková, A., Naesens, L., Keough, D. T., & Guddat, L. W. (2016). Crystal structures and inhibition of *Trypanosoma brucei* hypoxanthine-guanine phosphoribosyltransferase. *Scientific Reports*, 6(1), 35894.  
<https://doi.org/10.1038/srep35894>
- Tonge, P. J. (2018). Drug-Target Kinetics in Drug Discovery. *ACS Chemical Neuroscience*, 9(1), 29-39. <https://doi.org/10.1021/acschemneuro.7b00185>
- Torrico, F., Gascon, J., Ortiz, L., Alonso-Vega, C., Pinazo, M.-J., Schijman, A., Almeida, I. C., Alves, F., Strub-Wourgaft, N., Ribeiro, I., Santana, G., Blum, B., Correia, E., Garcia-Bournisen, F., Vaillant, M., Morales, J. R., Pinto Rocha, J. J., Rojas Delgadillo, G., Magne Anzoleaga, H. R., ... Jimenez, W. (2018). Treatment of adult chronic indeterminate Chagas disease with benznidazole and three E1224 dosing regimens: A proof-of-concept, randomised, placebo-controlled trial. *The Lancet Infectious Diseases*, 18(4), 419-430.  
[https://doi.org/10.1016/S1473-3099\(17\)30538-8](https://doi.org/10.1016/S1473-3099(17)30538-8)
- Torrie, L. S., Wyllie, S., Spinks, D., Oza, S. L., Thompson, S., Harrison, J. R., Gilbert, I. H., Wyatt, P. G., Fairlamb, A. H., & Frearson, J. A. (2009). Chemical Validation of Trypanothione Synthetase: A POTENTIAL DRUG TARGET FOR HUMAN TRYPANOSOMIASIS. *Journal of Biological Chemistry*, 284(52), 36137-36145. <https://doi.org/10.1074/jbc.M109.045336>
- Zingales, B. (2011). *Trypanosoma cruzi*: One parasite, two parasites or several parasites of chagas disease? <https://core.ac.uk/reader/268366460>

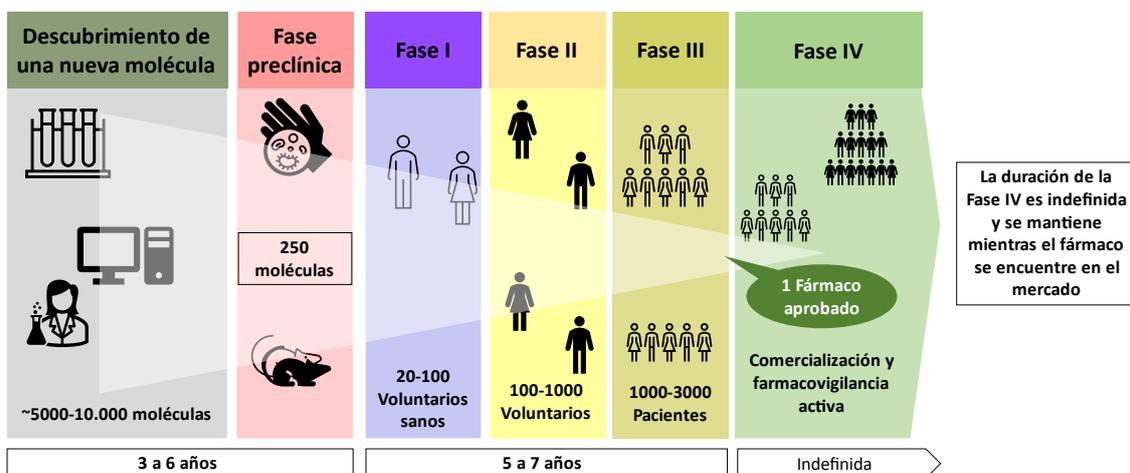
# Capítulo 2 | Descubrimiento de Fármacos



## 2.1 Descubrimiento de fármacos: Generalidades

El descubrimiento y desarrollo de un nuevo fármaco es un desafío complejo que involucra varias etapas; las mismas comprenden desde identificar la actividad biológica de una molécula hasta la realización de estudios clínicos que garanticen seguridad y eficacia a los pacientes (Schneider, 2018).

La introducción de un nuevo fármaco a la clínica debería traer aparejado al menos una mejora en el perfil de eficacia, reducir los tiempos de administración, espaciar los intervalos de administración, o minimizar la probabilidad de efectos adversos en el paciente, y en el mejor de los casos producir un cambio sustancial respecto a las terapias ya disponibles (Field et al., 2017). Sin embargo, la consecución de estos objetivos suele ser dificultosa: por cada nuevo fármaco que alcanza la aprobación se estudian, aproximadamente, 5.000 a 10.000 compuestos; el candidato debe atravesar asimismo una compleja batería de ensayos preclínicos y estudios clínicos de diferente complejidad y escala, lo que demora en promedio más de 12 años (ver **figura 2.1**) (Chan et al., 2019). En general los procesos implicados en el descubrimiento y desarrollo de fármacos de *novo* implican un costo promedio de alrededor de 2,6 billones de dólares (Chan et al., 2019; Papapetropoulos & Szabo, 2018), por lo cual las pequeñas empresas farmacéuticas e instituciones académicas se encuentran en una clara desventaja competitiva, que ocasionalmente puede subsanarse mediante distintos modelos colaborativos (Allarakhia, 2013). En general, el número de nuevos fármacos aprobados por las agencias regulatorias muestra una tendencia positiva en los últimos años, y a pesar de la pandemia de COVID-19 en 2020 FDA aprobó un total de 53 nuevos fármacos, lo que eleva el promedio de los últimos 5 años (2015-2020) a 46 fármaco por año (Mullard, 2021), observándose sin embargo importantes variaciones según el área terapéutica. En términos generales, las grandes compañías farmacéuticas priorizan aquellas áreas con mayor probabilidad de éxito o retorno de inversión (Dowden & Munro, 2019).



**Figura 2.1.** Esquema del proceso general de descubrimiento y desarrollo de fármacos.

En general, se pueden diferenciar dos grandes enfoques en el descubrimiento de nuevos tratamientos para enfermedades causadas por tripanosomátidos:

**Basado en blancos moleculares:** consiste en evaluar diferentes compuestos frente a una biomolécula blanco (proteínas, ácidos nucleicos, lípidos) específica, y analizar si su interacción modifica significativamente las funciones celulares, con el fin de obtener así la respuesta farmacológica deseada. Con el advenimiento de las tecnologías computarizadas y la robótica, las grandes compañías farmacéuticas se han apoyado en plataformas de cribado de alto rendimiento (*HTS*, del inglés *High Throughput Screening*) para llevar adelante esta primera fase de identificación de candidatos activos (Gilbert, 2013).

**Enfoque fenotípico:** este enfoque se focaliza en el cribado de compuestos sobre células o microorganismo completo, desconociendo a priori los mecanismos moleculares involucrado en el efecto terapéutico. Una vez ratificada la eficacia *in vitro*, se puede dilucidar el modo de acción del compuesto activo. Hoy en día, este tipo de cribado también puede realizarse mediante plataformas *HTS* (Gilbert, 2013). La metodología permite, además, predecir posibles eventos adversos o tóxicos frente a diversas líneas celulares (Krishna et al., 2021).

Un nuevo enfoque que permite integrar ambas estrategias es el **Cribado de Alto Contenido** (*HCS*, del inglés *High Content Screening*). Esta nueva aproximación genera un gran volumen de datos que es recopilado y analizado mediante softwares y algoritmos para el análisis de alto contenido. Dicha combinación de herramientas permite contemplar y evaluar cambios generados en varios niveles de los cultivos celulares, microorganismos o tejidos, procesando los datos desde una perspectiva más fisiológica sobre la modulación de un compuesto frente a un blanco determinado (Cornaglia et al., 2017; Lin et al., 2020).

Hasta principios de los años noventa la mirada fenotípica gobernaba la búsqueda de nuevos fármacos; sin embargo, con el avance logrado por la genómica en esos años comenzó a ganar terreno el descubrimiento de fármacos basado en blancos moleculares (D. Swinney, 2013). Este nuevo punto de vista se fundamentaba en una perspectiva reduccionista de los sistemas biológicos, en este marco, la búsqueda de fármacos se fundamenta en la premisa de que cada patología podía ser controlada a partir de la interacción de un fármaco con una única macromolécula implicada en la fisiopatología de la enfermedad (Margineanu, 2016; Talevi, 2016). El nuevo enfoque permitió descubrimientos y desarrollos de terapias específicas como en el caso de los anticuerpos monoclonales en la lucha contra el cáncer, o diversos agentes antimicrobianos a partir de la evaluación de candidatos contra blancos moleculares presentes en los microorganismos, pero sin homólogos en mamíferos (Margineanu, 2016; Swinney & Anthony, 2011). Típicamente, las nuevas entidades químicas que han ingresado al mercado luego de ciclos de optimización orientados a lograr un balance adecuado de potencia, selectividad y características farmacocinéticas se han obtenido siguiendo la óptica del descubrimiento basado en blancos moleculares (Swinney & Anthony, 2011; Talevi, 2016).

Ambos enfoques han sido cuestionados a lo largo del tiempo, sin embargo, diversos autores proponen que en la búsqueda de nuevos fármacos la perspectiva que se desea desarrollar debe ser planificada con anterioridad, y en consecuencia establecer la metodología de trabajo adecuada a la situación fisiopatológica que se desee modular. Por ejemplo, para desordenes mendelianos, que generalmente involucran un solo blanco, una aproximación mediante blancos moleculares resulta más adecuada que un cribado fenotípico. Mientras que para patologías más complejas como alteraciones del Sistema Nervioso Central (SNC) un plano

fenotípico parece ser más apropiado, debido a que estas patologías suelen estar asociadas con alteraciones que involucran diversos blancos moleculares como así también alteraciones de la concentración de ligandos endógenos. Cabe mencionar que, si bien el enfoque fenotípico manifiesta de manera más completa las características de la patología, la óptica centrada en blancos moleculares permite conocer en detalle el mecanismo de acción y la selectividad de un compuesto por una determinada diana farmacológica. El traslado de un fármaco de la fase preclínica a la clínica es un paso crítico, ya que la diferencia entre especies (rata/humano o cobayo/humano) ha dado lugar a pérdidas de nuevas moléculas por motivos de toxicidad, perfil farmacocinético inadecuado o falta de eficacia (Margineanu, 2016; Swinney & Anthony, 2011; Talevi, 2016).

Actualmente, con la continua integración entre las ciencias informáticas y la biología se está abriendo una nueva era en el descubrimiento de fármacos, dominada por los principios de la biología de sistemas y la farmacología de redes. Esta nueva mirada podría verse como una integración de los enfoques tradicionales, ya que intenta establecer relaciones directas fármaco-proteína, pero a su vez tiene en cuenta otros blancos presentes de las vías, diferentes interconexiones entre ellas, y como la modulación de un blanco repercute en el organismo, o al menos la red metabólica más cercana. Si bien la farmacología de redes parece ser una prometedora metodología en el descubrimiento de fármacos es evidente que falta mucho camino por recorrer para que se establezca como alternativa más accesible, ya que requiere de estudios genómicos, proteómicos y fenotípicos que puedan integrar y reconstruir la red metabólica interna de las células y el organismo en general, como así también herramientas computacionales que hagan énfasis en el nuevo paradigma donde una droga module múltiples blancos con baja afinidad, revirtiendo así la tendencia actual donde se busca que cada fármaco actúe, con alta afinidad, contra un blanco específico (Margineanu, 2016; Parvathaneni, 2019; Talevi, 2016).

Las estrategias de búsquedas de nuevos fármacos pueden ser netamente experimentales, iniciando los estudios directamente desde ensayos *in vitro* o contemplar una combinación de herramientas computacionales (modelado *in silico*, estrategias de *Machine Learning*, entre otras) con etapas de validación *in vitro*, donde solo se evaluarán las moléculas que fueron clasificadas como activas por las estrategias computacionales. Hoy en día, con el avance de la quimioinformática y las técnicas de modelado molecular la combinación de metodologías computacional-experimental se ha convertido en un enfoque clave para la búsqueda de nuevos fármacos, permitiendo un ahorro significativo de tiempo y dinero en comparación con las metodologías netamente experimentales (Prieto-Martínez et al., 2019).

En la **figura 2.2** se presenta un esquema de los cambios de paradigmas que fueron ocurriendo a lo largo de los años en el campo del descubrimiento de fármacos, los cuales van desde hallazgos fortuitos hasta la actual farmacología de redes.

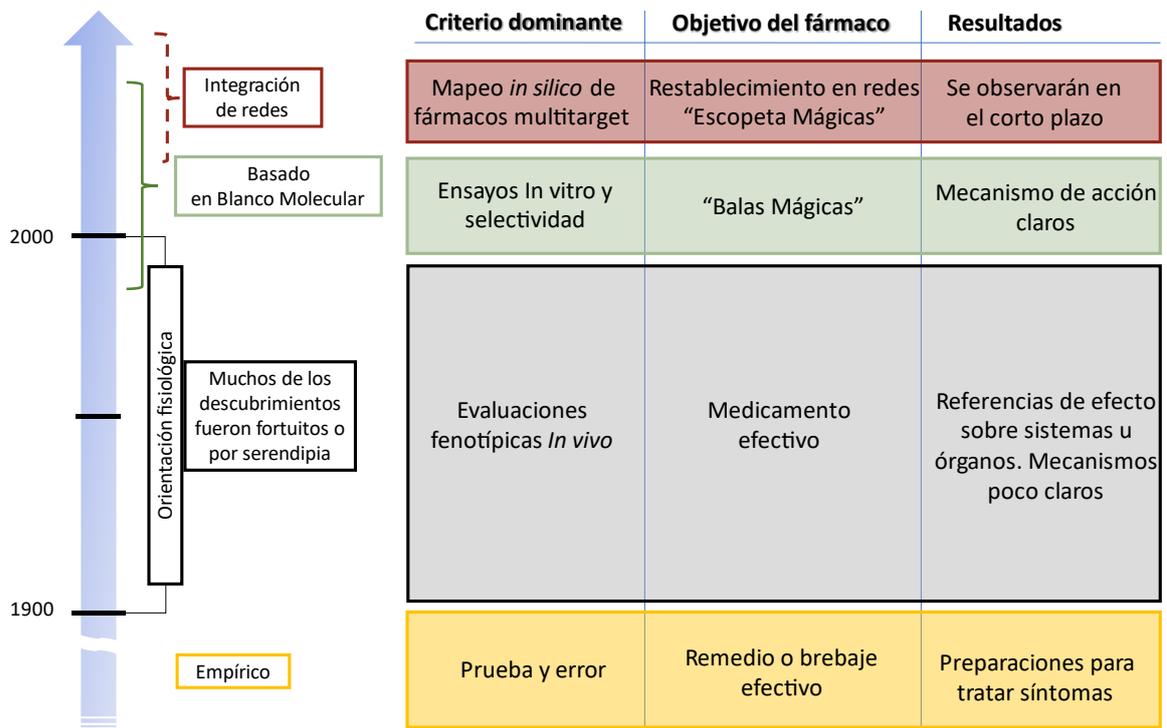


Figura 2.2. Evolución temporal de los diferentes enfoques en el campo del descubrimiento de fármacos. Adaptada de (Margeanu, 2016).

## 2.2 Cribado Virtual

La utilización de metodologías computacionales para a la identificación de nuevos fármacos puede abordarse desde dos estrategias yuxtapuestas, donde en una de ellas, el **diseño racional de fármacos**, se utiliza la metodología computacional como base para diseñar y posteriormente sintetizar nuevas entidades químicas; mientras que la otra, conocida como **búsqueda racional o búsqueda asistida por computadoras** consiste en cribar grandes quimiotecas de compuestos, identificando así nuevas actividades biológicas en entidades químicas ya conocidas (Talevi, 2016). Esta última estrategia se conoce como **cribado virtual (CV)** y se define como el conjunto de técnicas computacionales que permiten explorar la actividad biológica de compuestos químicos que se hallan almacenados en quimiotecas o bibliotecas digitales de estructuras químicas (Gavernet, 2021).

A continuación, se repasan las principales ventajas del cribado virtual:

**Ventajas bioéticas:** La utilización de metodologías computacionales reduce significativamente el número de compuestos a testear, priorizando la evaluación experimental de los hits seleccionados de manera *in silico*. El CV responde a la filosofía de las pautas bioéticas vigentes a nivel internacional que proponen utilizar modelos computacionales y/o *in vitro* previamente a los ensayos *in vivo*, con el objeto de minimizar los ensayos en modelos animales y maximizar la tasa de resultados positivos en los mismos.

**Carácter teórico:** El CV, por su carácter teórico, no requiere *a priori* de una muestra física de los compuestos de la quimioteca explorada. La adquisición o síntesis se realiza

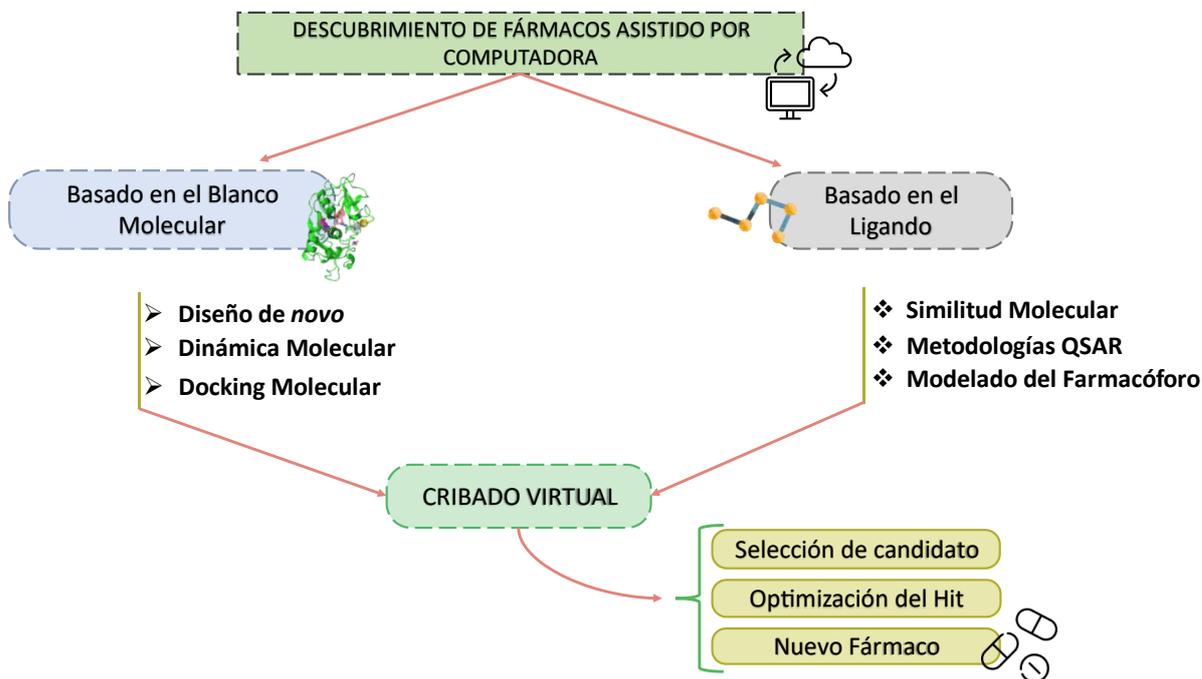
con posterioridad a la selección computacional de un número restringido de compuestos de la quimioteca; es decir, solo avanzarán a ensayos *in vitro/in vivo* compuestos que son señalados como promisorios candidatos por los modelos o algoritmos utilizados (McInnes, 2007).

**Ventajas económicas:** El carácter teórico del CV permite evaluar/estimar la actividad de compuestos químicos sin la necesidad de contar físicamente con las moléculas, por lo tanto, permite evaluar un gran número de compuestos y adquirir solo aquellos que resulten promisorios en las campañas de CV. Esto genera un importante ahorro económico (Neves et al., 2018). Adicionalmente existen hoy en día diversas herramientas con licencias libres, académicas o de código abierto, las cuales pueden ser aplicadas en diferentes etapas del CV, algunas de ellas son ChemAxon (<https://chemaxon.com/>); PaDEL (Yap, 2011); UCSF Chimera (Pettersen et al., 2004); paquetes de Rstudio (<https://rstudio.com/>), Open Babel (O'Boyle et al., 2011) o diferentes paquetes en el lenguaje Python (<https://www.python.org/>) entre otras que, combinadas con el acceso libre y gratuito a diversas bases de datos con información sobre compuestos químicos, contribuyen de manera eficiente a la reducción de los costos. Entre las bases de datos de acceso libre podemos mencionar: DrugBank (Wishart et al., 2018); SweetLead (Novick et al., 2013); PubChem (Kim et al., 2015); Zinc15 (Sterling & Irwin, 2015); o ChEMBL (Gaulton et al., 2018), las mismas compilan, desde propiedades fisicoquímicas hasta indicaciones clínicas y reportes de toxicidad.

**Reposicionamiento:** Debido al volumen de datos clínicos con los que se cuenta hoy en día, y que existen diversas quimiotecas/bibliotecas de compuestos químicos que recopilan, almacenan y clasifican datos de fármacos ya aprobados, retirados/discontinuados, o en fases de investigación avanzadas, es posible cribar dichas bibliotecas con el objetivo de poder reutilizar estos “viejos” fármacos en patologías que aún no cuenten con tratamientos enteramente satisfactorios.

Finalmente, dependiendo de la estrategia *in silico* empleada y de la cantidad de compuestos a cribar, es posible realizar el CV con computadoras de uso doméstico.

En el presente capítulo repasaremos brevemente las metodologías más relevantes para la identificación de fármacos basados en la estrategia de cribado virtual. Esta estrategia computacional puede ser abordada desde dos grandes perspectivas (**figura 2.3**), metodologías basadas en el blanco molecular (métodos directos) o metodologías basadas en el ligando (métodos indirectos).



**Figura 2.3.** Esquema de las principales metodologías utilizadas en la búsqueda de nuevos fármacos asistida por computadoras.

### 2.2.1 Métodos basados en el blanco molecular

Los métodos de **cribado virtual basados en el blanco molecular (CVBBM)** requieren necesariamente del conocimiento tridimensional de la estructura de la diana; la misma puede ser obtenida a través de cristalografía de rayos X, Resonancia Magnética Nuclear (RMN), Microscopía electrónica, mediante modelado virtual por homología o mediante modelado *ab initio* (Macalino et al., 2015; Talevi, 2018). Dentro de estos procedimientos podemos englobar la metodología de Docking molecular o acoplamiento molecular y la metodología de Dinámica molecular.

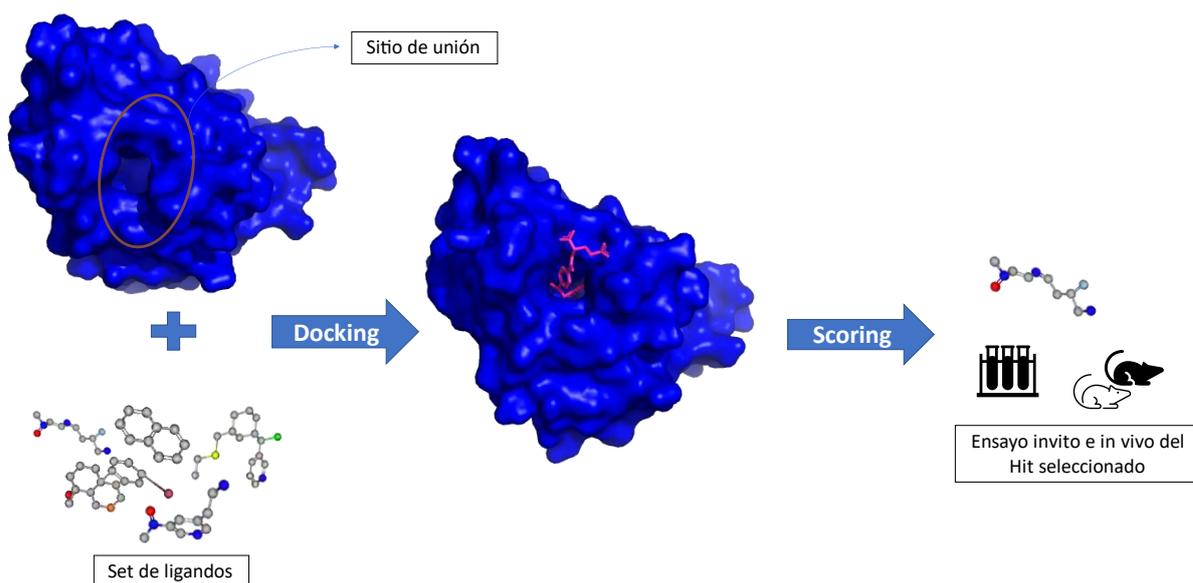
#### 2.2.1.1 Docking molecular

El **Docking molecular** es la metodología de tamizado virtual más explorada hasta el momento, y la misma permite estimar o predecir las interacciones que diferentes compuestos (ligandos) presentan en el sitio de unión de una proteína (blanco molecular). Este procedimiento podría resumirse en tres grandes pasos: *i*) elección y acondicionamiento del blanco molecular, *ii*) recopilación y optimización de un conjunto de ligandos, y *iii*) acoplamiento entre los ligandos y el blanco molecular. Al complejo formado en el paso *iii* (ligando-blanco molecular) se le calcula, mediante algoritmos matemáticos, la energía libre ( $\Delta G$ ) de unión y se realiza un ranking de prioridades con los complejos de mayor estabilidad ( $\Delta G$  más negativo) (Lionta et al., 2014; Macalino et al., 2015).

El procedimiento de Docking permite estimar el rol de los factores entálpicos y entrópicos que favorecen, o no, la formación del complejo ligando-proteína y las diferentes conformaciones que podría adoptar el ligando (ligando flexible) al unirse al blanco molecular o

el comportamiento del blanco (blanco flexible) al acoplarse el ligando (Lavecchia & Giovanni, 2013; Lionta et al., 2014).

Actualmente existen softwares de acceso libre y gratuitos que permiten realizar Docking con las condiciones que el operador desee setear. Algunos de estos softwares son: ZDOCK (ZDOCK: Protein Docking (umassmed.edu)), AUTODOCK (AutoDock — AutoDock (scripps.edu)) y DOCK (UCSF DOCK) (Lavecchia & Giovanni, 2013). A su vez cada operador podría utilizar diferentes funciones de score, ya sean funciones basadas en campos de fuerzas, funciones de base empírica, o funciones basadas en el conocimiento; esta elección se realiza teniendo en cuenta que el propósito del Docking es un resultado de puntuación o de clasificación. Independientemente de los resultados que se obtengan por uno u otro procedimiento, es manifiesto que, cuando hay consenso entre dos o más funciones los resultados predichos se corresponden con mayores probabilidades de aciertos experimentales (Macalino et al., 2015). En la **figura 2.4** se presenta un esquema básico de la técnica de Docking.



**Figura 2.4.** La figura representa un esquema básico de la técnica de *Docking*.

### 2.2.1.2 Dinámica molecular

La Dinámica molecular (DM) es una metodología computacionalmente más costosa que el Docking molecular, sin embargo, el constante desarrollo de herramientas informáticas ha permitido poder utilizar esta metodología para el estudio de pequeños conjuntos de ligandos (Gioia et al., 2017).

La DM parte de un sistema básico similar al Docking, ya que requiere de un blanco y, ocasionalmente, diferentes ligandos. Sin embargo, la metodología desarrollada difiere en que mientras cada experimento de acoplamiento es una evaluación estática (independientemente de que se muestreen distintas conformaciones del ligando y el blanco), la DM permite reconstruir con la mayor precisión posible el proceso completo de acoplamiento ligando-

receptor, teniendo en cuenta tanto los procesos mecánicos como energéticos, posibilitando a su vez el cálculo de diversas constantes cinéticas de unión y desacople, ya que como su nombre lo indica es una metodología de acoplamiento dinámico, es decir, estudiará la evolución del sistema durante un cierto tiempo de simulación (Gioia et al., 2017; Vivo & Cavalli, 2017).

### 2.2.1.3 TryS y CVBBM

Hasta el momento la estructura 3D de la enzima **TryS** ha sido resuelta de manera parcial y solo para la especie de *Leishmania* (Koch et al., 2013). En consecuencia, solo se han reportado dos artículos de CVBBM frente a la enzima **TryS** de *T. cruzi* (**TcTrys**), y en ambos casos el modelo de la enzima es generado a partir de homología con la proteína de *Leishmania* (Bombaça et al., 2019; Vázquez et al., 2017) (marzo-2021).

## 2.2.2 Métodos basados en ligando

A pesar de los avances en cristalografía, RMN, microscopía electrónica, o en modelados computacionales por homología aún se desconoce la estructura tridimensional de muchas proteínas. Cuando esto sucede no es posible realizar un cribado virtual basado en el blanco, y es allí donde la metodología de **cribado virtual basado en el ligando (CVBL)** comienza a ganar terreno, ya que no solo se independiza de la estructura del blanco, sino que además cualquiera de las técnicas utilizadas en el CVBL requiere un menor costo computacional en comparación con las técnicas basada en el blanco molecular.

La metodología de CVBL engloba diferentes técnicas entre las que se destacan la **Similitud Molecular**, la técnica de **Farmacóforo**, y la búsqueda basada en funciones químio-matemáticas basadas en **Descriptores Moleculares** (Li & Shah, 2017).

### 2.2.2.1 Métodos basados en Farmacóforo

El concepto Farmacóforo podría resumirse como el arreglo tridimensional de las características moleculares que garantizan una determinada actividad biológica; sin embargo, la IUPAC lo define de manera más específica como el conjunto de características estéricas y electrónicas que garantizan las interacciones supramoleculares óptimas entre la molécula y su blanco biológico permitiendo desencadenar una respuesta (Gavernet, 2021; Seidel et al., 2020).

La técnica comienza con la recopilación de un conjunto de entrenamiento, el cual se encuentra conformado por pequeñas moléculas que comparten una determinada actividad biológica, y la superposición de las moléculas que conforman dicho conjunto para extraer de allí un patrón estructural en común. Una vez obtenido el farmacóforo se realiza el cribado de un conjunto de moléculas con actividad biológica desconocida, y de esta manera se seleccionarán aquellas moléculas que contemplen el farmacóforo como parte de su estructura tridimensional (Yang, 2010). Este método presenta como principal desventaja que requiere de la optimización conformacional de todas las estructuras que se deseen cribar; no obstante, presenta una ventaja casi única de los métodos indirectos y es que permite, además del

cribado, el diseño de nuevos fármacos, ya que identificado el farmacóforo se pueden proponer optimizaciones moleculares a partir del mismo (Gavernet, 2021).

En ocasiones este arreglo farmacofórico se conserva aún en moléculas estructuralmente muy disimiles, y si bien las características principales podrían ser suficiente para una determinada actividad biológica, no se deben obviar características secundarias, ya que el tamaño, la lipofilia u otras propiedades intrínsecas de cada compuesto repercuten de manera directa o indirecta sobre las interacciones moleculares. Esta repercusión podría traducirse en ganancia o pérdida de la actividad deseada (Seidel et al., 2020; Talevi, 2016).

### 2.2.2.2 Similitud molecular

Entre las principales metodologías de cribado virtual no se puede obviar la similitud molecular, un enfoque basado en la premisa de que moléculas estructuralmente similares compartirán una actividad biológica determinada (García-Hernández et al., 2019).

La metodología comienza con la recopilación y clasificación de moléculas que poseen una determinada actividad biológica (Prieto-Martínez et al., 2019). Esta metodología requiere la representación de la estructura molecular, a partir de la cual se priorizarán las propiedades moleculares relevantes para la actividad biológica propuesta, y posteriormente, mediante diversos algoritmos, se realiza la evaluación y cuantificación del grado de similitud entre una molécula testeada biológicamente y diversos compuestos de la quimioteca seleccionada. (Sliwoski et al., 2014). A pesar de que existen diferentes algoritmos para búsqueda de similaridad el enfoque actual se centra en dos metodologías, una de ellas se basa en encontrar un subgrafo o subestructura máxima común, es decir, poder identificar la mayor subestructura compartida entre las moléculas comparadas, mientras que la otra mirada propone evaluar la presencia o ausencia de pequeños subgrafos comprendidos en las llamadas huellas dactilares o *fingerprints* moleculares. Existen vectores que codifican la presencia o ausencia de un dado subgrafo u, ocasionalmente, la frecuencia de cada uno de ellos (Nouheo Ilemo et al., 2019).

Si bien resulta sencillo decir cuáles compuestos comparten una subestructura común, no es tan simple poder determinar el grado de similitud entre ellos, ya que cada observador podría ponderar diferentes aspectos estructurales a la hora de comparar varias moléculas (Maggiore et al., 2014). Para evitar sesgos individuales, se puede recurrir a métodos que a partir de la representación estructural de las moléculas y de la ponderación de diversas características moleculares se podrá obtener un coeficiente de similitud que, generalmente, oscilará entre 0 y 1 (Maggiore et al., 2014).

Hoy en día existen diferentes coeficientes de similitud, los cuales difieren en la fórmula de cálculo del coeficiente, o en la manera de definir los átomos que componen cada subgrafo. En **figura 2.5** se puede apreciar un ejemplo para el cálculo del coeficiente de similitud de Tanimoto.

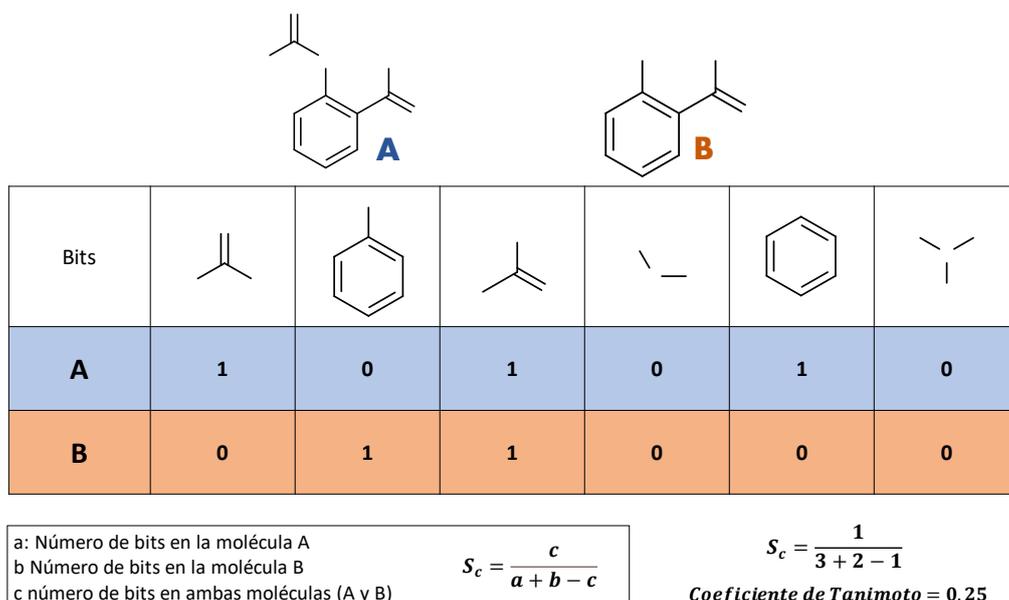


Figura 2.5. Ejemplo de cálculo del coeficiente de Tanimoto.

A diferencia del enfoque farmacofórico esta técnica considera la molécula en su totalidad, y en consecuencia no existe, por default, un fragmento de mayor relevancia como ocurre en el procedimiento descrito con anterioridad; sin embargo, la no identificación de fragmentos principales podría introducir ruido en la comparación, ya que partes de la estructura podría ser perjudicial en el balance de actividad (Sliwoski et al., 2014). En esta disyuntiva y dado que la técnica es computacionalmente económica, algunos autores sugieren que la misma podría ser utilizada como un prefiltro para metodologías más costosas como el *Docking* o la Dinámica molecular, reduciendo así el número de compuestos a cribar (Q. Li & Shah, 2017). Cabe mencionar que entre sus principales ventajas la técnica permite realizar cribados cuando se cuenta tan solo con una molécula activa de referencia, sobre todo en búsquedas de nuevos ligandos para blancos moleculares novedosos o poco estudiados.

### 2.2.2.3 Descriptores moleculares

Los descriptores moleculares se definen como índices numéricos que representan de manera matemática diversas características químicas o estructurales de una molécula (Consonni & Todeschini, 2012). En otras palabras, un descriptor es la transformación de una propiedad molecular en valor numérico. Por ejemplo, el peso molecular (PM) de un compuesto químico refleja su composición atómica (tipo y número de átomos que lo componen). Los descriptores moleculares pueden ser determinados mediante mediciones experimentales, o bien a partir de cálculos teóricos y algoritmos computacionales que permiten obtener información más compleja (Ferreira, 2017).

Debido a su naturaleza y complejidad, los descriptores moleculares pueden recoger información de propiedades básica como el PM hasta propiedades más complejas como la frecuencia de subestructuras o las propiedades electrónicas de una molécula. Esta gran diversidad de información permite clasificar a los descriptores en dos niveles o dimensiones, a)

descriptores moleculares de baja dimensionalidad o independientes de la conformación, y *b*) descriptores de alta dimensionalidad o dependiente de la conformación (Grisoni et al., 2018).

*a*) **Baja dimensionalidad:** Dentro de este grupo se encuentran descriptores que varían entre 0 y 2 dimensiones:

**0D:** Son descriptores fáciles de calcular, pero que brindan escasa información. Los mismos se obtienen a partir de la fórmula molecular, y contemplan propiedades básicas como PM o cantidad de átomos. Estos descriptores presentan un alto grado de degeneración, ya que muchas moléculas pueden presentar similares valores de PM y sin embargo no estar compuestas por los mismos átomos.

**1D:** Este conjunto de descriptores codifican la molécula mediante fragmentos o agrupamiento de átomos, indicando solo presencia, ausencia o frecuencia de grupos funcionales o subestructuras de una molécula sin describir la conectividad entre los diversos fragmentos que componen la molécula (Gavernet, 2021).

**2D:** Los descriptores 2D adicionan información a los niveles de 1 dimensión ya que codifican mediante gráficas de vértices (átomos) y ejes (enlaces) la conectividad y adyacencia entre los átomos, describiendo la estructura geométrica de la molécula en el plano. Adicionalmente permiten obtener información particular para cada átomo como su estado de hibridación electrónica, identidad atómica o valencias ocupadas.

*b*) **Alta dimensionalidad:** Son descriptores derivados de la conformación espacial de la molécula, este conjunto comprende los descriptores 3D y 4D. Al ser dependientes de la conformación se requiere de una optimización molecular previa al cálculo de los descriptores.

**3D:** Este conjunto corresponde al nivel más elevado de los descriptores moleculares, donde además de toda la información contenida en los de baja dimensionalidad se puede obtener información sobre la geometría molecular (por ejemplo, características geométricas de conformaciones estables de la molécula). Si bien estos descriptores proporcionan un alto nivel de información, son más difíciles de calcular, requieren de optimización molecular previa y son computacionalmente más costosos.

**4D:** En esta modalidad el cálculo es similar a los descriptores 3D, pero la información que codificada es un promedio ponderado de las conformaciones más probables de cada molécula (Ferreira, 2017).

Es importante mencionar que cuanto más alto es el nivel de información que brinda un conjunto de descriptores a evaluar mayor es la dificultad para interpretar los modelos generados.

Existen diferentes estrategias estadísticas para encontrar correlaciones entre una variable dependiente (propiedad o actividad a modelar) y uno o más descriptores moleculares; estas relaciones se conocen **Relaciones Cuantitativas Estructura-Actividad (QSAR)**, (del inglés *Quantitative Structure-Activity Relationship*).

Una relación estructura-actividad puede sintetizarse en la siguiente expresión:

$$A = f(d_1, d_2, d_3, \dots, d_n)$$

Donde  $A$  es la actividad biológica o propiedad a modelar, representada por un función  $f$  que contempla diversas propiedades moleculares codificadas, cada una, en diversos descriptores  $d$  (Grisoni et al., 2018). Cabe mencionar que el conjunto óptimo de descriptores que integren la función estará determinado por la actividad que se desea predecir como así también por la técnica de cribado que se implementará (Sliwoski et al., 2014).

A partir de los modelos QSAR se puede obtener como repuesta u *output* una variable dependiente continua, es decir, pueden emplearse para predecir el valor numérico de la propiedad modelada (por ejemplo,  $IC_{50}$ ,  $ED_{50}$ ), o pueden generar una respuesta discreta (modelos clasificatorios o discriminantes) cuyo valor se asocia a distintas categorías de objetos.

### 2.2.3 Métodos combinados

El enfoque de los métodos híbridos o combinados presenta dos aristas, una de ellas es que permite realizar mediante CVBL un prefiltro de bibliotecas químicas con grandes volúmenes de datos, reduciendo de esta manera el tamaño muestral de las moléculas a cribar mediante metodologías más costosas como Docking o Dinámica molecular (Q. Li & Shah, 2017). En estas situaciones es necesario utilizar un enfoque secuencial, es decir que primero se realiza la metodología de CVBL y luego se aplica, a los compuestos seleccionados, una técnica de CVBBM.

El segundo enfoque es la combinación en paralelo, donde el objetivo perseguido no es reducir la cantidad de moléculas a cribar mediante los métodos computacionalmente costosos, sino que se pretende obtener una mayor fiabilidad de los resultados computacionales (Lavecchia & Giovanni, 2013), considerando *hits* solo aquellos compuestos seleccionados por ambas técnicas de cribado; alternativamente, métodos distintos pueden conducir a conjuntos de *hits* altamente complementarios, ampliando la diversidad estructural del universo de candidatos. En general, sin embargo, las predicciones de consenso entre metodologías independientes suelen aumentar la tasa de éxitos durante las validaciones en el laboratorio húmedo.

## 2.3 Reposicionamiento de fármacos

El reposicionamiento de fármacos (también conocido en inglés como *drug repositioning*) es una estrategia que se focaliza en buscar una nueva indicación clínica a fármacos ya aprobados o al menos investigados para otras patologías (Sleire et al., 2017). Algunas de las ventajas que presenta esta estrategia frente al diseño de *novo*, es que la tasa de fracasos en fase clínica debido a problemas de seguridad y toxicidad disminuye casi un 50%, debido a que dichos estudios han sido realizados en el momento de evaluar la indicación original. Además, mitiga significativamente los elevados costos económicos del desarrollo de fármacos (se estima que podrían reducirse de US\$ 2-3 billones a US\$ 300 millones (Pushpakom et al., 2019)) y por último genera una disminución de más de 5 años en los tiempos de investigación y desarrollo comparando con los de una nueva molécula (Cha et al., 2018).

Un comentario aparte merece la reducción en los tiempos de investigación y desarrollo en el marco de la pandemia de COVID-19 declarada en marzo de 2020 por la OMS. Está claro que en este contexto debe actuarse rápido en el desarrollo de tratamientos seguros y efectivos, y el reposicionamiento de fármaco podría ser la estrategia más eficaz para hacer frente a la pandemia en el corto plazo, sobre todo garantizando seguridad a los pacientes. Es evidente que el diseño y desarrollo de una nueva entidad química, que conlleva más de una década, es una medida poco feliz para situaciones como estas ya que, dado los ciclos biológicos de los virus y sus derivadas genéticas es posible que en unos pocos años la enfermedad sea controlada y atenuada mediante epítopos del sistema inmune (Senanayake, 2020). La estrategia de reposicionamiento logró dar frutos en menos de un año, y tal es así que para noviembre de 2020 FDA aprobó dos fármacos para el tratamiento de pacientes hospitalizados por COVID-19. Uno de ellos el fármaco Remdesivir, el cual fue desarrollado originalmente contra el virus del Ébola, obteniendo su segundo uso en octubre 2020 para el virus SARCOV-2, mientras que en noviembre de 2020 la misma agencia aprobó el uso de Baricitinib (indicado inicialmente para tratamiento de artritis) en pacientes con COVID-19 (Bellera et al., 2021; Commissioner, 2020a, 2020b).

Históricamente el reposicionamiento de fármacos ha ocurrido de manera fortuita y en general debido a observaciones individuales del efecto de diferentes moléculas sobre objetivos no explorados como blanco molecular de dicha molécula (Pushpakom et al., 2019). Si bien hasta el momento los grandes logros de reposicionamiento han sido principalmente mediante serendipia, cabe mencionar que el estudio de los mecanismos de acción de los fármacos en conjunto con una mayor comprensión de las vías metabólicas involucradas en diversas patologías ha permitido identificar nuevas indicaciones clínicas, incluso frente a fármacos que habían fracasado en su indicación original (Cha et al., 2018). Este paradigma ha sentado bases sólidas dentro del sector académico y de la industria farmacéutica, quienes comparten el entusiasmo respecto a la reutilización de fármacos. Tal es así que en las últimas décadas se han desarrollado diferentes metodologías enfocadas en la reutilización de fármacos, dichos métodos comprenden desde cribados fenotípicos en laboratorios húmedos hasta algoritmos y aplicaciones computacionales para predecir nuevas actividades farmacológicas (Cha et al., 2018). Independientemente del enfoque metodológico la estrategia de reposicionamiento comprende tres pasos comunes a todas las técnicas: i) identificación de un fármaco candidato, ii) evaluación preclínica del efecto deseado, y iii) evaluación clínica de la posible nueva indicación (Pushpakom et al., 2019).

Si bien la mayor ventaja es la reducción de los tiempos y costos de inversión, las principales desventajas para el sector privado son las condiciones de patente y propiedad intelectual ya que el derecho de propiedad y explotación sobre un fármaco reposicionado es bastante más acotado que los derechos sobre una molécula novedosa (Bellera et al., 2019).

### 2.3.2 Métodos sistemáticos de reposicionamiento de fármaco

Una ventaja de sistematizar esta estrategia es que ha permitido ordenar la gran cantidad de datos sobre moléculas que se utilizan o se han utilizado como fármacos, permitiendo conocer propiedades fisicoquímicas, mecanismos de acción, efectos secundarios, propiedades farmacocinéticas y farmacodinámicas, perfiles de seguridad y eficacia, etc., y de

esta manera poder aprovechar diversas metodologías para la exploración de nuevas indicaciones clínicas (Cha et al., 2018).

### 2.3.2.1. Enfoques no computacionales

Con el avance científico sobre el conocimiento de los procesos fisiopatológicos de diversas enfermedades y la elucidación de los mecanismos de acción y efectos secundarios de muchos fármacos, se han propuesto nuevas relaciones entre patologías y mecanismos de acción, en otras palabras, predecir qué fármacos del mercado podrían resultar útiles en patologías que no cuentan todavía con tratamientos farmacológicos accesibles, o ampliar el abanico de tratamientos de una determinada patología (Naveja et al., 2016). Esta metodología se la conoce como “**estrategia de reposicionamiento basada en el conocimiento**”.

Otra metodología enmarcada en este enfoque es el cribado experimental, el cual puede realizarse mediante cribado fenotípico o cribado dirigido sobre un blanco. Independientemente de la técnica empleada, la metodología consiste en evaluar mediante ensayos *in vitro* el efecto de un fármaco sobre un microorganismo o modelo celular (cribado fenotípico) o sobre un blanco seleccionado (cribado dirigido), y testear si alguno de los fármacos resulta útil para esta nueva indicación (Pushpakom et al., 2019; Weng et al., 2018). Esta metodología tiene como desventaja que se requiere de insumos de laboratorio (reactivos, cepas o proteínas) y disponer físicamente de todos los fármacos a testear.

### 2.3.2.2 Enfoques computacionales

Los enfoques computacionales para la búsqueda de segundas indicaciones clínicas pueden englobarse en tres grandes grupos: quimioinformáticos, bioinformáticos y basado en análisis de literatura; el enfoque metodológico queda sujeto a la cantidad y calidad de los datos disponibles como así también a las herramientas computacionales de las que disponga el operador (Alberca et al., 2018).

#### 2.3.2.2.1 Metodologías quimioinformáticas

Las metodologías quimioinformáticas para el reposicionamiento de fármacos son, en líneas generales, las mismas que se discutieron con anterioridad para los métodos de cribado virtual.

Para enfocar el cribado virtual en el reposicionamiento se utilizan quimiotecas que recopilen información de moléculas tipo fármaco, algunas de las bibliotecas más conocidas son SweetLead (Novick et al., 2013), DrugBank (Wishart2017), DrugCentral (Ursu et al., 2017), y Nubbe (Pilon et al., 2017) entre otras. Además, para el caso CVBBM se pueden utilizar bases de datos que compilan blancos moleculares involucrados en diferentes patologías humanas, una de las más reconocidas es *Therapeutic Target Database* (Li et al., 2018).

#### 2.3.2.2.2 Metodologías bioinformáticas

El creciente desarrollo de los avances en biología molecular, y el aumento de disponibilidad de información genómica y transcriptómica, han permitido realizar técnicas de reposicionamiento enfocadas en la enfermedad (Cha et al., 2018), así como en

quimioinformática se parte de la premisa de que moléculas similares competirían por una determinada actividad biológica, en bioinformática se sugiere que situaciones fisiopatológicas (mutación de proteínas, alteración o disfunción de vías de señalización, etc.) similares podrían ser resueltas con los mismos fármacos (Bellera et al., 2019).

Hoy en día existen numerosos recursos de acceso público que permiten establecer conexiones fármaco-enfermedad, y entre los más interesantes podemos mencionar el proyecto *Connectivity Map* o *CMap* desarrollado por el Instituto Board (Lamb, 2006) y su versión ampliada, el proyecto *Library of integrated Network-Based Cellular Signatures (LINCS)* (Vidovic et al., 2014). Estos repositorios compilan una serie de perfiles de expresión génica derivados de tratamientos o exposiciones de las células a determinados fármacos. El algoritmo empleado intenta descubrir/inferir, mediante cambios en la expresión de genes, si existen conexiones entre genes, fármacos y enfermedades, generando en cada acierto (conexión) el inicio de una nueva etapa de estudio para un fármaco y una enfermedad determinada (Li et al., 2016).

Es importante mencionar que la finalización y publicación de los resultados del proyecto “*Human Genome Project*” han facilitado la identificación de variantes genéticas relacionadas con la aparición de diversas enfermedades, permitiendo además comprender cuáles son las vías o estructuras disfuncionales involucradas en un proceso patológico. Esto permite identificar un nuevo *target* y predecir qué fármacos podrían estar vinculados a dicho blanco por ser moduladores de una proteína similar que se haya vinculado con otras patologías humanas, o con proteínas homólogas de otras especies. Esta estrategia se la suele vincular con los estudios de fármacos sobre líneas celulares inmortalizadas (LCI), donde la observación del efecto que un fármaco ejerce sobre las células es producto de una red metabólica interna que la ciencia aun no conoce en su totalidad (Pushpakom et al., 2019).

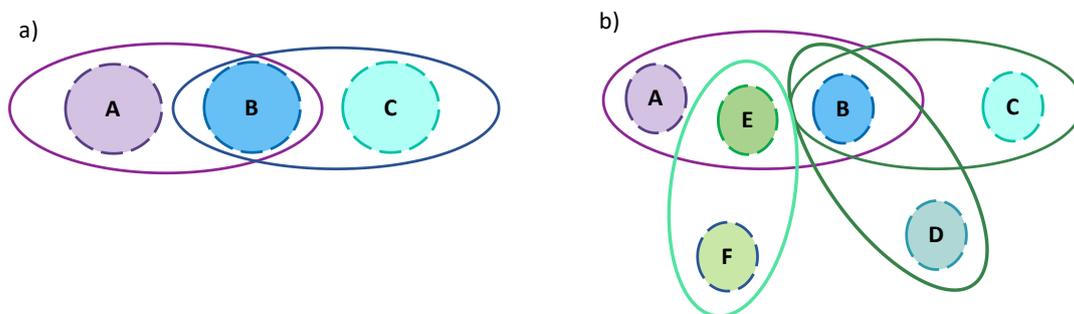
Por último, muchas proteínas que no presentan un ancestro común, o tiene un bajo coeficiente de homología estructural podrían ser útiles en técnicas de reposicionamiento de fármacos, ya que la interacciones fármaco-proteína ocurre en el sitio activo, y se ha sugerido que muchas proteínas presentan un alto grado de conservación de dicho sitio, aun cuando la homología global es muy baja. En este caso un reconocimiento bioinformático del sitio activo permitirá reconocer ligandos que interaccionen con un sitio activo similar al de otras proteínas (Bellera et al., 2019).

### 2.3.2.2.3 Métodos basados en análisis de literatura

La hipótesis publicada por Swanson en 1991 y conocida como “El modelo ABC” ha sido el puntapié inicial para el desarrollo de diversas metodologías de descubrimiento basadas en literatura (*literature-based discovery – LBD*) (Andronis et al., 2011; Deftereos et al., 2011).

El modelo propuesto establece que en el conocimiento biomédico el conjunto A y el conjunto C podrían estar interconectados de manera implícita por un conjunto de elementos o términos comunes a ambos (Swanson, 1991). El desafío de las estrategias de LDB es justamente desentrañar las conexiones implícitas entre los diferentes conjuntos, y de esta manera facilitar a la comunidad científica un método de búsqueda que le permita establecer de manera rápida las conexiones indirectas no establecidas hasta el momento (Andronis et al., 2011).

Actualmente el conocimiento biomédico es almacenado en formato de texto libre, y el continuo crecimiento de las publicaciones científicas ha conducido a la formación de “islas” de conocimiento; esto le dificulta cada vez más al investigador poder establecer relaciones directas sobre un tema entre las diferentes “islas”. Para ello se han desarrollado diversas estrategias de LBD que, mediante el análisis de grandes volúmenes de bibliografía, van reconociendo las relaciones implícitas entre la bibliografía analizada, hoy en día, esto es posible incluso cuando los artículos no utilizan exactamente la misma terminología (Andronis et al., 2011). El método de descubrimiento puede realizarse mediante un proceso cerrado o un proceso abierto. Mientras que en el primero de ellos se intenta establecer la relación que el conjunto B establece entre los conjuntos A y C conocidos, el segundo parte del conjunto A, pero desconoce el contenido del conjunto C (Deftereos et al., 2011). En la **figura 2.6** se representa un esquema básico del modelo de descubrimiento basado en literatura.



**Figura 2.6.** Modelo de descubrimiento basado en literatura. a) Modelo de *LBD* cerrado. El concepto A está relacionado con el concepto B mediante la elipse lila, el concepto B se relaciona con C mediante la elipse azul. Se puede inferir entonces que el concepto A y C están relacionados. El modelo parte de la hipótesis de que existe una relación entre A y C, e intenta validarla. Es decir, se parte de la hipótesis de dicha conexión. b) Modelo de *LBD* abierto. El concepto A se relaciona con B y E, e indirectamente se relaciona con los conceptos C, D y F mediante las otras relaciones. El modelo se llama abierto porque no se conoce de antemano las relaciones del concepto A con los conceptos C, D o F que van apareciendo a lo largo del proceso. Adaptada de (Deftereos et al., 2011).

### 2.3.3 Reposicionamiento de fármacos para ETD

Históricamente las ETD afectan a poblaciones y regiones de escasos recursos y con una alta inestabilidad política, además en muchos casos prima la escasez alimentaria y un inadecuado sistema sanitario. Este contexto, sumado a los elevados costos de inversión en el desarrollo de nuevas terapias farmacológicas permite comprender, en parte, la falta de políticas sobre investigación y desarrollo de las compañías farmacéuticas en fármacos para la prevención, diagnóstico y tratamientos de esta serie de enfermedades (Sbaraglini et al., 2018).

En el periodo de tiempo comprendido entre 1975 y 2014 se aprobaron o ingresaron a diversas fases clínicas más de 6000 moléculas tipo fármaco, sin embargo, solo 84 (<1,5%) de ellas fueron destinadas a ETD (Weng et al., 2018).

En muchos casos el campo de patologías y tratamientos abandonados por las grandes industrias farmacéuticas es abordado por el sector científico-académico, que no cuenta con el

mismo equipamiento, personal, recursos económicos y tecnológicos de las grandes industrias. Una manera de suplir estas diferencias es la búsqueda de nuevos tratamientos enfocada en la reutilización o reposicionamiento de fármacos (Klug et al., 2016). Hoy en día se estima que alrededor del 40% de los tratamientos para ETD son producto del reposicionamiento.

### 2.3.4 Casos prometedores y exitosos en tripanosomátidos.

Algunos casos de fármacos reposicionados o que se encuentran en estudios clínicos avanzados (fase II o III) para la enfermedad de Chagas han sido mencionados en el capítulo anterior. Por lo tanto, aquí solo repasaremos brevemente algunos casos exitosos del reposicionamiento de fármacos frente a otros tripanosomátidos.

En el caso de la enfermedad del sueño (causada por *T. Brucei*), muchos fármacos fueron originalmente propuestos para esta patología, pero el desarrollo de resistencias parasitarias o intolerancia de los pacientes a los tratamientos obligó al sector académico/farmacéutico a continuar con la búsqueda de nuevas alternativas. Fue así como en 1990 FDA aprobó el uso de Eflornitina, un fármaco originalmente indicado para el tratamiento antineoplásico. El fármaco Nifurtimox fue lanzado al mercado para el tratamiento de la enfermedad de Chagas, sin embargo, años después se demostró la efectividad del fármaco en infecciones causadas *T. brucei*; si bien este último no fue aprobado como monoterapia contra *T. brucei*, el mismo es utilizado en combinación con Eflornitina u otros agentes tripanocidas (Sbaraglini et al., 2016).

Otro ejemplo es el Fexinidazol, un antimicrobiano de amplio espectro que se comenzó a estudiar entre 1970-1980, luego por diversas causas se abandonó el estudio y años más tarde se rescató la droga para utilizarse contra *T. brucei*. La misma actividad se ha sugerido para el fármaco Orlistat, indicado originalmente para el tratamiento de la obesidad (Sbaraglini et al., 2016).

Varios fármacos aprobados continúan siendo evaluados frente a *Leishmania*. El primer antifúngico aprobado por la FDA para el tratamiento de *Leishmaniasis* fue la Anfotericina B (1997), un fármaco que se utilizaba en pacientes inmunosuprimidos para evitar infecciones fúngicas generalizadas; también se destacan derivados azólicos como Itraconazol y Posaconazol que demostraron eficacia frente a *Leishmania amazonensis*. Adicionalmente se han evaluado fármacos de otros grupos terapéuticos como antibióticos (Azitromicina y Paramomicina), antineoplásicos (Miltefosina, inhibidores de quinasa), inmunomoduladores (Imiquimod), Bifosfonatos e incluso algunos antiparasitarios como el Sitamaquina, el cual fue propuesto originalmente como tratamiento para la malaria (Sbaraglini et al., 2016).

## 2.4 Estrategia utilizada en el trabajo de tesis

En el presente trabajo hemos desarrollado modelos computacionales enfocados en la búsqueda de nuevos inhibidores de la enzima **TryS**. Se realizaron dos campañas de CV, las cuales se basan en la estrategia de búsqueda indirecta mediante modelos computacionales que comprenden descriptores 0-2D (QSAR-2D).

Una vez obtenidos y validados los modelos computacionales y teniendo en cuenta los exitosos resultados de reposicionamiento de fármacos (**sección 2.3.3.1**) y el contexto socioeconómico y cultural de la patología en estudio, se cribaron diferentes quimiotecas, las cuales contienen información química, estructural y clínica de fármacos aprobados y/o discontinuados, y de compuestos químicos candidatos a fármacos.

Adicionalmente, debido a los resultados obtenidos experimentalmente y a la recopilación bibliográfica sobre el blanco molecular hemos desarrollado un modelo computacional para evaluar la factibilidad de modular farmacológicamente la proteína. (**ver capítulo 5**).

### 2.5 Referencias Bibliográficas

- Alberca, L. N., Sbaraglini, M. L., Morales, J. F., Dietrich, R., Ruiz, M. D., Pino Martínez, A. M., Miranda, C. G., Fraccaroli, L., Alba Soto, C. D., Carrillo, C., Palestro, P. H., & Talevi, A. (2018). Cascade Ligand- and Structure-Based Virtual Screening to Identify New Trypanocidal Compounds Inhibiting Putrescine Uptake. *Frontiers in Cellular and Infection Microbiology*, 8, 173. <https://doi.org/10.3389/fcimb.2018.00173>
- Allarakhia, M. (2013). Open-source approaches for the repurposing of existing or failed candidate drugs: Learning from and applying the lessons across diseases. *Drug Design, Development and Therapy*, 753. <https://doi.org/10.2147/DDDT.S46289>
- Andronis, C., Sharma, A., Virvilis, V., Deftereos, S., & Persidis, A. (2011). Literature mining, ontologies and information visualization for drug repurposing. *Briefings in Bioinformatics*, 12(4), 357-368. <https://doi.org/10.1093/bib/bbr005>
- Bellera, C. L., Sbaraglini, M. L., Alberca, L. N., Alice, J. I., & Talevi, A. (2019). In Silico Modeling of FDA-Approved Drugs for Discovery of Therapies Against Neglected Diseases: A Drug Repurposing Approach. En *In Silico Drug Design* (pp. 625-648). Elsevier. <https://doi.org/10.1016/B978-0-12-816125-8.00021-3>
- Bellera, C. L., Llanos, M., Gantner, M. E., Rodriguez, S., Gavernet, L., Comini, M., & Talevi, A. (2021). Can drug repurposing strategies be the solution to the COVID-19 crisis? *Expert Opinion on Drug Discovery*, 16(6), 605-612. <https://doi.org/10.1080/17460441.2021.1863943>
- Bombaça, A. C. S., Viana, P. G., Santos, A. C. C., Silva, T. L., Rodrigues, A. B. M., Guimarães, A. C. R., Goulart, M. O. F., da Silva Júnior, E. N., & Menna-Barreto, R. F. S. (2019). Mitochondrial dysfunction and ROS production are essential for anti-Trypanosoma cruzi activity of  $\beta$ -lapachone-derived naphthoimidazoles. *Free Radical Biology and Medicine*, 130, 408-418. <https://doi.org/10.1016/j.freeradbiomed.2018.11.012>
- Cha, Y., Erez, T., Reynolds, I. J., Kumar, D., Ross, J., Koytiger, G., Kusko, R., Zeskind, B., Risso, S., Kagan, E., Papapetropoulos, S., Grossman, I., & Laifenfeld, D. (2018). Drug repurposing from the perspective of pharmaceutical companies. *British Journal of Pharmacology*, 175(2), 168-180. <https://doi.org/10.1111/bph.13798>
- Chan, H. C. S., Shan, H., Dahoun, T., Vogel, H., & Yuan, S. (2019). Advancing Drug Discovery via Artificial Intelligence. *Trends in Pharmacological Sciences*, 40(8), 592-604. <https://doi.org/10.1016/j.tips.2019.06.004>
- Commissioner, O. of the. (2020a, octubre 22). *FDA Approves First Treatment for COVID-19*. FDA; FDA. <https://www.fda.gov/news-events/press-announcements/fda-approves-first-treatment-covid-19>
- Commissioner, O. of the. (2020b, noviembre 23). *Coronavirus (COVID-19) Update: FDA Authorizes Drug Combination for Treatment of COVID-19*. FDA; FDA. <https://www.fda.gov/news-events/press-announcements/coronavirus-covid-19-update-fda-authorizes-drug-combination-treatment-covid-19>
- Consonni, V., & Todeschini, R. (2012). Multivariate Analysis of Molecular Descriptors. En M. Dehmer, K. Varmuza, & D. Bonchev (Eds.), *Statistical Modelling of Molecular Descriptors in QSAR/QSPR* (pp. 111-147). Wiley-VCH Verlag GmbH & Co. KGaA. <https://doi.org/10.1002/9783527645121.ch4>

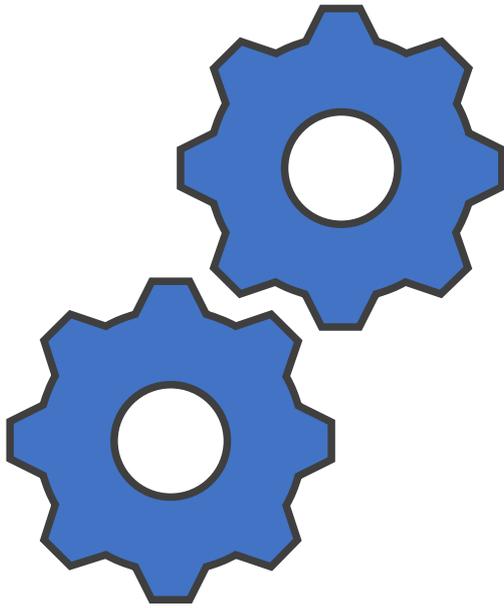
- Cornaglia, M., Lehnert, T., & Gijs, M. A. M. (2017). Microfluidic systems for high-throughput and high-content screening using the nematode *Caenorhabditis elegans*. *Lab on a Chip*, 17(22), 3736-3759. <https://doi.org/10.1039/C7LC00509A>
- Deftereos, S. N., Andronis, C., Friedla, E. J., Persidis, A., & Persidis, A. (2011). Drug repurposing and adverse event prediction using high-throughput literature analysis: Drug repurposing and adverse event prediction. *Wiley Interdisciplinary Reviews: Systems Biology and Medicine*, 3(3), 323-334. <https://doi.org/10.1002/wsbm.147>
- Dowden, H., & Munro, J. (2019). Trends in clinical success rates and therapeutic focus. *Nature Reviews Drug Discovery*, 18(7), 495-496. <https://doi.org/10.1038/d41573-019-00074-z>
- Ferreira, M. M. C. (2017). Multivariate QSAR. *Journal of the Brazilian Chemical Society*, 13(6). <https://doi.org/10.1590/S0103-50532002000600004>
- Field, M. C., Horn, D., Fairlamb, A. H., Ferguson, M. A. J., Gray, D. W., Read, K. D., De Rycker, M., Torrie, L. S., Wyatt, P. G., Wyllie, S., & Gilbert, I. H. (2017). Anti-trypanosomatid drug discovery: An ongoing challenge and a continuing need. *Nature Reviews Microbiology*, 15(4), 217-231. <https://doi.org/10.1038/nrmicro.2016.193>
- Garcia-Hernandez, C., Fernández, A., & Serratosa, F. (2019). Ligand-Based Virtual Screening Using Graph Edit Distance as Molecular Similarity Measure. *Journal of Chemical Information and Modeling*, 59(4), 1410-1421. <https://doi.org/10.1021/acs.jcim.8b00820>
- Gaulton, A., Hersey, A., Nowotka, M., Bento, A. P., Chambers, J., Mendez, D., Motow, P., Atkinson, F., & Bellis, L. J. (2018). *The ChEMBL database in 2017*. 10.
- Gavernet, L. (2021). *Introducción a la Química Medicinal*. Editorial de la Universidad Nacional de La Plata (EDULP). <http://sedici.unlp.edu.ar/handle/10915/114312>
- Gilbert, I. H. (2013). Drug Discovery for Neglected Diseases: Molecular Target-Based and Phenotypic Approaches: Miniperspectives Series on Phenotypic Screening for Antiinfective Targets. *Journal of Medicinal Chemistry*, 56(20), 7719-7726. <https://doi.org/10.1021/jm400362b>
- Gioia, D., Bertazzo, M., Recanatini, M., Masetti, M., & Cavalli, A. (2017). Dynamic Docking: A Paradigm Shift in Computational Drug Discovery. *Molecules*, 22(11), 2029. <https://doi.org/10.3390/molecules22112029>
- Grisoni, F., Ballabio, D., Todeschini, R., & Consonni, V. (2018). Molecular Descriptors for Structure–Activity Applications: A Hands-On Approach. En O. Nicolotti (Ed.), *Computational Toxicology* (Vol. 1800, pp. 3-53). Springer New York. [https://doi.org/10.1007/978-1-4939-7899-1\\_1](https://doi.org/10.1007/978-1-4939-7899-1_1)
- Kim, S., Thiessen, P. A., Bolton, E. E., Chen, J., Fu, G., Gindulyte, A., Han, L., He, J., He, S., Shoemaker, B. A., Wang, J., Yu, B., Zhang, J., & Bryant, S. H. (2015). *PubChem Substance and Compound databases*. 12.
- Klug, D. M., Gelb, M. H., & Pollastri, M. P. (2016). Repurposing strategies for tropical disease drug discovery. *Bioorganic & Medicinal Chemistry Letters*, 26(11), 2569-2576. <https://doi.org/10.1016/j.bmcl.2016.03.103>
- Koch, O., Cappel, D., Nocker, M., Jäger, T., Flohé, L., Sotriffer, C. A., & Selzer, P. M. (2013). Molecular Dynamics Reveal Binding Mode of Glutathionylspermidine by Trypanothione Synthetase. *PLoS ONE*, 8(2). <https://doi.org/10.1371/journal.pone.0056788>
- Krishna, S., Berridge, B., & Kleinstreuer, N. (2021). High-Throughput Screening to Identify Chemical Cardiotoxic Potential. *Chemical Research in Toxicology*, 34(2), 566-583. <https://doi.org/10.1021/acs.chemrestox.0c00382>

- Lamb, J. (2006). The Connectivity Map: Using Gene-Expression Signatures to Connect Small Molecules, Genes, and Disease. *Science*, 313(5795), 1929-1935. <https://doi.org/10.1126/science.1132939>
- Lavecchia, A., & Giovanni, C. (2013). Virtual Screening Strategies in Drug Discovery: A Critical Review. *Current Medicinal Chemistry*, 20(23), 2839-2860. <https://doi.org/10.2174/09298673113209990001>
- Li, J., Zheng, S., Chen, B., Butte, A. J., Swamidass, S. J., & Lu, Z. (2016). A survey of current trends in computational drug repositioning. *Briefings in Bioinformatics*, 17(1), 2-12. <https://doi.org/10.1093/bib/bbv020>
- Li, Q., & Shah, S. (2017). Structure-Based Virtual Screening. En C. H. Wu, C. N. Arighi, & K. E. Ross (Eds.), *Protein Bioinformatics* (Vol. 1558, pp. 111-124). Springer New York. [https://doi.org/10.1007/978-1-4939-6783-4\\_5](https://doi.org/10.1007/978-1-4939-6783-4_5)
- Li, Y. H., Yu, C. Y., Li, X. X., Zhang, P., Tang, J., Yang, Q., Fu, T., Zhang, X., Cui, X., Tu, G., Zhang, Y., Li, S., Yang, F., Sun, Q., Qin, C., Zeng, X., Chen, Z., Chen, Y. Z., & Zhu, F. (2018). Therapeutic target database update 2018: Enriched resource for facilitating bench-to-clinic research of targeted therapeutics. *Nucleic Acids Research*, 46(D1), D1121-D1127. <https://doi.org/10.1093/nar/gkx1076>
- Lin, S., Schorpp, K., Rothenaigner, I., & Hadian, K. (2020). Image-based high-content screening in drug discovery. *Drug Discovery Today*, 25(8), 1348-1361. <https://doi.org/10.1016/j.drudis.2020.06.001>
- Lionta, E., Spyrou, G., Vassilatis, D., & Cournia, Z. (2014). Structure-Based Virtual Screening for Drug Discovery: Principles, Applications and Recent Advances. *Current Topics in Medicinal Chemistry*, 14(16), 1923-1938. <https://doi.org/10.2174/1568026614666140929124445>
- Macalino, S. J. Y., Gosu, V., Hong, S., & Choi, S. (2015). Role of computer-aided drug design in modern drug discovery. *Archives of Pharmacal Research*, 38(9), 1686-1701. <https://doi.org/10.1007/s12272-015-0640-5>
- Maggiora, G., Vogt, M., Stumpfe, D., & Bajorath, J. (2014). Molecular Similarity in Medicinal Chemistry: Miniperspective. *Journal of Medicinal Chemistry*, 57(8), 3186-3204. <https://doi.org/10.1021/jm401411z>
- Margineanu, D. G. (2016). Neuropharmacology beyond reductionism – A likely prospect. *Biosystems*, 141, 1-9. <https://doi.org/10.1016/j.biosystems.2015.11.010>
- McInnes, C. (2007). Virtual screening strategies in drug discovery. *Current Opinion in Chemical Biology*, 11(5), 494-502. <https://doi.org/10.1016/j.cbpa.2007.08.033>
- Mullard, A. (2021). 2020 FDA drug approvals. *Nature Reviews Drug Discovery*, 20(2), 85-90. <https://doi.org/10.1038/d41573-021-00002-0>
- Naveja, J. J., Dueñas-González, A., & Medina-Franco, J. L. (2016). Drug Repurposing for Epigenetic Targets Guided by Computational Methods. En *Epi-Informatics* (pp. 327-357). Elsevier. <https://doi.org/10.1016/B978-0-12-802808-7.00012-5>
- Neves, B. J., Braga, R. C., Melo-Filho, C. C., Moreira-Filho, J. T., Muratov, E. N., & Andrade, C. H. (2018). QSAR-Based Virtual Screening: Advances and Applications in Drug Discovery. *Frontiers in Pharmacology*, 9, 1275. <https://doi.org/10.3389/fphar.2018.01275>
- Nouleho llemo, S., Barth, D., David, O., Quessette, F., Weisser, M.-A., & Watel, D. (2019). Improving graphs of cycles approach to structural similarity of molecules. *PLOS ONE*, 14(12), e0226680. <https://doi.org/10.1371/journal.pone.0226680>

- Novick, P. A., Ortiz, O. F., Poelman, J., Abdulhay, A. Y., & Pande, V. S. (2013). SWEETLEAD: An In Silico Database of Approved Drugs, Regulated Chemicals, and Herbal Isolates for Computer-Aided Drug Discovery. *PLoS ONE*, 8(11), e79568. <https://doi.org/10.1371/journal.pone.0079568>
- O'Boyle, N. M., Banck, M., James, C. A., Morley, C., Vandermeersch, T., & Hutchison, G. R. (2011). Open Babel: An open chemical toolbox. *Journal of Cheminformatics*, 3(1), 33. <https://doi.org/10.1186/1758-2946-3-33>
- Papapetropoulos, A., & Szabo, C. (2018). Inventing new therapies without reinventing the wheel: The power of drug repurposing: Editorial. *British Journal of Pharmacology*, 175(2), 165-167. <https://doi.org/10.1111/bph.14081>
- Parvathaneni, V. (2019). Drug repurposing: A promising tool to accelerate the drug discovery process. *Drug Discovery Today*, 24(10), 10.
- Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., & Ferrin, T. E. (2004). UCSF Chimera?A visualization system for exploratory research and analysis. *Journal of Computational Chemistry*, 25(13), 1605-1612. <https://doi.org/10.1002/jcc.20084>
- Pilon, A. C., Valli, M., Dametto, A. C., Pinto, M. E. F., Freire, R. T., Castro-Gamboa, I., Andricopulo, A. D., & Bolzani, V. S. (2017). NuBBEDB: An updated database to uncover chemical and biological information from Brazilian biodiversity. *Scientific Reports*, 7(1), 7215. <https://doi.org/10.1038/s41598-017-07451-x>
- Prieto-Martínez, F. D., López-López, E., Eurídice Juárez-Mercado, K., & Medina-Franco, J. L. (2019). Computational Drug Design Methods—Current and Future Perspectives. En *In Silico Drug Design* (pp. 19-44). Elsevier. <https://doi.org/10.1016/B978-0-12-816125-8.00002-X>
- Pushpakom, S., Iorio, F., Eyers, P. A., Escott, K. J., Hopper, S., Wells, A., Doig, A., Guilliams, T., Latimer, J., McNamee, C., Norris, A., Sanseau, P., Cavalla, D., & Pirmohamed, M. (2019). Drug repurposing: Progress, challenges and recommendations. *Nature Reviews Drug Discovery*, 18(1), 41-58. <https://doi.org/10.1038/nrd.2018.168>
- Sbaraglini, M., Cristina Vanrell, M., Leticia Bellera, C., Benaim, G., Carrillo, C., Talevi, A., & Silvia Romano, P. (2016). Neglected Tropical Protozoan Diseases: Drug Repositioning as a Rational Option. *Current Topics in Medicinal Chemistry*, 16(19), 2201-2222. <https://doi.org/10.2174/1568026616666160216154309>
- Sbaraglini, M. L., Bellera, C., & Talevi, A. (2018). The Challenge of Finding New Therapies for Sleeping Sickness. En *The Microbiology of Central Nervous System Infections* (pp. 279-303). Elsevier. <https://doi.org/10.1016/B978-0-12-813806-9.00014-7>
- Schneider, G. (2018). Automating drug discovery. *Nature Reviews Drug Discovery*, 17(2), 97-113. <https://doi.org/10.1038/nrd.2017.232>
- Seidel, T., Wieder, O., Garon, A., & Langer, T. (2020). Applications of the Pharmacophore Concept in Natural Product inspired Drug Design. *Molecular Informatics*, 39(11), 2000059. <https://doi.org/10.1002/minf.202000059>
- Senanayake, S. L. (2020). Drug repurposing strategies for COVID-19. *Future Drug Discovery*, 2(2), fdd-2020-0010. <https://doi.org/10.4155/fdd-2020-0010>
- Sleire, L., Førde, H. E., Netland, I. A., Leiss, L., Skeie, B. S., & Enger, P. Ø. (2017). Drug repurposing in cancer. *Pharmacological Research*, 124, 74-91. <https://doi.org/10.1016/j.phrs.2017.07.013>

- Sliwoski, G., Kothiwale, S., Meiler, J., & Lowe, E. W. (2014). Computational Methods in Drug Discovery. *Pharmacological Reviews*, 66(1), 334-395. <https://doi.org/10.1124/pr.112.007336>
- Sterling, T., & Irwin, J. J. (2015). ZINC 15 – Ligand Discovery for Everyone. *J. Chem. Inf. Model.*, 14.
- Swanson, D. R. (1991). Complementary structures in disjoint science literatures. *Proceedings of the 14th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval - SIGIR '91*, 280-289. <https://doi.org/10.1145/122860.122889>
- Swinney, D. (2013). Phenotypic vs. Target-Based Drug Discovery for First-in-Class Medicines. *CLINICAL PHARMACOLOGY*, 93(4), 3.
- Swinney, D., & Anthony, J. (2011). How were new medicines discovered? *DRUG DISCOVERY*, 13.
- Talevi, A. (2016). Tailored Multi-Target Agents. Applications and Design Considerations. *Current Pharmaceutical Design*, 22(21), 3164-3170.
- Talevi, A. (2018). Computer-Aided Drug Design: An Overview. En M. Gore & U. B. Jagtap (Eds.), *Computational Drug Discovery and Design* (Vol. 1762, pp. 1-19). Springer New York. [https://doi.org/10.1007/978-1-4939-7756-7\\_1](https://doi.org/10.1007/978-1-4939-7756-7_1)
- Ursu, O., Holmes, J., Knockel, J., Bologa, C. G., Yang, J. J., Mathias, S. L., Nelson, S. J., & Oprea, T. I. (2017). DrugCentral: Online drug compendium. *Nucleic Acids Research*, 45(D1), D932-D939. <https://doi.org/10.1093/nar/gkw993>
- Vázquez, C., Mejía-Tlachi, M., González-Chávez, Z., Silva, A., Rodríguez-Zavala, J. S., Moreno-Sánchez, R., & Saavedra, E. (2017). Buthionine sulfoximine is a multitarget inhibitor of trypanothione synthesis in *Trypanosoma cruzi*. *FEBS Letters*, 591(23), 3881-3894. Scopus. <https://doi.org/10.1002/1873-3468.12904>
- Vidovic, D., Koletić, A., & Schirer, S. C. (2014). Large-scale integration of small molecule-induced genome-wide transcriptional responses, Kinome-wide binding affinities and cell-growth inhibition profiles reveal global trends characterizing systems-level drug action. *Frontiers in Genetics*, 5. <https://doi.org/10.3389/fgene.2014.00342>
- Vivo, M. D., & Cavalli, A. (2017). Recent advances in dynamic docking for drug discovery. *WIREs Computational Molecular Science*, 7(6), e1320. <https://doi.org/10.1002/wcms.1320>
- Weng, H.-B., Chen, H.-X., & Wang, M.-W. (2018). Innovation in neglected tropical disease drug discovery and development. *Infectious Diseases of Poverty*, 7(1), 67. <https://doi.org/10.1186/s40249-018-0444-1>
- Wishart, D. S., Feunang, Y. D., Guo, A. C., Lo, E. J., Marcu, A., Grant, J. R., Sajed, T., Johnson, D., Li, C., Sayeeda, Z., Assempour, N., Iynkkaran, I., Liu, Y., Maciejewski, A., Gale, N., Wilson, A., Chin, L., Cummings, R., Le, D., ... Wilson, M. (2018). DrugBank 5.0: A major update to the DrugBank database for 2018. *Nucleic Acids Research*, 46(D1), D1074-D1082. <https://doi.org/10.1093/nar/gkx1037>
- Yang, S.-Y. (2010). Pharmacophore modeling and applications in drug discovery: Challenges and recent advances. *Drug Discovery Today*, 15(11-12), 444-450. <https://doi.org/10.1016/j.drudis.2010.03.013>
- Yap, C. W. (2011). PaDEL-descriptor: An open source software to calculate molecular descriptors and fingerprints. *Journal of Computational Chemistry*, 32(7), 1466-1474. <https://doi.org/10.1002/jcc.21707>

# Capítulo 3 | Metodología



### 3.1 Introducción

La metodología QSAR emplea una serie de pasos secuenciales para la construcción de modelos. Los mismos podrían sintetizarse en: a) **Recopilación y curado de un set de datos**, que consiste en un conjunto de compuestos químicos con sus correspondientes valores de la actividad o propiedad que se desea modelar; b) **Partición del set de datos en al menos dos subconjuntos representativos**, un conjunto de entrenamiento (o *training set*) y un conjunto de prueba (o *test set*); c) **Cálculo de descriptores moleculares** que potencialmente se incorporarán al modelo como variables independientes; d) **Construcción de modelos**, que consiste en la selección del subconjunto de descriptores que mejor correlacionan con la propiedad de interés, y la ponderación de su contribución a la explicación de la varianza de los datos experimentales. Superados esos pasos se procede a la **Validación y aplicación de los modelos**. Cada una de las etapas mencionadas se desarrollarán con mayor profundidad en el presente capítulo.

En la **figura 3.1** se puede observar un diagrama general de la metodología que empleamos para la construcción y aplicación de los modelos QSAR.

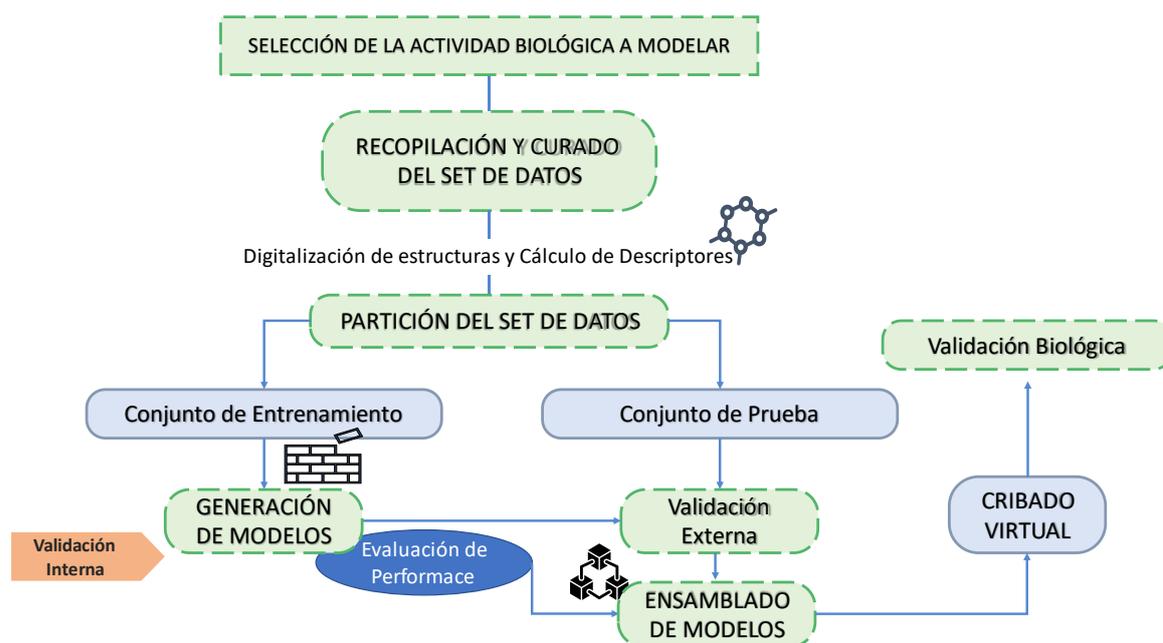


Figura 3.1 Flujo de trabajo general empleado en metodologías QSAR.

## Primera campaña de cribado virtual y evaluación experimental

### 3.2 Construcción de modelos QSAR – Funciones Discriminantes

#### 3.2.1 Recopilación y curado del set de datos

El primer paso para la construcción de modelos QSAR es la recopilación de un set de datos con valores experimentales de la actividad que se desea modelar.

Comenzado el proceso de búsqueda bibliográfica nos encontramos con que, pese a tratarse de un blanco molecular establecido para la búsqueda de nuevos agentes tripanocidas, existía un muy bajo número de compuestos químicos (menos de 30) con reportes de evaluaciones bioquímicas frente a la enzima TcTryS. Por lo tanto, y dado que para la especie *Trypanosoma brucei* (*T. brucei*) existe un mayor número de reportes de evaluaciones bioquímicas, decidimos desarrollar nuestros modelos a partir de un conjunto de entrenamiento con datos de actividad sobre esta última. La decisión se fundamenta en la gran similitud de secuencia y función que presenta *T. brucei* con los homólogos de otros tripanosomátidos (Comini et al., 2003; Oza, 2003). Debido a ello, cualquier *hit* identificado contra *T. brucei* constituye un punto de partida interesante para inhibir la enzima de *T. cruzi* y de distintas especies de *Leishmania*, estrategia que ha ido ganando impulso en los últimos años y que se conoce como reposicionamiento de blancos moleculares (Klug et al., 2016).

Consideramos, en nuestra búsqueda, aquellos artículos donde los datos reportados cumplieran con alguno de los criterios mencionados a continuación:

- ✓ Reportes del valor de IC<sub>50</sub> o Ki.
- ✓ Reporte de porcentaje de inhibición o porcentaje de actividad residual enzimática en ensayos unipuntuales (realizados a una única concentración).

Según el valor de IC<sub>50</sub> o porcentaje de inhibición reportado, los compuestos incluidos en el set de datos fueron clasificados como **ACTIVOS** (con actividad inhibitoria sobre la enzima) cuando tuvieran una IC<sub>50</sub> menor a 20 μM, o, en caso de actividad reportada a partir de un ensayo unipuntual, cuando al menos un 50% de inhibición hubiera sido reportado a concentraciones 20 μM o menores. Por otro lado, se clasificaron como **INACTIVOS** (sin actividad inhibitoria) a los compuestos que no cumplen con alguna de las condiciones anteriormente mencionadas. La búsqueda se realizó en los repositorios de artículos científicos Scopus<sup>3</sup> y PubMed<sup>4</sup>.

El curado de la base de datos es un paso crítico para desarrollar modelos con alta capacidad predictiva. Un error en la representación de la estructura será disruptivo para la relación estructura-actividad (en tanto el dato de actividad reportado se asignará a una estructura molecular errónea) y se traducirá, por tanto, en ruido durante el proceso de

---

<sup>3</sup> <https://www.scopus.com>

<sup>4</sup> <https://www.ncbi.nlm.nih.gov/pubmed>

construcción de modelos. En el presente trabajo de tesis, se recurrió a un procedimiento de estandarización para erradicar incongruencias en la representación de las estructuras moleculares. Todas las estructuras que integran la base de datos fueron digitalizadas en 2D mediante el programa Marvin Sketch (V. 6.2.2), perteneciente al paquete ChemAxon 2014, y posteriormente estandarizadas utilizando Standardizer (V. 16.9.12.0, Instant JChem). Se realizaron las siguientes operaciones descritas en el orden de aplicación:

- ✓ **Eliminar la configuración 2D** (*clean 2D*), este proceso comienza eliminando la coordenada z de cualquiera de los átomos de la molécula, y luego recalcula las coordenadas de los átomos en el plano para representar una estructura menos tensionada, con longitudes de enlace probables.
- ✓ **Aromatizar** (*aromatize*), este procedimiento convierte las diferentes formas de resonancia de sistemas aromáticos en una representación aromática canónica al convertir enlaces simples y dobles alternados en una representación aromática representándola como electrones deslocalizados.
- ✓ **Agregar los hidrógenos explícitos** (*add explicit hydrogens*) este procedimiento incorpora de manera explícita los hidrógenos que no han sido esquematizados durante el proceso de digitalización.
- ✓ **Neutralizar** (*neutralize*) esta acción neutraliza las cargas moleculares.

Luego de estas acciones se procedió al análisis y remoción de compuestos duplicados, mediante el programa JChem V.6 (ChemAxon2013) y finalmente a una inspección visual para corroborar la bondad de las representaciones.

Para caracterizar la diversidad molecular del set de datos final se recurrió al cálculo del coeficiente de Tanimoto, utilizando para ello el programa JChem V.6 del paquete ChemAxon (2013); eligiendo el sistema de fingerprinting ECFP\_4 y calculando el coeficiente de Tanimoto en la forma binaria. Se construyeron mapas de calor para ilustrar la diversidad molecular intra- e inter-categorías, utilizando para ello el programa Gitoools v. 2.3.1 (Perez-Llamas & Lopez-Bigas, 2011).

### 3.2.2 Partición del set de datos

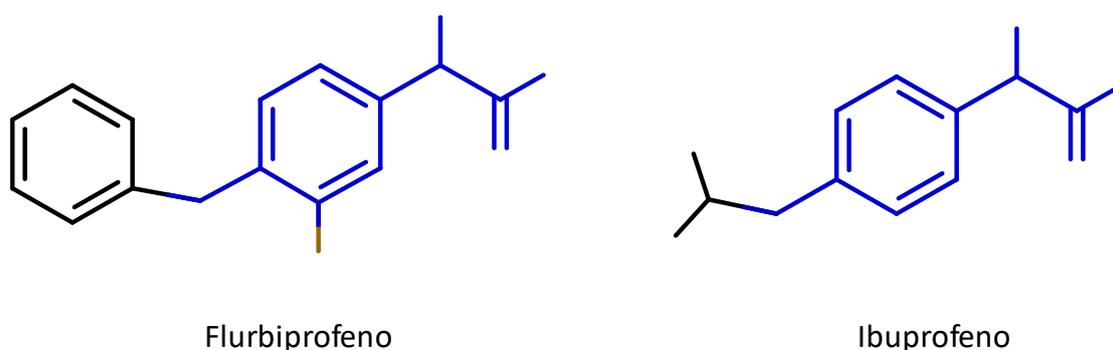
La partición del set de datos en conjuntos de entrenamiento y conjunto de prueba debe llevarse a cabo de manera representativa y estratificada, de modo de abarcar, con los subconjuntos, el mayor espacio químico posible (Moreno-Torres et al., 2012). La partición representativa (habitualmente denominada “selección racional”) provee, de este modo, modelos con mayor capacidad predictiva (Leonard & Roy, 2006; Martin et al., 2012).

La técnica de agrupamiento (o *clustering*) tiene como objetivo reunir elementos con la mayor similitud posible dentro de un mismo grupo, maximizando la variabilidad entre los elementos de grupos diferentes (Bansal et al., 2017). Diversos autores señalan, que, si bien las técnicas de agrupamiento suelen ser bastantes robustas, no dejan de presentar un factor de subjetividad cuando en la selección inicial del número e identidad de las semillas de los grupos considerados interviene un operador humano; por ello, se recomienda la combinación de al menos dos técnicas de *clustering* diferentes (Nazeer & Sebastian, 2009).

Para minimizar un posible sesgo vinculado a esta problemática, realizamos la partición del set de datos combinando de manera secuencial dos técnicas de agrupamiento. En primer término, ejecutamos un agrupamiento jerárquico; de los grupos resultantes de esa primera etapa se tomaron, aleatoriamente, las semillas que se utilizarían en la segunda técnica, basada en un método de agrupamiento no jerárquico. Este procedimiento se llevó a cabo solo para el grupo de compuestos **ACTIVOS**; dado que los compuestos **INACTIVOS** fueron reservados en su totalidad para conformar el conjunto de entrenamiento. La incorporación de todos los **INACTIVOS** en dicho conjunto se fundamenta en la escasez de datos experimentales correspondientes a esa categoría que encontramos reportados en literatura. Esta escasez relativa de datos de compuestos **INACTIVOS** resulta, al menos en parte, del habitual sesgo hacia el reporte de datos positivos que se observa en la comunidad científica (Mlinarić et al., 2017; Montori et al., 2000; Sharma & Verma, 2019).

### 3.2.2.1 Agrupamiento Jerárquico: Library MCS

El agrupamiento jerárquico se realizó mediante el programa Library MCS, disponible en el paquete JChem 6.3.0 (ChemAxon 2014). Esta técnica calcula la similitud molecular de todos los pares de moléculas presentes en el conjunto analizado (en nuestro caso, el conjunto de moléculas de la categoría **ACTIVOS**); luego, busca la máxima subestructura común (MCS) del par de moléculas con mayor similitud. Todas las moléculas que comparten esa máxima subestructura común se agrupan en el mismo cluster; el procedimiento se repite con las estructuras remanentes hasta que el algoritmo no encuentre compuestos que compartan una subestructura en común con el mínimo número de átomos especificado por el usuario (Zhang et al., 2015), que en nuestro caso se definió en 7 átomos, permitiéndole al algoritmo utilizar fragmentos de anillos, sin necesidad de considerar la totalidad de éstos. En la **figura 3.2** se presenta, a modo de ejemplo, dos moléculas que comparten una MCS.



**Figura 3.2** Ejemplo de Máxima Subestructura Común entre Ibuprofeno y el Flurbiprofeno. Puede observarse que la subestructura de color azul es común a ambos compuestos.

### 3.2.2.2 Agrupamiento No Jerárquico: K-means

A partir del agrupamiento inicial mediante el método jerárquico, se implementó el método de optimización *K-means*; incorporado en el software Statistica 10 (Statsoft, 2011). Se trata de un algoritmo iterativo, para generar  $K$  grupos de  $n$  elementos; para ello, el primer paso consiste en la asignación de elementos (semillas) que servirán para la primera ronda de agrupamiento. En esta ronda, el algoritmo asigna cada elemento a agrupar al grupo correspondiente a la semilla más cercana, utilizando un criterio de distancia para la asignación. Una vez que todos los elementos son incluidos en alguno de los grupos, se calcula el centroide de cada grupo y se reagrupan los elementos, ahora, conforme a la distancia a los centroides. Este procedimiento se repite hasta que en dos rondas sucesivas de agrupamiento no se modifica la composición de los grupos, cuando se considera que el algoritmo ha logrado la convergencia (Bansal et al., 2017). En su defecto, el usuario puede definir el número máximo de ciclos utilizados por el algoritmo. El cálculo de distancia aplicado se explicita en la **ecuación 3.1**.

$$\sum_{m=1}^k \sum_{l=1}^{nm} (X_{ml} - \check{X}_m)^2 \quad \text{Ec. 3.1}$$

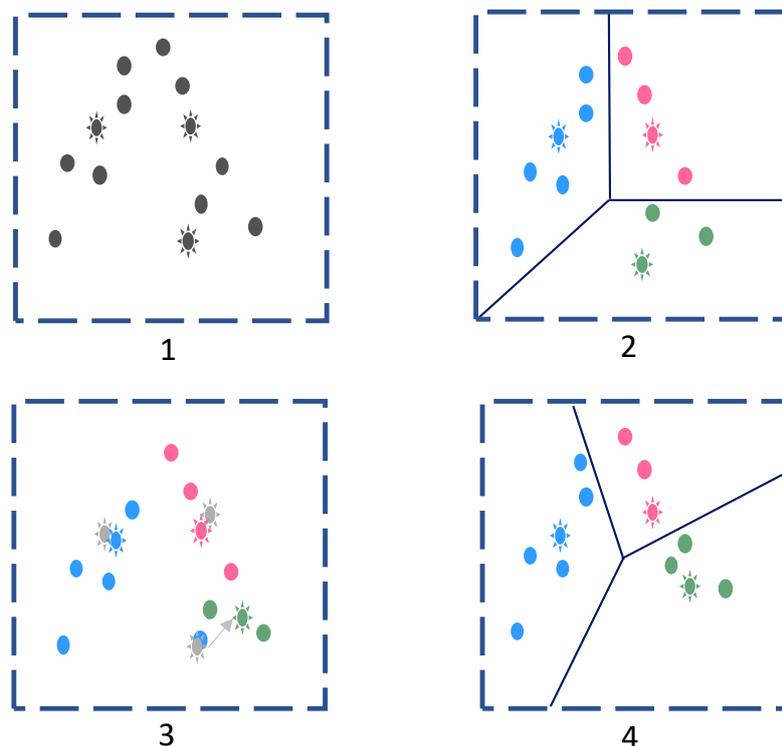
Donde  $k$  representa el número de grupos en los que se particionan los objetos,  $nm$  es el número ( $n$ ) de objetos en el grupo  $m$ ,  $x_{ml}$  denota al vector p-dimensional (considerando que se utilizan p variables para caracterizar cada objeto) del objeto  $l$  del grupo  $m$ , y  $\check{X}_m$  es el vector que representa el vector correspondiente al centroide de los elementos del grupo  $m$  en el espacio p-dimensional.

Se seleccionaron siete descriptores para caracterizar las moléculas a agrupar: peso molecular (**MW**); suma de los radios atómicos de Van der Waals (**Sv**); índice de contenido atómico total (**IAC**); número de átomos donores de hidrógenos (**nHDon**); número de átomos aceptores de hidrógenos (**nHDAcc**); área topológica de superficie polar (**TPSA (Tot)**) y el coeficiente de partición octanol-agua estimado teóricamente según el método de Moriguchi (**MlogP**). En esta campaña limitamos a 50 el número de iteraciones realizadas por el algoritmo *k-means*. En la **figura 3.3** podemos observar una representación gráfica del proceso matemático que lleva a cabo este algoritmo. Los siete descriptores fueron normalizados según la **ecuación 3.2**.

$$Z = \frac{X-N}{\sigma} \quad \text{Ec. 3.2}$$

Donde  $Z$  es el valor del descriptor normalizado,  $X$  representa el valor del descriptor molecular (sin normalizar),  $N$  es la media aritmética del descriptor para el conjunto de compuestos considerados, y  $\sigma$  denota la desviación estándar.

Los descriptores fueron calculados con el programa Dragon V.6.0 (Milano Chemometrics, 2011).



**Figura 3.3** Esquema simplificado de *K-means*. 1) Es el primer paso y corresponde a la elección de la semilla, que se realiza una única vez. 2) El método agrupa los compuestos según la distancia a la semilla más cercana. 3) El método calcula el centroide de cada grupo. 4) Nuevo reagrupamiento de moléculas en función de la distancia al centroide más cercano. La secuencia de los pasos 2, 3 y 4 se realiza de manera iterativa hasta convergencia. Las semillas y centroides se representan como puntos estrellados.

### 3.2.3 Selección de los conjuntos de entrenamiento y de prueba

Una vez ejecutado el agrupamiento se procedió a la conformación del conjunto de entrenamiento y del conjunto de prueba.

El conjunto de entrenamiento se conformó con el 70 % de los compuestos **ACTIVOS** tomando, de manera representativa, moléculas de cada uno de los grupos, e incorporándose a este conjunto la totalidad de los compuestos **INACTIVOS**. El conjunto de prueba se conformó con **ACTIVOS** remanentes y, dado que no se contaba con compuestos **INACTIVOS** más que los asignados al conjunto de entrenamiento, se generó un conjunto de señuelos (presuntos inactivos) utilizando la plataforma online *Database of Useful Decoys: Enhanced* (DUD-e, Mysinger et al., 2012). Los señuelos (*decoys*) DUD-e, presentan características fisicoquímicas similares a los compuestos **ACTIVOS** (por ejemplo, log P, MW, composición atómica, etc.), pero son topológicamente diferentes a estos. Las propiedades fisicoquímicas similares entre los señuelos y los compuestos del conjunto de prueba generan un desafío más riguroso para los modelos, mientras que las diferencias en la topología se requieren para minimizar la probabilidad de que los *decoys* se unan al mismo sitio activo que el compuesto activo en el blanco molecular de interés. La herramienta DUD-e genera 50 señuelos por cada molécula cargada a la plataforma y dado que cada señuelo proviene de un compuesto **ACTIVO**, la

estructura obtenida por la aplicación se considera presunto **INACTIVO**. De esta manera el conjunto de prueba quedó conformado por 19 **ACTIVOS** y 950 **señuelos**.

### 3.2.4 Cálculo de descriptores

Con el programa Dragon V.6.0 (Milano Chemometrics, 2011) se calculó un pool de descriptores moleculares para cada uno de los compuestos. Los cálculos se realizaron de manera independiente para el conjunto de entrenamiento y el conjunto de prueba. Se calcularon exclusivamente los descriptores de DRAGON independientes de la conformación, obteniéndose en total 3668 descriptores moleculares.

### 3.2.5 Construcción de los modelos

Durante el proceso de modelado se seleccionan del pool de variables independientes aquellas que mejor reflejen la relación entre la actividad biológica y la estructura molecular de los compuestos del conjunto de entrenamiento. Se pretende, sin embargo, que la ecuación resultante corresponda a una relación general, aplicable a compuestos externos al conjunto de entrenamiento, y que no capture particularidades no generalizables de los ejemplos de entrenamiento (Roy & Mitra, 2011).

Dado que los datos recopilados suelen provenir de ensayos desarrollados en distintos laboratorios, no siempre bajo condiciones experimentales estrictamente idénticas, como se mencionó en secciones anteriores, es importante poder diferenciar la señal del ruido (Vyškovský et al., 2016), y en este sentido los modelos clasificadores resultan útiles ya que permiten mitigar el ruido de datos experimentales recabados mediante ensayos no estandarizados (Talevi et al., 2012). Hemos asignado un valor numérico arbitrario (etiqueta de clase) a cada una de las categorías (**ACTIVOS** e **INACTIVOS**): a los compuestos **ACTIVOS** se asignó el valor "1", y a los **INACTIVOS** el valor "0".

### 3.2.6 Aleatorización de subespacios

La metodología de subespacios aleatorios consiste en aleatorizar el pool de variables independientes y a partir de allí tomar pequeños subconjuntos de variables, de modo tal de diversificar los descriptores moleculares incorporados a los modelos y evitar que los algoritmos de selección de descriptores tiendan a concentrarse, reiteradamente, en unos pocos descriptores con alta capacidad explicativa (Daho & Chikh, 2015; Vyškovský et al., 2016). Esta estrategia se conoce, también, como *feature bagging*. Se trata de una técnica que explora el espacio de las posibles variables independientes de manera estocástica.

En este trabajo, aleatorizamos y particionamos 3668 descriptores moleculares independientes de la conformación en 1000 subconjuntos de 200 descriptores cada uno. Luego, se generó un modelo clasificador lineal por cada subconjunto.

### 3.2.7 Selección de descriptores moleculares y generación de modelos

Las metodologías QSAR utilizan descriptores moleculares para relacionar la actividad biológica o propiedad de interés con características moleculares de un compuesto determinado. Actualmente, se dispone de una gran cantidad y variedad de descriptores moleculares. Este hecho da lugar a algunas problemáticas. Por un lado, la redundancia de información extraída, dado que muchos descriptores presentan un alto grado de correlación entre sí. La segunda dificultad que se debe controlar es que muchas de estas variables son irrelevantes para el estudio de una actividad o propiedad particular (González et al., 2008).

La incorporación de una variable que resulte relevante y no redundante para un modelo es un tema de estudio y discusión constante, ya que no existe un algoritmo matemático que optimice ambos parámetros en simultáneo (González et al., 2008). Sin embargo, existen varios métodos que permiten un balance “óptimo” entre estos parámetros. Entre algunos de los más utilizados, por cuestiones de costo computacional, encontramos las metodologías de regresión progresiva (por pasos sucesivos, hacia atrás o hacia adelante). La selección por **pasos hacia atrás** comienza con la incorporación de todas las variables a la función matemática, retirando una variable en cada paso, lo cual resulta útil para el manejo de pequeños pools de descriptores. El procedimiento opuesto, la selección **por pasos hacia adelante** comienza con una función desierta de variables, sumando un descriptor en cada paso. Este procedimiento se detiene cuando no quedan variables fuera de la función que modifiquen significativamente el rendimiento del modelo, o cuando la incorporación de una nueva variable produce un descenso en el desempeño de la función (Celeux et al., 2019).

En este trabajo, y dada la cantidad de variables en cada subconjunto de descriptores (200, según se discutió en la **sección 3.2.6**), hemos aplicado la técnica de semicorrelación utilizando la modalidad de selección por pasos hacia adelante. En la **figura 3.4** se puede apreciar gráficamente la diferencia entre las técnicas de **correlación tradicional** y las técnicas de **semicorrelación**, como la que empleamos en esta tesis.

## CORRELACIÓN CLÁSICA

## SEMI-CORRELACIÓN

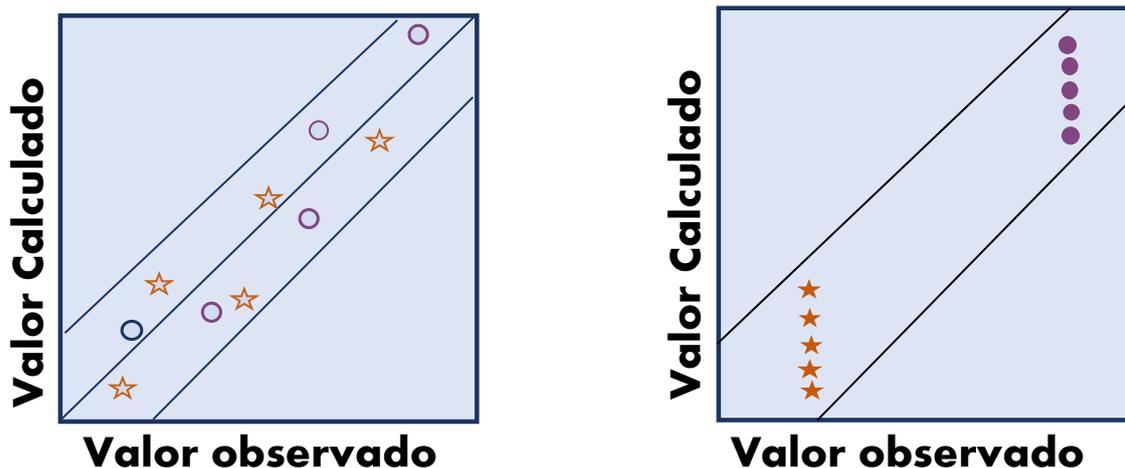


Figura 3.4 Gráfica ilustrativa de una correlación clásica vs una semi-correlación. Adaptado de Toropova et al., 2019.

Dado que la cantidad de pasos está directamente relacionada con el número de descriptores incorporados en el modelo, es fundamental el control de este parámetro para reducir la probabilidad de sobreajuste en la función. Babyak pone en evidencia cómo la incorporación de un número excesivo de variables, incluso siendo aleatorias, podrían generar un modelo con capacidad explicativa / descriptiva. Se desprende que, de excedernos en el número de variables independientes, la función generada tendrá buen desempeño dentro del conjunto de entrenamiento, pero fallará al intentar predecir los compuestos que se encuentren fuera del mismo (Babyak, 2004). Debido a esto, en la presente campaña hemos limitado el número máximo de pasos a un paso por cada diez compuestos del conjunto de entrenamiento.

Otro aspecto crítico en los métodos de análisis por regresión es la colinealidad o correlación de las variables involucradas en un modelo, evaluado generalmente por medio de la tolerancia  $\beta$  del modelo o su inversa, **el factor de inflación de la varianza (FIV)** (Daoud, 2017; García et al., 2015). Muchas veces es difícil encontrar un modelo de regresión que incluya variables absolutamente independientes entre sí, pero, aun así, es posible construir una función que minimice la colinealidad de variables, y, por lo tanto, no se afecte de manera significativa el desempeño del modelo). La tolerancia  $\beta$  se define en la **ecuación 3.3**.

$$\beta_j = 1 - R_j^2 \quad \text{Ec. 3.3}$$

Donde  $R_j$  es el coeficiente de correlación entre la variable independiente  $j$  y el resto de las variables independientes incluidas en el modelo. En el presente trabajo hemos utilizamos un valor de tolerancia  $\beta = 0,5$ , lo que se corresponde con un FIV = 2. De esta manera, cuando un par de descriptores presenta una correlación mayor a 0,5, el modelo incluye solo uno de ellos.

Según el método de semi-correlación utilizado, construimos funciones con la forma general que se muestra en la **ecuación 3.4**, para un modelo de  $k$  variables independientes.

$$FD = a_0 + \sum_{i=1}^k a_i d_i \quad \text{Ec. 3.4}$$

Donde **FD** representa el valor que asume la función discriminante;  $a_0$  es la ordenada al origen y  $a_i$  el coeficiente asociado al descriptor  $d_i$ .

### 3.3 Validación de los modelos

La validación es uno de los pasos críticos de las metodologías QSAR, ya que este aspecto nos garantizará confiabilidad y relevancia de los procesos realizados (Ravichandran Veerasamy et al., 2011). Dado que los modelos desarrollados mediante estas metodologías son, en general, aplicados en la selección de moléculas para ensayos experimentales, un aspecto fundamental de ser evaluado es la capacidad de generalización del modelo; por otra parte, se estudiará su robustez (es decir, que pequeñas perturbaciones en el procedimiento de modelado no impacten significativamente en el desempeño del modelo) y se evaluará si el procedimiento utilizado es factible de resultar en correlaciones espurias.

Los métodos de validación se pueden clasificar según el conjunto con el que se esté realizando la prueba. En este aspecto, se llama **validación interna**, a los procedimientos de validación que reutilizan el propio conjunto de entrenamiento, mientras que llamamos **validación externa** a los procedimientos que recurren a uno o más conjuntos de prueba, independientes de los ejemplos de entrenamiento (Roy & Mitra, 2011).

#### 3.3.1 Validación Interna

Las instancias de validación interna apuntan a evaluar la robustez del modelo, excluyendo o minimizando, por otra parte, la posibilidad de que el modelo obtenido sea producto del azar.

En el presente trabajo, evaluamos los modelos mediante las técnicas de **Aleatorización de Fisher** (también conocida como *Y-Randomization*) y **Validación Cruzada** (en inglés, *Cross Validation*).

##### 3.3.1.1 Aleatorización de Fisher

El carácter predictivo de un modelo puede resultar escaso cuando el número de ejemplos de entrenamiento es bajo y el número de variables independientes que se utilizan para describir dichos elementos es alto (Babyak, 2004; Kaneko, 2019). Es bien sabido que el uso de un pool excesivamente grande de variables independientes puede conducir a correlaciones espurias. Para descartar la probabilidad de generación azarosa de los modelos, se puede evaluar a los mismos mediante la técnica de **Aleatorización de Fisher**. Dicha técnica consiste en reasignar las categorías del conjunto de entrenamiento, de manera aleatoria (rompiendo así la relación estructura-actividad observada), y generar nuevos modelos a partir

de esta reasignación de categorías. Este proceso se repite numerosas veces, comparando los resultados obtenidos por los nuevos modelos con respecto a los modelos originales (los que se obtuvieron a partir de una variable dependiente sin aleatorizar). Se espera que los modelos aleatorizados presenten un pobre desempeño. Cuanto más disímil sean los resultados de los modelos obtenidos con y sin aleatorización, más confiable resultará el modelo verdadero (Veerasamy et al., 2011).

Para realizar la validación mediante esta técnica, hemos aleatorizado las etiquetas de los compuestos del conjunto de entrenamiento. Por ende, algunos de los compuestos **ACTIVOS** se etiquetaron como **INACTIVOS** y viceversa. Con estas nuevas etiquetas reconstruimos los modelos, y evaluamos el desempeño de estos nuevos modelos como porcentajes de buenas clasificaciones (%BC). Este procedimiento se repitió 50 veces, y el promedio de los %BC se comparó con los resultados obtenidos para los modelos originales.

### 3.3.1.2 Validación Cruzada

La **Validación cruzada** es una técnica que nos permite evaluar la posibilidad de sobreajuste de los modelos, así como la eventual presencia de datos excesivamente influyentes. El término sobreajuste hace referencia a que el modelo podrá, dentro del conjunto de entrenamiento, relacionar correctamente la categoría real de un compuesto con la predicción del modelo, pero fallará al intentar dar una predicción sobre un compuesto externo a dicho conjunto (Veerasamy et al., 2011). Es decir, que el modelo presenta capacidad descriptiva / explicativa pero escasa capacidad predictiva.

El método de Validación cruzada consiste en remover una molécula (modalidad conocida en inglés como *Leave One Out*), o un grupo de **n** moléculas (*Leave Group Out*) del conjunto de entrenamiento y generar nuevamente modelos utilizando, para ello, el conjunto de entrenamiento que se genera luego de la remoción de moléculas. La remoción de moléculas en grupos se puede efectuar de manera totalmente aleatorizada o estratificada (Roy & Ambure, 2016; Varoquaux et al., 2017). El procedimiento se repite de manera iterativa, según los criterios del operador.

En esta campaña realizamos la validación cruzada bajo modalidad LGO, removiendo 5 moléculas de cada categoría por ronda de validación. Con el conjunto restante de ejemplos de entrenamiento (80 compuestos) generamos nuevos modelos clasificatorios, los cuales fueron evaluados frente al grupo de moléculas retiradas. Este procedimiento se repitió 50 veces de manera sistemática, habiendo removido cada compuesto del conjunto de entrenamiento al menos una vez. Los resultados se expresaron como el promedio del %BC obtenidos.

### 3.3.2 Validación externa

Diversos autores coinciden en que la manera más rigurosa de estimar la capacidad predictiva de un modelo QSAR, es evaluar dicha capacidad sobre un conjunto externo al conjunto de entrenamiento (Gramatica, 2013; Veerasamy et al., 2011).

Como hemos mencionado en la **sección 3.2.3**, nuestra metodología contemplaba desde el inicio dos conjuntos, uno de entrenamiento y otro de prueba. Este último conjunto se

ha reservado para evaluar la capacidad predictiva de los modelos generados. Debido al desbalance de clases, menos del 2% de **ACTIVOS**, se ha evaluado el desempeño de los modelos mediante %BC global, es decir porcentaje total de compuestos bien clasificados por los modelos. Cabe mencionar que, en las instancias de validación, el desbalance de clases representa un desafío real y más riguroso para los modelos, ya que, en general, en una aplicación de cribado virtual prospectivo, la cantidad de compuestos **ACTIVOS** será notablemente inferior al número total de compuestos químicos en la quimioteca cribada (Kubát & Miroslav, 2017).

### 3.4 Comportamiento y desempeño de los modelos

Los resultados de los modelos QSAR binarios (diseñados para predecir dos clases o categorías) pueden analizarse recurriendo a una matriz de confusión o contingencia, en la que se pueden identificar 4 categorías: **Verdaderos Positivos (VP)**, **Verdaderos Negativos (VN)**, **Falsos Positivos (FP)** y **Falsos Negativos (FN)**. En la categoría verdaderos se incluyen aquellos elementos que fueron bien clasificados por el modelo, es decir, un VP es un elemento positivo real que el modelo lo clasificó como tal. Mientras que la categoría falsa corresponde al caso de un elemento positivo o negativo que el modelo clasifica en la categoría opuesta (Gramatica, 2013). Una vez generada esta matriz de confusión, podemos estimar el desempeño o poder predictivo de los modelos a partir de dos parámetros fundamentales: Sensibilidad (**Se**) y Especificidad (**Sp**). (Ballabio et al., 2018).

**Sensibilidad:** Es la tasa o proporción de **Verdaderos Positivos (VP)**. En nuestro trabajo, este parámetro corresponde a los compuestos **ACTIVOS** que el modelo clasificó como tales. La **Se** se puede calcular a partir de la **ecuación 3.5**.

$$Se = \frac{VP}{VP+FN} \quad \text{Ec.3.5}$$

**Especificidad:** Es la tasa o proporción de **Verdaderos Negativos (VN)**. En nuestro trabajo, este parámetro corresponde a los compuestos **INACTIVOS** que el modelo clasificó en dicha categoría. La **Sp** se puede calcular utilizando la **ecuación 3.6**.

$$Sp = \frac{VN}{VN+FP} \quad \text{Ec.3.6}$$

Observando las **ecuaciones 3.5** y **3.6**, podemos deducir que estos dos parámetros (**Se** y **Sp**) se comportan de manera inversa, es decir que cuando uno crece el otro disminuye. Por lo tanto, si modificáramos las condiciones del modelo clasificador para mejorar una de las métricas, indefectiblemente estaríamos “empeorando” la otra. En consecuencia, es importante poder encontrar un balance adecuado tal que nos permita alcanzar un equilibrio apropiado entre ambos parámetros. El estudio de las Curvas de caracterización de Funcionamiento del Receptor suele ser de gran utilidad para este tipo de análisis.

### 3.4.1 Análisis de las curvas ROC

Las curvas *Receiver Operating Characteristics* (ROC), resultan una herramienta visual útil para estudiar el desempeño de un clasificador y, ocasionalmente, comparar el desempeño de distintos clasificadores. Estas curvas permiten evaluar cómo se modifica el desempeño del modelo en base al valor de corte seleccionado para el *score*, ya que en el eje **Y** se grafica la **Se**, y en el eje **X**, **1-Especificidad (1-Sp)**. Por lo tanto, para un modelo clasificador esperaríamos, en términos ideales, obtener un punto de corte tal que se encuentre localizado en el extremo superior izquierdo; donde la relación **VP/FP** es mayor; es decir una alta tasa de **VP** y baja cantidad de **FP** (Gramatica, 2013; Roy, Kar, & Das, 2015).

Otra métrica interesante es el valor del **Área Bajo la Curva ROC (AUCROC)**, ya que nos posibilita evaluar si un modelo se comporta significativamente mejor que una clasificación al azar, permitiendo, también, la comparación estadística de entre los distintos modelos generados (Triballeau et al., 2005).

El AUCROC se mueve entre **0** y **1**. El máximo valor indica que el modelo se comporta como un clasificador perfecto, es decir, que recupera el 100% de los **ACTIVOS** y ningún **INACTIVO**, mientras que un valor de 0,5 se corresponde a una clasificación al azar. En la **figura 3.5** podemos observar cómo se mueven en direcciones opuestas los valores de **Se** y **Sp**, y en la misma gráfica (B) podemos visualizar el gráfico de AUCROC, donde la diagonal de pendiente=1, representa la clasificación aleatoria, es decir, la ausencia de capacidad predictiva.

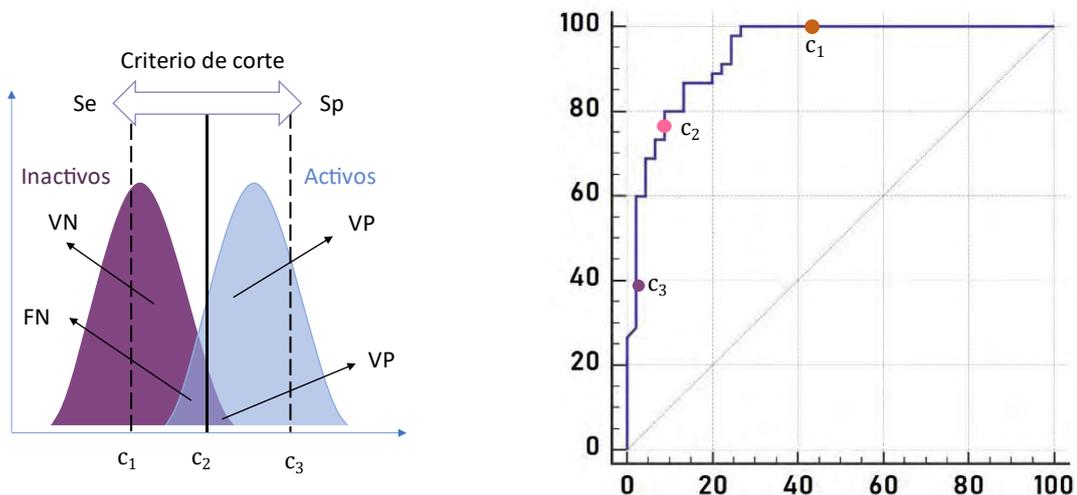


Figura 3.5 Esquema de construcción de una curva ROC.

En el presente trabajo se utilizó el paquete *pROC* del entorno R para calcular el AUCROC de los 1000 modelos generados (Robin et al., 2011); de esta manera evaluamos dicha métrica en el conjunto de prueba. Adicionalmente realizamos los gráficos de curvas con el programa MedCalc (MedCalc Software 2011), que nos permite, además, su comparación estadística mediante el método no paramétrico propuesto por De Long (DeLong et al., 1988).

### 3.4.2 Campaña retrospectiva de cribado virtual

Mediante diferentes experimentos de simulación de cribado retrospectivo, Truchon y colaboradores han demostrado que las métricas de AUCROC y sus desviaciones dependen fuertemente de la composición de la base de datos en estudio, en particular de la relación **ACTIVOS/INACTIVOS** de la misma. Los mismos autores proponen que cuando la proporción de **ACTIVOS** es menor a 0,05 (<5%) del total de la base de compuestos, entonces las desviaciones estándar de las métricas convergen a un valor constante, evitando además la probabilidad de saturación, lo que contribuye a un resultado más robusto (Truchon & Bayly, 2007).

Es importante señalar este concepto, dado que los modelos clasificatorios generados, serán aplicados al cribado virtual sobre bases de compuestos de gran tamaño, donde, sabemos, la proporción de **ACTIVOS** suele ser muy pequeña.

La validación externa se llevó a cabo realizando un cribado retrospectivo frente al conjunto de prueba, generado como se mencionó en la **sección 3.2.3**, y donde la proporción de **ACTIVOS** es menor a 0,02.

## 3.5 Combinación de Modelos

La combinación o ensamblado de modelos clasificadores, más conocido en inglés como *Ensemble Learning*, es una técnica de aprendizaje automático que se basa en combinar, de diversas maneras, el *score* o puntaje asignado por modelos individuales para generar un meta-clasificador. Se ha observado que el ensamblado de modelos mejora la capacidad predictiva (Alberca et al., 2018; Krawczyk et al., 2017); incluso se ha sugerido que podría expandir el dominio de aplicación (Zhu et al., 2008), además de mitigar la influencia de datos ruidosos (Polanski et al., 2006).

Diversos autores coinciden en que un ensamble presenta mayor robustez a la hora de predecir las categorías de un espacio muestral; esta mayor robustez está dada debido a que un modelo puede presentar sesgo de características o ser deficiente en la información incorporada por sus variables, pero difícilmente todos los modelos presenten el mismo tipo de sesgo o falta de consistencia en la información de las variables que lo componen (Wang et al., 2014).

Por otro lado, Wang y col. y Krawczyk y col., en diversos trabajos, han demostrado que si bien el mejor modelo individual suele presentar un valor relativamente alto en alguna de las métricas utilizadas para su evaluación, también es común que presente sobreajuste o caída en alguna otra métrica de desempeño; por lo tanto, aunque la combinación de modelos podría exhibir una ligera disminución en algunas métricas, en relación al mejor modelo clasificador, siempre presentará mayor robustez, garantizando que ante una pequeña modificación de alguna de las variables que lo conforman no se modifique el desempeño del clasificador de manera significativa (Krawczyk et al., 2017; Wang et al., 2014).

Para mejorar el desempeño de los modelos obtenidos recurrimos a la técnica de ensamblado mediante 4 operadores diferentes: MÍNIMO; PROMEDIO, PROMEDIO del

*RANKING* y PROMEDIO del VOTO. El operador **MÍNIMO** devuelve como salida (*output*), para cada molécula, el menor valor de *score* entre todos los asignados por los modelos individuales que componen el ensamblado. En el caso del operador **PROMEDIO** la salida es el promedio de los valores de *score* de los modelos individuales que componen el ensamble. El **PROMEDIO de RANKING** consiste en ordenar de manera decreciente a los compuestos, según el valor de *score* de cada modelo individual, y luego obtener el valor medio de dicho *Ranking*. El *output* del **PROMEDIO del VOTO** es el valor que se obtiene a partir de la **ecuación 3.7**:

$$Voto_j = \text{máx}(0, \text{int}(11 - \frac{Rank_j}{0,02N_{BD}})) \quad \text{Ec.3.7}$$

Donde  $Rank_j$  corresponde a la posición que ocupa el compuesto  $j$  en el *ranking* cuando los mismos se ordenan de manera decreciente según el valor de *score*, mientras que  $N_{BD}$  representa el número de compuestos totales en la quimioteca cribada. El valor *Voto* se obtiene al promediar los votos obtenido por cada compuesto (Alberca et al., 2019).

En esta campaña de CV se aplicaron los cuatro operadores antes mencionados a los modelos que mejores valores de AUCROC presentaron frente al conjunto de prueba. Así construimos, para cada operador, cuatro ensambles diferentes en los cuales se incluyeron 5, 10, 15 y 20 modelos individuales de manera secuencial; es decir, que el ensamble de 10 modelos contenía los cinco modelos incluidos en el primer ensamble más las siguientes cinco funciones en orden decreciente de AUCROC.

### 3.6 Selección del valor de corte

Un paso crítico en la aplicación de modelos QSAR es la selección de un adecuado valor de corte del *score*.

Como se discutió más arriba, la selección del valor de corte o *score* no depende puramente de cuestiones numéricas o estadísticas, sino que debe analizarse el contexto en el que se aplicarán los modelos. En nuestro caso, considerando el escenario económico y la limitada disponibilidad de recursos a nivel local, hemos optado por priorizar la **Sp** sobre la **Se**, reduciendo, de esta manera, la probabilidad de adquirir falsos positivos, es decir compuestos que fueran clasificados como **ACTIVOS**, pero que en la práctica resultarán **INACTIVOS**; cabe mencionar que esta decisión tiene como costo la pérdida de verdaderos positivos, que el modelo estará clasificando como falsos negativos, con la posible desestimación de andamiajes activos valiosos.

Para realizar una adecuada selección del valor de corte se construyeron curvas ROC con el software MedCalc (MedCalc Software, 2011); dicho software genera como *output* una tabla con diferentes valores (**Se**, **Sp**, y *score*, entre otros) a partir de la cual puede elegirse el valor de corte que provee una adecuada relación entre **Se** y **Sp**.

### 3.7 Cribado Virtual

Una vez seleccionado el ensamble a utilizar y el valor de *score* para discernir entre compuestos **ACTIVOS** e **INACTIVOS**, aplicamos el modelo clasificador para cribar dos bibliotecas de compuestos químicos. Una de las bibliotecas cribadas fue **DrugBank 3.0**. Esta es una biblioteca *online* que en la versión empleada presenta más de 6600 entidades químicas tipo fármaco, compilando información de compuestos que se encuentran atravesando estudios clínicos, que han sido aprobados por la *Food and Drug Administration* (FDA) de Estados Unidos, así como compuestos de uso veterinario, fármacos discontinuados y nutracéuticos (Knox et al., 2011).

También realizamos el cribado de la biblioteca **SweetLead**, que contiene derivados de productos naturales, nutracéuticos y algunas drogas aprobadas por otras agencias regulatorias, como China o India, para uso en medicina humana. La versión utilizada en este trabajo de tesis cuenta con más de 4000 entidades químicas (Novick et al., 2013).

Las bibliotecas fueron acondicionadas para el cálculo de descriptores moleculares y posterior cribado virtual. Con ambas bibliotecas se procedió de manera similar a la metodología implementada para el set de datos inicial, por lo que mediante el programa Standardizer V 16.9.12.0 (ChemAxon2016) se estandarizaron las estructuras y luego se calcularon los descriptores correspondientes.

### 3.8 Dominio de aplicación de los modelos

Los modelos desarrollados mediante las técnicas QSAR “aprenden” a identificar características relevantes para la propiedad o actividad modelada según el conjunto con el que se los haya entrenado, por lo que diferentes conjuntos de entrenamiento generarán modelos diferentes, aun cuando el procedimiento de obtención sea el mismo. Dado que estos modelos serán utilizados como predictores de actividad en conjuntos externos al de entrenamiento, es necesario definir un Dominio de Aplicación o Dominio de Aplicabilidad (DA). Este dominio es un espacio en el cual las predicciones del modelo se considerarán confiables (Roy, Kar, & Ambure, 2015).

Existen diversos métodos para evaluar el DA. Entre ellos, **la regla de la palanca**, la cual supone una aproximación robusta para la evaluación del DA de los modelos desarrollados mediante Regresiones Lineales (Roy, Kar, & Ambure, 2015).

La técnica se basa en obtener el valor de palanca ( $h_i$ ) para cada uno de los compuestos cribados. Este valor se obtiene a partir de la **ecuación 3.8**.

$$h_i = X(X^T X)^{-1} * X^T \quad \text{Ec. 3.8}$$

Donde  $X$  es una matriz compuesta por los  $k$  parámetros (descriptores incluidos en el modelo) para los  $n$  elementos (moléculas del conjunto de entrenamiento) y  $X^T$  representa la matriz transpuesta (Eslamimanesh et al., 2012).

Se considera que un compuesto se encuentra dentro del DA del modelo si el mismo presenta un valor  $h_i$  por debajo del valor crítico ( $h^*$ ). En nuestro caso, basándonos en la experiencia de nuestro grupo como así también de diversos autores (Eslamimanesh et al., 2012; Roy, Kar, & Ambure, 2015), consideramos el valor de  $h^*$  según se define en la **ecuación 3.9**

$$h^* = \frac{3k}{n} \quad \text{Ec. 3.9}$$

Se realizó el análisis de DA por la técnica antedicha para cada modelo incluido en nuestro ensamble y para todos los compuestos predichos como **ACTIVOS** por el ensamble.

Actualmente la técnica QSAR se puede llevar adelante casi de manera automatizada, pero esto no quita que se pueda prescindir de los conocimientos de expertos en la metodología, ya que además de los parámetros estadísticos y de validación, conviene realizar una inspección visual / manual, caso por caso, en la que se contemple diversos aspectos según la aplicación que se le dará al modelo (Roy et al., 2017). En nuestro caso el análisis de los *hits*, dado que nos enfocamos en el reposicionamiento de drogas con fines tripanocidas, contemplaba diferentes características farmacológicas y farmacocinéticas observando principalmente la dosis máxima permitida, el mecanismo de acción, el metabolismo y vías de excreción, como así también la unión del fármaco a las proteínas plasmáticas. Adicionalmente se evaluó también la disponibilidad en proveedores habituales y la accesibilidad a cada una de las drogas que fueron predichas como inhibidores de la enzima **TryS** por el meta-clasificador.

### 3.9 Sección Experimental

En esta sección desarrollaremos las metodologías de validación biológica / bioquímica, para los compuestos señalados por los modelos como potenciales inhibidores de **TryS**.

#### 3.9.1 Evaluaciones biológicas

Las evaluaciones biológicas (fenotípicas) se realizaron frente a dos especies de tripanosomátidos, *T. cruzi* y *T. brucei*.

En el caso de *T. cruzi* se realizó la evaluación de los candidatos frente a los estadios epimastigote y tripomastigote, mientras para *T. brucei* se realizó en el estadio sanguíneo, denominado en inglés *bloodstream*.

Los ensayos se realizaron a concentración definida según la especie o estadio parasitario con el que realiza el estudio. Este modo de trabajo permite obtener una estimación de la actividad tripanocida de cada una de las drogas seleccionadas. En todos los casos se realizaron los respectivos controles con Benznidazol o Nifurtimox (según el ensayo) a concentraciones cercanas a la  $EC_{50}$  de cada droga control; y controles de crecimiento sin droga, pero con Dimetilsulfóxido (DMSO 1 o 2%). El control con DMSO se incorpora debido a que

todas las drogas ensayas fueron disueltas en DMSO; y a partir de estas soluciones madre se realizaban las diluciones correspondientes para conseguir la concentración a ensayar.

### 3.9.1.1 Ensayos de proliferación frente a epimastigotes de *T. cruzi*

Para evaluar si los de los candidatos seleccionados interfieren en la proliferación de los parásitos, se realizó un ensayo a cuatro concentraciones fijas (25, 50, 100 y 200  $\mu\text{M}$ ) frente la cepa Y-GFP de *T. cruzi* en estadio epimastigotes.

El ensayo se realizó utilizando parásitos en la fase de crecimiento exponencial del cultivo ( $2 \times 10^6$  parásitos/mL). Los parásitos fueron cultivados en medio BHT suplementado con 10% suero fetal bovino (SFB) (inactivado por calor) y Penicilina (100 U/mL) más Estreptomicina (100  $\mu\text{g}/\text{mL}$ ). Bajo estas condiciones se adicionaron los fármacos, en concentraciones tales que por dilución en placa se obtiene la concentración final deseada. Luego se incubaron a 28 °C. Los días 4 y 8 posteriores a la siembra se realizó la determinación de proliferación mediante densidad óptica (DO) a 630 nm (Mosquillo et al., 2018).

La determinación de  $\text{EC}_{50}$  se realizó de igual manera que en el párrafo anterior, pero las concentraciones evaluadas fueron las siguientes: 0,01; 0,5; 1; 2,5; 5; 7,5; 15; 25; 50 y 100  $\mu\text{M}$ ; y el tiempo de análisis fue 72 h post siembra.

### 3.9.1.2 Ensayo de viabilidad frente a tripomastigotes de *T. cruzi*

En el presente ensayo se utilizó la cepa RA de *T. cruzi* en estadio tripomastigote. Como medio de cultivo se empleó caldo RPMI suplementado con 10% de SFB y Penicilina (100 U/mL) más Estreptomicina (100  $\mu\text{g}/\text{mL}$ ) como antibióticos.

En placa de 96 pocillos y por triplicado, se adicionaron 100  $\mu\text{L}$  de una suspensión que contenía aproximadamente  $1 \times 10^6$  parásitos/mL. Se adicionaron 100  $\mu\text{L}$  de solución de droga a evaluar o control y se incubó a 37 °C. Transcurrida las 24 h de incubación, se realizó el conteo de parásitos vivos, dicho conteo se llevó a cabo en cámara de Neubauer mediante microscopio óptico; de esta manera se estima la cantidad de parásitos vivos por mL. Todos los datos son normalizados respecto a control negativo (DMSO) (Miranda et al., 2015).

### 3.9.1.3 Ensayo de viabilidad frente a *T. brucei*

Para este ensayo se utilizó la línea celular 449 de *T. brucei* (cepa 427). Esta cepa expresa un biosensor redox (hGrx-roGFP2) fluorescente que permite el análisis de viabilidad mediante citometría de flujo (Gutscher et al., 2008).

Los parásitos, mantenidos en fase exponencial de la curva de crecimiento se mantienen en botellas de cultivo en medio HMI-9 suplementado con 10% de SFB, Higromicina (5  $\mu\text{g}/\text{mL}$ ) y Floemicina (0.2  $\mu\text{g}/\text{mL}$ ) como antibióticos, en estufa a 37 °C con corriente de 5% de  $\text{CO}_2$  y atmosfera humedecida. Oxitetraciclina fue incorporada al cultivo 24 horas antes del ensayo a fin de generar la expresión del biosensor en los parásitos (parásitos inducidos). De igual manera, pero sin agregado de Oxitetraciclina se procedió para un cultivo de parásitos al que no se le indujo la expresión del biosensor (parásitos no inducidos) (Franco et al., 2017).

La evaluación de las drogas se realizó en placa de 96 pocillos. En cada pocillo se colocó 2  $\mu\text{L}$  de la solución de droga a ensayar y 200  $\mu\text{L}$  de parásitos, los cuales previamente se resuspendieron en medio de cultivo fresco y a densidad de  $5 \times 10^5$  parásitos/mL. La placa se incubó por 24 horas en las condiciones antes mencionadas. Durante este ensayo se incluyeron los siguientes controles: control de viabilidad en DMSO 1% para los parásitos inducidos y no inducidos, y control de sensibilidad frente a Nifurtimox para el caso de los parásitos inducidos.

Previo al análisis por citometría, 24 horas posterior a la siembra, los parásitos fueron incubados por 1 minuto con Yoduro de Propidio (PI) en una concentración final de 2  $\mu\text{g}/\text{mL}$ . Las muestras se analizaron con láser de 488 nm en presencia de los filtros  $\lambda_{em}=530/40$  y  $\lambda_{em}=6130/30$  nm en el citómetro de flujo, modelo C6Accuri (BD). La presencia de ambos filtros permite identificar la señal debida al biosensor o PI respectivamente.

Todas las drogas fueron evaluadas a concentración fija de 30  $\mu\text{M}$ , y por triplicado.

Para la determinación del valor de  $EC_{50}$  se aplicó la misma metodología que para el estudio de viabilidad, pero utilizando, para este ensayo, el siguiente esquema de concertaciones: 1, 4, 8, 16, 18, 24 y 48  $\mu\text{M}$

### 3.9.2 Evaluaciones bioquímicas

Las evaluaciones bioquímicas frente a la enzima se llevaron a cabo contra la enzima **TryS** de tres especies de tripanosomátidos (*T. cruzi*, *T. brucei* y *L. infantum*). Estas enzimas se obtuvieron a partir de cultivos recombinantes de *Escherichia coli* (*E.coli*) (Benítez et al., 2016). Para cada enzima se realizaron tres evaluaciones independientes y a concentración fija de 30  $\mu\text{M}$ ; cada ensayo se ejecutó por cuadruplicado. Este primer cribado tiene como objetivo verificar si alguna de las drogas presentaba efecto inhibitorio significativo frente a esta la enzima.

Los ensayos se llevaron a cabo en placa de 96 pocillos. En cada pocillo se colocaron 5  $\mu\text{L}$  de la solución de droga a evaluar en concentración de 300  $\mu\text{M}$ , luego se agregó 40  $\mu\text{L}$  de la mezcla de reacción, la cual está compuesta por ATP, Espermidina, DTT, y Glutación en una composición definida para cada una de las enzimas de las diferentes especies de tripanosomátidos (Krauth-Siegel & Comini, 2008). Por último, se agregó 5  $\mu\text{L}$  de enzima, y se incubó a temperatura ambiente por 15 minutos. Transcurrido este tiempo se interrumpió la reacción con el agregado de 200  $\mu\text{L}$  de Verde de Malaquita (VM). 20 minutos posteriores al agregado de VM se procedió a la lectura de la placa mediante espectrofotometría, cuantificándose la actividad residual de la enzima en base a la señal de absorbancia a 650 nm. El ensayo cuenta con dos controles, un control de inhibición positivo consistente en 5  $\mu\text{L}$  de una solución de inhibidor enzimático, Mol-2008 (30  $\mu\text{M}$ ) o J-18 (1,5  $\mu\text{M}$ ), según la enzima de la especie de tripanosomátidos con la que se esté trabajando. El control de actividad enzimática (control negativo) se realiza reemplazando los 5  $\mu\text{L}$  de solución de droga por el mismo volumen de DMSO. Adicionalmente se realizan blancos de todas las drogas, sin enzima, para evaluar posibles interferencias de color durante la lectura (Benítez et al., 2016).

Dado que la actividad residual se cuantifica mediante una reacción colorimétrica medida por espectrofotometría, por cada droga se debe realizar un control (blanco) para corregir los posibles desvíos correspondientes a la absorbancia de cada droga.

Segunda campaña de cribado virtual y evaluación experimental

En el transcurso de la tesis doctoral, se estableció una colaboración con el Dr. Marcelo Comini, responsable del Laboratorio de Biología Redox de Tripanosomátidos del Instituto Pasteur Montevideo, quien nos facilitó el acceso a un set de datos propio, a fin de entrenar un nuevo conjunto de modelos e implementar una segunda campaña de cribado virtual. Adicionalmente se incorporaron, nuevas metodologías computacionales para el desarrollo y validación de los modelos *in silico*.

A continuación, se presentan los cambios de metodología computacional y de validación experimental incorporados en la segunda campaña.

Con el fin de facilitar al lector la búsqueda de alguna metodología en particular dentro de la segunda campaña se decidió conservar los títulos de cada sección.

### 3.10 Construcción de modelos QSAR – Funciones Discriminantes

#### 3.10.1 Recopilación y curado del set de datos

El set de datos conformado para la primera campaña de CV fue ampliado con el set de datos facilitado por el Dr. Comini, dicho set contempla una gran diversidad de compuestos químicos con sus respectivos valores de actividad frente a la enzima **TryS** de diferentes tripanosomátidos, entre los que se incluyen *Leishmania*, *T. brucei* y *T. cruzi*.

Para ser consistentes con el cribado experimental primario desarrollado en el Instituto Pasteur, elevamos, para esta campaña, el valor de corte de 20  $\mu\text{M}$  a 30  $\mu\text{M}$ , considerando **ACTIVOS** a aquellos compuestos con  $\text{IC}_{50}$  menor a 30  $\mu\text{M}$ , y para el caso de los ensayos unipuntuales aquellas moléculas que inhibían al menos el 50% de la actividad enzimática a concentraciones 30  $\mu\text{M}$  o menores. Se consideraron **INACTIVOS** a los compuestos que no cumplen con alguna de estas condiciones.

Para evaluar la diversidad estructural del nuevo set de datos se utilizó el programa JChemV.6 del paquete ChemAxon 2013, en esta campaña se volvió a seleccionar el sistema de *fingerprinting* ECFP\_4 y se calculó el Coeficiente de Tanimoto. Adicionalmente construimos mapas de calor con el programa Gitoools (Perez-Llamas & Lopez-Bigas, 2011), tal como se realizó para la primera campaña.

#### 3.10.2 Partición del set de datos

Para la partición del set de datos se aplicó la misma combinación de técnicas de agrupamiento que la empleada en la campaña previa. Las técnicas de agrupamiento utilizadas fueron: Library MCS, disponible en el JChem V6 (ChemAxon 2014) y agrupamiento no jerárquico mediante *k-means*, incorporado en el software Statistica 10 (statsoft, 2011).

### 3.10.3 Selección de los conjuntos de entrenamiento y de prueba

La conformación de los conjuntos se realizó de manera similar a las particiones de la primera campaña, pero en esta ocasión se generaron tres conjuntos, un conjunto de entrenamiento y dos conjuntos de prueba (denominados, DUD-e A y DUD-e B) para diferentes etapas de validación: el conjunto DUD-e A se utilizó para validar los modelos individuales y entrenar el ensamble (eligiendo esquema de combinación y número de modelos combinados); mientras que el conjunto DUD-e B se reservó para la validación del ensamble generado. Los conjuntos de prueba fueron conformados con un bajo número de **ACTIVOS** (~20 moléculas), compuestos **INACTIVOS** y adicionalmente se incorporaron señuelos o presuntos inactivos generados mediante la herramienta *online* DUD-e (Mysinger et al., 2012). El conjunto de entrenamiento, en cambio, quedó estrictamente balanceado con igual cantidad de compuestos **ACTIVOS** e **INACTIVOS**.

### 3.10.4 Digitalización de las estructuras y cálculo de descriptores

Se digitalizaron todas las estructuras con el programa Marvin Sketch V 6.2.2 del paquete ChemAxon 2014, la estandarización de estas se realizó mediante el programa Standardizer (V. 16.9.12.0, Instant JChem) de igual manera y con los mismos filtros que se utilizaron para la primera campaña de CV. Para el cálculo de descriptores moleculares se utilizó el programa Dragon V 6.0 (Milano 2011).

### 3.10.5 Construcción de los modelos

Metodología descrita en sección correspondiente a la primera campaña.

### 3.10.6 Aleatorización de subespacios

Metodología descrita en sección correspondiente a la primera campaña.

### 3.10.7 Selección de descriptores moleculares y generación de los modelos

La selección de descriptores y la conformación de los modelos discriminantes se realizó exactamente igual que en la primera campaña.

## 3.11 Validación de los modelos

Aquí se emplearon las mismas técnicas de validación que las aplicadas durante la campaña previa, tanto para instancias de **validación interna** como **validación externa**, para más detalles consultar las **secciones 3.3.1 y 3.3.2** respectivamente.

## 3.12 Comportamiento y desempeño de los modelos

Para evaluar el desempeño de los modelos individuales se utilizó uno de los conjuntos de prueba (DUD-e A). En la presente campaña se ha evaluado el AUCROC tal como se describe en la campaña previa. Adicionalmente, en esta segunda campaña hemos incorporado dos métricas para evaluar de manera más exhaustiva el desempeño de los modelos. Las mismas se detallan a continuación.

### 3.12.1 Factor de enriquecimiento

El **Factor de Enriquecimiento** (EF, *Enrichment Factor*) es una métrica que permite establecer cuánto se enriquece, en **ACTIVOS**, la primera fracción de la lista de compuestos ordenada según los modelos (Truchon & Bayly, 2007). El EF1% indica la cantidad de activos recuperados cuando se recorre de manera decreciente el 1% de la lista ordenada o se alcanza el 1% de los señuelos. El EF se calcula según la **ecuación 3.10**.

$$EF_{x\%} = \frac{N_A \text{ en el } x\% \text{ mejor rankeado de la quimioteca}}{N_A \text{ en el } x\% \text{ de la quimioteca ordenada al azar}} \quad \text{Ec. 3.10}$$

Es importante recordar que la métrica no reconoce el orden en que se encuentran los **ACTIVOS** en el porcentaje de la lista analizada, sin embargo; una alta tasa de **ACTIVOS** en el 1% de la base aumenta significativamente la probabilidad de éxitos en la etapa de cribado experimental de los compuestos seleccionados.

En la presente campaña hemos utilizado la métrica con dos valores de corte; uno de ellos es cuando se recorre el 1% de la base ( $EF_{1\% \text{base}}$ ), mientras que el segundo valor propuesto es cuando se recupera el 1% de los señuelos ( $EF_{1\% \text{señuelos}}$ ).

### 3.12.2 Métrica BEDROC

BEDROC (en inglés, *Boltzmann-enhanced discrimination of ROC*) es una métrica que permite evaluar el enriquecimiento temprano en las campañas de cribado virtual. La contribución de cada activo, rankeado según el *score* del modelo correspondiente, al valor de la métrica se asigna según una función exponencial que depende de la posición de un dado **ACTIVO** en el ranking ordenado, según se aprecia en la **ecuación 3.11**. A diferencia de las curvas ROC, que refleja **la posición promedio de los activos** en el *ranking* provisto por el modelo, BEDROC pondera preferencialmente el peso de los compuestos en las primeras posiciones del *ranking*. Dicha métrica adopta valores entre 0 y 1, y se interpreta como la probabilidad de que un compuesto **ACTIVO** clasificado por el modelo evaluado se encuentre entre los primeros puestos del *ranking* de la lista ordenada. Como se observa de la **ecuación 3.11** el término exponencial está delimitado / definido por el parámetro  $\alpha$ . Lo que demuestra que cuando mayor es el valor de  $\alpha$  la curva decae más rápido, es decir se están priorizando las primeras porciones del *ranking*. El grueso del valor de la métrica estará provisto por los

**ACTIVOS** localizados en el  $1/\alpha$  % de los compuestos mejores rankeados de la quimioteca sometida a cribado.

$$BEDROC = \left( \frac{\sum_{i=1}^{A_{tot}} e^{-\alpha r_i/N}}{\frac{A_{tot}}{N} \frac{1-e^{-\alpha}}{e^{(\alpha/N)}-1}} \right) \left( \frac{\frac{A_{tot}}{N} \operatorname{Senh} \frac{\alpha}{2}}{\operatorname{Cosh} \frac{\alpha}{2} - \cosh \left( \frac{\alpha}{2} - \alpha \frac{A_{tot}}{N} \right)} \right) + \frac{1}{1-e^{-\alpha \left( 1 - \frac{A_{tot}}{N} \right)}} \quad \text{Ec 3.11}$$

Donde  $A_{tot}$  es el número total de **ACTIVOS**,  $r_i$  la posición en el *ranking* del *i*-ésimo **ACTIVO**,  $N$  el número de compuestos totales en la base cribada, y  $\alpha$  el factor exponencial.

Dado que  $\alpha$  es un parámetro exponencial no es posible comparar valores de la métrica BEDROC correspondientes a distintos valores de  $\alpha$ . Esto sería equivalente a comparar el factor de enriquecimiento a diferentes porcentajes. También es importante recordar que no se puede asociar el valor de la BEDROC con el de las curvas ROC, ya que la segunda es una función lineal, y en consecuencia un valor de AUCROC = 0,5 se correlaciona con una clasificación aleatoria, mientras que, dado el carácter exponencial de la métrica BEDROC, un valor de 0,5 de la misma se considera muy favorable (Truchon & Bayly, 2007).

En la presente campaña se ha utilizado un  $\alpha=20$ , lo que indica que el 8% del inicio del *ranking* contribuye al 80% del total de la curva BEDROC (Ntie-Kang, 2020). Para la realización de la curva BEDROC se utilizó la herramienta *on line* Rocker (<http://www.medchem.fi/rocker/>).

### 3.12.3 Campaña retrospectiva de cribado virtual

Para la campaña retrospectiva de los modelos individuales se utilizó el conjunto Duda A, el mismo quedó conformado, por compuestos por 20 compuestos **ACTIVOS**, 110 **INACTIVOS** y un conjunto de 979 señuelos generado con la herramienta DUD-e (Mysinger et al., 2012). De esta manera la proporción de **ACTIVOS** es menor al 0,05, y está en concordancia con lo propuesto por Truchon y colaboradores (Truchon & Bayly, 2007).

## 3.13 Combinación de modelos

Para la combinación de modelos de la presente campaña utilizamos la misma técnica y operadores aplicados en la campaña previa.

En esta oportunidad se realizó la combinación acumulada de los 100 mejores modelos, es decir, se combinaron los dos mejores modelos, luego los tres mejores modelos y así sucesivamente hasta combinar los mejores cien modelos. En el capítulo correspondiente a resultados el lector podrá observar las curvas AUCROC vs Cantidad de modelos que hemos realizado con el Software R (<https://cran.r-project.org/package=plotly>) con el fin de visualizar y evaluar el comportamiento de las diferentes combinaciones.

### 3.14 Selección del valor de corte

Como se mencionó en la **sección 3.6** de la primera campaña de CV, para poder conocer si la predicción del modelo o la combinación de modelos es válida, es decir, el modelo predice un determinado compuesto como **ACTIVO**, es necesario establecer un valor de *score*, que se utilizará en el cribado prospectivo. Una posibilidad es definir el valor de corte como aquel que conduce a una relación **Se - Sp** adecuada, a través de las gráficas de las curvas ROC como se realizó en la campaña previa. Otra alternativa, es la que se empleó en esta segunda campaña de CV, mediante la construcción y el análisis de gráficos de superficie del Valor Predictivo Positivo (**PPV**, por sus siglas en inglés, *Positive Predictive Value*).

En las campañas de CV es importante predecir la probabilidad real con la que un *hit* pueda confirmar su actividad predicha cuando se someta a pruebas experimentales. La estimación de esta probabilidad es, sin embargo, imposible de calcular para las quimiotecas sometidas a CV prospectivo dado que en su cálculo aparece la proporción de compuestos **ACTIVOS** de la biblioteca (*Ya*), la cual se desconoce *a priori*. El valor de PPV se puede obtener a partir de la **ecuación 3.12**.

$$PPV = \frac{Se Ya}{SeYa + (1 - Sp)(1 - Ya)} \quad \text{Ec. 3.12}$$

Donde *Ya* es la proporción de activos, también llamada prevalencia de actividad o rendimiento de activos. **Se** y **Sp** son los valores de sensibilidad y especificidad.

Para construir la gráfica de superficie PPV se calcula, utilizando el conjunto DUD-e A, la **Se** y **Sp** para todo el rango de valores de *score* posibles para el ensamble de modelos seleccionado, y se observa la relación entre el valor de PPV en función de la relación **Se/Sp**, variando la proporción de compuestos **ACTIVOS** (*Ya*) entre 0,01 y 0,001, que corresponde el rango de proporciones de **ACTIVOS** que habitualmente se identifican en un cribado farmacológico de alto rendimiento (Alberca et al., 2018; Morales et al., 2020).

Para la obtención del gráfico en 3D del **PPV** utilizamos el paquete de gráficos del software R (<https://cran.r-project.org/package=plotly>).

El comportamiento del ensamble seleccionado para la posterior etapa de CV fue testeado frente al conjunto DUD-e B.

### 3.15 Cribado virtual

El cribado virtual para la selección de candidatos a evaluar experimentalmente se realizó frente a las bibliotecas **SweetLead** (Novick et al., 2013) y **DrugBank 4.0** (Law et al., 2014). La versión 4.0 de DrugBank ha aumentado en más del 40% sus campos de información, más de un 15% el número de fármacos aprobados por FDA y en más de 1200 nuevas entidades químicas en fases de investigación clínica con respecto a la versión 3.0 (Knox et al., 2011) pero lo más interesante es que incorpora datos ADME (Absorción, Distribución Metabolismo y Excreción), como así también rutas de síntesis, y links a patentes y trabajos científicos sobre

nuevas entidades químicas. Ambas bibliotecas fueron acondicionadas con el mismo programa y bajo los mismos parámetros que se utilizaron para el set de datos.

### 3.16 Dominio de Aplicación de los modelos

La selección del espacio químico en el cual pueden aplicarse los modelos fue realizada bajo el método de la **regla de la palanca** tal como se describe en la **sección 3.8**.

### 3.17 Sección experimental

#### 3.17.1 Evaluaciones biológicas

Las evaluaciones biológicas se realizaron frente a *T. brucei* y *T. cruzi*. Los estadíos evaluados para *T. cruzi* fueron epimastigote y tripomastigote.

##### 3.17.1.1 Ensayos de proliferación frente a epimastigotes de *T. cruzi*

El efecto que las drogas podrían ejercer sobre la proliferación de los parásitos se evaluó realizando una serie de ensayos por triplicado a concentración fija (50  $\mu\text{M}$ ), sobre la cepa Y de *T. cruzi* en estadio epimastigotes.

El ensayo se realizó utilizando parásitos en fase exponencial del cultivo ( $2 \times 10^6$  parásitos/mL). Los parásitos fueron cultivados en medio BHT suplementado con 10% SFB (inactivado por calor) y Penicilina (100 U/mL) más Estreptomicina (100  $\mu\text{g/mL}$ ). Bajo estas condiciones se adicionaron las drogas, en las concentraciones tales que por dilución en placa se obtiene la concentración final deseada. Luego se incubaron a 28°C. Los días 4 y 8 posteriores a la siembra se realizó la determinación de proliferación mediante conteo de parásitos en cámara de Neubauer y microscopio óptico (Bellera et al., 2013).

##### 3.17.1.2 Ensayo de viabilidad frente a tripomastigote de *T. cruzi*

Se aplicó el mismo protocolo que el descrito para la primera campaña.

##### 3.17.1.3 Ensayo de viabilidad frente a *T. brucei*

Para la evaluación del efecto de los diferentes fármacos frente a *T. brucei* se utilizó la cepa BB3, la cual expresa una enzima dependiente ATP que utiliza Luciferina como sustrato.

Los parásitos se cultivaron de igual manera que en la campaña previa, pero sin agregado de Oxitetraciclina, ya que en este caso nos independizamos de la expresión del biosensor Redox (Franco et al., 2017)

El protocolo de siembra y evaluación de los compuestos fue similar la que se aplicó en la primera campaña, modificando solo el método de lectura y análisis de resultados.

La lectura de resultados se llevó a cabo en un luminómetro, el cual transforma la señal lumínica de una muestra en una variable numérica. Una vez colocado el fármaco en contacto con los parásitos y transcurrido el tiempo de incubación necesario (24 h), la solución de parásitos junto con el fármaco es transferida a una placa negra de 96 pocillos y a continuación se procede al agregado de 20  $\mu\text{L}$  de una solución de Luciferina (1.5 mg/mL) y Triton (0.05 %V/V) en PBS-Glucosa, posteriormente se lee la señal emitida en cada pocillo (Benítez et al., 2020). El protocolo de ensayo se ajustó a los detalles publicados por Benítez y Col (Benítez et al., 2020). Finalmente se utilizaron diferentes softwares para el análisis de la señal lumínica obtenida.

### 3.17.2 Evaluaciones bioquímicas

Las evaluaciones bioquímicas frente a la enzima se llevaron a cabo contra la enzima *TryS* de tres especies de tripanosomátidos (*T. cruzi*, *T. brucei* y *L. Infantum.*). Estas enzimas se obtuvieron a partir de cultivos recombinantes de *Escherichia coli* (*E.coli*) (Benítez et al., 2016). Para todas las enzimas se realizaron tres ensayos independientes y a concentración fija de 25  $\mu\text{M}$ ; cada ensayo se ejecutó por cuadruplicado. Este primer cribado tiene como objetivo evidenciar si alguna de las drogas presentaba un efecto inhibitorio significativo frente a esta la enzima.

Los ensayos se llevaron a cabo en placa de 384 pocillos. En cada pocillo se colocó 1  $\mu\text{L}$  de la solución de droga a evaluar en concentración de 400  $\mu\text{M}$ , luego se agregó 10  $\mu\text{L}$  del buffer que contiene la enzima a evaluar y se incubó por 1 hora a 28°C. Transcurrido este tiempo se incorporó 5  $\mu\text{L}$  de la mezcla de reacción, la cual está compuesta por ATP, Espermidina, DTT, y Glutatión en una composición definida para cada una de las enzimas de las diferentes especies de tripanosomátidos (Krauth-Siegel & Comini, 2008). La placa fue incubada nuevamente por 1 hora a 28°C y por último se interrumpió la reacción con el agregado de 60  $\mu\text{L}$  de Verde de Malaquita (VM). Veinte minutos posteriores al agregado de VM se procedió a la lectura de la placa mediante espectrofotometría, cuantificándose la actividad residual de la enzima en base a la señal de absorbancia a 620 nm. El ensayo cuenta con dos controles. El control positivo, el cual indica que la enzima es susceptible de ser inhibida, se realiza reemplazando la droga en estudio por 1  $\mu\text{L}$  de una solución de inhibidor enzimático, a diferencia de la campaña previa en esta oportunidad se utiliza el mismo inhibidor para la enzima de las tres especies de tripanosomátidos (AUC). Adicionalmente se realiza un control negativo en el cual el volumen de droga es remplazado por 1  $\mu\text{L}$  de DMSO 1% a fin de mantener constante el volumen final.

### 3.18 Referencias Bibliográficas

- Alberca, L. N., Chuguransky, S. R., Álvarez, C. L., Talevi, A., & Salas-Sarduy, E. (2019). In silico Guided Drug Repurposing: Discovery of New Competitive and Non-competitive Inhibitors of Falcipain-2. *Frontiers in Chemistry*, 7, 534. <https://doi.org/10.3389/fchem.2019.00534>
- Alberca, L. N., Sbaraglini, M. L., Morales, J. F., Dietrich, R., Ruiz, M. D., Pino Martínez, A. M., Miranda, C. G., Fraccaroli, L., Alba Soto, C. D., Carrillo, C., Palestro, P. H., & Talevi, A. (2018). Cascade Ligand- and Structure-Based Virtual Screening to Identify New Trypanocidal Compounds Inhibiting Putrescine Uptake. *Frontiers in Cellular and Infection Microbiology*, 8, 173. <https://doi.org/10.3389/fcimb.2018.00173>
- Babyak, M. A. (2004). What You See May Not Be What You Get: A Brief, Nontechnical Introduction to Overfitting in Regression-Type Models. *Psychosomatic Medicine*, 66(3), 411-421. <https://doi.org/10.1097/00006842-200405000-00021>
- Ballabio, D., Grisoni, F., & Todeschini, R. (2018). Multivariate comparison of classification performance measures. *Chemometrics and Intelligent Laboratory Systems*, 174, 33-44. <https://doi.org/10.1016/j.chemolab.2017.12.004>
- Bansal, A., Sharma, M., & Goel, S. (2017). Improved K-mean Clustering Algorithm for Prediction Analysis using Classification Technique in Data Mining. *International Journal of Computer Applications*, 157(6), 35-40. <https://doi.org/10.5120/ijca2017912719>
- Bellera, C. L., Balcazar, D. E., Alberca, L., Labriola, C. A., Talevi, A., & Carrillo, C. (2013). Application of Computer-Aided Drug Repurposing in the Search of New Cruzipain Inhibitors: Discovery of Amiodarone and Bromocriptine Inhibitory Effects. *Journal of Chemical Information and Modeling*, 53(9), 2402-2408. <https://doi.org/10.1021/ci400284v>
- Benítez, D., Medeiros, A., Fiestas, L., Panozzo-Zenere, E. A., Maiwald, F., Prousis, K. C., Roussaki, M., Calogeropoulou, T., Detsi, A., Jaeger, T., Šarlauskas, J., Peterlin Mašič, L., Kunick, C., Labadie, G. R., Flohé, L., & Comini, M. A. (2016). Identification of Novel Chemical Scaffolds Inhibiting Trypanothione Synthetase from Pathogenic Trypanosomatids. *PLOS Neglected Tropical Diseases*, 10(4), e0004617 <https://doi.org/10.1371/journal.pntd.0004617>
- Benítez, D., Dibello, E., Bonilla, M., & Comini, M. A. (2020). A simple, robust, and affordable bioluminescent assay for drug discovery against infective African trypanosomes. *Drug Development Research*, ddr.21634. <https://doi.org/10.1002/ddr.21634>
- Celeux, G., Maugis-Rabusseau, C., & Sedki, M. (2019). Variable selection in model-based clustering and discriminant analysis with a regularization approach. *Advances in Data Analysis and Classification*, 13(1), 259-278. <https://doi.org/10.1007/s11634-018-0322-5>
- Comini, M., Menge, U., & Flohé, L. (2003). Biosynthesis of Trypanothione in *Trypanosoma brucei brucei*. *Biological Chemistry*, 384(4). <https://doi.org/10.1515/BC.2003.072>
- Daho, M. E. H., & Chikh, M. A. (2015). Combining Bootstrapping Samples, Random Subspaces and Random Forests to Build Classifiers. *Journal of Medical Imaging and Health Informatics*, 5(3), 539-544. <https://doi.org/10.1166/jmih.2015.1423>
- Daoud, J. I. (2017). Multicollinearity and Regression Analysis. *Journal of Physics: Conference Series*, 949, 012009. <https://doi.org/10.1088/1742-6596/949/1/012009>
- Eslamimanesh, A., Gharagheizi, F., Mohammadi, A. H., & Richon, D. (2012). A statistical method for evaluation of the experimental phase equilibrium data of simple clathrate hydrates. *Chemical Engineering Science*, 80, 402-408. <https://doi.org/10.1016/j.ces.2012.06.019>

- Franco, J., Sardi, F., Szilágyi, L., Kövér, K. E., Fehér, K., & Comini, M. A. (2017). Diglycosyl diselenides alter redox homeostasis and glucose consumption of infective African trypanosomes. *International Journal for Parasitology: Drugs and Drug Resistance*, 7(3), 303-313. <https://doi.org/10.1016/j.ijpddr.2017.08.001>
- García, C. B., García, J., López Martín, M. M., & Salmerón, R. (2015). Collinearity: Revisiting the variance inflation factor in ridge regression. *Journal of Applied Statistics*, 42(3), 648-661. <https://doi.org/10.1080/02664763.2014.980789>
- Gonzalez, M., Teran, C., Saiz-Urra, L., & Teijeira, M. (2008). Variable Selection Methods in QSAR: An Overview. *Current Topics in Medicinal Chemistry*, 8(18), 1606-1627. <https://doi.org/10.2174/156802608786786552>
- Gramatica, P. (2013). On the Development and Validation of QSAR Models. En B. Reisfeld & A. N. Mayeno (Eds.), *Computational Toxicology* (Vol. 930, pp. 499-526). Humana Press. [https://doi.org/10.1007/978-1-62703-059-5\\_21](https://doi.org/10.1007/978-1-62703-059-5_21)
- Gutscher, M., Pauleau, A.-L., Marty, L., Brach, T., Wabnitz, G. H., Samstag, Y., Meyer, A. J., & Dick, T. P. (2008). Real-time imaging of the intracellular glutathione redox potential. *Nature Methods*, 5(6), 553-559. <https://doi.org/10.1038/nmeth.1212>
- Kaneko, H. (2019). Estimation of predictive performance for test data in applicability domains using  $\gamma$ -randomization. *Journal of Chemometrics*, 33(9). <https://doi.org/10.1002/cem.3171>
- Klug, D. M., Gelb, M. H., & Pollastri, M. P. (2016). Repurposing strategies for tropical disease drug discovery. *Bioorganic & Medicinal Chemistry Letters*, 26(11), 2569-2576. <https://doi.org/10.1016/j.bmcl.2016.03.103>
- Knox, C., Law, V., Jewison, T., Liu, P., Ly, S., Frolkis, A., Pon, A., Banco, K., Mak, C., Neveu, V., Djoumbou, Y., Eisner, R., Guo, A. C., & Wishart, D. S. (2011). DrugBank 3.0: A comprehensive resource for «Omics» research on drugs. *Nucleic Acids Research*, 39(Database), D1035-D1041. <https://doi.org/10.1093/nar/gkq1126>
- Krauth-Siegel, R. L., & Comini, M. A. (2008). Redox control in trypanosomatids, parasitic protozoa with trypanothione-based thiol metabolism. *Biochimica et Biophysica Acta (BBA) - General Subjects*, 1780(11), 1236-1248. <https://doi.org/10.1016/j.bbagen.2008.03.006>
- Krawczyk, B., Minku, L. L., Gama, J., Stefanowski, J., & Woźniak, M. (2017). Ensemble learning for data stream analysis: A survey. *Information Fusion*, 37, 132-156. <https://doi.org/10.1016/j.inffus.2017.02.004>
- Kubát, & Miroslav. (2017). *An Introduction to Machine Learning* (2nd ed.). Springer International Publishing: Imprint: Springer. <http://gen.lib.rus.ec/book/index.php?md5=68fca0936c73f900735cb3e40ba83717>
- Law, V., Knox, C., Djoumbou, Y., Jewison, T., Guo, A. C., Liu, Y., Maciejewski, A., Arndt, D., Wilson, M., Neveu, V., Tang, A., Gabriel, G., Ly, C., Adamjee, S., Dame, Z. T., Han, B., Zhou, Y., & Wishart, D. S. (2014). DrugBank 4.0: Shedding new light on drug metabolism. *Nucleic Acids Research*, 42(D1), D1091-D1097. <https://doi.org/10.1093/nar/gkt1068>
- Leonard, J. T., & Roy, K. (2006). On Selection of Training and Test Sets for the Development of Predictive QSAR models. *QSAR & Combinatorial Science*, 25(3), 235-251. <https://doi.org/10.1002/qsar.200510161>
- Martin, T. M., Harten, P., Young, D. M., Muratov, E. N., Golbraikh, A., Zhu, H., & Tropsha, A. (2012). Does Rational Selection of Training and Test Sets Improve the Outcome of QSAR Modeling? *Journal of Chemical Information and Modeling*, 52(10), 2570-2578. <https://doi.org/10.1021/ci300338w>

- Miranda, C. G., Solana, M. E., Curto, M. de los A., Lammel, E. M., Schijman, A. G., & Alba Soto, C. D. (2015). A flow cytometer-based method to simultaneously assess activity and selectivity of compounds against the intracellular forms of *Trypanosoma cruzi*. *Acta Tropica*, *152*, 8-16. <https://doi.org/10.1016/j.actatropica.2015.08.004>
- Mlinarić, A., Mlinarić, A., Horvat, M., Horvat, M., Smolčić, V. Š., Smolčić, V. Š., & Smolčić, V. Š. (2017). Dealing with the positive publication bias: Why you should really publish your negative results. *Biochemia Medica*, *27*(3), 0-0. <https://doi.org/10.11613/BM.2017.030201>
- Montori, V. M., Smieja, M., & Guyatt, G. H. (2000). Publication Bias: A Brief Review for Clinicians. *Mayo Clinic Proceedings*, *75*(12), 1284-1288. <https://doi.org/10.4065/75.12.1284>
- Moreno-Torres, J. G., Saez, J. A., & Herrera, F. (2012). Study on the Impact of Partition-Induced Dataset Shift on  $\$k\$$ -Fold Cross-Validation. *IEEE Transactions on Neural Networks and Learning Systems*, *23*(8), 1304-1312. <https://doi.org/10.1109/TNNLS.2012.2199516>
- Morales, J. F., Chuguransky, S., Alberca, L. N., Alice, J. I., Goicoechea, S., Ruiz, M. E., Bellera, C. L., & Talevi, A. (2020). Positivity Predictive Value surfaces as a complementary tool to assess the performance of virtual screening methods. *Mini-Reviews in Medicinal Chemistry*, *20*. <https://doi.org/10.2174/1871525718666200219130229>
- Mosquillo, M. F., Bilbao, L., Hernández, F., Machado, I., Gambino, D., Garat, B., & Pérez-Díaz, L. (2018). Effect of a new anti-T. cruzi metallic compound based on palladium. *BioMetals*, *31*(6), 961-974. <https://doi.org/10.1007/s10534-018-0140-4>
- Mysinger, M. M., Carchia, M., Irwin, John. J., & Shoichet, B. K. (2012). Directory of Useful Decoys, Enhanced (DUD-E): Better Ligands and Decoys for Better Benchmarking. *Journal of Medicinal Chemistry*, *55*(14), 6582-6594. <https://doi.org/10.1021/jm300687e>
- Nazeer, K. A. A., & Sebastian, M. P. (2009). *Improving the Accuracy and Efficiency of the k-means Clustering Algorithm*. 6.
- Ntie-Kang, F. (2020). *Fundamental Concepts*. Walter de Gruyter GmbH & Co KG.
- Novick, P. A., Ortiz, O. F., Poelman, J., Abdulhay, A. Y., & Pande, V. S. (2013). SWEETLEAD: An In Silico Database of Approved Drugs, Regulated Chemicals, and Herbal Isolates for Computer-Aided Drug Discovery. *PLoS ONE*, *8*(11), e79568. <https://doi.org/10.1371/journal.pone.0079568>
- Oza, S. (2003). Properties of trypanothione synthetase from *Trypanosoma brucei*. *Molecular and Biochemical Parasitology*, *131*(1), 25-33. [https://doi.org/10.1016/S0166-6851\(03\)00176-2](https://doi.org/10.1016/S0166-6851(03)00176-2)
- Perez-Llamas, C., & Lopez-Bigas, N. (2011). Gitools: Analysis and Visualisation of Genomic Data Using Interactive Heat-Maps. *PLoS ONE*, *6*(5), e19541. <https://doi.org/10.1371/journal.pone.0019541>
- Polanski, J., Bak, A., Gieleciak, R., & Magdziarz, T. (2006). Modeling Robust QSAR. *Journal of Chemical Information and Modeling*, *46*(6), 2310-2318. <https://doi.org/10.1021/ci050314b>
- Ravichandran Veerasamy, Rajak, H., Jain, A., Sivadasan, S., Christopher, P. V., & Agrawal, R. (2011). Validation of QSAR Models—Strategies and Importance. *Int J Drug Design and Discov*, *2*, 511-519.
- Robin, X., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sanchez, J.-C., & Müller, M. (2011). pROC: An open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics*, *12*(1), 77. <https://doi.org/10.1186/1471-2105-12-77>

- Roy, K., & Ambure, P. (2016). The “double cross-validation” software tool for MLR QSAR model development. *Chemometrics and Intelligent Laboratory Systems*, 159, 108-126. <https://doi.org/10.1016/j.chemolab.2016.10.009>
- Roy, K., Ambure, P., & Aher, R. B. (2017). How important is to detect systematic error in predictions and understand statistical applicability domain of QSAR models? *Chemometrics and Intelligent Laboratory Systems*, 162, 44-54. <https://doi.org/10.1016/j.chemolab.2017.01.010>
- Roy, K., Kar, S., & Ambure, P. (2015). On a simple approach for determining applicability domain of QSAR models. *Chemometrics and Intelligent Laboratory Systems*, 145, 22-29. <https://doi.org/10.1016/j.chemolab.2015.04.013>
- Roy, K., Kar, S., & Das, R. N. (2015). Statistical Methods in QSAR/QSPR. En K. Roy, S. Kar, & R. N. Das, *A Primer on QSAR/QSPR Modeling* (pp. 37-59). Springer International Publishing. [https://doi.org/10.1007/978-3-319-17281-1\\_2](https://doi.org/10.1007/978-3-319-17281-1_2)
- Roy, K., & Mitra, I. (2011). On Various Metrics Used for Validation of Predictive QSAR Models with Applications in Virtual Screening and Focused Library Design. *Combinatorial Chemistry & High Throughput Screening*, 14(6), 450-474. <https://doi.org/10.2174/138620711795767893>
- Sharma, H., & Verma, S. (2019). Is positive publication bias really a bias, or an intentionally created discrimination toward negative results? *Saudi Journal of Anaesthesia*, 13(4), 352. [https://doi.org/10.4103/sja.SJA\\_124\\_19](https://doi.org/10.4103/sja.SJA_124_19)
- Talevi, A., L. Bellera, C., Di Ianni, M., R. Duchowicz, P., E. Bruno-Blanch, L., & A. Castro, E. (2012). An Integrated Drug Development Approach Applying Topological Descriptors. *Current Computer Aided-Drug Design*, 8(3), 172-181. <https://doi.org/10.2174/157340912801619076>
- Toropova, A. P., Toropov, A. A., Veselinović, A. M., Veselinović, J. B., Leszczynska, D., & Leszczynski, J. (2019). Semi-correlations combined with the index of ideality of correlation: A tool to build up model of mutagenic potential. *Molecular and Cellular Biochemistry*, 452(1-2), 133-140. <https://doi.org/10.1007/s11010-018-3419-4>
- Triballeau, N., Acher, F., Brabet, I., Pin, J.-P., & Bertrand, H.-O. (2005). Virtual Screening Workflow Development Guided by the “Receiver Operating Characteristic” Curve Approach. Application to High-Throughput Docking on Metabotropic Glutamate Receptor Subtype 4. *Journal of Medicinal Chemistry*, 48(7), 2534-2547. <https://doi.org/10.1021/jm049092j>
- Truchon, J.-F., & Bayly, C. I. (2007). Evaluating Virtual Screening Methods: Good and Bad Metrics for the “Early Recognition” Problem. *Journal of Chemical Information and Modeling*, 47(2), 488-508. <https://doi.org/10.1021/ci600426e>
- Varoquaux, G., Raamana, P. R., Engemann, D. A., Hoyos-Idrobo, A., Schwartz, Y., & Thirion, B. (2017). Assessing and tuning brain decoders: Cross-validation, caveats, and guidelines. *NeuroImage*, 145, 166-179. <https://doi.org/10.1016/j.neuroimage.2016.10.038>
- Vyškovský, R., Schwarz, D., Janoušová, E., & Kašpárek, T. (2016). *Random Subspace Ensemble Artificial Neural Networks for First-episode Schizophrenia Classification*. 317-321. <https://doi.org/10.15439/2016F333>
- Wang, G., Sun, J., Ma, J., Xu, K., & Gu, J. (2014). Sentiment classification: The contribution of ensemble learning. *Decision Support Systems*, 57, 77-93. <https://doi.org/10.1016/j.dss.2013.08.002>
- Zhang, B., Vogt, M., Maggiora, G. M., & Bajorath, J. (2015). Design of chemical space networks using a Tanimoto similarity variant based upon maximum common substructures. *Journal of*

*Computer-Aided Molecular Design*, 29(10), 937-950. <https://doi.org/10.1007/s10822-015-9872-1>

Zhu, H., Tropsha, A., Fourches, D., Varnek, A., Papa, E., Gramatica, P., Öberg, T., Dao, P., Cherkasov, A., & Tetko, I. V. (2008). Combinatorial QSAR Modeling of Chemical Toxicants Tested against *Tetrahymena pyriformis*. *Journal of Chemical Information and Modeling*, 48(4), 766-784. <https://doi.org/10.1021/ci700443v>

# Capítulo 4 | Resultados



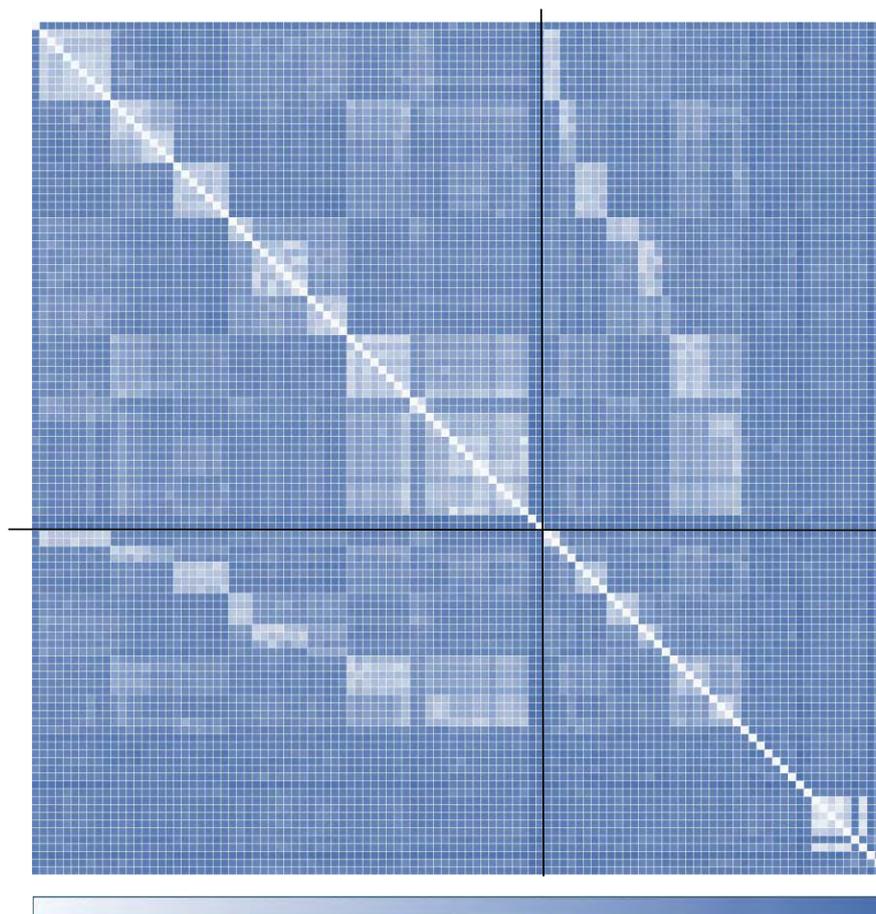
### 4.1 Resultados de la primera campaña: Estrategia Computacional

#### 4.1.1 Compilado del set de datos

Mediante búsqueda bibliográfica se recopilaron 10 trabajos científicos (9 artículos y 1 patente) en los cuales se reportan ensayos de actividad de diferentes compuestos químicos frente a la enzima **TryS** de *T. brucei* (Daunes et al., 2001; D'Silva et al., 2000; Koch et al., 2013; Leroux & Krauth-Siegel, 2016; Linares et al., 2006; Oza et al., 2008; Spinks et al., 2012; Stuhlmann et al., 2007; Torrie et al., 2009; Zimmermann et al., 2013). Debido a la escasez de datos experimentales de actividad inhibitoria reportados para **TcTryS** el modelado se llevó a cabo considerando inhibidores de la enzima de *T. brucei*, dado que es la especie de tripanosomátidos para la cual se ha reportado mayor cantidad de inhibidores. Cabe mencionar que para aquellos compuestos que se evaluaron frente a **TbTryS** y **TcTryS** se consideró solo el valor reportado frente a *T. brucei*, ya que los nueve compuestos para los que se reportaron de valores de  $IC_{50}$  frente a ambas enzimas quedaron etiquetados en diferentes categorías (**ACTIVO/INACTIVO**) según la especie de tripanosomátidos que se observara.

Con los datos disponibles en literatura y luego de los procedimientos de curado, estandarización y eliminación de moléculas duplicadas se generó un set de datos conformado por 109 compuestos químicos.

Aplicando los criterios de corte para discernir entre compuestos **ACTIVOS** e **INACTIVOS** (mencionados en la **sección 3.2.1**) los 109 compuestos del set de datos quedaron divididos en 64 **ACTIVOS** y 45 **INACTIVOS**. Ambos subconjuntos fueron estandarizados con el programa Standardizer Instant JCHEM v. 16.9.12.0, luego se procedió a realizar un mapa de calor o *heatmap* con el programa Gitoools v. 2.3.1. (Perez-Llamas & Lopez-Bigas, 2011), este mapa permite explorar la similitud/disimilitud de los compuestos que conforman el set de datos. El resultado del cálculo de similitud se expresa como 1 menos el valor del coeficiente de similitud; por lo tanto, el máximo valor (1) corresponde a pares de moléculas totalmente disimiles, mientras que el valor mínimo (0) es la situación de mayor similitud. En la **figura 4.1** se puede observar el mapa de calor generado para este set de datos.



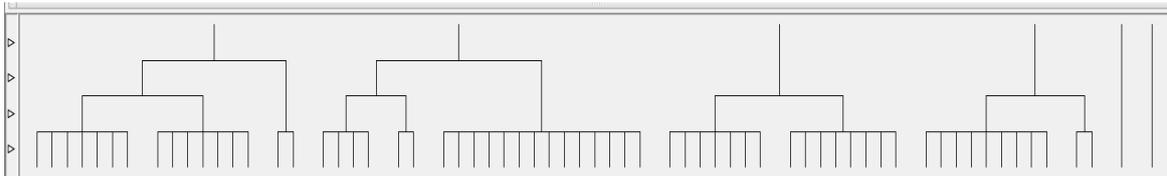
**Figura 4.1.** Mapa de calor para la totalidad de los compuestos incluidos en el set de datos.

En el mapa de calor cada fila o columna está asociada a un determinado compuesto del set de datos. Mientras que cada celda del mapa representa la disimilitud molecular entre el par de compuestos asociados a la fila y columna correspondientes (se trata de un gráfico simétrico respecto a la diagonal). Las celdas de la diagonal (blancas) representan la disimilitud de una molécula consigo misma. La figura muestra la diversidad estructural del set de datos completo, y como se puede apreciar, la mayoría del mapa se encuentra en el tono más intenso de color azul, lo que denota una amplia diversidad estructural de los compuestos. Esta diversidad permite inferir que la región química que se explorará a lo largo del trabajo mediante la construcción de los modelos será lo más extensa posible, y por lo tanto el dominio de aplicación abarcará un gran espacio químico.

### 4.1.2 Partición del set de datos

El procedimiento de agrupamiento jerárquico se realizó solo para la categoría de compuestos **ACTIVOS**, y la técnica generó cuatro agrupamientos conformados por 20, 16, 15 y 11 moléculas cada uno, mientras que dos compuestos fueron considerados *outliers*. El proceso de agrupamiento se llevó a cabo con el programa Library MCS donde se seleccionó un mínimo de siete átomos en la mínima subestructura común, permitiendo la apertura de anillos. Los

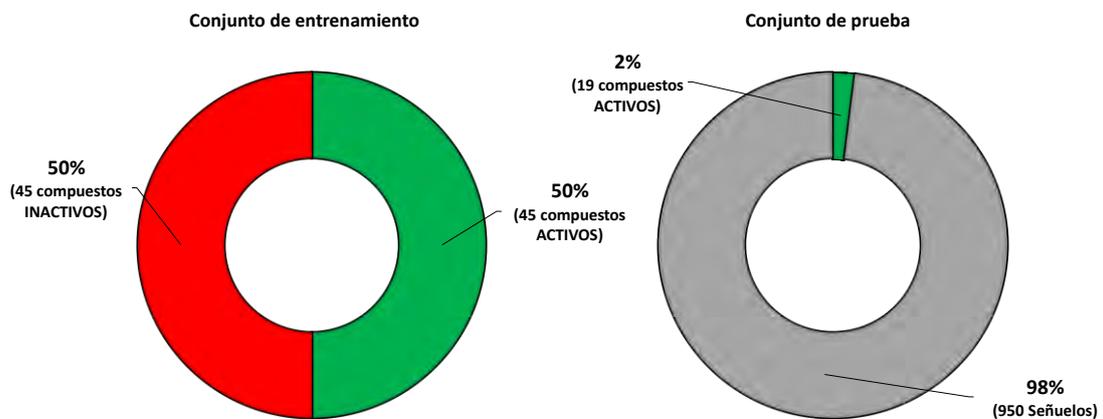
compuestos de la categoría **INACTIVOS** no fueron sometidos a técnicas de agrupamiento, debido a que los mismos se incorporaron en su totalidad al conjunto de entrenamiento. En la **figura 4.2** se observa el dendograma obtenido para el grupo de compuestos **ACTIVOS**.



**Figura 4.2.** Dendograma obtenido para la categoría **ACTIVOS**. Cada nivel corresponde a una subestructura compartida entre las moléculas de un mismo *cluster*.

A partir de cada uno de los agrupamientos se tomó al azar una semilla para realizar un *clustering* bajo el algoritmo *K-means* (ver sección 3.2.2.2). Mediante esta técnica se obtuvieron cuatro nuevos agrupamientos, uno por cada semilla, que contenían 16, 12, 15 y 21 moléculas cada uno. Como se puede apreciar el proceso de agrupación por la técnica de *K-means* modificó la composición de los conjuntos provenientes del *clustering* jerárquico. Las metodologías de agrupamiento permiten una partición representativa del set de datos, minimizando la posibilidad de particiones sesgadas.

Una vez realizada la secuencia de *clustering* se procedió a la conformación del conjunto de entrenamiento y del conjunto de prueba. El primero de los conjuntos quedó conformado por un total de 90 compuestos, incorporando en el mismo la totalidad de las moléculas **INACTIVOS**, y el 70% de los compuestos **ACTIVOS** de cada *cluster*. En consecuencia, el conjunto de prueba constituido por los 19 **ACTIVOS** restantes y un conjunto de señuelos obtenidos a partir de la aplicación *on-line* DUD-e (Mysinger et al., 2012a). En la **figura 4.3** se indica la cantidad de compuestos que integran cada uno de los conjuntos como así también la composición porcentual de cada categoría en los respectivos conjuntos.



**Figura 4.3.** Composición de los conjuntos conformados.

Como se puede ver en la **figura 4.3**, el conjunto de entrenamiento quedó estrictamente balanceado entre compuestos de ambas categorías; mientras que en el conjunto de prueba los **ACTIVOS** representaban algo menos del 2% del total de compuestos, lo que resulta adecuado para calcular métricas de enriquecimiento (Truchon & Bayly, 2007).

### 4.1.3 Generación y validación de los modelos

#### 4.1.3.1 Construcción de los modelos

Se calcularon 3668 descriptores moleculares de baja dimensionalidad (0-2D). Con ellos se generaron 1000 combinaciones aleatorias para conformar subconjuntos de 200 descriptores cada uno.

Cada subconjunto se utilizó para entrenar 1 modelo clasificador, obteniendo así 1000 modelos clasificadores independientes. La construcción de estos modelos se llevó a cabo en el entorno *R studio* (<https://cran.r-project.org/>) aplicando la metodología de pasos hacia adelante (*forward stepwise*), permitiendo un valor de tolerancia  $\beta = 0,5$ , de manera tal que se excluyan de los modelos aquellos descriptores altamente correlacionados. La relación de variables independientes máxima fue establecida en 1:10 entre el número de compuestos del conjunto de entrenamiento y el número de descriptores incluidos en el modelo. Todos estos parámetros fueron configurados en el *script* que se utilizó para generar/entrenar los modelos discriminantes. A continuación, se presentan las ecuaciones de los mejores cinco modelos individuales según el AUCROC obtenida frente al conjunto de prueba.

##### Modelo 321

$$\text{Score} = 0,193 - 0,2199 * NssO + 0,8548 * MATS5m + 3,6289 * RBF$$

##### Modelo 744

$$\begin{aligned} \text{Score} = & -0,12786 - 0,6876 * B10[C - O] + 0,72926 * S-110 + 0,34613 * B09[N - F] \\ & + 0,37486 * C-004 + 0,10336 * SpMAD_{Dz}(e) - 0,04838 \\ & * CATS2D_{(07\_AL)} \end{aligned}$$

##### Modelo 628

$$\begin{aligned} \text{Score} = & 0,30265 - 0,01026 * P_{(VSA_{(LogP_5)})} + 37,08550 * JGI8 + 0,22687 \\ & * B01[C - F] - 0,12146 * NsssCH \end{aligned}$$

##### Modelo 429

$$\text{Score} = -2,0646 - 0,6556 * Eig04_{EA(dm)} - 0,4912 * SpDiam_{EA} + 0,3155 * B10[N - N]$$

##### Modelo 470

$$\begin{aligned} \text{Score} = & -0,3005 - 0,0510 * Eig13_{EA(ed)} - 1,1307 * GATS7i - 0,4879 * B10[C - O] \\ & + 0,1952 * TI2_L + 0,5214 * B05[S - S] - 0,3847 * O-059 - 0,926 \\ & * GATS6p - 0,4252 * B03[N - S] - 0,6932 * nRSR \end{aligned}$$

Los descriptores que conforman los mejores cinco modelos individuales se describen en el **Anexo I**.

Puede observarse que solo uno de estos cinco modelos (modelo 470) presenta una relación de 1 descriptor por cada 10 moléculas del conjunto de entrenamiento, y teniendo en cuenta que dicho conjunto cuenta con un total de 90 moléculas, la relación entre el número de descriptores y el número de moléculas en el conjunto de entrenamiento que presentan el resto de los modelos se encuentra entre 1 a 15 y 1 a 30, lo que permite inferir una baja probabilidad de sobreajuste.

#### 4.1.3.2 Validación de los modelos

La validación de los modelos se realizó a través de distintas técnicas las cuales incluyen tanto validaciones internas como externas. Este procedimiento tiene como objetivo estudiar diversas métricas que permiten evaluar la robustez y la capacidad predictiva de los modelos generados, evaluando además la posibilidad de correlaciones espurias.

En la **tabla 4.1** se presentan los resultados de la validación interna para los mejores cinco modelos, allí se informan los %BC obtenidos para el modelo original, la validación cruzada y la aleatorización de Fisher. En todas las pruebas se utilizó “0,5” como valor de corte para discernir entre compuestos **ACTIVOS** e **INACTIVOS**. Dicho valor surge del promedio de los valores asignados a cada una de las categorías que componen el conjunto de entrenamiento. (**INACTIVOS** = 0; **ACTIVOS** = 1) (Toropova & Toropov, 2017).

Recordemos que en la prueba de validación cruzada, se evaluó el desempeño de los modelos generados a partir de la remoción de subconjuntos aleatorios (pero estratificados por clase) de 10 moléculas del conjunto de entrenamiento original; mientras que en la aleatorización de Fisher se construyeron los modelos a partir de la asignación al azar las categorías **ACTIVOS** e **INACTIVOS**. Estos procedimientos fueron descriptos con mayor detalle en la **sección 3.3.1.2**.

**Tabla 4.1.** Resultados de las diferentes técnicas de validación interna para los mejores cinco modelos individuales

| MODELO | Original | Validación Cruzada |      | Aleatorización de Fisher |                    |                    |
|--------|----------|--------------------|------|--------------------------|--------------------|--------------------|
|        | %BC      | %BC                | DE   | %BC                      | Intervalo Inferior | Intervalo Superior |
| 321    | 74,4     | 64,2               | 15,6 | 58,9                     | 53,9               | 63,9               |
| 744    | 80,0     | 63,8               | 15,3 | 59,6                     | 54,8               | 64,3               |
| 628    | 75,6     | 63,6               | 15,8 | 60,4                     | 56,8               | 64,0               |
| 429    | 72,6     | 59,8               | 14,8 | 63,1                     | 58,0               | 68,2               |
| 470    | 80,0     | 72,2               | 14,0 | 60,3                     | 55,2               | 65,4               |

Como puede observarse en la **tabla 4.1** los modelos originales han conseguido, como es de esperar, mejores %BC respecto a la prueba de aleatorización de Fisher; ya que el valor superior del intervalo de confianza que se alcanza con esta validación se encuentra siempre por debajo de los resultados obtenidos para cada uno de los modelos originales. Sin embargo, los %BC obtenidos mediante validación cruzada se encuentran siempre por debajo (alrededor de 10%) del valor reportado para los modelos originales. Este dato permite inferir que existe

un pequeño sobreajuste de los modelos, lo cual parece lógico si pensamos en la escasez de compuestos con la que se contaba al momento de realizar la presente campaña de modelado.

La capacidad predictiva de los modelos fue evaluada frente al conjunto de prueba . Los resultados obtenidos se expresan en la **tabla 4.2**, donde se presentan los valores de AUCROC y los %BC globales para el conjunto de prueba.

**Tabla 4.2.** Resultados de los mejores cinco modelos individuales frente a los diferentes conjuntos

| MODELO | Conjunto de Prueba |      | Conjunto de Entrenamiento |
|--------|--------------------|------|---------------------------|
|        | AUCROC             | %BC  | %BC                       |
| 321    | 0,870 (0,025)      | 74,1 | 74,4                      |
| 744    | 0,831 (0,035)      | 76,3 | 80,0                      |
| 628    | 0,807 (0,040)      | 77,8 | 75,6                      |
| 429    | 0,793 (0,049)      | 71,3 | 72,2                      |
| 470    | 0,790 (0,049)      | 61,4 | 80,0                      |

Como se desprende de la **tabla 4.2** los modelos presentan un desempeño aceptable frente al conjunto de entrenamiento, ya que todos los resultados muestran %BC por encima de 70%. Se puede asumir, entonces, que estos resultados se ven bien reflejados en el conjunto de prueba, ya que los %BC son similares a los valores observados para el conjunto de entrenamiento, además los valores de AUCROC frente al conjunto de prueba son siempre mayores a 0,75 lo que indica una buena capacidad predictiva de los modelos individuales.

#### 4.1.4 Evaluación del desempeño de los modelos

Luego de las validaciones realizadas frente a los diferentes conjuntos (entrenamiento y prueba) se analizaron otros parámetros que permiten estimar el desempeño de cada modelo; en este sentido se evaluó la Sensibilidad (tasa de verdaderos positivos), la Especificidad (tasa de verdaderos negativos), el %BC, la Precisión y el parámetro *F-measure*, este último permite estimar la superposición entre la clase real y la clase predicha para un mismo compuesto. El parámetro *F-measure* se puede calcular como la media armónica entre la *Se* y la precisión (**ecuación 4.1**), donde la *Se* es el valor que se obtiene de dividir la cantidad de **ACTIVOS** correctamente clasificados sobre la sumatoria de todos los **ACTIVOS** del conjunto, mientras que la precisión es el resultado que se obtiene al dividir los **ACTIVOS** clasificados correctamente por la cantidad de **ACTIVOS** que predijo el modelo (esto corresponde a la sumatoria de VP + FN). Este parámetro permite evaluar el grado de clasificaciones incorrectas que realizó el modelo y toma valores entre 0 y 1, donde el máximo valor (1) indica que la precisión y recuperación del modelo son perfectas (Roy & Mitra, 2011).

$$F - measure = \frac{2Se * Precisión}{Se + Precisión} \quad \text{Ec. 4.1}$$

Como se desprende de la **ecuación 4.1**, o de la definición antes mencionada, el parámetro *F-measure* ignora la tasa de **INACTIVOS** que fueron correctamente clasificados, en tanto se enfoca en la tasa de buenas clasificaciones dentro de los conjuntos de los verdaderos **ACTIVOS** y de los compuestos predichos por el modelo como tales (aunque no necesariamente lo sean).

En las **tablas 4.3** y **4.4** se presentan los valores obtenidos para cada uno de los parámetros antes mencionados.

**Tabla 4.3.** Parámetros de desempeño evaluados frente al conjunto de entrenamiento

| Modelo | Sensibilidad | Especificidad | %BC  | Precisión | <i>F-measure</i> |
|--------|--------------|---------------|------|-----------|------------------|
| 321    | 0,867        | 0,622         | 74,4 | 0,696     | 0,772            |
| 744    | 0,756        | 0,829         | 80,0 | 0,829     | 0,791            |
| 628    | 0,911        | 0,695         | 75,6 | 0,695     | 0,789            |
| 429    | 0,933        | 0,656         | 72,2 | 0,656     | 0,771            |
| 470    | 0,867        | 0,765         | 80,0 | 0,765     | 0,813            |

**Tabla 4.4.** Parámetros de desempeño evaluados frente al conjunto de prueba

| Modelo | Sensibilidad | Especificidad | %BC  | Precisión | <i>F-measure</i> |
|--------|--------------|---------------|------|-----------|------------------|
| 321    | 0,842        | 0,740         | 74,1 | 0,061     | 0,113            |
| 744    | 0,737        | 0,763         | 76,3 | 0,059     | 0,109            |
| 628    | 0,632        | 0,780         | 77,7 | 0,054     | 0,100            |
| 429    | 0,895        | 0,710         | 71,3 | 0,058     | 0,109            |
| 470    | 0,790        | 0,611         | 61,4 | 0,039     | 0,074            |

Observando la **tabla 4.3** puede apreciarse que los modelos han demostrado un buen desempeño frente al conjunto de entrenamiento, dicha inferencia se fundamenta en los valores que se observan para cada uno de los parámetros informados en la tabla. Sin embargo al observar la **tabla 4.4**, correspondiente al conjunto de prueba, se puede ver una significativa caída de los resultados obtenidos en la precisión de los modelos, estos valores pueden ser explicados a partir del desbalance de clases que presenta dicho conjunto (<2% **ACTIVOS**), es decir, a pesar de que los modelos cometen relativamente pocos errores cuando clasifican a los compuestos **INACTIVOS**, la acumulación de estos errores en un conjunto con una elevada proporción de **INACTIVOS** conduce a un bajo valor de precisión. Este resultado, a su vez, se traduce en un bajo puntaje de *F-measure*, ya que como se mencionó anteriormente este parámetro no le da peso a los verdaderos **INACTIVOS** correctamente clasificados.

A pesar de los buenos resultados obtenidos hasta aquí, se ha recurrido a la técnica de combinación de modelos (o aprendizaje por ensamblado) mediante combinaciones de dos o más modelos con el objetivo de mejorar la robustez e incrementar la capacidad predictiva.

### 4.1.5 Combinación de modelos

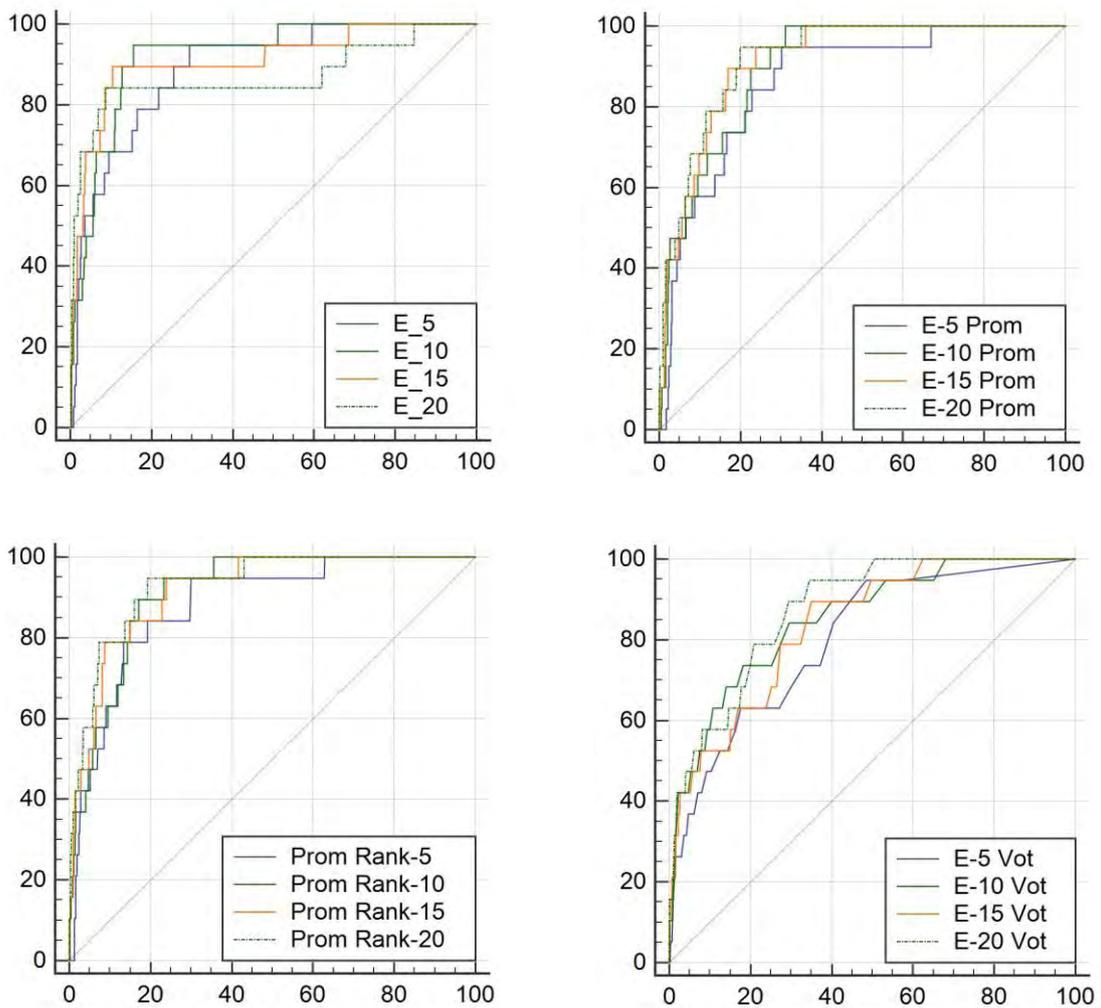
Para cada operador de ensamblado (MÍNIMO; PROMEDIO, PROMEDIO del *RANKING* y PROMEDIO del *VOTO*) se generaron cuatro combinaciones en series de 5, 10, 15 y 20 modelos para cada ensamble o meta-clasificador, y se evaluó el AUCROC en el conjunto de prueba para

cada una de las combinaciones generadas. Los resultados obtenidos se resumen en la **tabla 4.5**, donde también se presentan las métricas para el mejor modelo individual. Adicionalmente, en la **figura 4.4** pueden observarse las curvas ROC obtenidas para las diferentes combinaciones.

**Tabla 4.5.** AUCROC frente al mejor modelo individual y a cada uno de los ensambles

**AUCROC CONJUNTO DE PRUEBA**

| Clasificador       | Mínimo        | Promedio      | Ranking       | Voto          |
|--------------------|---------------|---------------|---------------|---------------|
| <b>Modelo 321</b>  | 0,870 (0,025) | 0,870 (0,025) | 0,870 (0,025) | 0,870 (0,025) |
| <b>Ensamble 5</b>  | 0,889 (0,034) | 0,863 (0,037) | 0,881 (0,036) | 0,802 (0,050) |
| <b>Ensamble 10</b> | 0,919 (0,027) | 0,899 (0,024) | 0,912 (0,022) | 0,851 (0,044) |
| <b>Ensamble 15</b> | 0,909 (0,041) | 0,914 (0,023) | 0,918 (0,025) | 0,837 (0,043) |
| <b>Ensamble 20</b> | 0,869 (0,610) | 0,921 (0,022) | 0,931 (0,025) | 0,876 (0,033) |



**Figura 4.4.** Curvas ROC para cada uno de los operadores. a) Mínimo, b) Promedio, c) Promedio del *Ranking* y d) Promedio del Voto

A partir de los valores observados en la **tabla 4.5** y de la evaluación analítica y visual de las curvas ROC para cada uno de los operadores se ha optado por evaluar los cuatros ensambles que se obtuvieron mediante el operador mínimo. Para ello se calcularon los mismos parámetros que se utilizaron para estimar el desempeño de los modelos individuales, dichos parámetros fueron evaluados frente al conjunto de entrenamiento (**tabla 4.6**) y al conjunto de prueba (**tabla 4.7**).

**Tabla 4.6.** Parámetros de desempeño de los ensambles generados mediante el operador mínimo evaluados frente al conjunto de entrenamiento

| Clasificador       | Sensibilidad | Especificidad | %BC  | Precisión | <i>F-measure</i> |
|--------------------|--------------|---------------|------|-----------|------------------|
| <b>Modelo 321</b>  | 0,867        | 0,696         | 74,4 | 0,696     | 0,772            |
| <b>Ensamble 5</b>  | 0,556        | 0,956         | 75,6 | 0,926     | 0,694            |
| <b>Ensamble 10</b> | 0,444        | 0,952         | 71,1 | 0,952     | 0,606            |
| <b>Ensamble 15</b> | 0,378        | 0,944         | 67,8 | 0,944     | 0,540            |
| <b>Ensamble 20</b> | 0,333        | 1,000         | 66,7 | 1,000     | 0,500            |

**Tabla 4.7.** Parámetros de desempeño de los ensambles generados mediante el operador mínimo evaluados frente al conjunto de prueba

| Clasificador       | Sensibilidad | Especificidad | %BC  | Precisión | <i>F-measure</i> |
|--------------------|--------------|---------------|------|-----------|------------------|
| <b>Modelo 321</b>  | 0,842        | 0,739         | 74,1 | 0,061     | 0,113            |
| <b>Ensamble 5</b>  | 0,368        | 0,985         | 97,3 | 0,333     | 0,350            |
| <b>Ensamble 10</b> | 0,158        | 0,998         | 98,1 | 0,600     | 0,250            |
| <b>Ensamble 15</b> | 0,158        | 0,999         | 98,2 | 0,750     | 0,261            |
| <b>Ensamble 20</b> | 0,105        | 0,999         | 98,1 | 0,667     | 0,182            |

La evaluación de los parámetros frente a cada uno de los conjuntos demostró que la combinación de modelos condujo a mejores valores para cada parámetro respecto al mejor modelo individual. Se puede observar que todos los ensambles generados han elevado en más de 20% el porcentaje de buenas clasificaciones frente al conjunto de prueba, como así también han mejorado notoriamente el resto de los parámetros. Aplicando el principio de parsimonia, el cual establece que ante dos modelos que expliquen un mismo resultado se debe optar por aquel que resulte menos complejo, se ha seleccionado para realizar el cribado virtual el ensamble **E-10** generado mediante el operador Mínimo, ya que para el mismo se obtienen mejores valores de AUCROC en el conjunto de prueba, y a su vez un alto %BC (>98%) en dicho conjunto.

### 4.1.6 Selección del valor de corte para cribado virtual

El ensamblado de los mejores modelos individuales conduce, generalmente, a una mejora de los resultados de clasificación en término de **Sp**.

Es importante optimizar la **Sp** ya que esta optimización se ve reflejada como una reducción de falsos positivos, es decir, compuestos **INACTIVOS** que el modelo/ensamble reconoce como **ACTIVOS** y en consecuencia pasarían a la etapa de evaluaciones *in vitro*. En el contexto de la investigación pública es imprescindible minimizar el número de falsos positivos,

ya que existen amplias limitaciones de recursos para adquirir y evaluar candidatos; cabe recordar que la ganancia en términos de **Sp** trae aparejado una pérdida de la **Se**.

En esta campaña el valor de corte se obtuvo a partir del análisis de la curva ROC, buscando así el mejor balance entre la **Se** y la **Sp**. El *score* seleccionado para discernir entre compuestos **ACTIVOS** e **INACTIVOS** fue de 0,3026, donde la **Sp** fue de 94,7% y la **Se** 84,2% frente al conjunto prueba.

### 4.1.7 Cribado virtual y selección de compuestos

El cribado virtual se llevó a cabo frente a las bibliotecas DrugBank 3.0 (Knox et al., 2011) y SweetLead (Novick et al., 2013). En esta oportunidad se utilizó el meta-clasificador **E-10**, generado mediante el operador mínimo, y aplicando el *score* seleccionado (0,3026) recogimos un total de 169 compuestos químicos que el ensamble predijo como **ACTIVOS**, de los cuales solo 44 corresponden a fármacos aprobados. El resto de los compuestos (125) pertenecían a las categorías de experimentación/investigación, o biotecnológicos, por lo que fueron descartados para este trabajo de reposicionamiento.

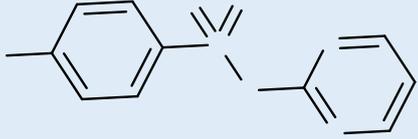
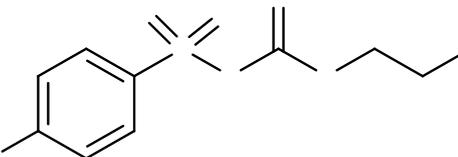
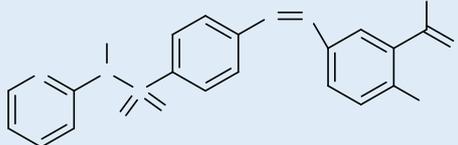
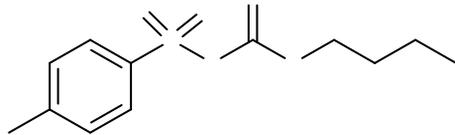
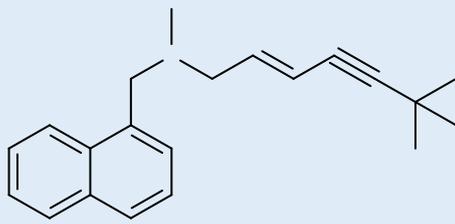
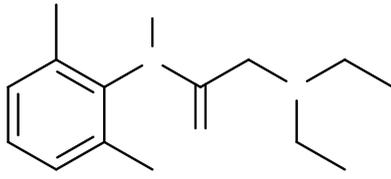
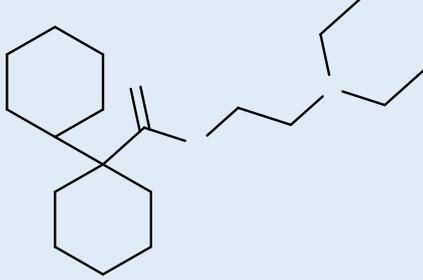
Para conocer si un compuesto predicho como **ACTIVO** se encuentra dentro del dominio de aplicación (DA) del meta-clasificador se aplicó la técnica de la palanca, y se comprobó que los 44 fármacos se encontraron dentro del mismo.

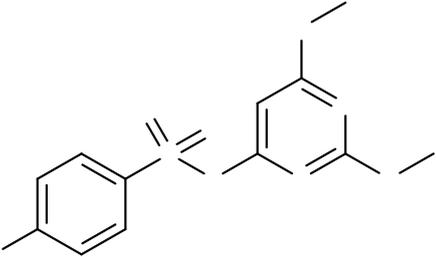
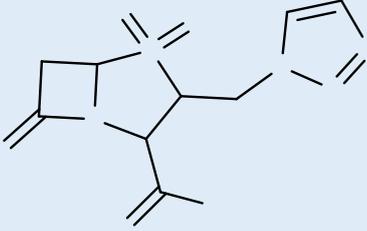
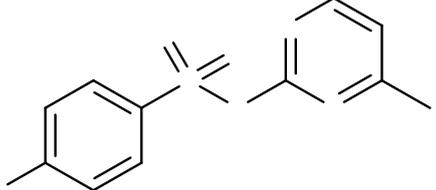
Para seleccionar qué compuestos adquirir se procedió a un curado manual de los 44 fármacos seleccionados. En primer lugar, se verificó que ninguno de los *hits in silico* tuviera actividad reportada frente a la enzima Tc**TryS**. Seguidamente se evaluó la disponibilidad y factibilidad para adquirir los candidatos seleccionados, y en este paso quedaron excluidos siete fármacos debido a que seis de ellos no estaban disponibles en los proveedores habituales y el sexto es un analgésico perteneciente a la familia de los opioides (Dextropropoxifeno) por lo que su comercialización es rigurosamente controlada.

Dado el entorno socioeconómico y cultural que viven las poblaciones más afectadas por esta patología, y teniendo en cuenta que el parásito es de vida intracelular en mamíferos, aplicamos un análisis manual adicional, en el que se evaluó la vía de administración, indicación clínica original, mecanismo de acción y dosis máxima recomendada por día para cada uno de los 37 candidatos restantes. En esta etapa se excluyeron otros 8 candidatos, tres de ellos debido a que son fármacos de uso tópico (Hexetindina, Dapiparazol y Butenafina), Clopidogrel por ser un profármaco (fármaco que debe metabolizarse en el hígado para transformarse en un compuesto activo), Fentolamina debido a que se aplica por vía intravenosa o intramuscular, lo que dificultaría un tratamiento ambulatorio, y los 3 fármacos restantes debido a sus mecanismos de acción, ya que pertenecen a la familia de bloqueantes de receptores ganglionares y en consecuencia presentan un perfil seguridad cuestionable para el uso ambulatorio.

Finalmente y mediante una inspección visual detallada de las 29 moléculas restantes se decidió evaluar los fármacos que se presentan en la **tabla 4.8**.

Tabla 4.8 Candidatos seleccionados para ser evaluados en ensayos *in vitro*

| Fármaco          | Indicación original                  | Estructura   |
|------------------|--------------------------------------|--|
| Sulfadiazina     | Antibiótico                          |    |
| Clorpropamida    | Hipoglucemiante                      |    |
| Sulfasalazina    | Antiinflamatorio/<br>Inmunomodulador |    |
| Tolbutamida      | Hipoglucemiante                      |   |
| Terbinafina      | Antifúngico                          |  |
| Lidocaína (base) | Anestésico local/<br>antiarrítmico   |  |
| Diciclomina      | Anticolinérgico                      |  |

|                        |  |   |
|------------------------|--|---|
| <b>Sulfadimetoxina</b> | Antibiótico                                    |  <p>The chemical structure of Sulfadimetoxina (Sulfamethoxazole) consists of a benzothiazine ring system. It features a benzene ring fused to a five-membered thiazine ring. The benzene ring has a methyl group at the 6-position and a methoxy group at the 7-position. The thiazine ring has a methyl group at the 5-position and a methoxy group at the 2-position.</p> |
| <b>Tazobactam</b>      | Inhibidor suicidad de le enzima beta-lactamasa |  <p>The chemical structure of Tazobactam is a beta-lactamase inhibitor. It features a beta-lactam ring fused to a five-membered thiazolidine ring. The thiazolidine ring has a methyl group at the 4-position and a propyl group at the 5-position. The beta-lactam ring has a methyl group at the 2-position and a propyl group at the 3-position.</p>                     |
| <b>Sulfamerazina</b>   | Antibiótico                                    |  <p>The chemical structure of Sulfamerazina (Sulfamethazine) consists of a benzothiazine ring system. It features a benzene ring fused to a five-membered thiazine ring. The benzene ring has a methyl group at the 6-position and a methoxy group at the 7-position. The thiazine ring has a methyl group at the 5-position and a methoxy group at the 2-position.</p>    |

### 4.2 Primera Campaña de Evaluación Experimental

Los ensayos experimentales se realizaron en colaboración con el Instituto de Ciencia y Tecnología Cesar Milstein (CONICET- Fundación Pablo Cassará); el Instituto de Investigaciones en Microbiología y Parasitología Médica -IMPAM- (UBA-CONICET) y el Laboratorio de Biología Redox de Tripanosomátidos, Instituto Pasteur de Montevideo -IPM- (Uruguay).

Como hemos mencionado en el capítulo anterior, la validación biológica se llevó a cabo frente a dos especies de tripanosomátidos.

#### 4.2.1 Evaluación biológica frente a *T. cruzi*

##### 4.2.1.1 Evaluación frente a epimastigotes

La evaluación biológica de los candidatos seleccionados se llevó a cabo frente a epimastigotes de la cepa “Y-GFP” de *T. cruzi*. Se realizó un único ensayo donde se evaluó cada uno de los fármacos a concentraciones fijas de 25, 50, 100 y 200  $\mu\text{M}$  con dos réplicas para cada concentración. En la **figura 4.5** se presentan las curvas de proliferación parasitaria obtenidas para cada uno de los fármacos y el control negativo con DMSO. Las curvas muestran la media de la variación de la densidad óptica (DO) a los días cuatro y ocho posterior a la siembra y sus respectivos desvíos en función de la cantidad de parásitos a cada una de las concentraciones.

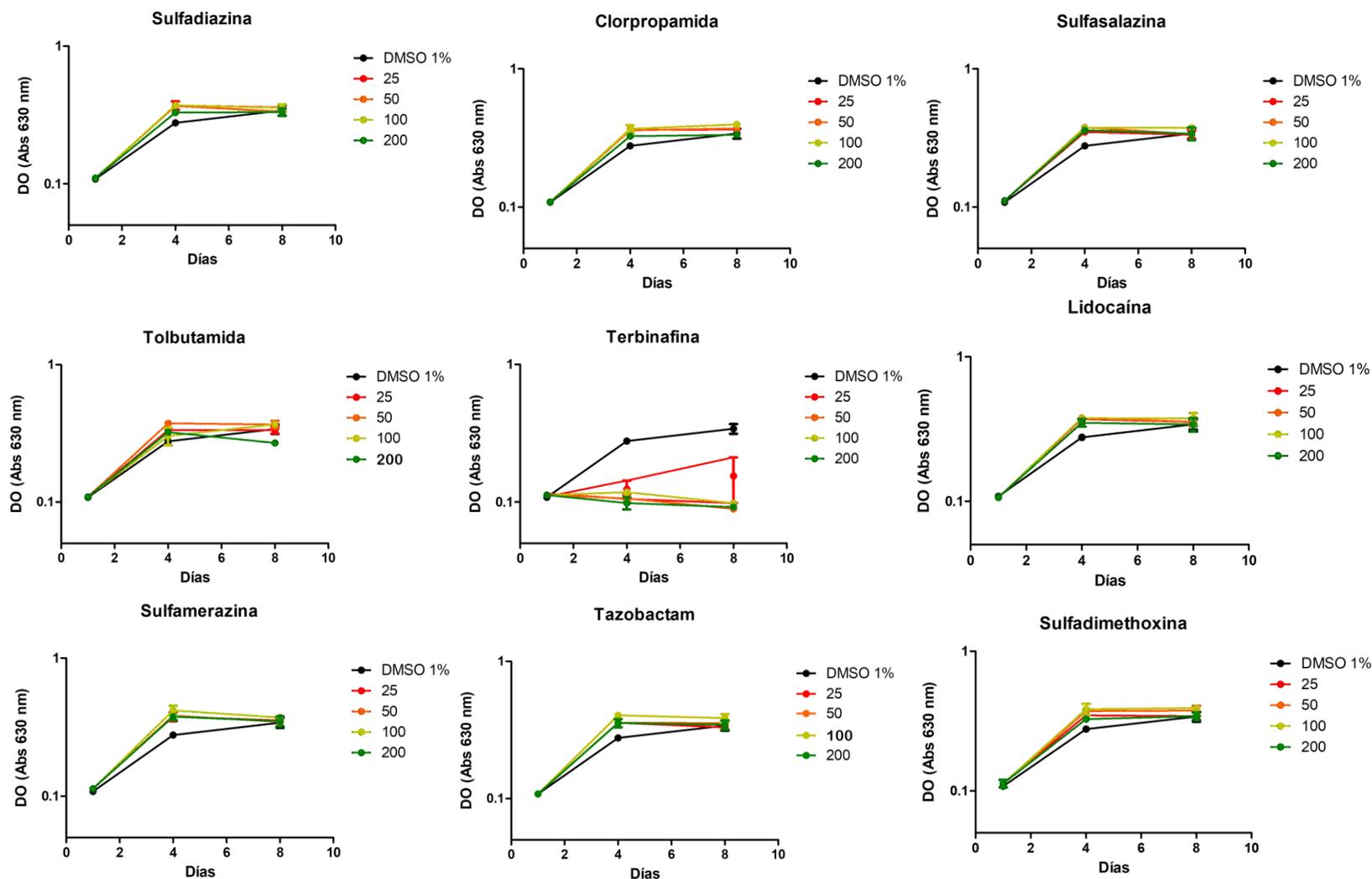
Los resultados obtenidos se analizaron en el programa GraphPad<sup>5</sup>, para ello se realizó una prueba de Análisis de la varianza (ANOVA) de una vía, seguida de la prueba estadística de Dunnett (\*p < 0,05). Esta prueba se llevó adelante utilizando el % de proliferación obtenido para cada pocillo a los días cuatro y ocho posterior a la siembra. Los resultados mostraron que solo Terbinafina presenta diferencias significativas (p-valor < 0,05) en la proliferación parasitaria respecto al control negativo (DMSO) para las cuatro concentraciones evaluadas, mientras que el resto de los fármacos no presentaron diferencias significativas respecto al mismo control para ninguna de las concentraciones ensayadas.

A partir de los resultados obtenidos se realizó la evaluación de la  $\text{EC}_{50}$  para el fármaco Terbinafina frente a epimastigotes, encontrándose un valor de 7,5  $\mu\text{M}$  ( $\pm 1,8$ ).

Diciclomina no fue evaluado en esa etapa, debido a que el mismo fue reportado por nuestro grupo de investigación como modulador de otro blanco terapéutico. Para esta droga los ensayos reportaron un valor de  $\text{EC}_{50} = 4,22 \mu\text{M}$ ; también sobre la cepa “Y” (Alberca et al., 2018).

---

<sup>5</sup> <https://www.graphpad.com/>

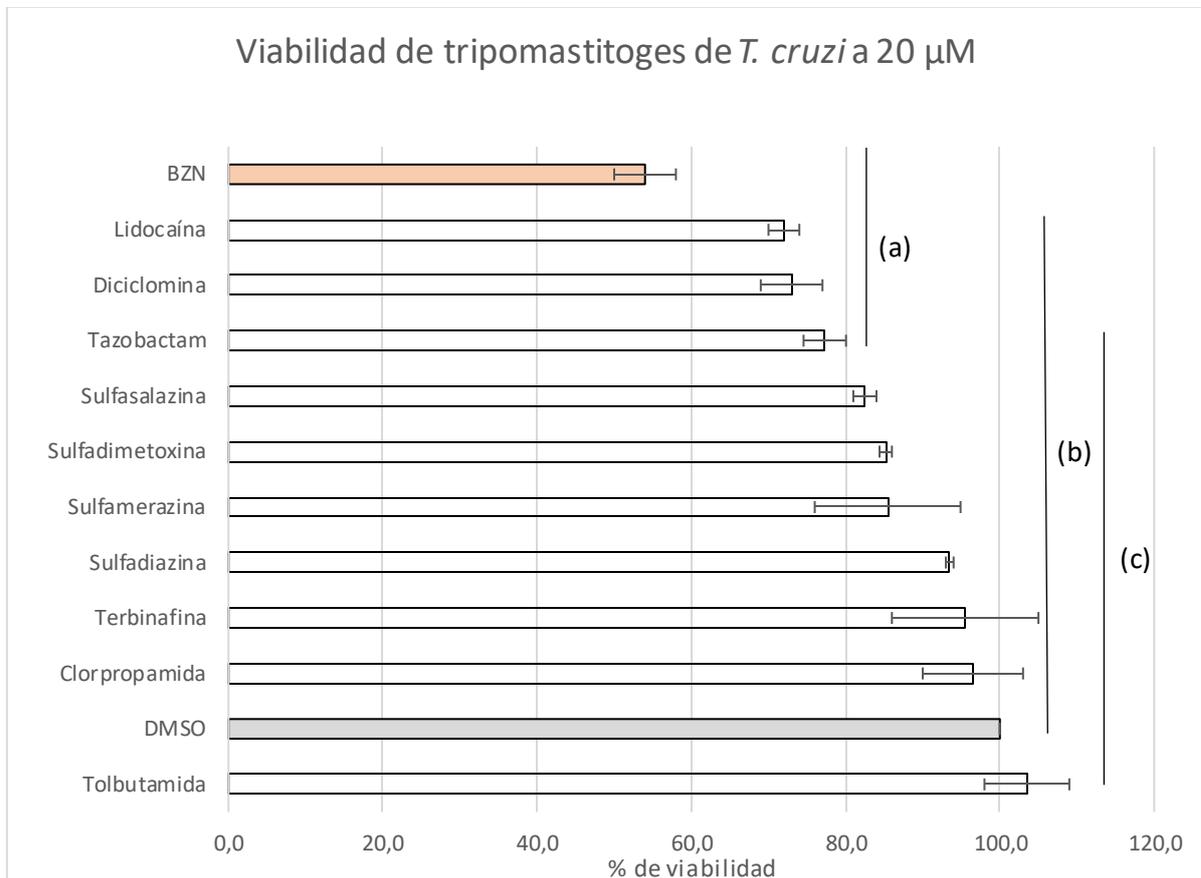


**Figura 4.5.** Curvas de proliferación de epimastigotes de *T. cruzi* frente a cada fármaco evaluado. Los resultados son expresados como la media  $\pm$  DE de un ensayo por duplicado. El análisis estadístico se realizó mediante una prueba de ANOVA de una vía seguida de una prueba de comparación múltiple post hoc de Dunnett (\* $p < 0,05$ ).

### 4.2.1.2 Evaluación frente a tripomastigote

Para conocer si los candidatos presentaban efecto tripanocida frente al estadio tripomastigote, se ensayaron los fármacos frente a la cepa “RA” de *T. cruzi*. Aquí se evaluó la viabilidad de los parásitos frente a una concentración fija de 20  $\mu\text{M}$  para cada candidato.

En la **figura 4.6** se presenta la media del porcentaje de viabilidad obtenido a partir de dos ensayos independientes con su desviación estándar asociada.



**Figura 4.6.** Porcentaje de viabilidad de tripomastitoges. La condición de control negativo se realizó con 1% DMSO (100 % de viabilidad). Los resultados se expresan como % de la media  $\pm$  DE. Las medias de los porcentajes con letras iguales indican que no existen diferencias significativas ( $p$ -valor  $>$  0,05).

El análisis de los resultados se realizó mediante un análisis ANOVA y posteriormente una comparación de medias mediante la prueba de Tukey con un  $p$ -valor  $<$  0,05, utilizando el software Infostat<sup>6</sup>. Los resultados estadísticos demuestran que ninguno de los fármacos ensayados presentó diferencias significativas respecto al control negativo (DMSO).

En general, se espera de un *hit* validado una  $EC_{50}$  por debajo de 10  $\mu\text{M}$ , diversos autores sugieren normalizar los criterios de aceptabilidad para un nuevo fármaco o compuesto líder, y consideran que estos compuestos deben denotar su actividad moduladora (inhibitoria, activadora, antimicrobiana, etc.) en ensayos celulares donde la concentración de los

<sup>6</sup> <https://www.infostat.com.ar/>

compuestos evaluados se encuentre entre 1 y 10  $\mu\text{M}$  como máximo, sobre todo cuando el agente patógeno invade tejidos, y en consecuencia el nuevo fármaco debe atravesar diversas membranas biológicas (células humanas y parásitos y cada una con diferentes características biológicas) para acceder al *target* bioquímico del parásito (Hughes et al., 2011; Katsuno et al., 2015). El valor de concentración que se ha utilizado duplica la cota de concentración sugerida en literatura, y aun así no se observaron efectos significativos frente a tripomastigotes (la única forma clínicamente relevante de *T. cruzi* contra la que se probaron los candidatos), en tanto a concentración 20  $\mu\text{M}$  ninguno de los fármacos evidenció efecto tripanocida por encima del 50%.

Por otra parte, se analizaron cuestiones farmacocinéticas y farmacológicas para decidir si valía la pena proseguir con la evaluación de los candidatos seleccionados. Este análisis permitió establecer que algunos de los fármacos presentan concentraciones plasmáticas máximas ( $C_{p_{\max}}$ ) en el rango bajo micromolar,  $C_{p_{\max}} < 10 \mu\text{M}$ . Para el fármaco Terbinafina se obtienen  $C_{p_{\max}}$  del orden de 3  $\mu\text{M}$ <sup>7</sup>, mientras que para Sulfasalazina (Schröder & Campbell, 1972) y Lidocaína<sup>8</sup> dichos valores se encuentran alrededor de 5,8 y 6,4  $\mu\text{M}$  respectivamente, estas concentraciones se encuentran entre 3 y 6 veces por debajo de la concentración evaluada (20  $\mu\text{M}$ ), lo que sugiere que sería poco viable establecer una posología adecuada, ya que para obtener una  $C_p$  próxima a 20  $\mu\text{M}$  debería al menos (considerando relación lineal entre de  $C_p$  y la dosis) triplicarse la dosis original de Lidocaína y sextuplicar la dosis de Terbinafina.

Concentraciones plasmáticas elevadas se han reportado para los hipoglucemiantes Clorpropamida y Tolbutamida. Sin embargo, estos valores provienen de medidas con gran variabilidad interindividual que van desde prácticamente cero para ambos fármacos hasta valores de 800  $\mu\text{M}$  y 370  $\mu\text{M}$  para Clorpropamida y Tolbutamida, respectivamente (Melander et al., 1978); esta variabilidad en las concentraciones plasmáticas, en combinación con el mecanismo de acción (liberación de insulina) establecen condiciones poco seguras para el tratamiento de pacientes no diabéticos, sobre todo por la dificultad de establecer un rango de dosis óptimo, pero también porque frente a pacientes no diabéticos estos fármacos podrían generar, a concentraciones terapéuticas, episodios de hipoglucemia (Contreras et al., 2002).

Por los motivos antes expuestos se decidió finalizar aquí las evaluaciones biológicas de estos fármacos frente a *T. cruzi*.

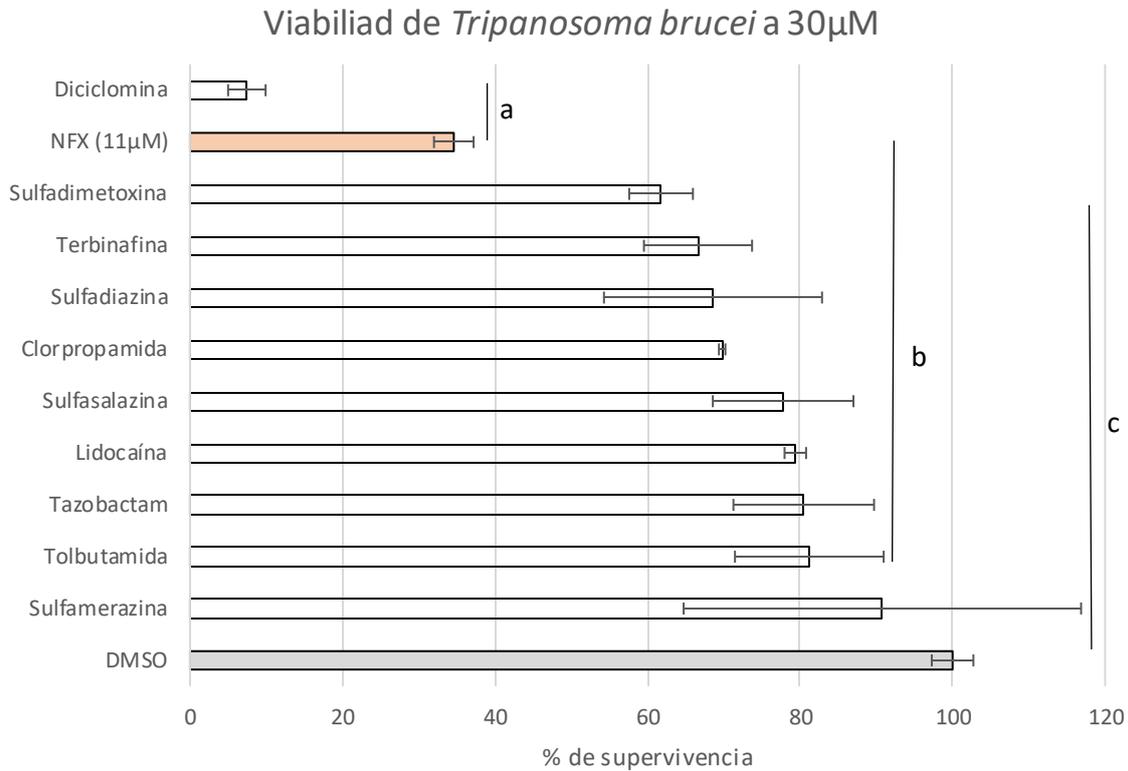
### 4.2.2 Evaluación biológica frente a *T. brucei*

Para el presente ensayo se utilizó la línea celular 449 de la cepa 427 para determinar la viabilidad del estadio sanguíneo de *T. brucei* frente a concentración fija de 30  $\mu\text{M}$  para cada uno de los fármacos. Para el análisis estadístico de los resultados se realizó un análisis de la varianza, y posteriormente se compararon las medias mediante la prueba de Tukey utilizando 0,05 como punto de corte para el p-valor, este procedimiento se llevó a cabo en software Infostat. En la **figura 4.7** se observan los resultados de viabilidad obtenidos luego de dos ensayos independientes.

---

<sup>7</sup> <https://ar.privademeccum.com/>

<sup>8</sup> <https://cima.aemps.es/>



**Figura 4.7.** Porcentaje de viabilidad de *T. brucei* frente a cada uno de los fármacos seleccionados. La condición de control negativo se realizó con 1% DMSO (100% de viabilidad) y mientras que como control positivo se utilizó NFX a 11 µM. Los resultados se expresan como % de la media ± DE. Las medias de los porcentajes con letras iguales indican que no existen diferencias significativas (p-valor > 0,05). Prueba de Tukey (p-valor < 0,05).

Como puede observarse en la **figura 4.7** y según se desprende del análisis estadístico, los tratamientos agrupados con la misma letra no presentan diferencias estadísticamente significativas en relación con su capacidad tripanocida, Diccilomina fue el único compuesto que presentó un efecto tripanocida similar al control positivo (naranja), mientras que el resto de los fármacos no mostraron diferencias significativas respecto al control negativo (gris). Dado que en este ensayo el valor de concentración utilizado triplica el límite superior del rango de concentraciones sugeridas, se ha determinado finalizar en este punto el estudio de los fármacos que no difieren del control negativo, mientras tanto para Diccilomina se realizó la curva de concentración-respuesta a fin de obtener un valor de EC<sub>50</sub> para este fármaco frente a *T. brucei*.

Con el fin de seleccionar un rango de concentraciones óptimo para realizar la determinación de EC<sub>50</sub>, se procedió a un primer barrido de concentraciones distribuidas de la siguiente manera: 2,5, 5, 10, 20 y 40 µM. En la **figura 4.8** se presenta el gráfico de la curva obtenida.

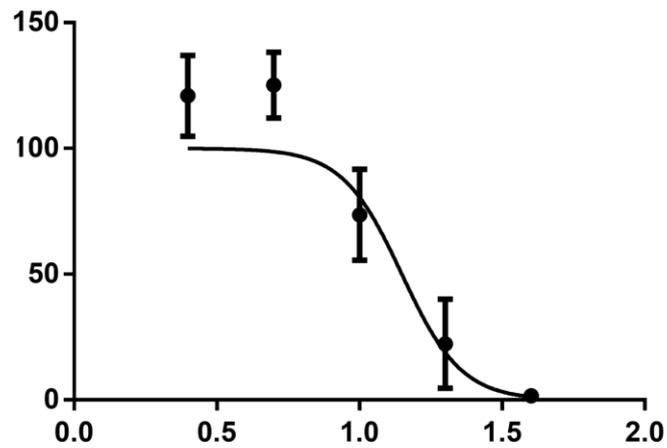


Figura 4.8. Curva para optimización de concentraciones de EC50

Una vez obtenidos los resultados del primer barrido de concentraciones se procedió a realizar la curva de viabilidad en función de la concentración de Diclomina bajo el siguiente rango de concentraciones: 1, 4, 8, 16, 24, y 48  $\mu\text{M}$ , el control positivo con NFX (11  $\mu\text{M}$ ) y el negativo con DMSO (1%). La curva resultante del presente ensayo se muestra en la figura 4.9.

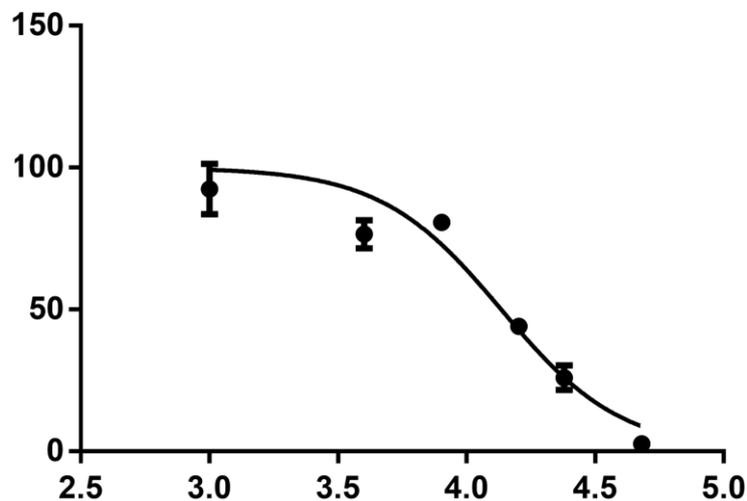


Figura 4.9. Curva de determinación de  $\text{EC}_{50}$  para Diclomina. El análisis estadístico se realizó mediante una prueba de ANOVA de una vía seguido de una prueba de comparación de Dunnett ( $p\text{-valor} < 0,05$ ). Los resultados se expresan como % de la media  $\pm$  DE.

El análisis de la varianza y posterior comparación de medias mediante la prueba de Dunnett demostró que existen diferencias significativas ( $p\text{-valor} < 0,05$ ) para el rango de concentración comprendido entre 8 y 48  $\mu\text{M}$ , mientras que no se reflejaron dichas diferencias

para las concentraciones de 1 y 4  $\mu\text{M}$  respecto al control negativo. A partir de este análisis se determinó, para Diclomina, un valor de  $\text{EC}_{50}$  de 13,6  $\mu\text{M}$  ( $\pm 3,1$ ).

### 4.2.3 Evaluación bioquímica frente a *Tripanotión sintetasa*

Como se mencionó en el capítulo anterior la evaluación bioquímica se llevó a cabo frente a la enzima **TryS** de tres especies de tripanosomátidos (*T. cruzi*, *T. brucei* y *L. Infantum*). Todos los candidatos fueron evaluados a concentración fija de 30  $\mu\text{M}$ . Este primer cribado tiene como objetivo verificar si alguno de los fármacos presenta efecto inhibitorio relevante frente a la enzima, y luego a partir del efecto observado diseñar un ensayo guiado por resultados.

Los resultados obtenidos se presentan, de manera resumida, en la **tabla 4.9** donde puede apreciarse el promedio de actividad residual expresado como porcentaje (%AR) y su respectivo coeficiente de variación porcentual (CV%) para cada una de las enzimas frente a los diferentes candidatos. Para cada uno de los ensayos se realizaron controles de actividad (DMSO-1%) e inhibición (**J18** para la Tc**TryS**, y **Mol2008** para Tb**TryS** y Li**TryS**).

**Tabla 4.9.** Actividad residual de la enzima **TryS** frente a diferentes fármacos

| Fármaco                 | TryS            |      |                  |      |                    |      |
|-------------------------|-----------------|------|------------------|------|--------------------|------|
|                         | <i>T. cruzi</i> |      | <i>T. brucei</i> |      | <i>L. infantum</i> |      |
|                         | % AR            | CV%  | % AR             | CV%  | % AR               | CV%  |
| Sulfadiazina            | 87,5            | 7,3  | 97,4             | 5,2  | 104,0              | 7,1  |
| Clorpropamida           | 75,9            | 7,8  | 76,3             | 11,7 | 86,3               | 3,4  |
| Sulfasalazina           | 78,8            | 2,1  | 81,4             | 5,3  | 105,7              | 4,0  |
| Tolbutamida             | 80,9            | 2,0  | 79,7             | 9,2  | 91,4               | 2,8  |
| Terbinafina             | 89,8            | 4,6  | 97,3             | 5,5  | 100,0              | 8,2  |
| Lidocaína               | 74,3            | 9,7  | 84,4             | 0,7  | 93,5               | 1,5  |
| Diciclomina             | 72,4            | 6,9  | 94,8             | 6,5  | 119,6              | 7,1  |
| Sulfadimetoxina         | 148,7           | 1,1  | 91,6             | 10,2 | 102,8              | 6,6  |
| Tazobactam              | 105,3           | 2,8  | 85,9             | 2,3  | 99,6               | 8,1  |
| Sulfamerazina           | 111,8           | 7,1  | 91,1             | 5,5  | 95,5               | 7,3  |
| DMSO                    | 100             | 2,3  | 100              | 2,5  | 100                | 4,2  |
| Inhibidor (J18/Mol2008) | 32,7            | 12,5 | 34,8             | 17,8 | 13,4               | 53,7 |

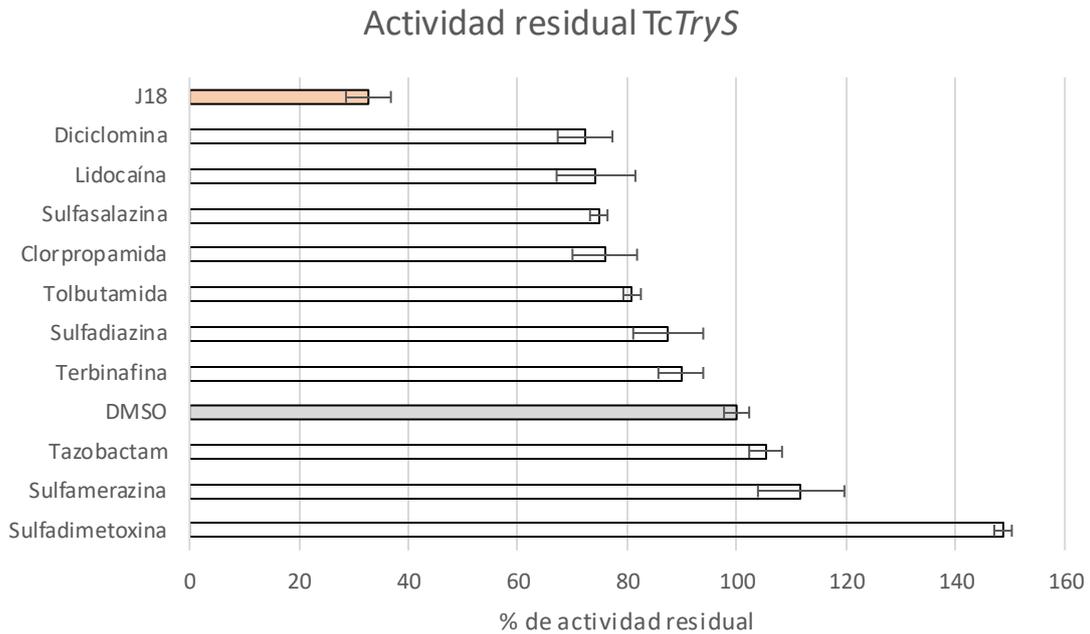
Adicionalmente se presentan las **figuras 4.10, 4.11 y 4.12** donde puede observarse de manera gráfica el %AR y la desviación estándar para cada uno de los candidatos frente a la enzima evaluada (Tc**TryS**, Tb**TryS** y Li**TryS** respectivamente). Para las tres enzimas se realizó la prueba del estadístico t, donde se comparó el fármaco que presentó mayor efecto inhibitorio con el control negativo (DMSO 1%).

Para cada una de las enzimas se plantearon las siguientes hipótesis de trabajo:

$$H_0: \mu_{\text{control}} = \mu_{\text{fármaco}}$$

$$H_1: \mu_{\text{control}} \neq \mu_{\text{fármaco}}$$

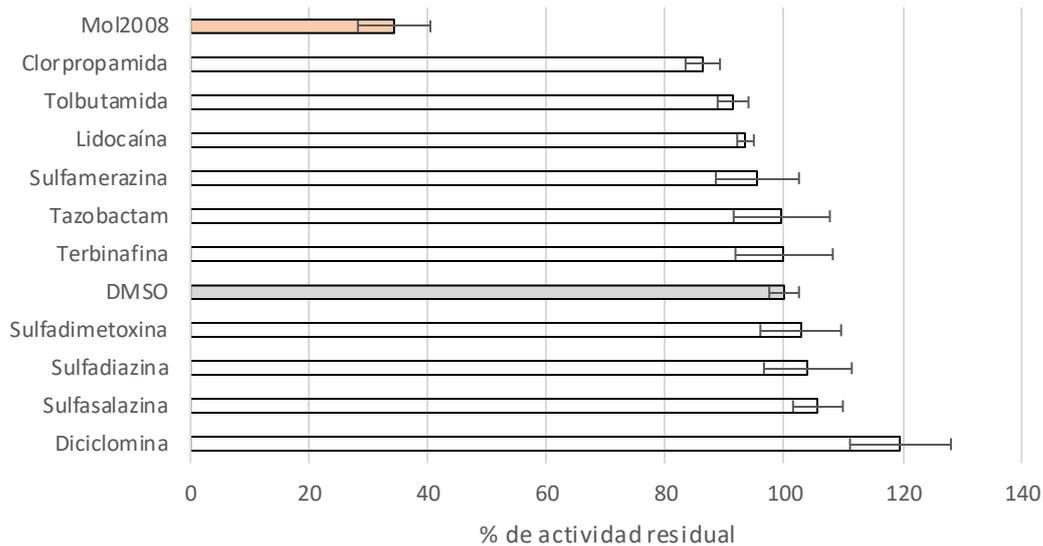
Donde  $\mu_{\text{control}}$  representa el %AR de la enzima frente al control con DMSO mientras que  $\mu_{\text{fármaco}}$  es el %AR de cada enzima frente al fármaco considerado.



**Figura 4.10.** Actividad residual de TcTryS frente a cada uno de los candidatos evaluados. Los resultados se expresan como % de la media  $\pm$  DE. Control negativo DMSO 1% y control positivo J18 30  $\mu$ M.

La prueba del estadístico t para TcTryS indicó que se debe rechazar la hipótesis nula ( $H_0$ ) al comparar Dicolomina (%AR=72,4%) contra DMSO (1%) (%AR= 100%), es decir que este fármaco presenta un efecto diferente al control negativo. Sin embargo, tampoco se obtuvo un % de inhibición próximo al 50% a 30  $\mu$ M para ninguno de los candidatos, estos resultados sugieren que ninguno de los fármacos presenta un efecto inhibitorio similar al que se obtiene con el compuesto J18, el cual a 30  $\mu$ M produce una inhibición próxima al 70%, mientras que Dicolomina (fármaco con mayor inhibición) solo logra una reducción de la actividad enzimática de apenas un 30% a igual concentración que el inhibidor de referencia. Estos resultados incitan a pensar que los candidatos no son capaces de perturbar la actividad enzimática de manera significativa, y en consecuencia se ha decidido la finalización de las evaluaciones bioquímicas de los fármacos frente a TcTryS.

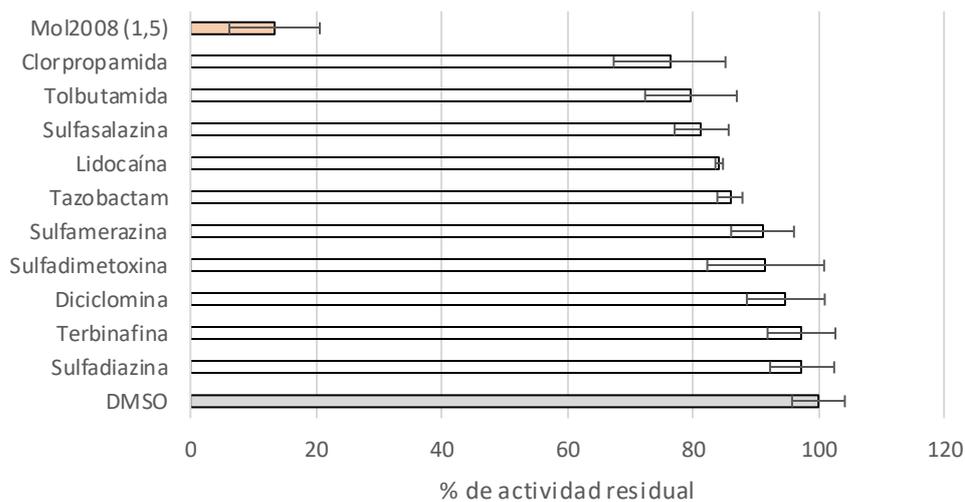
Actividad residual TbTryS



**Figura 4.11.** Actividad residual de TbTryS frente a cada uno candidatos evaluados. Los resultados se expresan como % de la media  $\pm$  DE. Control negativo DMSO 1% y control positivo Mol2008 a 30  $\mu$ M.

Los resultados del análisis estadístico de Clorpropamida frente a DMSO en evaluaciones de la TbTryS denotaron que se debe rechazar la  $H_0$ , es decir que Clorpropamida no se comporta como DMSO, sin embargo como se puede apreciar de la **tabla 4.7** y la **figura 4.11**, tampoco presentó un efecto similar al inhibidor Mol2008, ya que a 30  $\mu$ M este compuesto produce una inhibición mayor al 65 %, mientras que Clorpropamida solo logra reducir la actividad enzimática en un 25%; en consecuencia, puede plantearse que este fármaco exhibe poca afinidad frente a la enzima, lo que sugiere una baja probabilidad de modular la actividad enzimática a concentraciones relevantes.

Actividad residual LiTryS



**Figura 4.12.** Actividad residual de LiTryS frente a cada uno candidatos evaluados. Los resultados se expresan como % de la media  $\pm$  DE. Control negativo DMSO 1% y control positivo Mol2008 1,5  $\mu$ M.

Para la enzima Li**TryS** la prueba del estadístico t del fármaco Clorpropamida contra DMSO demostró que se debe rechazar  $H_0$ , es decir que Clorpropamida demuestra un leve efecto inhibitorio cuando se compara este fármaco frente a DMSO, sin embargo, el efecto no es comparable con la inhibición producida el compuesto Mol2008. Este último presenta a 1,5  $\mu\text{M}$  una inhibición mayor al 85%, por lo que dicho compuesto exhibe una potencia casi 20 veces mayor a Clorpropamida, es decir que ninguno de los candidatos evaluados frente a Li**TryS** es capaz de ofrecer resultados que justifiquen profundizar las evaluaciones bioquímicas frente a esta enzima.

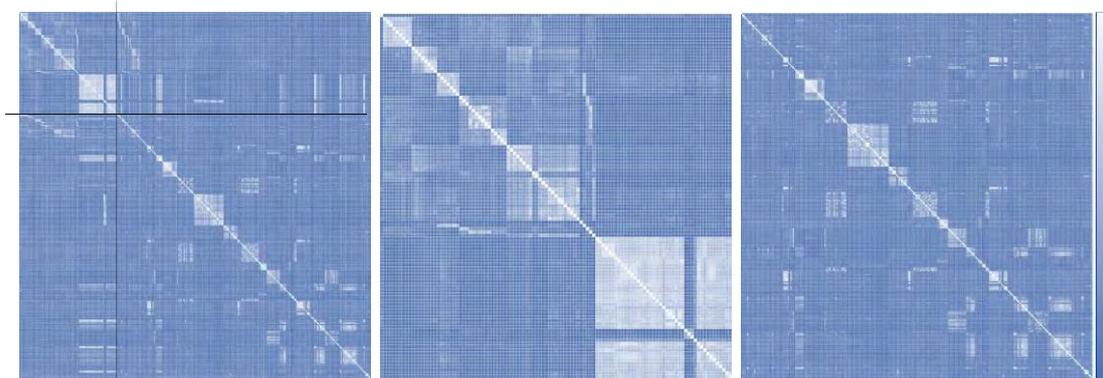
Como se desprende de los párrafos anteriores, ninguno de los fármacos evaluados frente a las diversas enzimas presentó, a 30  $\mu\text{M}$ , una inhibición similar a la que se obtuvo con el inhibidor de referencia. También puede observarse, a partir de la **tabla 4.7** y las **figuras 4.10 - 4.12**, que los candidatos no fueron capaces de reducir la actividad enzimática en un valor próximo al 50%. Con estos resultados, y dado que, como ya se discutió, diversos autores sugieren que el rango de concentraciones para evaluar un *hit* validado debe estar comprendido entre 1 y 10  $\mu\text{M}$  (Hughes et al., 2011; Katsuno et al., 2015) es que se ha decidido finalizar la evaluación bioquímica de estos fármacos frente a la enzima **TryS** de las tres especies de tripanosomátidos, ya que el *screening* primario se realizó a una concentración que triplica el valor recomendado.

## 4.3 Resultados de la segunda campaña: Estrategia computacional

### 4.3.1 Compilado de un nuevo set de datos

El set de datos compilado para la primera campaña de CV fue ampliado a mediados del año 2017. Dicha ampliación se llevó a cabo gracias a una biblioteca de compuestos químicos facilitada por el Laboratorio de Biología Redox de Tripanosomátidos del Instituto Pasteur de Montevideo. El nuevo set quedó compuesto por un total de 401 moléculas; las cuales se repartieron en 111 **ACTIVOS** (Inhibidores) y 290 **INACTIVOS** (no inhibidores), utilizando para la partición los mismos criterios ( $IC_{50}$  o % de inhibición) que en la campaña anterior, pero modificando el valor de corte de 20 a 30  $\mu\text{M}$  para ser consistente con el cribado experimental realizado en el Instituto Pasteur. Para esta campaña se optó por categorizar los compuestos en **ACTIVOS** / **INACTIVOS** según el valor reportado frente a TbTryS, debido a que el set de datos de la campaña anterior está basado en dicha enzima. Sin embargo, cabe destacar que en esta oportunidad más del 80% de los compuestos que se evaluaron frente a TcTryS y TbTryS comparten la categoría según los valores de actividad reportados por el Dr. Comini, ya que el 2,2 % de los compuestos totales resultaron **ACTIVOS** frente TcTryS y TbTryS, mientras que un 78,9% de los mismos se consideraron **INACTIVOS** para ambas enzimas.

A continuación, se presenta un mapa de calor (ver **figura 4.13**) realizado con el programa Gitoools v. 2.3.1 (Perez-Llamas & Lopez-Bigas, 2011), en el mismo se aprecia la similitud/disimilitud entre compuestos que conforman el set de datos.



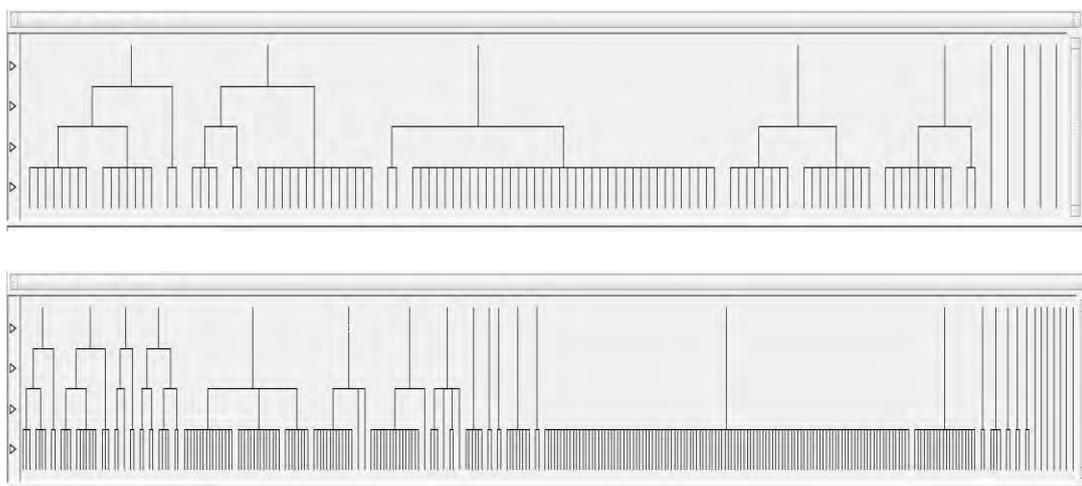
**Figura 4.13.** Mapas de calor de los sets de datos. Set de datos completo (izquierda), Conjunto de compuestos **ACTIVOS** (centro), Conjunto de compuestos **INACTIVOS** (derecha).

Como se puede ver en la **figura 4.13** el set de datos está conformado por una amplia diversidad estructural (mayor cantidad de celdas de color azul intenso). Esta observación nos permite inferir que los modelos clasificatorios construidos a partir de este set de datos contemplarán un espacio químico amplio y diverso.

### 4.3.2 Partición set de datos

Para la partición del set de datos se empleó la misma metodología de agrupamiento secuencial utilizada en la campaña previa. En esta oportunidad, y dado el aumento significativo

en la cantidad de compuestos **INACTIVOS** (de 45 a 290), las técnicas de agrupamiento fueron aplicadas a las dos categorías de compuestos. El agrupamiento jerárquico para la categoría de compuestos **ACTIVOS** arrojó cinco *clusters* con 40, 21, 17 (x2) y 11 moléculas en cada subconjunto y 5 moléculas atípicas (*outliers*), mientras que para los compuestos **INACTIVOS** se obtuvieron trece agrupamientos, donde el *cluster* más poblado contenía 116 compuestos y los doce subgrupos restantes quedaron conformados por un número de entre 4 y 38 moléculas cada uno; 21 moléculas fueron consideradas *outliers*. En la **figura 4.14** se presentan los dendogramas obtenido para cada una de las categorías.



**Figura 4.14.** Dendogramas obtenidos para a) Conjunto de compuestos **ACTIVOS**. b) Conjunto de compuestos **INACTIVOS**.

La técnica de *cluster K-means* (*clustering* no jerárquico) reordenó los subgrupos de ambas categorías de la siguiente manera: para la categoría **ACTIVOS** el método reasignó los 111 compuestos en cinco subgrupos que quedaron conformados por 37, 18, 34, 1 y 21 compuestos en cada subgrupo, mientras que para el conjunto de **INACTIVOS** se generaron trece *clusters* resignando las moléculas en subgrupos de 32, 11, 24, 10, 17, 21, 39, 13, 22, 35, 33, 11 y 22 en cada uno de los agrupamientos.

Para conformar el conjunto de entrenamiento se tomó de cada *cluster* un 65% de los compuestos **ACTIVOS** y el 25% de los **INACTIVOS**. Este conjunto quedó conformado con un total de 140 compuestos. El resto de las moléculas se repartió de manera representativa y equitativa en dos conjuntos de prueba (DUD-e A y DUD-e B), adicionalmente, y al igual que lo realizado en la campaña previa, se generó para cada conjunto de prueba un set de señuelos mediante la aplicación *on line* DUD-e (Mysinger et al., 2012b). En la **figura 4.15** se puede observar la cantidad y categoría de los compuestos en cada uno de los conjuntos, como así también la distribución porcentual de cada categoría en los diferentes conjuntos.

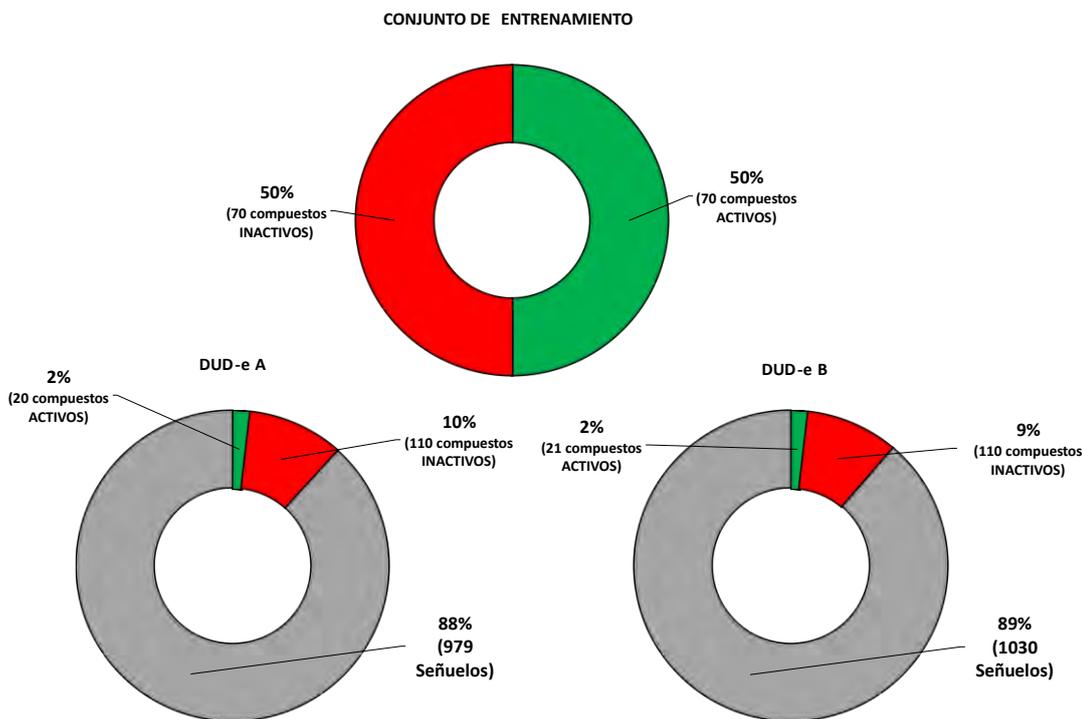


Figura 4.15. Composición de los tres conjuntos generados.

### 4.3.3 Generación y validación de modelos

#### 4.3.3.1 Construcción de los modelos

A partir de los 3.668 descriptores moleculares de baja dimensionalidad (0-2D) calculados para cada molécula del conjunto de entrenamiento, se generaron 1.000 subconjuntos de 200 descriptores cada uno, y con ellos se construyó un modelo clasificador para cada subconjunto. La construcción de los modelos se llevó a cabo en entorno *R studio* aplicando la metodología de pasos hacia adelante (*forward stepwise*) y limitando a catorce la cantidad máxima de pasos. Las ecuaciones de los mejores cinco modelos individuales, según el AUCROC obtenida frente al conjunto DUD-e A, se detallan a continuación.

#### Modelo 319

$$\begin{aligned}
 \text{Score} = & -1,6391 + 0,68769 * B10[Cl - Br] + 0,4902 * B01[N - N] + 2,92669 \\
 & * X2Av + 0,26707 * B01[C - F] + 0,61351 * GATS3s - 0,36813 \\
 & * CAST2d_09_DD + 0,18382 * N - 067 - 0,02424 * CATS2D_07_LL \\
 & + 0,02869 * nCbH + 0,16267 * SpMax_AEA(dm)
 \end{aligned}$$

### Modelo 31

$$\begin{aligned} \text{Score} = & -0,306446 + 0,488922 * F10[Cl - Br] + 4,106990 * Eta\_sh\_x + 0,270569 \\ & * B03[N - N] - 0,002262 * P\_VSA\_v\_2 + 0,761790 * GATS3s \\ & + 0,364899 * B03[C - F] - 0,2878166 * C - 029 - 0,063467 \\ & * CATS2D\_AA \end{aligned}$$

### Modelo 633

$$\begin{aligned} \text{Score} = & 0,52321 + 0,33715 * N - 073 + 0,67804 * B01[O - S] + 0,67847 * B10[Cl \\ & - Br] + 0,27487 * B10[C - F] - 0,09849 * GGI3 - 0,08707 \\ & * Chi\_D/Dt - 0,04398 * H - 050 \end{aligned}$$

### Modelo 845

$$\begin{aligned} \text{Score} = & -0,1700 + 0,7666 * F10[Cl - Br] + 0,3756 * nPyrazoles + 0,1766 \\ & * MAXDN - 1,0189 * SM1Dz(e) + 0,1590 * B08[C - S] + 0,2075 \\ & * F05[N - F] - 0,1569 * CATS2D\_00\_PP - 1,1833 * MATS3e + 0,2126 \\ & * HVcpx \end{aligned}$$

### Modelo 74

$$\begin{aligned} \text{Score} = & 0,141035 + 0,279086 * F10[Cl - Br] - 8,046257 * PW4 - 0,467357 \\ & * CATS2D\_09\_DD + 0,057833 * SpDiam\_EA(ed) - 0,001586 * SAacc \\ & + 0,825908 * GATS3s - 0,515516 * C - 029 + 0,191897 * B04[N - N] \\ & + 0,006213 * SsF - 0,054241 * F04[N - O] \end{aligned}$$

Los descriptores contemplados en el mejor modelo individual se encuentran descritos en el **Anexo I**.

Si bien se permitió que los modelos incorporaran hasta catorce descriptores, ninguna de las ecuaciones aquí plasmadas ha incorporado más de diez descriptores, lo que permite inferir que la relación real de descriptor/moléculas estará siempre por debajo del límite permitido y en consecuencia se verá reducida la probabilidad de sobreajuste.

### **4.3.3.2 Validación de los modelos**

En la **tabla 4.10** se presentan los resultados obtenidos para las dos técnicas de validación interna (aleatorización de Fisher y validación cruzada) frente a los mejores cinco modelos individuales. En ambas pruebas se utilizó “0,5” como valor de corte para diferenciar entre compuestos **ACTIVOS** e **INACTIVOS**, tal como surge del promedio de los valores asignados a cada una de las categorías que componen el conjunto de entrenamiento. (**INACTIVOS** = 0; **ACTIVOS** = 1) (Toropova & Toropov, 2017).

Para la presente campaña en la prueba de validación cruzada la remoción de subconjuntos se realizó de manera aleatoria y balanceada retirando un total de 14 moléculas (7 de cada categoría) del conjunto de entrenamiento original, este procedimiento se realizó

200 veces. Mientras que la prueba de aleatorización de Fisher se realizó aleatorizando las clases tal como se describió en la campaña previa.

**Tabla 4.10.** Resultados de las métricas de validación interna

| MODELO | Original | Validación Cruzada |      | Aleatorización de Fisher |                    |                    |
|--------|----------|--------------------|------|--------------------------|--------------------|--------------------|
|        | %BC      | %BC                | DE   | %BC                      | Intervalo Inferior | Intervalo Superior |
| 319    | 92,9     | 86,7               | 8,5  | 60,0                     | 57,1               | 62,9               |
| 31     | 92,9     | 87,3               | 8,7  | 60,0                     | 56,3               | 63,8               |
| 633    | 89,3     | 86,4               | 9,0  | 60,1                     | 56,2               | 63,9               |
| 845    | 89,3     | 82,9               | 10,2 | 59,5                     | 55,1               | 63,9               |
| 74     | 90,7     | 80,2               | 10,3 | 61,4                     | 57,6               | 65,3               |

A partir de los datos volcados en la **tabla 4.10** puede apreciarse que, para todos los modelos informados, la prueba de aleatorización muestra una notable caída en los %BC respecto a los valores obtenidos para los modelos originales, este comportamiento afianza los resultados de los modelos originales, dado que puede asumirse minimizada o prácticamente descartada la probabilidad de generar buenos modelos clasificadores a partir del azar o correlaciones fortuitas.

En cuanto a la técnica de validación cruzada, en la que se evalúa el sobreajuste de los modelos, se puede apreciar que los %BC son similares a los obtenidos para los modelos originales, y en consecuencia se infiere que los modelos no presentan inconvenientes en este aspecto.

Para estimar la capacidad predictiva de todos los modelos recurrimos a evaluar el %BC y el AUCROC en el conjunto de prueba DUD-e A, en la **tabla 4.11** se informan los resultados para los mejores cinco modelos. Además, se anexa en la **figura 4.16** las curvas ROC obtenidas para cada modelo frente al mismo conjunto.

**Tabla 4.11.** Evaluación de la capacidad predictiva de los modelos

| MODELO | Conjunto de Entrenamiento |      | DUD-e A       |
|--------|---------------------------|------|---------------|
|        | %BC                       | %BC  | AUCROC        |
| 319    | 92,9                      | 83,1 | 0,941 (0,017) |
| 31     | 92,9                      | 76,6 | 0,930 (0,024) |
| 633    | 89,3                      | 80,8 | 0,921 (0,020) |
| 845    | 89,3                      | 82,4 | 0,921 (0,020) |
| 74     | 90,7                      | 76,1 | 0,921 (0,025) |

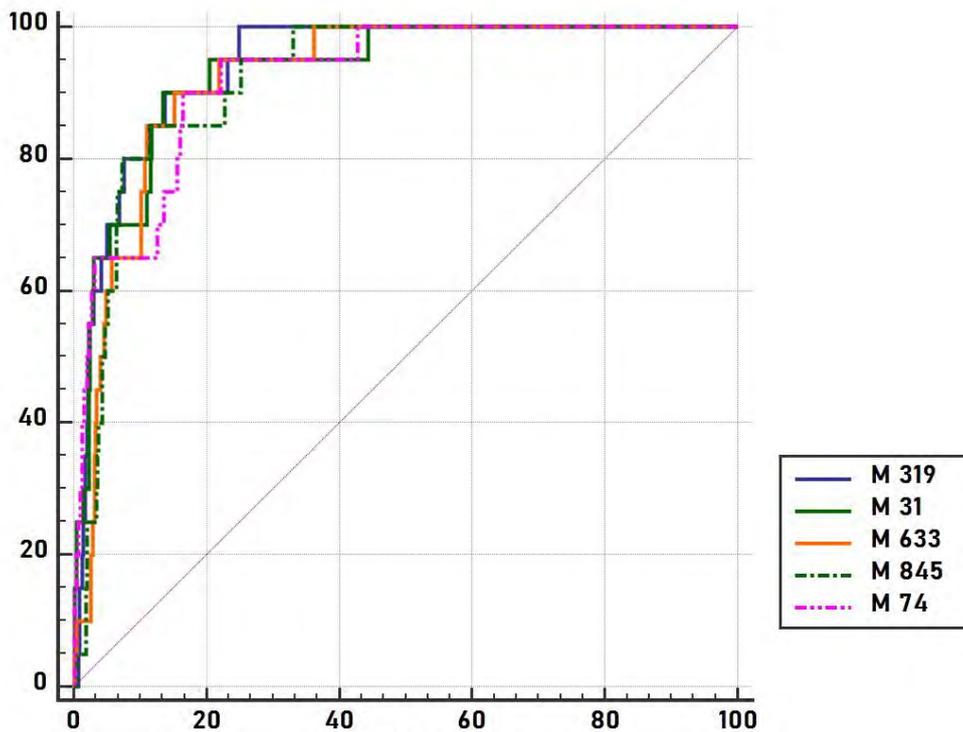


Figura 4.16. Curvas ROC para los mejores cinco modelos individuales frente al conjunto DUD-e A.

Los resultados obtenidos para la validación externa permiten estimar que los modelos presentan una buena capacidad predictiva, ya que el %BC para cada uno de ellos frente al conjunto DUD-e A es mayor al 75%, es decir que cada modelo es capaz de reconocer en la categoría correcta a más del 75% de las moléculas que componen dicho conjunto. Adicionalmente el valor del AUCROC frente a este mismo conjunto se mantiene siempre por encima de 0,92, un valor muy próximo al límite superior (1) para esta métrica.

#### 4.3.4 Evaluación del desempeño de los modelos

Una vez aplicadas las diferentes técnicas de validación se realizó la evaluación de otros parámetros que permiten estimar (y reforzar) el desempeño de cada modelo. En este paso se evaluó la *Se*, la *Sp*, el %BC, la Precisión y el valor *F-measure*. En las tablas 4.12 y 4.13 se presentan los valores obtenidos para cada uno de estos parámetros frente al conjunto de entrenamiento y al conjunto DUD-e A respectivamente.

Tabla 4.12. Parámetros evaluados al conjunto de entrenamiento

| Modelo | Sensibilidad | Especificidad | %BC  | Precisión | F-measure |
|--------|--------------|---------------|------|-----------|-----------|
| 319    | 0,971        | 0,886         | 92,9 | 0,895     | 0,932     |
| 31     | 0,971        | 0,895         | 92,9 | 0,895     | 0,932     |
| 633    | 0,914        | 0,877         | 89,3 | 0,877     | 0,895     |
| 845    | 0,943        | 0,857         | 89,3 | 0,857     | 0,898     |
| 74     | 0,971        | 0,861         | 90,7 | 0,861     | 0,913     |

**Tabla 4.13.** Parámetros evaluados frente al conjunto DUD-e A

| Modelo | Sensibilidad | Especificidad | %BC  | Precisión | F-measure |
|--------|--------------|---------------|------|-----------|-----------|
| 319    | 0,900        | 0,830         | 83,1 | 0,089     | 0,161     |
| 31     | 0,950        | 0,763         | 76,7 | 0,069     | 0,128     |
| 633    | 0,900        | 0,806         | 80,8 | 0,079     | 0,145     |
| 845    | 0,850        | 0,824         | 82,4 | 0,081     | 0,149     |
| 74     | 0,950        | 0,758         | 76,1 | 0,067     | 0,125     |

Con los valores que se obtuvieron para los diferentes parámetros puede asumirse que los modelos presentan un desempeño aceptable, ya que en general exhiben una buena tasa reconocimientos para cada una de las categorías como así también para cada conjunto en su totalidad. También es posible observar que en todos los modelos reportados aquí frente al conjunto DUD-e A el %BC es mayor al 75%, incluso algunos de ellos (3/5) presentan valores por encima del 80%. En cuando a los resultados de precisión y *F-measure* los bajos valores pueden explicarse debido al desbalance de clases del conjunto DUD-e A tal como se explicó para la primera campaña de CV (ver **sección 4.1.4**).

### 4.3.5 Combinación de modelos

Como se mencionó con anterioridad la combinación de modelos genera un metaclasificador con mayor robustez en comparación con los modelos individuales.

Para seleccionar qué combinación de modelos resultaría más adecuada para una etapa de cribado prospectivo, se construyó un gráfico de AUCROC vs. cantidad de modelos combinados para cada operador. La gráfica contiene combinaciones desde 2 a 100 modelos. En la **figura 4.17** se muestran los gráficos obtenidos para cada operador.

La evaluación de la métrica debe incluir, además del valor de AUCROC en sí, un análisis de la desviación estándar (estimada mediante *bootstrapping*) y, eventualmente, una comparación estadística de los valores obtenidos a través de distintos esquemas de combinación de modelos. Como se observa de los gráficos solo el operador MÍNIMO mantuvo todos los valores de AUCROC por encima de 0,95 y con una desviación estándar aceptable, mientras que el resto de los operadores fueron desestimados para evaluar un posible ensamble, debido a que muestran una tendencia hacia la baja en los valores de AUCROC a medida que se incorporan modelos al ensamble y presentan desvíos estándar que en algunos casos duplican o triplican las desviaciones encontradas para el operador seleccionado.

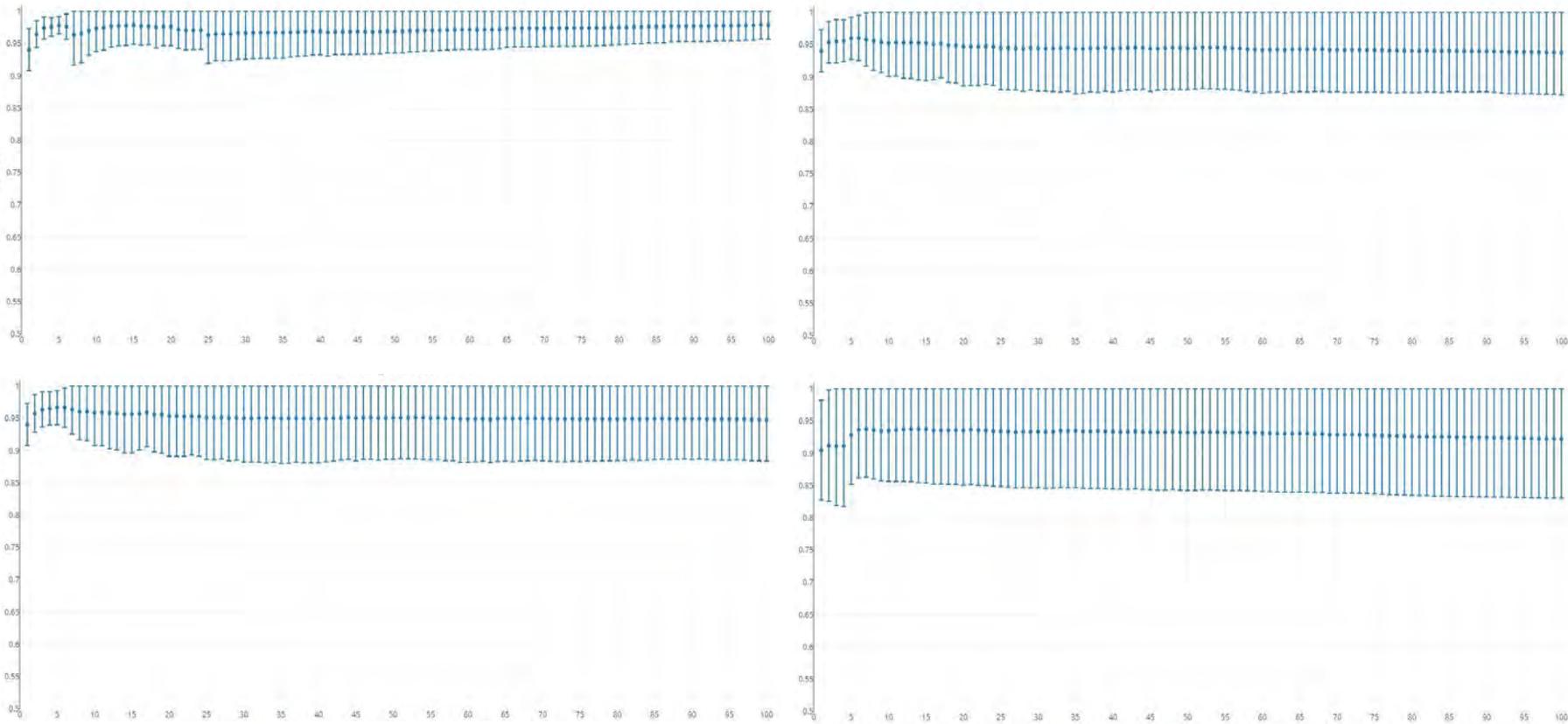


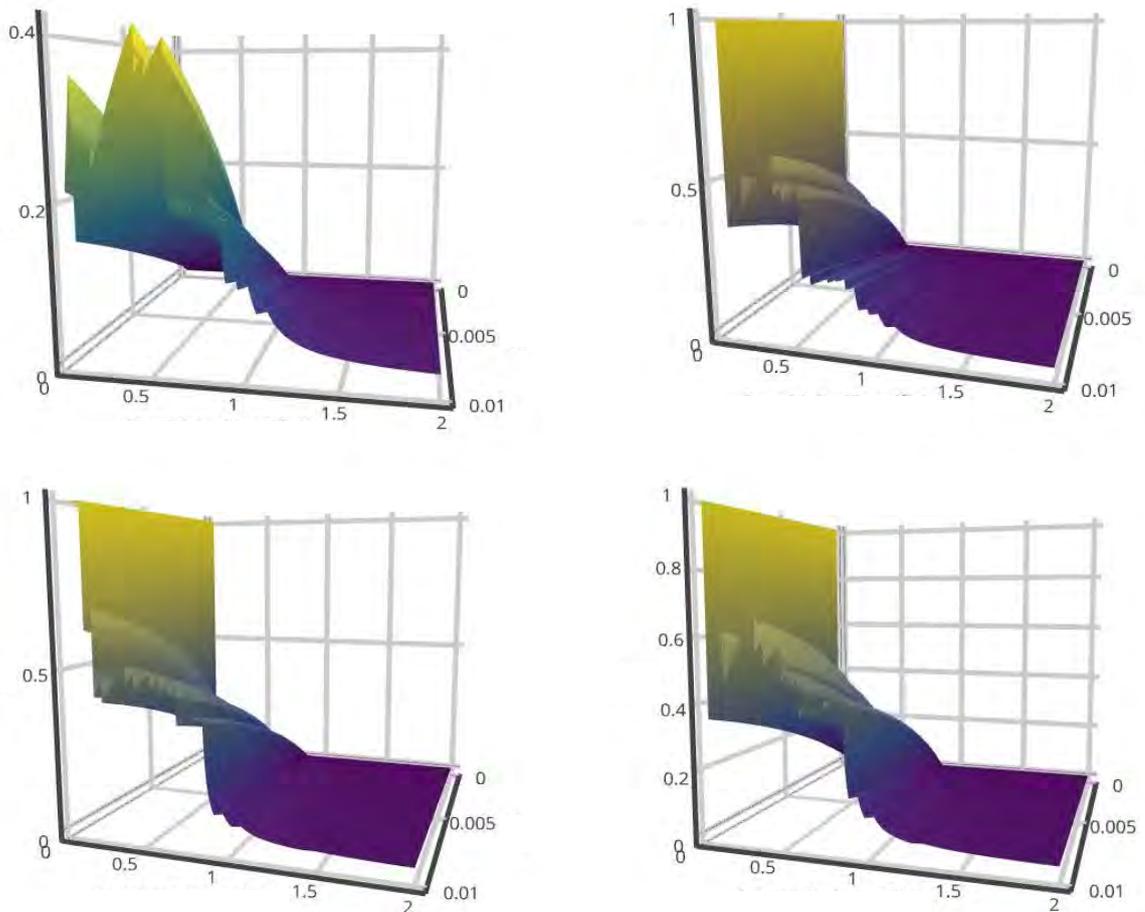
Figura 4.17. Gráficos de AUCROC vs. Cantidad de modelos combinados: a) Operador MÍNIMO, b) Operador PROMEDIO, c) Operador promedio del RANKING d) Operador promedio del VOTO.

Luego del análisis de la gráfica, y evaluando el valor de AUCROC con sus respectivas desviaciones estándar se seleccionaron cuatro posibles combinaciones, los valores del AUCROC para cada posible ensamble se informan en la **tabla 4.14**.

**Tabla 4.14.** Valores de AUCROC para las diferentes combinaciones de modelos mediante el operador mínimo evaluadas sobre DUD-e A. El asterisco (\*) denota diferencias estadísticamente significativas respecto al mejor modelo individual.

| Ensamble          | AUCROC         |
|-------------------|----------------|
| <b>Modelo 319</b> | 0,941 (0,017)  |
| <b>Mínimo 4</b>   | 0,975 (0,014)* |
| <b>Mínimo 5</b>   | 0,978 (0,014)* |
| <b>Mínimo 50</b>  | 0,969 (0,031)* |
| <b>Mínimo 90</b>  | 0,977 (0,024)* |

Los ensambles mostraron diferencias significativas de AUCROC respecto al mejor modelo individual (M-319) con un p-valor  $< 0,05$  según el método aplicado por el programa MedCalc (MedCalc Software 2011); asimismo, no se encontraron diferencias significativas entre los valores de AUCROC obtenidos para los cuatros ensambles seleccionados. Debido a esto y con el objetivo de refinar la selección del mejor ensamble, se recurrió a la evaluación del comportamiento de las diferentes combinaciones a partir de la construcción y análisis de superficies PPV para cada ensamble. En la **figura 4.18** se presentan las superficies 3D obtenidas para cada una de las combinaciones seleccionadas.



**Figura 4.18.** Superficies PPV para las diferentes combinaciones mediante el operador mínimo de: a) 4 modelos, b) 5 modelos, c) 50 modelos y d) 90 modelos frente al conjunto DUD-eA.

Además de construir las superficies PPV, y a fin de realizar un análisis más riguroso se procedió a la evaluación de la métrica BEDROC y del factor de enriquecimiento (EF). Estas métricas ponen énfasis en los compuestos **ACTIVOS** de las primeras posiciones del ranking ordenado según el ensamble de modelos. Para obtener los valores de las diferentes métricas se recurrió a la herramienta *online* Rocker (<http://www.medchem.fi/rocker/>). En la **tabla 4.15** se informan los resultados obtenidos para cada uno de los ensambles frente al conjunto DUD-eA.

**Tabla 4.15.** Resultados de las métricas calculadas para los ensambles seleccionados  
**Ensamblés (operador mínimo)**

|                        | E-4   | E-5   | E-50  | E-90  |
|------------------------|-------|-------|-------|-------|
| EF1%base               | 20    | 40    | 35    | 40    |
| EF1%señuelos           | 60    | 55    | 70    | 60    |
| BEDROC ( $\alpha=20$ ) | 0,726 | 0,757 | 0,805 | 0,824 |

Del análisis de la superficie PPV se observa que el ensamble de 90 modelos (**E-90**) conduce a una superficie con valores de PPV por encima de las otras tres combinaciones. Esta observación es consistente con los valores informados en la **tabla 4.15**, donde se puede ver

que el meta-clasificador de 90 modelos revela mejores resultados de las métricas evaluadas. En consecuencia, dicho ensamble fue seleccionado para la etapa de cribado prospectivo.

En las **tablas 4.16** y **4.17** se informan los resultados obtenidos para los diferentes ensambles según los parámetros evaluados frente al conjunto de entrenamiento y conjunto DUD-e A, donde también se consigan los mismos parámetros para el mejor modelo individual.

**Tabla 4.16.** Parámetros de desempeño de los ensambles generados mediante el operador mínimo evaluados frente al conjunto de entrenamiento

| Modelo       | Sensibilidad | Especificidad | %BC  | Precisión | <i>F-measure</i> |
|--------------|--------------|---------------|------|-----------|------------------|
| <b>M-319</b> | 0,971        | 0,886         | 92,9 | 0,895     | 0,932            |
| <b>E-4</b>   | 0,871        | 0,897         | 88,6 | 0,897     | 0,884            |
| <b>E-5</b>   | 0,857        | 0,909         | 88,6 | 0,909     | 0,882            |
| <b>E-50</b>  | 0,557        | 0,975         | 77,1 | 0,975     | 0,709            |
| <b>E-90</b>  | 0,471        | 0,971         | 72,9 | 0,971     | 0,635            |

**Tabla 4.17.** Parámetros de desempeño de los ensambles generados mediante el operador mínimo evaluados frente al conjunto de DUD-e A

| Modelo       | Sensibilidad | Especificidad | %BC  | Precisión | <i>F-measure</i> |
|--------------|--------------|---------------|------|-----------|------------------|
| <b>M-319</b> | 0,900        | 0,830         | 83,1 | 0,089     | 0,161            |
| <b>E-4</b>   | 0,850        | 0,952         | 95,0 | 0,246     | 0,382            |
| <b>E-5</b>   | 0,850        | 0,957         | 95,5 | 0,266     | 0,405            |
| <b>E-50</b>  | 0,600        | 0,993         | 98,6 | 0,600     | 0,600            |
| <b>E-90</b>  | 0,500        | 0,994         | 98,6 | 0,625     | 0,556            |

### 4.3.6 Selección del valor de corte para cribado virtual

Con el objeto de seleccionar y optimizar un valor de corte adecuado, desafiamos el meta-clasificador seleccionado (combinación de 90 modelos por el operador mínimo) frente a un segundo conjunto de prueba (DUD-e B). Para este conjunto se realizó una nueva superficie PPV, la cual se muestra en la **figura 4.19**. En la **tabla 4.18** se presentan los resultados obtenidos frente al conjunto DUD-e B para las métricas AUCROC, EF y BEDROC, mientras que en la **tabla 4.19** se pueden apreciar los parámetros de **Se**, **Sp**, %BC, precisión y *F-measure* logrados por el meta clasificador frente al modelo 319 y al conjunto DUD-e B.

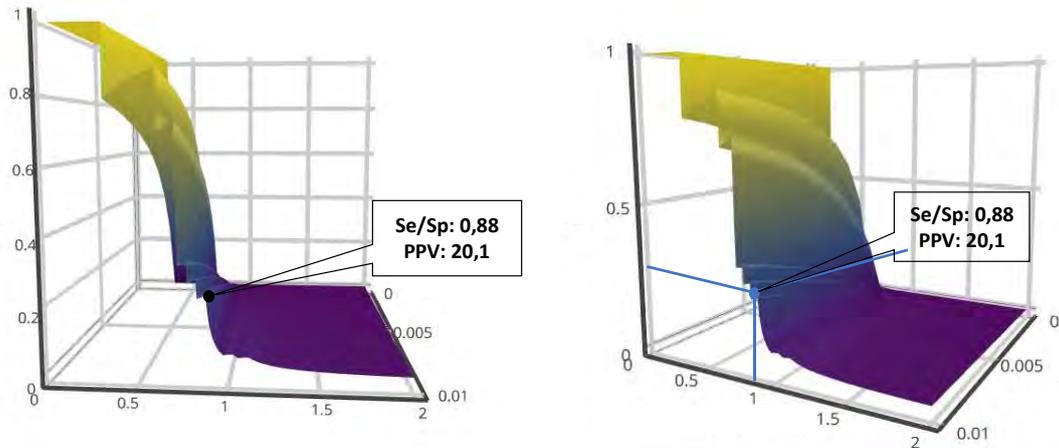


Figura 4.19. Superficie PPV frente a DUD-e B vista de frente y levemente rotada.

Tabla 4.18. Métricas evaluadas frente a DUD-e B para el mejor modelo individual y el ensamble de 90 modelos generado por el operador mínimo

|  | M-319         | E-90          |
|--|---------------|---------------|
| <b>AUCROC</b>                          | 0,951 (0,033) | 0,939 (0,036) |
| <b>EF<sub>1%base</sub></b>             | 18,4          | 49,7          |
| <b>EF<sub>1%señuelos</sub></b>         | 19,0          | 71,4          |
| <b>BEDROC (<math>\alpha=20</math>)</b> | 0,540         | 0,833         |

Tabla 4.19. Parámetros de desempeño de los ensambles generado mediante el operador mínimo evaluados frente al conjunto de DUD-e B

| Modelo       | Sensibilidad | Especificidad | %BC  | Precisión | F-measure |
|--------------|--------------|---------------|------|-----------|-----------|
| <b>M-319</b> | 0,952        | 0,846         | 84,8 | 0,103     | 0,185     |
| <b>E-90</b>  | 0,333        | 0,999         | 98,7 | 0,875     | 0,483     |

La **tabla 4.18** permite ver claramente que las métricas evaluadas mejoraron significativamente cuando se realizó la combinación de modelos, en comparación con el mejor modelo individual (M-319). Como puede observarse a partir del valor de del  $EF_{1\%base}$  el mejor modelo individual presenta una probabilidad 18 veces mayor de encontrar un **ACTIVO** en el 1% mejor rankeado de la base, cuando la misma es ordenada según el valor de score, en comparación con un ordenamiento aleatorio. Para el meta-clasificador seleccionado (90 modelos), la probabilidad de encontrar un *hit* en el 1% mejor rankeado asciende a 50 veces mayor respecto al ordenamiento aleatorio. Un comportamiento similar se observa cuando se analiza el cribado de la lista ordenada en el que se toma como punto de corte el tamizado del primer 1% de los señuelos, donde el mejor modelo presenta una probabilidad 19 veces mayor respecto a un ordenamiento al azar, y el ensamble de 90 modelos mejora dicha probabilidad, ya que este valor es casi 70 veces mayor que en un ordenamiento aleatorio. Estos resultados son reafirmados con los valores obtenidos para la métrica BEDROC. Adicionalmente, la **tabla 4.19** permite observar que los parámetros evaluados reflejan una fuerte mejoría de sus resultados cuando se compara el modelo 319 frente al E-90 modelos en relación con la clasificación del conjunto DUD-e B. Es notorio cómo aumentan la especificidad, el %BC, la

precisión y el valor de *F-measure*, lo que indica que esta combinación aumenta la probabilidad de éxitos a la hora de evaluar un *hit* seleccionado en ensayos húmedos.

A partir del análisis de la superficie PPV frente al conjunto DUD-e B se seleccionó un valor de corte de 0,131, obteniendo un valor de 97,3 % para la **Sp** y de 85,7% de **Se**.

### 4.3.7 Cribado virtual y selección de compuestos

El cribado prospectivo se realizó frente a las bibliotecas DrugBank 5.0 (Wishart et al., 2018) y SweetLead (Novick et al., 2013). Los compuestos de ambas bibliotecas fueron previamente estandarizados mediante el programa Standardizer (V. 16.9.12.0, InstantJChem), con las mismas acciones que se describen en la sección metodología para la primera campaña de CV.

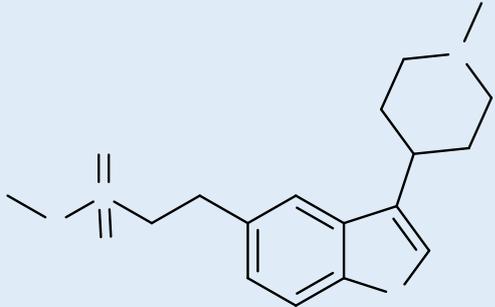
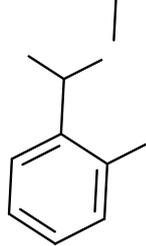
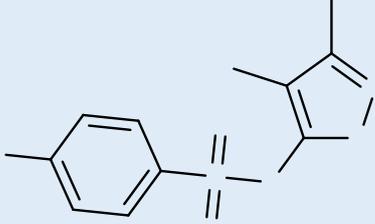
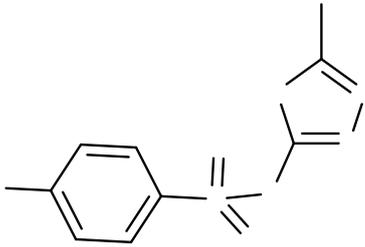
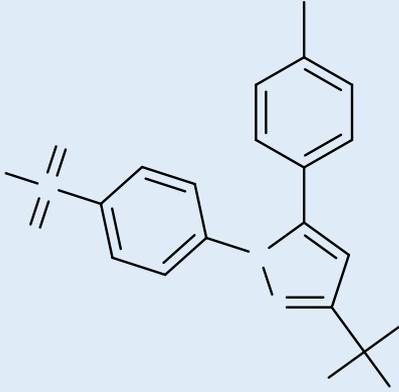
Luego de realizar el cribado de ambas bases de datos aplicando el meta-clasificador de 90 modelos se seleccionaron 82 compuestos, 18 provenientes de la biblioteca SweetLead y 64 de DrugBank, que superaron el umbral de corte propuesto (0,131).

Dado que varios compuestos se encuentran presentes en ambas bibliotecas, el primer paso fue eliminar de una de ellas los compuestos duplicados. Luego se restringió la búsqueda solo a moléculas que se encuentran aprobadas para uso clínico o en fases avanzadas de evaluaciones clínicas, obteniendo así un conjunto de 30 candidatos. Cuando se evaluó el dominio de aplicación (DA) del ensamble se eliminó un compuesto más, debido a que el mismo no estaba dentro del DA. Finalmente se descartaron los fármacos Tolbutamida y Clorpropamida debido a que los mismos habían sido evaluados en la campaña previa obteniéndose resultados negativos. El conjunto de candidatos quedó entonces conformado por un total de 27 compuestos.

Aplicando criterios de disponibilidad y accesibilidad comercial tal como se realizó en la campaña previa, y no habiéndose reportados ensayos bioquímicos frente a **TryS** de ninguno de los 27 candidatos, se decidió adquirir cuatro fármacos: Naratriptán, Metil antranilato, Sulfisoxazol, y Sulfametizol. Adicionalmente y dado que contábamos con el compuesto Celecoxib en nuestro laboratorio se decidió evaluar el mismo frente a **TryS** y a *T. brucei*, pero no frente a *T. cruzi*, ya que el compuesto se encontraba reportado como inactivo para dicha especie (Hideko Tatakihara et al., 2008; Malvezi et al., 2014; Planer et al., 2014).

En la **tabla 4.20** se presenta la estructura y la indicación clínica original para cada uno de los fármacos seleccionado.

Tabla 4.20. Candidatos seleccionados, indicación original y estructura química

| Fármaco           | Indicación original                 | Estructura   |
|-------------------|-------------------------------------|--|
| Naratriptán       | Antimigrañoso                       |    |
| Metil Antranilato | Aromatizante/saborizante comestible |    |
| Sulfisoxazol      | Antibiótico                         |   |
| Sulfametizol      | Antibiótico                         |  |
| Celecoxib         | Inhibidor selectivo de COX-2        |  |

## 4.4 Segunda campaña de evaluación experimental

Los ensayos experimentales se realizaron en colaboración con el Instituto de Ciencia y Tecnología Cesar Milstein (CONICET- Fundación Pablo Cassará); el Instituto de Investigaciones en Microbiología y Parasitología Médica -IMPAM- (UBA -CONICET) y el Instituto Pasteur de Montevideo -IPM- (Uruguay).

Tal como se realizó en la primera campaña, la validación biológica se llevó a cabo frente a 2 especies de tripanosomátidos y a la enzima **TryS** de *T. cruzi*, *T. brucei* y *L. infantum*.

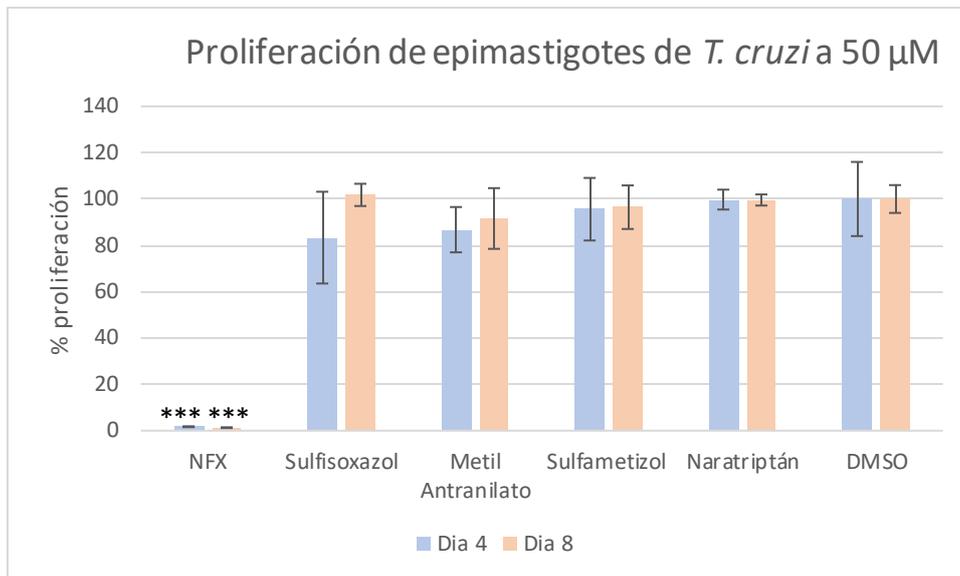
### 4.4.1. Evaluación biológica frente a *T. cruzi*

#### 4.4.1.1 Evaluación frente a epimastigotes

El primer ensayo biológico de los fármacos seleccionados se realizó frente a la cepa “Y” de *T. cruzi* en el estadio epimastigote, en esta oportunidad se procedió con un screening a concentración fija de 50  $\mu$ M y por triplicado con el fin de comprobar si alguno de los candidatos presentaba actividad inhibitoria frente a la proliferación del parásito. Cabe mencionar que epimastigote es uno de los estadios replicativos del parásito, por lo que el recuento se realizó el día cuatro y ocho posterior a la siembra. En la **tabla 4.21** se muestra el promedio de los porcentajes de proliferación obtenidos para cada día, los mismos resultados pueden observarse de manera grafica en la **figura 4.20**. Los resultados fueron normalizados respecto a DMSO (0,01%), considerando para este una proliferación de 100%. Como control positivo se utilizó NFX a 25  $\mu$ M.

**Tabla 4.21** Resultados de proliferación normalizado respecto a DMSO  
% de proliferación

| Fármaco           | Día 4       | Día 8       |
|-------------------|-------------|-------------|
| NFX               | 1,7 (0,18)  | 1,3 (0,17)  |
| Sulfisoxazol      | 83,3 (19,8) | 101,7 (4,8) |
| Metil Antranilato | 86,7 (13,5) | 91,6 (13,1) |
| Sulfametizol      | 95,6 (13,5) | 96,5 (9,4)  |
| Naratriptán       | 99,7 (4,3)  | 99,6 (2,4)  |
| DMSO              | 100 (16)    | 100 (6)     |



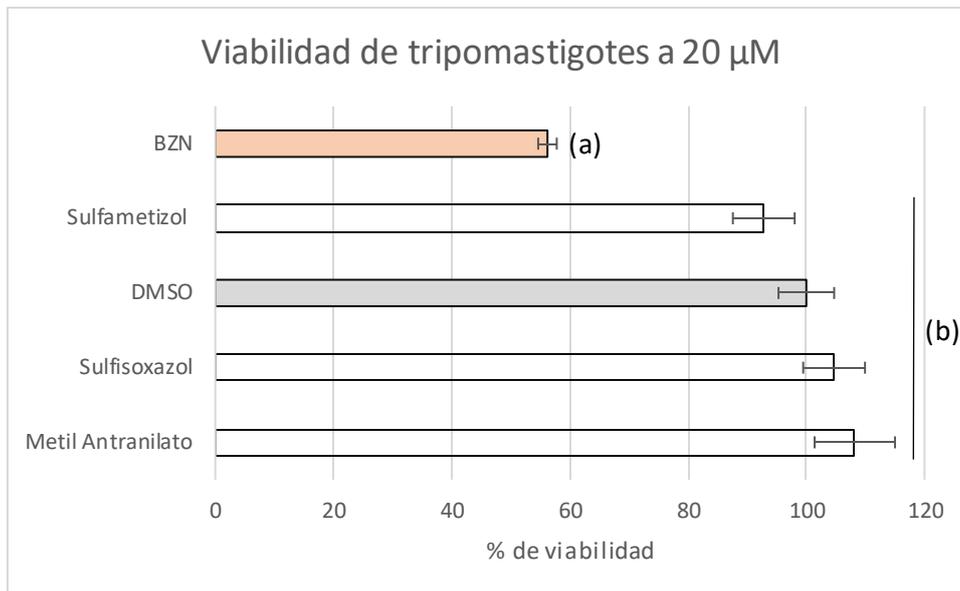
**Figura 4.20.** Gráfico de la proliferación de epimastigotes. Los asteriscos indican diferencias significativas para la prueba de Tukey con p-valor < 0,001.

Como se puede observar a partir de la **tabla 4.21** todos los candidatos han permitido una proliferación parasitaria por encima del 80% para el día cuatro y por arriba del 90% para el día ocho. El fármaco Celecoxib no se evaluó frente a *T. cruzi* ya que fue reportado como inactivo por diversos autores (Hideko Tatakahara et al., 2008; Malvezi et al., 2014; Planer et al., 2014).

Posteriormente los resultados fueron analizados mediante la prueba de Tukey aplicando previamente un análisis de la varianza del recuento obtenido. Este análisis demostró que ninguno de los fármacos evaluados presentó diferencias significativas con respecto al control negativo (DMSO), mientras que el control positivo (NFX 25 µM) mostró, para la misma prueba, diferencias significativas con un p-valor < 0,001 con respecto al resto de los candidatos. Es decir que ninguno de estos fármacos fue capaz de interferir en la proliferación de epimastigotes. Con estos resultados se da por finalizada la exploración de los fármacos frente a epimastigotes de *T. cruzi*.

#### 4.4.1.2 Evaluación frente a tripomastigotes

Para estimar el posible efecto tripanocida de los candidatos frente al estadio tripomastigote se realizó un ensayo de viabilidad a 20 µM con tres réplicas para cada fármaco. A continuación, se presenta un gráfico de barras en el que se puede observar el promedio del % de viabilidad obtenido para cada uno de los candidatos (ver **figura 4.21**), los datos se encuentran normalizados respecto a los valores obtenidos para DMSO (100% de viabilidad).



**Figura 4.21.** Porcentaje de viabilidad de tripomastigote. Las medias con una letra en común no presentan diferencias estadísticamente significativas para la prueba de Tukey ( $p$ -valor  $>0,05$ ).

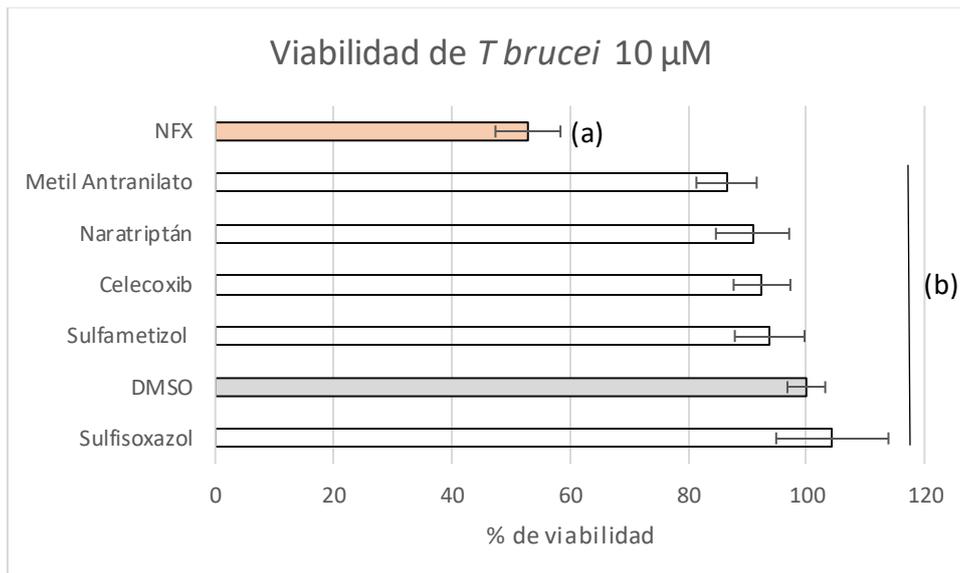
La prueba de Tukey aplicada al análisis de la varianza de los resultados obtenidos reveló que no existen diferencias significativas entre los candidatos y el control negativo (DMSO 1%), y que además ningún fármaco es capaz de disminuir el número de parásitos de manera similar al control positivo (BZN 20 µM). Aquí se decidió no continuar con los ensayos sobre este estadio. Considerando que la enzima de interés está presente a lo largo de todo el ciclo de vida el parásito, los estadios más críticos en relación con el estrés oxidativo que debe sortear *T. cruzi* es cuando se encuentra en la forma epimastigote (estadio replicativo) o tripomastigote (estadio infectivo). En el estadio epimastigote, además de observarse un elevado metabolismo, el parásito se encuentra sometido a condiciones externas agresivas, dado que debe sobrevivir a las secreciones gástricas del insecto vector (Ariyanayagam et al., 2003; Mesías et al., 2019). Mesías y col. sugieren que tripomastigote es el estadio con mayor demanda de  $[TS]_2$ , ya que en esta etapa del ciclo *T. cruzi* debe afrontar elevadas concentraciones de radicales libres y sustancias mediadoras de la inflamación y activación del sistema inmune del huésped mamífero (Mesías et al., 2019). Adicionalmente es importante recordar que concentraciones 10 µM deberían ser suficientes para identificar compuestos con actividad relevante frente a un blanco molecular (Hughes et al., 2011), y en esta oportunidad no se observó actividad incluso a concentraciones que duplican dicho valor. A partir de los resultados de los ensayos fenotípicos en los estadios de *T. cruzi* sujetos a mayor estrés oxidativo, es claro que los candidatos evaluados no parecen prometedores en relación a su capacidad inhibitoria sobre el blanco molecular de interés, como se corroboró más adelante en los estudios enzimáticos.

Los resultados del fármaco Naratriptán no pudieron ser cuantificados debido a inconvenientes con la placa de cultivo al momento del recuento (bajo volumen). Sin embargo, y solo como comentario informativo, se puede mencionar que el recuento de parásitos pareciera no ser diferente a los recuentos obtenidos para el resto de los compuestos, por lo que sospechamos que el fármaco podría no presentar efectos tripanocidas. Este resultado no se ha podido comprobar debido a la pandemia declarada en marzo de 2020 por la OMS, y

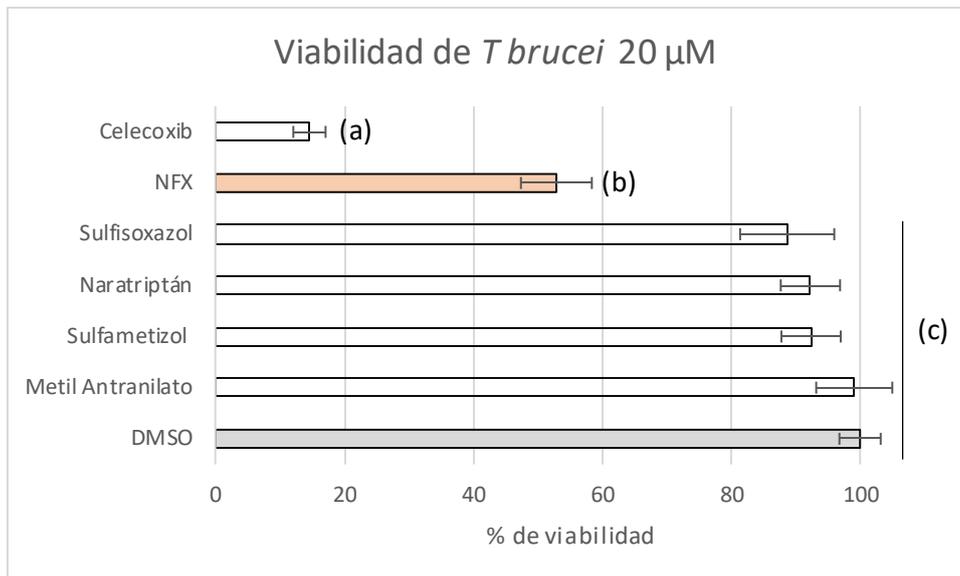
posterior aislamiento social, preventivo y obligatorio dispuesto por la Presidencia de la Nación Argentina.

#### 4.4.2 Evaluación biológica frente a *T. brucei*

En el presente ensayo se determinó la viabilidad de los parásitos en estadio *bloodstream* frente la línea celular 427 cepa BB3 que expresa la enzima Luciferasa. Los candidatos fueron evaluados a concentraciones de 10 y 20  $\mu\text{M}$  en placa de 96 pocillos y por triplicado. El efecto que ejercen los fármacos sobre la viabilidad de los parásitos fue cuantificado mediante el análisis de la intensidad de la señal lumínica, la cual es determinada a partir de la actividad enzimática. A mayor intensidad, más cantidad de enzima expresada, y en consecuencia es mayor la cantidad de parásitos viables. Los resultados volcados en la figura 4.22 representan la media de la determinación de viabilidad en tres pocillos para el ensayo a 10  $\mu\text{M}$  y las barras representan la desviación estándar para cada fármaco, mientras que la figura 4.23 muestra similares resultados, pero para el ensayo realizado a 20  $\mu\text{M}$ .



**Figura 4.22.** Viabilidad de *T. brucei* frente los diferentes candidatos. Las medias con una letra en común no son significativamente diferentes para la prueba de Tukey ( $p$ -valor  $>0,05$ ). Control negativo DMSO 1%, control positivo NFX 6  $\mu\text{M}$ .



**Figura 4.23.** Viabilidad de *T. brucei* frente los diferentes candidatos. Las medias con una letra en común no son significativamente diferentes para la prueba de Tukey ( $p$ -valor  $>0,05$ ). Control negativo DMSO 1%, control positivo NFX 6  $\mu$ M.

Los resultados fueron analizados mediante una prueba ANOVA de una vía y posterior prueba estadística de Tukey, este análisis reveló que ninguno de los candidatos presentó, a 10  $\mu$ M, efecto significativamente diferente al control negativo (DMSO 1%), mientras que a 20  $\mu$ M el fármaco Celecoxib mostró una actividad tripanocida significativamente mayor que el control positivo (NFX 6  $\mu$ M). Como se puede apreciar de la **figura 4.23**, Celecoxib produjo una reducción de más del 85% en la viabilidad del parásito. Estos datos sugieren que la  $EC_{50}$  de este fármaco se encuentra entre 10 y 20  $\mu$ M.

A partir de los datos recabados anteriormente se realizó un curva dosis-respuesta, a fin de poder determinar el valor de  $EC_{50}$  para Celecoxib frente a *T. brucei*. Las concentraciones variaron desde 30 hasta 5  $\mu$ M descendiendo en 5 unidades cada punto de la curva (30, 25, 20, 15, 10 y 5  $\mu$ M), para cada concentración se sembraron tres pocillos. En la **figura 4.24** se presenta la curva obtenida, donde cada punto representa el promedio del % de viabilidad de los tres pocillos sembrados, y las barras muestran la desviación estándar alrededor del promedio.

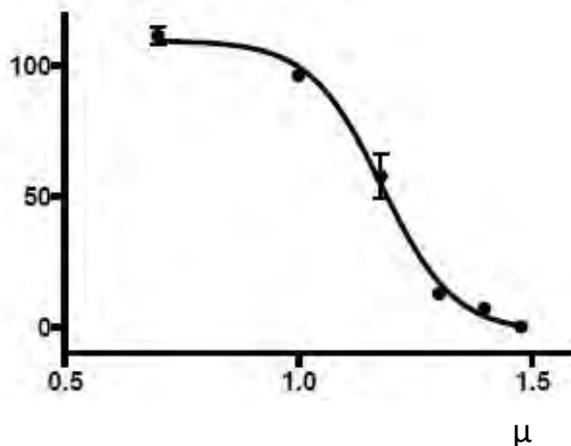


Figura 4.24. Curva Supervivencia vs Concentración.

Con el software GraphPad<sup>1</sup>, y realizando las transformaciones correspondientes se determinó que el valor de  $EC_{50}$  fue de 17,54  $\mu$ M, con un  $IC_{95}$  (11,28  $\mu$ M; 27,26  $\mu$ M).

Estos resultados podrían ser alentadores cuando se observan solo los ensayos *in vitro*, sin embargo el tratamiento de una patología humana requiere indefectiblemente la administración de una (o reiteradas) dosis del fármaco al paciente y en esa situación es donde el posible tratamiento innovador se detiene, ya que los datos bibliográficos reportan que Celecoxib alcanza una concentración plasmática máxima ( $C_{p_{max}}$ ) de 1,85  $\mu$ M luego de una dosis de 200 mg (Wishart et al., 2018). Estos datos farmacocinéticos sugieren que para alcanzar una concentración plasmática próxima a la  $EC_{50}$  de Celecoxib frente a *T. brucei* se debe administrar a los pacientes una dosis de 1500 mg, es decir que dicho tratamiento requeriría una dosis de carga de fármaco que se encuentra entre 7 y 8 veces por encima de dosis recomendada (200 mg). En este aspecto, resulta poco ético profundizar el estudio de este fármaco frente a modelos animales, ya que con los valores de la dosis que se estimó para Celecoxib, a partir de la  $EC_{50}$ , será poco probable que se pueda realizar un estudio clínico frente a pacientes portadores de *T. brucei*. Sin embargo, desde la óptica experimental este fármaco puede ser el puntapié inicial para la optimización y diseño de compuestos que, manteniendo el núcleo base del fármaco, mejoren la potencia de este frente *T. brucei*.

### 4.4.3 Evaluación bioquímica frente a TryS

Los candidatos seleccionados fueron evaluados frente a la enzima **TryS** de las especies *T. cruzi*, *T. Brucei* y *L. infantum*. La concentración del screening primario fue de 25  $\mu$ M en placa, y los resultados pueden observarse en la **tabla 4.22**, donde se informa el promedio del %AR de los 4 pocillos con su respectivo CV%. En esta campaña se utilizó para las tres enzimas DMSO 1% como control negativo y el inhibidor Ácido aurintricarboxílico (AUC) a 25  $\mu$ M como control positivo.

Tabla 4.22. Actividad residual de la enzima *TryS* frente a diferentes fármacos

| Fármaco                       | <i>Trys</i>     |      |                  |      |                    |      |
|-------------------------------|-----------------|------|------------------|------|--------------------|------|
|                               | <i>T. cruzi</i> |      | <i>T. brucei</i> |      | <i>L. infantum</i> |      |
|                               | % AR            | CV%  | % AR             | CV%  | % AR               | CV%  |
| Metil Antranilato             | 95,1            | 3,6  | 96,0             | 2,2  | 91,8               | 12,6 |
| Sulfametizol                  | 99,4            | 4,0  | 96,9             | 3,6  | 99,4               | 6,0  |
| Sulfisoxazol                  | 100,6           | 1,8  | 98,8             | 2,3  | 97,9               | 3,7  |
| Naratriptán                   | 100,1           | 1,3  | 100,7            | 2,6  | 96,1               | 1,0  |
| Celecoxib                     | 108,1           | 1,1  | 98,3             | 16,6 | 92,2               | 11,3 |
| DMSO 1%                       | 100,0           | 5,1  | 99,8             | 2,3  | 100,0              | 5,4  |
| AUC 25 $\mu$ M<br>(inhibidor) | 48,7            | 11,9 | 54,9             | 9,5  | 36,2               | 10,2 |

En las figuras 4.25 a 4.27 que se presentan a continuación se puede observar la media del %AR para cada uno de los fármacos evaluados frente a las diferentes enzimas, indicando además el desvío estándar de cada uno.

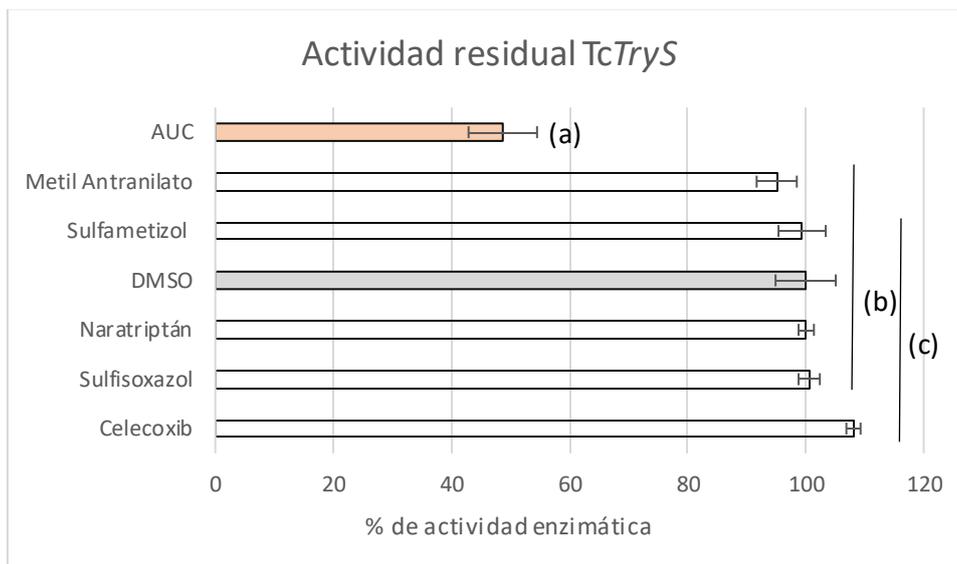
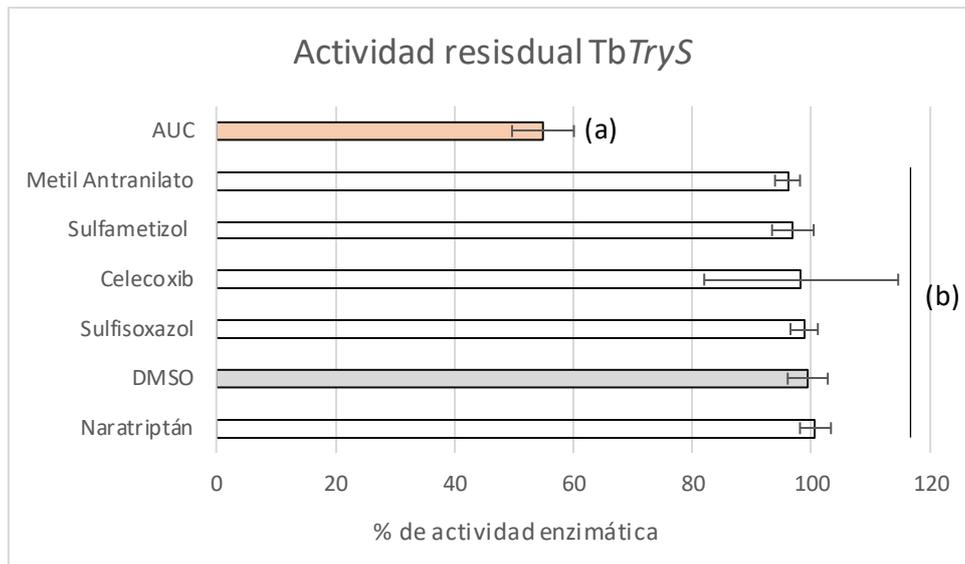


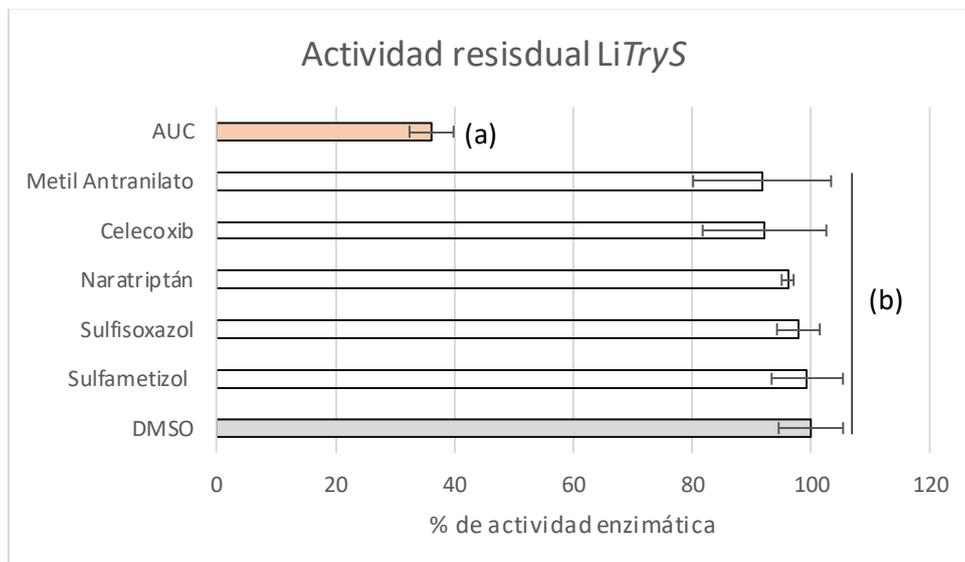
Figura 4.25. Actividad residual de *TcTryS* frente a cada fármaco. Los compuestos incluidos en grupos agrupados con una misma letra no presentan diferencias significativas entre las medias de actividad residual ( $p$ -valor > 0,05). Control negativo DMSO 1%, control positivo AUC 25  $\mu$ M.

Los candidatos evaluados frente a *TcTryS* no mostraron diferencias significativas respecto al control negativo cuando fueron analizados estadísticamente bajo una prueba ANOVA y posterior prueba de Tukey con un  $p$ -valor crítico de 0,05. Estos resultados indican que los fármacos no interfieren en la actividad enzimática de *TcTryS*, aun cuando la concentración evaluada es 2,5 veces mayor a la que se sugiere en bibliográfica (10  $\mu$ M) (Hughes et al., 2011).



**Figura 4.26.** Actividad residual de TbTryS frente a cada fármaco. Los compuestos incluidos en grupos agrupados con una misma letra no presentan diferencias significativas entre las medias de actividad residual ( $p$ -valor  $> 0,05$ ). Control negativo DMSO 1%, control positivo AUC 25  $\mu$ M.

Los datos recogidos en los ensayos frente a TbTryS fueron sometidos a un análisis de la varianza de una vía seguido de una prueba de Tukey; este análisis demostró que ninguno de los candidatos se diferencia significativamente del control negativo, es decir que de estos cinco fármacos ninguno es capaz de reducir la actividad enzimática. Dado que las evaluaciones fueron diseñadas siguiendo el esquema de un cribado primario, la concentración que se ha testeado se encuentra en el valor máximo de lo que se establece para este protocolo frente a la TryS de las diversas especies de tripanosomátidos.



**Figura 4.27.** Actividad residual de LiTryS frente a cada fármaco. Los compuestos incluidos en grupos agrupados con una misma letra no presentan diferencias significativas entre las medias de actividad residual ( $p$ -valor  $> 0,05$ ). Control negativo DMSO 1%, control positivo AUC 25  $\mu$ M.

El análisis estadístico mediante la prueba de Tukey aplicado luego del ANOVA de una vía para los resultados frente a LiTryS estableció con un 95% de confianza que ninguno de los

cinco fármacos testeados es capaz de reducir la actividad enzimática a porcentajes significativamente inferiores al 100% (DMSO), y como ya se ha planteado en párrafos anteriores, la concentración que se utilizó es superior al límite superior del rango aconsejado en bibliografía.

Como se desprende de los análisis estadísticos realizados para cada una de las enzimas (Tc**TryS**, Tb**TryS** y Li**TryS**) ninguno de los fármacos ha producido una inhibición considerable (al menos 50%) a 25  $\mu\text{M}$ , y en consonancia con las recomendaciones de cribado experimental, donde se sugiere que para un *hit* validado la  $\text{IC}_{50}$  no debería ser mayor a 10  $\mu\text{M}$  (Hughes et al., 2011; Katsuno et al., 2015), se ha decidido finalizar los estudios experimentales de estos cinco candidatos frente a las **TryS** ya evaluadas.

Luego de dos campañas de cribado virtual, y habiéndose evaluado experimentalmente un total de 15 compuestos provenientes de dichas campañas, se nos plantea un nuevo interrogante: si una enzima que posee un rol clave en el metabolismo del parásito (**versección 1.3**), ¿es plausible de ser fácilmente modulada de manera exógena?, ¿o es probable que presente, a nivel estructural, resistencia frente a la modulación farmacológica?

En el siguiente capítulo se abordan metodologías complementarias a las presentadas hasta aquí con el objetivo de encontrar respuestas a los comportamientos observados.

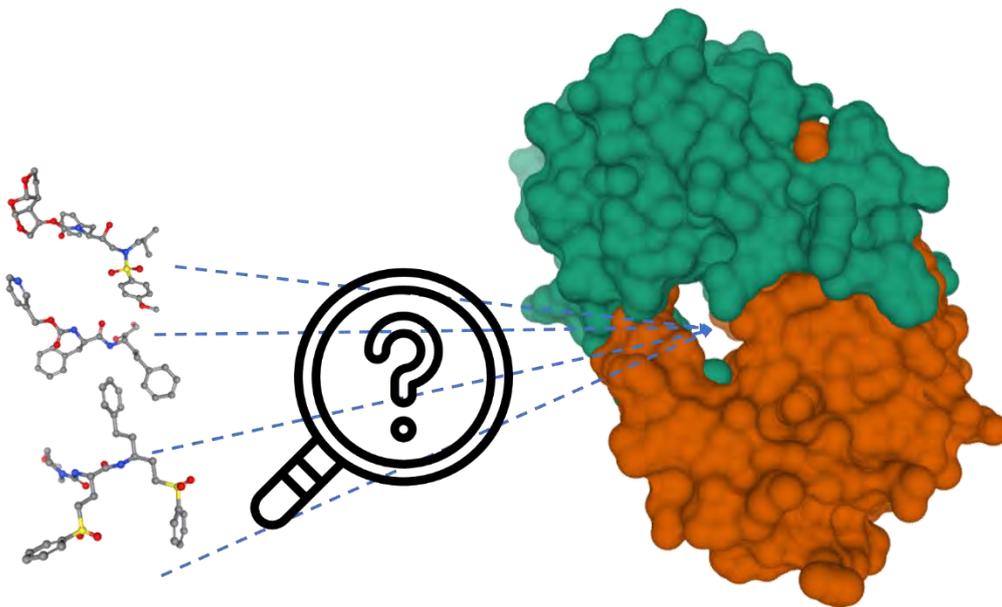
## 4.5 Referencias Bibliográficas

- Alberca, L., Sbaraglini, M., Ruiz, D., Larrea, S. V., Villamil, S. F., Carrillo, C., & Talevi, A. (2018). Computer-aided search of novel inhibitors of n-myristoyl transferase with trypanocidal effects. *International Journal of Infectious Diseases*, 73, 311. <https://doi.org/10.1016/j.ijid.2018.04.4121>
- Ariyanayagam, M. R., Oza, S. L., Mehlert, A., & Fairlamb, A. H. (2003). Bis(glutathionyl)spermine and Other Novel Trypanothione Analogues in *Trypanosoma cruzi*. *Journal of Biological Chemistry*, 278(30), 27612-27619. <https://doi.org/10.1074/jbc.M302750200>
- Contreras, F., Romero, B., Suárez, N., González, M., Fouillioux, C., Guevara, E., Betancourt, M. C., Torres, D., & Velasco, M. (2002). Receptores Sur y Sulfonilureas en el tratamiento de la Diabetes Mellitus Tipo 2. *Archivos Venezolanos de Farmacología y Terapéutica*, 21(2), 148-155.
- Daunes, S., D'Silva, C., Kendrick, H., Yardley, V., & Croft, S. L. (2001). QSAR Study on the Contribution of Log P and  $E_s$  to the in Vitro Antiprotozoal Activity of Glutathione Derivatives. *Journal of Medicinal Chemistry*, 44(18), 2976-2983. <https://doi.org/10.1021/jm000502n>
- D'Silva, C., Daunes, S., Rock, P., Yardley, V., & Croft, S. L. (2000). Structure-Activity Study on the in Vitro Antiprotozoal Activity of Glutathione Derivatives. *Journal of Medicinal Chemistry*, 43(10), 2072-2078. <https://doi.org/10.1021/jm990259w>
- Hideko Tatakihara, V. L., Cecchini, R., Borges, C. L., Malvezi, A. D., Graça-de Souza, V. K., Yamada-Ogatta, S. F., Rizzo, L. V., & Pinge-Filho, P. (2008). Effects of cyclooxygenase inhibitors on parasite burden, anemia and oxidative stress in murine *Trypanosoma cruzi* infection. *FEMS Immunology & Medical Microbiology*, 52(1), 47-58. <https://doi.org/10.1111/j.1574-695X.2007.00340.x>
- Hughes, J., Rees, S., Kalindjian, S., & Philpott, K. (2011). Principles of early drug discovery: Principles of early drug discovery. *British Journal of Pharmacology*, 162(6), 1239-1249. <https://doi.org/10.1111/j.1476-5381.2010.01127.x>
- Katsuno, K., Burrows, J. N., Duncan, K., van Huijsduijnen, R. H., Kaneko, T., Kita, K., Mowbray, C. E., Schmatz, D., Warner, P., & Slingsby, B. T. (2015). Hit and lead criteria in drug discovery for infectious diseases of the developing world. *Nature Reviews Drug Discovery*, 14(11), 751-758. <https://doi.org/10.1038/nrd4683>
- Knox, C., Law, V., Jewison, T., Liu, P., Ly, S., Frolkis, A., Pon, A., Banco, K., Mak, C., Neveu, V., Djoumbou, Y., Eisner, R., Guo, A. C., & Wishart, D. S. (2011). *DrugBank 3.0: A comprehensive resource for 'Omics' research on drugs*. 39(November 2010), 1035-1041. <https://doi.org/10.1093/nar/gkq1126>
- Koch, O., Jäger, T., Flohé, L., & Selzer, P. M. (2013). Inhibition of Trypanothione Synthetase as a Therapeutic Concept. En T. Jäger, O. Koch, & L. Flohé (Eds.), *Trypanosomatid Diseases* (pp. 429-443). Wiley-VCH Verlag GmbH & Co. KGaA. <https://doi.org/10.1002/9783527670383.ch23>

- Leroux, A. E., & Krauth-Siegel, R. L. (2016). Thiol redox biology of trypanosomatids and potential targets for chemotherapy. *Molecular and Biochemical Parasitology*, 206(1-2), 67-74. <https://doi.org/10.1016/j.molbiopara.2015.11.003>
- Linares, G., Ravaschino, E., & Rodriguez, J. (2006). Progresses in the Field of Drug Design to Combat Tropical Protozoan Parasitic Diseases. *Current Medicinal Chemistry*, 13(3), 335-360. <https://doi.org/10.2174/092986706775476043>
- Malvezi, A. D., Panis, C., da Silva, R. V., de Freitas, R. C., Lovo-Martins, M. I., Tatakihara, V. L. H., Zanluqui, N. G., Neto, E. C., Goldenberg, S., Bordignon, J., Yamada-Ogatta, S. F., Martins-Pinge, M. C., Cecchini, R., & Pinge-Filho, P. (2014). Inhibition of Cyclooxygenase-1 and Cyclooxygenase-2 Impairs *Trypanosoma cruzi* Entry into Cardiac Cells and Promotes Differential Modulation of the Inflammatory Response. *Antimicrobial Agents and Chemotherapy*, 58(10), 6157-6164. <https://doi.org/10.1128/AAC.02752-14>
- Melander, A., Sartor, G., Wahlin, E., Schersten, B., & Bitzen, P.-O. (1978). Serum tolbutamide and chlorpropamide concentrations in patients with diabetes mellitus. *BMJ*, 1(6106), 142-144. <https://doi.org/10.1136/bmj.1.6106.142>
- Mesías, A. C., Sasoni, N., Arias, D. G., Pérez Brandán, C., Orban, O. C. F., Kunick, C., Robello, C., Comini, M. A., Garg, N. J., & Zago, M. P. (2019). Trypanothione synthetase confers growth, survival advantage and resistance to anti-protozoal drugs in *Trypanosoma cruzi*. *Free Radical Biology and Medicine*, 130, 23-34. <https://doi.org/10.1016/j.freeradbiomed.2018.10.436>
- Mysinger, M. M., Carchia, M., Irwin, John. J., & Shoichet, B. K. (2012a). Directory of Useful Decoys, Enhanced (DUD-E): Better Ligands and Decoys for Better Benchmarking. *Journal of Medicinal Chemistry*, 55(14), 6582-6594. <https://doi.org/10.1021/jm300687e>
- Mysinger, M. M., Carchia, M., Irwin, John. J., & Shoichet, B. K. (2012b). Directory of Useful Decoys, Enhanced (DUD-E): Better Ligands and Decoys for Better Benchmarking. *Journal of Medicinal Chemistry*, 55(14), 6582-6594. <https://doi.org/10.1021/jm300687e>
- Novick, P. A., Ortiz, O. F., Poelman, J., Abdulhay, A. Y., & Pande, V. S. (2013). SWEETLEAD: An In Silico Database of Approved Drugs, Regulated Chemicals, and Herbal Isolates for Computer-Aided Drug Discovery. *PLoS ONE*, 8(11), e79568. <https://doi.org/10.1371/journal.pone.0079568>
- Oza, S. L., Chen, S., Wyllie, S., Coward, J. K., & Fairlamb, A. H. (2008). ATP-dependent ligases in trypanothione biosynthesis - kinetics of catalysis and inhibition by phosphinic acid pseudopeptides: Kinetics and inhibition of ATP-dependent ligases. *FEBS Journal*, 275(21), 5408-5421. <https://doi.org/10.1111/j.1742-4658.2008.06670.x>
- Perez-Llamas, C., & Lopez-Bigas, N. (2011). Gitools: Analysis and Visualisation of Genomic Data Using Interactive Heat-Maps. *PLoS ONE*, 6(5), e19541. <https://doi.org/10.1371/journal.pone.0019541>
- Planer, J. D., Hulverson, M. A., Arif, J. A., Ranade, R. M., Don, R., & Buckner, F. S. (2014). Synergy Testing of FDA-Approved Drugs Identifies Potent Drug Combinations against *Trypanosoma cruzi*. *PLoS Neglected Tropical Diseases*, 8(7), e2977. <https://doi.org/10.1371/journal.pntd.0002977>
- Ravaschino, E. L., Docampo, R., & Rodriguez, J. B. (2006). Design, Synthesis, and Biological Evaluation of Phosphinopeptides against *Trypanosoma cruzi* Targeting Trypanothione

- Biosynthesis. *Journal of Medicinal Chemistry*, 49(1), 426-435.  
<https://doi.org/10.1021/jm050922i>
- Roy, K., & Mitra, I. (2011). On Various Metrics Used for Validation of Predictive QSAR Models with Applications in Virtual Screening and Focused Library Design. *Combinatorial Chemistry & High Throughput Screening*, 14(6), 450-474.  
<https://doi.org/10.2174/138620711795767893>
- Schröder, H., & Campbell, D. E. S. (1972). Absorption, metabolism, and excretion of salicylazosulfapyridine in man. *Clinical Pharmacology & Therapeutics*, 13(4), 539-551.  
<https://doi.org/10.1002/cpt1972134539>
- Spinks, D., Torrie, L. S., Thompson, S., Harrison, J. R., Frearson, J. A., Read, K. D., Fairlamb, A. H., Wyatt, P. G., & Gilbert, I. H. (2012). Design, Synthesis and Biological Evaluation of Trypanosoma brucei Trypanothione Synthetase Inhibitors. *ChemMedChem*, 7(1), 95-106.  
<https://doi.org/10.1002/cmdc.201100420>
- Stuhlmann, F. D., Jäger, T. D., Flohé, L. P. D., & Schinzer, D. P. D. (2007). *N5-substituted benzo-2,3/azepino-4,5-b/indol-6-ones for treating tropical diseases* (European Union Patent N.º EP1757607A1). <https://patents.google.com/patent/EP1757607A1/un>
- Toropova, A. P., & Toropov, A. A. (2017). CORAL: Binary classifications (active/inactive) for drug-induced liver injury. *Toxicology Letters*, 268, 51-57.  
<https://doi.org/10.1016/j.toxlet.2017.01.011>
- Torrie, L. S., Wyllie, S., Spinks, D., Oza, S. L., Thompson, S., Harrison, J. R., Gilbert, I. H., Wyatt, P. G., Fairlamb, A. H., & Frearson, J. A. (2009). Chemical Validation of Trypanothione Synthetase: A POTENTIAL DRUG TARGET FOR HUMAN TRYPANOSOMIASIS. *Journal of Biological Chemistry*, 284(52), 36137-36145. <https://doi.org/10.1074/jbc.M109.045336>
- Truchon, J.-F., & Bayly, C. I. (2007). Evaluating Virtual Screening Methods: Good and Bad Metrics for the “Early Recognition” Problem. *Journal of Chemical Information and Modeling*, 47(2), 488-508. <https://doi.org/10.1021/ci600426e>
- Wishart, D. S., Feunang, Y. D., Guo, A. C., Lo, E. J., Marcu, A., Grant, J. R., Sajed, T., Johnson, D., Li, C., Sayeeda, Z., Assempour, N., Iynkkaran, I., Liu, Y., Maciejewski, A., Gale, N., Wilson, A., Chin, L., Cummings, R., Le, D., ... Wilson, M. (2018). DrugBank 5.0: A major update to the DrugBank database for 2018. *Nucleic Acids Research*, 46(D1), D1074-D1082.  
<https://doi.org/10.1093/nar/gkx1037>
- Zimmermann, S., Oufir, M., Leroux, A., Krauth-Siegel, R. L., Becker, K., Kaiser, M., Brun, R., Hamburger, M., & Adams, M. (2013). Cynaropicrin targets the trypanothione redox system in Trypanosoma brucei. *Bioorganic & Medicinal Chemistry*, 21(22), 7202-7209.  
<https://doi.org/10.1016/j.bmc.2013.08.052>

# Capítulo 5 | Drogabilidad



### 5.1 Introducción

El proceso de descubrimiento de fármacos centrado en el blanco molecular comienza con la identificación y validación de una o más biomoléculas cuya modulación interferirá en la progresión de la enfermedad, o al menos aliviará los síntomas. Además de verificar la asociación entre tales blancos moleculares y la enfermedad, la validación del blanco a menudo implica la evaluación de propiedades relevantes, entre ellas la drogabilidad (Agoni et al., 2020).

Para quienes trabajan en el descubrimiento de fármacos es manifiesto que no toda molécula pequeña podrá ser considerada un fármaco, ya que para que esto ocurra la molécula debe presentar ciertas propiedades fisicoquímicas, farmacológicas o farmacocinéticas, entre otras. Este mismo razonamiento es aplicable a los blancos moleculares. Es decir, que para que una proteína sea considerada un blanco molecular atractivo debe reunir al menos dos requisitos, por un lado presentar un rol importante en la aparición o desarrollo de la enfermedad, mientras que el segundo requisito está vinculado con la drogabilidad de la proteína en estudio (Barril, 2013; Volkamer, Kuhn, Grombacher, et al., 2012). Si bien en la sección 1.4.2 se hace referencia a la vulnerabilidad del blanco molecular, este concepto es aplicable solo si el blanco molecular es drogable.

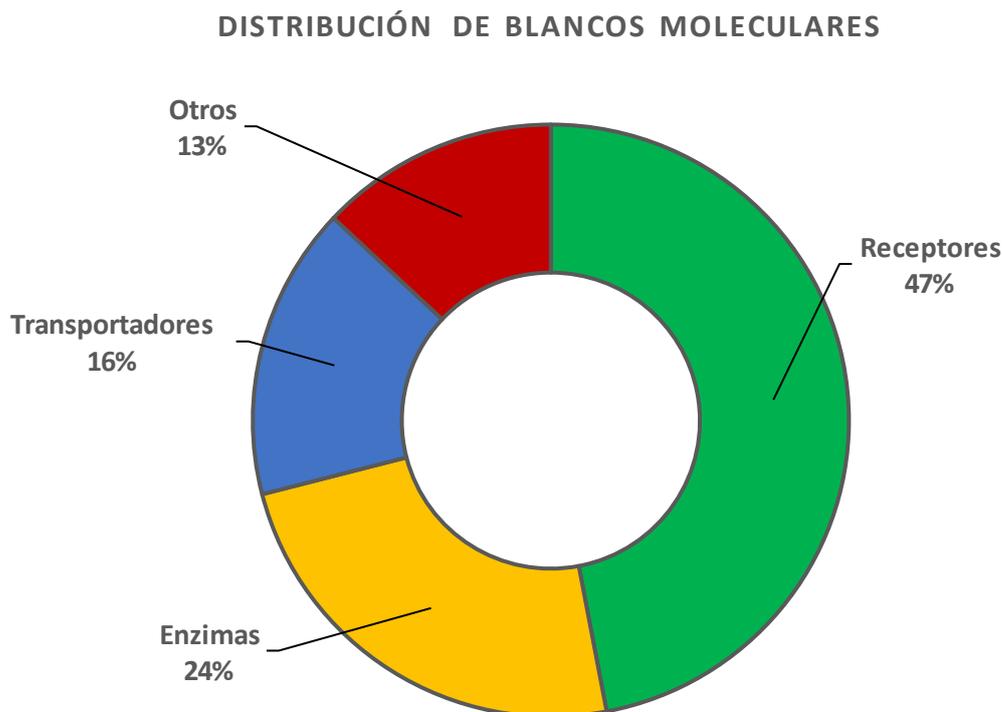
El término **drogabilidad** (en inglés *druggability*) fue acuñado por Hopkins y Groom en 2003 para referirse a la factibilidad con la que una macromolécula, generalmente una proteína, puede ser modulada farmacológicamente mediante pequeños compuestos tipo fármaco con alta afinidad y especificidad (Hopkins & Groom, 2003).

Si una molécula pequeña se une a un blanco molecular de una manera específica, este se considera tradicionalmente como drogable (recientemente, sin embargo, se ha propuesto que diferentes dianas moleculares pueden presentar diferentes grados de drogabilidad, es decir que algunas macromoléculas son "fáciles de modular" y que otras resultan "difíciles de modular"). Esto no significa de ninguna manera que la ausencia de un ligando conocido de alta afinidad sea necesariamente equivalente a la falta de drogabilidad del blanco molecular. En consecuencia, se requieren herramientas de predicción de drogabilidad para evaluar prospectivamente si un blanco molecular es drogable o no.

La drogabilidad también ha sido asumida de manera implícita en base a relaciones evolutivas (usando una especie de principio de transitividad o principio de "culpabilidad por asociación", es decir, si una macromolécula evolutivamente cercana al blanco molecular de interés se ha demostrado drogable, se asume que el blanco molecular de interés es drogable). Por otra parte, es posible predecir de manera directa los bolsillos drogables a partir de diferentes niveles de información estructural, sin comparación con ninguna macromolécula de referencia (Wang et al., 2013).

Si bien el concepto de drogabilidad puede ser aplicable a todas las macromoléculas biológicas, las herramientas que predicen drogabilidad se enfocan esencialmente en proteínas. El sesgo del enfoque está estrechamente vinculado a la distribución de blancos moleculares identificados hasta el momento, donde más del 85 % de las dianas son proteínas, mientras que el porcentaje restante se reparte entre ácidos nucleicos y lípidos (Barril, 2013). En la **figura 5.1**

se observa la distribución de los blancos moleculares, donde se puede ver que la mayoría corresponden a proteínas.



**Figura 5.1** Distribución de blancos moleculares según su función. Adaptado de (Barril, 2013)

En el **cuadro 5.1** se presentan algunas definiciones que podrían facilitar la lectura del presente capítulo.

**Cuadro 5.1.** Definiciones útiles, adaptado de (Hajduk et al., 2005)

- **Bolsillo:** Es una cavidad en la proteína que presenta un volumen y profundidad determinados por la composición y arreglo de los aminoácidos que lo circundan.
- **Bolsillo drogable:** Es un bolsillo al cual se pueden unir con alta afinidad y especificidad pequeñas moléculas tipo fármaco.
- **Bolsillo no drogable:** Es un bolsillo al cual, por sus características y/o tamaño, no pueden unirse moléculas tipo fármaco.
- **Bolsillo señuelo:** Es un bolsillo que los modelos computacionales predicen como drogable, pero experimentalmente no se conocen ligandos que se puedan unir allí.

En los últimos años el estudio de drogabilidad en proteínas se ha incrementado gracias al avance de metodologías bioinformáticas; estas técnicas permiten simulaciones y estudios de drogabilidad de proteínas aun cuando no se conozcan homólogos o cuando no se hayan realizado validaciones experimentales con la proteína de interés (Agoni et al., 2020). Es

importante mencionar que la bioinformática ha sido un aporte valioso para superar las limitaciones que plantean los ensayos húmedos como el cribado de fragmentos por RMN, el Cribado Cristalográfico de Solventes Múltiples (*MSCS*, por sus siglas en Inglés), el Cribado de Alto Rendimiento *HTS* u otros métodos experimentales (Barril, 2013; Volkamer et al., 2012). En la **tabla 5.1** se resumen las principales ventajas y limitaciones de algunos métodos de cribado húmedo.

**Tabla 5.1.** Ventajas y limitaciones de métodos de cribado húmedo

| Método experimental                  | Ventajas   | Limitaciones  |
|--------------------------------------|--|---|
| <b>Cribado de fragmentos por RMN</b> | Más sencillo que realizar pruebas bioquímicas.<br>Requiere poca cantidad de proteínas.                                   | La identificación del sitio de unión requiere ensayos adicionales (competencia de ligandos, información estructural, etc.).   |
| <b>HTS</b>                           | Las tasas de aciertos en cribados de <i>HTS</i> anteriores pueden brindar información sobre drogabilidad y compendiarse. | Sesgo debido al andamiaje químico de la biblioteca química cribada.<br>Costoso.   |
| <b>MSCS</b>                          | Identificación directa del sitio de unión.   | El blanco debe formar cristales que resistan altas concentraciones de solventes orgánicos.<br>Requiere mapas de densidad eléctricas de alta calidad que permitan diferenciar las moléculas de solvente. |

Actualmente existen diversos enfoques que permiten estimar la drogabilidad de los bolsillos de las proteínas. A modo de ejemplo, repasaremos brevemente las características principales de cada uno de ellos:

**Enfoque basado en análisis de secuencia:** Así como los aminoácidos determinan la funcionalidad y estructura de la proteína que conforman, también pueden considerarse responsables de la drogabilidad de la misma, ya que las propiedades hidrofóbicas, electroestáticas, etc. de los aminoácidos que componen el bolsillo determinarán sus propiedades. Este enfoque se basa en la identificación de bolsillos drogables a partir de la estructura primaria mediante alineación de secuencia con una proteína patrón que ya ha sido identificada como drogable. El método identifica las regiones conservadas en ambas macromoléculas, y por lo tanto si un bolsillo de la proteína de interés conserva la composición aminoacídica de la macromolécula de referencia, entonces puede asumirse que dicho bolsillo presenta una elevada probabilidad de ser también drogable. La principal desventaja del método es que requiere la comparación con otras proteínas que hayan demostrado

drogabilidad, ya que la secuencia por sí sola no aporta ninguna información relevante (Agoni et al., 2020).

**Enfoque basado en energía:** Estos métodos utilizan una sonda química (pequeñas moléculas o fracciones de ellas) para identificar regiones en las que se manifiestan interacciones del tipo puentes de hidrógeno o acoplamientos electrónicos  $\pi$  entre la proteína y la sonda química. Los sitios donde ocurren dichas interacciones se considerarán probablemente drogables (Clark et al., 2020).

**Enfoque basado en aprendizaje automático (*machine learning*):** Los métodos de aprendizaje automático han sido un gran avance en las técnicas para la identificación de blancos moleculares drogables. Estas metodologías facilitan la integración de un gran número de descriptores de bolsillos drogables de diversa naturaleza, previamente validados como relevantes, utilizándolos para entrenar algoritmos de decisión que pueden ir desde simples árboles de decisión hasta complejas redes neuronales. Este enfoque tolera, también, la combinación de diferentes estrategias y el desarrollo de métodos de consenso, donde la resolución del problema se basa en la superposición de predicciones realizadas por diferentes algoritmos que se encuentran en servidores web (Clark et al., 2020).

**Enfoque basado en geometría:** En este enfoque se engloban tres métodos diferentes, uno basado en cuadrícula, otro basado en esferas y el último basado en formas  $\alpha$ ; sin embargo, en los tres métodos las decisiones sobre si un bolsillo es o no drogable se han basado principalmente en dos propiedades del bolsillo: profundidad y tamaño (Kawabata & Go, 2007). La profundidad está relacionada con que tan enterrado está el bolsillo en relación con la superficie de la proteína, mientras que el tamaño infiere la apertura mínima que debe poseer una cavidad para permitir el ingreso del ligando (Laskowski et al., 1996). Los métodos basados en cuadrículas generan una superficie cuadriculada o grilla sobre la superficie 3D de la proteína y consideran que una cuadrícula vacía corresponde a un bolsillo, siempre y cuando dicho punto cumpla con determinadas propiedades geométricas o energéticas. En los métodos de esferas se recorre la superficie proteica con sondas esféricas de distinto tamaño, los resultados obtenidos son agrupados y posteriormente analizados para estimar posibles cavidades o sitios de unión. Los métodos de formas  $\alpha$  se basan en los subconjuntos de teselaciones de Delaunay, y se identifica un bolsillo cuando un tetraedro de Delaunay se encuentra vacío.

### 5.1.1 TryS: una enzima para estudiar

Como se mencionó en el capítulo 3 (sección 3.2.1) el número de compuestos moduladores informados para la enzima Tc**TryS** es sumamente escaso considerando el tiempo transcurrido desde que se propuso como posible blanco molecular de fármacos. Por otra parte, el número de inhibidores reportados para Tb**TryS** es bastante bajo en relación con la cantidad de compuestos que fueron testeados contra esta enzima. Esta observación, en conjunto con los resultados experimentales obtenidos luego de las dos campañas de cribado virtual utilizando aproximaciones idénticas a las que, para otros blancos moleculares, han demostrado ser exitosas en nuestro equipo de trabajo (Alberca et al., 2016, 2018, 2019; Bégamo et al., 2020) sugieren que la enzima **TryS** podría presentar dificultades para su modulación farmacológica: en otras palabras, que podría ser poco drogable.

La hipótesis planteada en el párrafo anterior fue el puntapié inicial para comenzar a desarrollar modelos basados en descriptores que permitieran predecir, *in silico*, la factibilidad de modular la enzima **TryS**. Pero que además los modelos generados resulten útiles, en el futuro, para ser aplicados frente a otras proteínas. En términos generales, los modelos desarrollados pueden clasificarse como modelos basados en aprendizaje automático.

## 5.2 Metodología

Algunas de las herramientas computacionales que existen actualmente permiten identificar cavidades o bolsillos expuestos sobre una proteína (Ghecom<sup>9</sup>, Cavity Plus<sup>10</sup>), mientras que otras van más allá y permiten predecir el grado de la drogabilidad de los diferentes bolsillos de manera individual (DoGSiteScorer<sup>11</sup>, COACH-D<sup>12</sup>, PuckDrug<sup>13</sup>). Sin embargo, en este trabajo desarrollamos modelos para clasificar a las proteínas en drogables o no drogables basados en la proteína en su conjunto.

### 5.2.1 Conformación del Set de datos

A partir de la base de datos de macromoléculas *Protein Data Bank*<sup>14</sup> (PDB), que compila información y la estructura 3D de diversas macromoléculas, mayormente proteínas, se conformó un set de proteínas para entrenar y validar nuestros modelos de drogabilidad. Las proteínas que integran dicho set se clasificaron en dos categorías: **Drogable** y **No Drogable**. Los criterios que se utilizaron para clasificar a las proteínas en cada categoría fueron los siguientes:

- ❖ Que se hayan reportado al menos 100 moduladores con  $IC_{50} < 20 \mu M$ .
- ❖ Que en campañas de cribado de alto rendimiento contra dicha proteína se hayan reportado al menos un 5% de *hits* respecto a los compuestos escrutados.
- ❖ Que se reporte un mínimo de cinco publicaciones por año reportando pequeñas moléculas moduladoras de su actividad (este criterio nos parece particularmente relevante para proteínas propuestas como blancos moleculares recientemente).

Una proteína se consideró **Drogable** si cumple con al menos dos de los tres criterios propuestos, de lo contrario se consideró **No Drogable**. Aquellas proteínas que por diversos

---

<sup>9</sup> <https://pdj.org/ghecom/>

<sup>10</sup> <http://www.pkumdl.cn:8000/cavityplus/index.php>

<sup>11</sup> <https://proteins.plus/#dogsite>

<sup>12</sup> <https://yanglab.nankai.edu.cn/COACH-D/>

<sup>13</sup> <http://pockdrug.rpbs.univ-paris-diderot.fr/cgi-bin/index.py?page=Druggability>

<sup>14</sup> <https://www.rcsb.org/>

motivos resultaban dudosas a la hora de ser clasificadas en una de las dos categorías fueron excluidas para esta primera etapa de desarrollo y validación de los modelos.

Adicionalmente se reforzaron ambas categorías (**Drogable** y **No Drogable**) con proteínas seleccionadas del directorio *Druggable Cavity Directory*<sup>15</sup> (DCD), una plataforma que compila proteínas analizadas en términos de drogabilidad e informa un *score* de drogabilidad y un valor de confianza para dicho *score*. En esta oportunidad se utilizó un *score* igual o mayor a 8 para considerar que la proteína es **Drogable**, y por debajo de 4 para considerar que la proteína es **No Drogable**; para los dos conjuntos se seleccionaron *scores* cuyo valor de confianza estaba por encima de 7 puntos, sugiriendo que la clase asignada por el sitio es confiable.

Finalmente, se incorporaron al conjunto **No Drogable** proteínas desordenadas, las cuales por su estructura desordenada son consideradas difíciles de modular con pequeñas moléculas tipo fármaco, ya que no cuentan con bolsillos definidos como en el caso de proteínas estructuralmente estables. Dichas proteínas se obtuvieron de la plataforma DisProt<sup>16</sup>, una página web que compila información y estructura 3D de proteínas desordenadas.

Para acondicionar cada una de las proteínas del Set se utilizó el programa PyMOL 2.3.2, y se aplicó el siguiente protocolo:

- ❖ Eliminación del solvente de cristalización.
- ❖ Eliminación de iones y ligando/s de cristalización.
- ❖ Eliminación de las secuencias o residuos de aminoácidos no relevantes debido a que algunas herramientas admiten un número limitado de aminoácidos.
- ❖ Todas las proteínas acondicionadas se guardaron en formato pdb.

### 5.2.2 Partición del set de datos

Para realizar una partición representativa del set de datos se calcularon 121 descriptores para cada proteína; dichos descriptores fueron obtenidos a partir de la plataforma 3D-surfer<sup>17</sup>. A este conjunto de descriptores se le aplicó un Análisis de Componentes Principales (PCA, del inglés, *Principal Component Analysis*), mediante el cual es posible obtener variables latentes ortogonales correspondientes a las direcciones de máxima variabilidad de los datos mediante combinación lineal de variables incluidas en el pool de descriptores. Así, la primera componente principal se orienta en la dirección de mayor variabilidad de los datos; la segunda componente se orienta en la segunda dirección de mayor variabilidad de los datos perpendicular a la primera componente, y así sucesivamente (Geladi & Linderholm, 2020). En este caso, utilizamos las componentes principales obtenidas de ese modo como variables para implementar metodologías de agrupamiento y, posteriormente, particionar representativamente el set de datos en conjuntos de entrenamiento y de

---

<sup>15</sup> <http://fpocket.sourceforge.net/>

<sup>16</sup> <https://disprot.org/>

<sup>17</sup> <https://kiharalab.org/3d-surfer>

validación. El procedimiento fue implementado en lenguaje Python utilizando la biblioteca Scikit-learn<sup>18</sup>.

Se aplicó la técnica *k-means* y se utilizó la semilla 123457 para que a partir de ella el algoritmo seleccione centroides de manera pseudo-aleatoria (y por lo tanto reproducible) y proponga diferentes agrupamientos para cada categoría. Para estos subgrupos se evaluó la métrica **Silhouette**, la cual permite estimar la calidad de los agrupamientos obtenidos en función de las distancias intra-grupo e inter-grupo para cada elemento agrupado (**ecuación 5.1**) (Dinh et al., 2019).

$$S(i) = \frac{b-a}{\max(a,b)} \quad \text{Ec. 5.1}$$

Donde  $S(i)$  es el valor que toma el coeficiente de Silhouette,  $a$  representa el promedio de distancias del elemento  $i$  con el resto de los elementos del mismo grupo, y  $b$  corresponde a la distancia mínima del elemento  $i$  con un elemento de otro grupo.

El valor  $S(i)$  se puede obtener a partir de la siguiente expresión:

$$S(i) = \begin{cases} 1 - \frac{a}{b}; & \text{Si } a < b \\ 0; & \text{Si } a = b \\ \frac{a}{b} - 1; & \text{Si } a > b \end{cases} \quad \text{Ec. 5.2}$$

Dado que  $-1 \leq S(i) \leq 1$ , se asume que el elemento  $i$  está bien clasificado en el grupo cuando  $S(i)$  es próximo a 1, mientras que debería ser removido al grupo vecino cuando dicho coeficiente se encuentra cercano a  $-1$ . Si  $S(i)=0$  entonces elemento se encuentra sobre la frontera de ambos grupos.

En esta oportunidad se analizó de manera sistemática cómo evolucionaba el coeficiente Silhouette variando el número de agrupamientos generados ( $k$ ) entre dos y diez subgrupos, definiendo el espacio químico mediante la primera componente principal (PC1) y la combinación de la primera y la segunda componente principal (PC1 y PC2). Ese procedimiento se implementó de manera independiente para los conjuntos de proteínas Drogables y No Drogables.

### 5.2.3 Conformación de los conjuntos de entrenamiento y validación

Para la conformación de los conjuntos se utilizó una partición estratificada por *cluster*, reservando el 70 % de cada categoría para el conjunto de entrenamiento, y el 30% restante para el conjunto de validación externa.

### 5.2.4 Obtención de descriptores

Se obtuvieron 147 descriptores de tipo CTD (*Composition, Transition and Distribution*) mediante el módulo PyProtein. El mismo se encuentra dentro del paquete PyBioMed para Python (Dong et al., 2018). Adicionalmente se utilizaron cinco herramientas *online* que

---

<sup>18</sup> [www.scikit-learn.org](http://www.scikit-learn.org)

predicen un *score* de drogabilidad o el volumen de la cavidad para cada bolsillo de la proteína, con el fin de generar descriptores útiles para el algoritmo en desarrollo se calcularon para cada una de las herramientas tres descriptores, los cuales fueron generados mediante el procesamiento manual de los resultados obtenidos a partir de cada plataforma. Los descriptores generados se describen a continuación.

**D1:** Cantidad de bolsillos drogables determinados por cada herramienta *i*.

**D2:** *Score* del mejor bolsillo drogable asignado por la herramienta *i*.

**D3:** Promedio del *score* asignado por la herramienta *i* a los tres mejores bolsillos drogables identificados por la misma.

De este modo, el total de descriptores obtenidos fue de 161, de los cuales 147 corresponden al tipo CTD y 14 fueron generados como se explicó en el párrafo precedente a partir de las cinco herramientas *on line*. Cabe mencionar que el descriptor **D1** no pudo ser calculado para los resultados provenientes de la plataforma FTMap, ya que dicha herramienta no indica de manera explícita la cantidad de bolsillos drogables identificados.

Las herramientas *online* utilizadas para el análisis de proteínas y para el cálculo de descriptores se listan y describen a continuación; también se informa, brevemente, la información de salida que ofrece cada herramienta:

**PockDrug**<sup>19</sup>: La plataforma predice el *score* de drogabilidad de cada bolsillo con valores que oscilan entre 0 y 1. Este algoritmo consta de una combinación de modelos lineales discriminantes, los cuales contemplan un total de 52 descriptores de bolsillos de tipo fisicoquímicos y geométricos. Para su desarrollo y validación se utilizó la base *non-redundant set of druggable and less druggable binding sites (NRDL)* (Krasowski et al., 2011) para generar los conjuntos de entrenamiento y de prueba. El resultado devuelto por la herramienta es una tabla en la que se observa la cantidad de bolsillos identificados, el número de aminoácidos que componen cada bolsillo y el *score* de drogabilidad predicho para cada bolsillo, como así también la fracción de residuos aromáticos y polares, entre otros resultados. Se considera drogable a los bolsillos con puntaje mayor a 0,5 y que involucran más de 14 residuos, mientras que todos los bolsillos con menos de 10 aminoácidos se consideran no drogables independientemente del *score* predicho (Hussein et al., 2015). Los bolsillos que comprenden entre 10 y 14 residuos deben ser evaluados individualmente. Adicionalmente se posibilita la descarga de un archivo tipo .pdb en el que se identifican todos bolsillos reconocidos por la herramienta.

**DoGSiteScorer**<sup>20</sup>: Esta herramienta se basa en métodos de aprendizaje automático. El método extiende una cuadrícula sobre la superficie 3D de la proteína y de esta manera identifica los posibles bolsillos presentes en la misma; una vez identificados dichos bolsillos el algoritmo predice un valor de drogabilidad para cada uno utilizando para ellos descriptores fisicoquímicos y propiedades geométricas. El valor de predicción se encuentra acotado entre 0 y 1 y se considera que un bolsillo es drogable si el mismo

---

<sup>19</sup> <http://pockdrug.rpbs.univ-paris-diderot.fr/cgi-bin/index.py?page=home>

<sup>20</sup> <https://proteins.plus/#dogsite>

obtiene un puntaje mayor a 0,5. El *output* del software consta de una tabla que permite observar la proporción de residuos no polares, el volumen y la profundidad del bolsillo evaluado. Estas tres características resultan relevantes para la interacción de las proteínas con pequeños ligandos (Volkamer, Kuhn, Rippmann, et al., 2012).

**COACH-D<sup>21</sup>**: La plataforma integra cinco métodos independientes para identificar y predecir bolsillos de unión en la proteína de interés. Cuatro de los métodos están basados en *templates* (proteínas con bolsillos drogables validados que se toman como referencia). El quinto método es una metodología que combina el análisis de secuencia, evaluando la conservación de esta, y propiedades geométricas de los sitios de unión. Finalmente, los resultados individuales son combinados bajo el algoritmo COACH (Wu et al., 2018; Yang et al., 2013).

**FTMap<sup>22</sup>**: Este servidor permite identificar sitios de unión mediante el mapeo computacional de macromoléculas (proteínas y ácidos nucleicos). Para ello el algoritmo comienza distribuyendo 16 sondas orgánicas (ligandos) sobre la superficie de la macromolécula, luego identifica las regiones en las que se unen mayor cantidad de grupos químicos diferentes de cada sonda. Estas regiones, donde se unen mayor cantidad de grupos químicos son identificadas como regiones o sitios de consenso. El sitio con mayor número de interacciones químicas se lo denomina punto crítico principal, mientras que el resto de los sitios de consenso se denominan puntos críticos secundarios. El análisis sobre la puntuación y posición de los grupos de las sondas se realiza mediante cálculos energéticos que contemplan fuerzas de van der Waals, interacciones electrostáticas de Poission-Boltzman, y potencial de interacción por pares basados en estructuras. Una vez finalizado el cálculo energético el algoritmo minimiza las energías de las poses de los grupos funcionales de las sondas, de manera tal que optimiza las interacciones registradas. Adicionalmente los autores proponen que un blanco molecular resulta drogable si el mismo presenta un punto crítico (sitio de unión) principal que interacciona con al menos 16 grupos funcionales diferentes y un punto crítico secundario con no menos de 13 grupos a una distancia menor a 8 Å del bolsillo principal (Kozakov et al., 2015).

**GHECOM<sup>23</sup>**: El algoritmo permite identificar sitios de unión utilizando sondas esféricas de distintos tamaños para trazar la forma y profundidad del bolsillo. Si bien la plataforma no ofrece un *score* de drogabilidad, la información brindada permite, mediante un pequeño análisis de los datos, identificar diferentes sitios de unión para ligandos (Kawabata, 2010). En este trabajo analizamos los volúmenes de los diferentes bolsillos, considerando no drogables aquellos bolsillos cuyo volumen resulta inferior a 100 Å<sup>3</sup>.

---

<sup>21</sup> <https://yanglab.nankai.edu.cn/COACH-D/>

<sup>22</sup> <https://ftmap.bu.edu/login.php>

<sup>23</sup> <https://pd bj.org/ghecom/>

### 5.2.5 Generación de modelos

Se procedió a la generación de modelos computacionales empleando dos estrategias: una de ellas utilizando solo los 14 descriptores obtenidos a partir de las herramientas *online* y realizando la búsqueda exhaustiva de todas las combinaciones de descriptores. El procedimiento se llevó a cabo en el entorno Python, y se incluyeron en los modelos desde un mínimo de dos hasta un máximo de cinco descriptores.

En la segunda estrategia se realizó una combinación de la metodología de subespacios aleatorizados (Daho & Chikh, 2015; Vyškovský et al., 2016) con la técnica de pasos hacia adelante (Celeux et al., 2019). Esta conjunción de metodologías se aplicó al conjunto de descriptores en el que se combinaron los 161 descriptores totales (147 CDT + 14 de herramientas *online*). Los subespacios aleatorizados contenían un máximo de 20 descriptores, y se limitó en cinco el número máximo de variables que podía incorporar cada modelo. Además, se estableció un valor de tolerancia  $\beta=0,5$  para evitar la incorporación de descriptores con alto grado de correlación (Daoud, 2017).

### 5.2.6. Validación de los modelos obtenidos

Se aplicaron diferentes técnicas de validación estándar con el objetivo de evaluar la robustez (presencia/ausencia de sobreajuste) y la capacidad predictiva de los modelos.

#### 5.2.6.1 Validación interna

Las técnicas contempladas dentro de las metodologías de validación interna utilizan al propio conjunto de entrenamiento para inferir la capacidad predictiva del modelo y su robustez, descartando por otro lado correlaciones espurias. La mayoría de estas técnicas fueron oportunamente presentadas en el capítulo 3.

##### 5.2.6.1.2 Validación cruzada

Esta técnica permite evaluar si los modelos presentan algún grado de sobreajuste, es decir, si existe algún sesgo hacia una región del espacio químico o si existen casos de entrenamientos excesivamente influyentes (Ravichandran Veerasamy et al., 2011). La metodología consiste en rehacer los modelos clasificatorios a partir de la remoción de una molécula (*LOO*) o de pequeños subgrupos de moléculas (*LGO*) del conjunto de entrenamiento, y posteriormente evaluar el comportamiento de los nuevos modelos obtenidos a partir del conjunto de entrenamiento reducido, frente al subgrupo removido (Roy & Ambure, 2016; Varoquaux et al., 2017).

Aquí se aplicó la técnica de remoción en subgrupo, de manera estratificada y aleatorizada, removiendo el 10% del total del conjunto de entrenamiento (es decir, 16 proteínas) en cada ronda de validación. El procedimiento se realizó 1000 veces, comprobando que todas las estructuras hubieran sido removidas al menos una vez.

### 5.2.6.1.3 Validación de Fisher

La *Y-randomization* o aleatorización de Fisher consiste en aleatorizar los valores de las variables dependientes en el conjunto de entrenamiento, cancelando por tanto cualquier relación que pudiera existir entre estructura y propiedad, y generar nuevos modelos para evaluar si existe probabilidad de correlación azarosa entre las variables independientes y la variable dependiente, (Ravichandran Veerasamy et al., 2011); en el presente trabajo algunas proteínas Drogables fueron etiquetadas, aleatoriamente, como No Drogables y viceversa. Dado este carácter de asignación de categorías librada al azar, se espera que los modelos obtenidos luego de este procedimiento presenten valores de métricas sustancialmente inferiores a las de los modelos originales.

En esta oportunidad se realizaron 1000 rondas de asignación aleatoria de clases, y para cada ronda se generaron nuevos modelos aplicando los mismos procedimientos que se utilizaron para la generación de los modelos originales.

### 5.2.6.4 Bootstrapping

La técnica de *Bootstrapping* es otra alternativa de validación que se basa en la remoción aleatoria y reemplazo de los elementos o valores de un conjunto. Esta metodología es útil especialmente para conjuntos pequeños o con distribución de datos anormal o desconocida (Egbert & Plonsky, 2020). Para la aplicación de esta metodología se debe prever que el conjunto de elementos sea representativo del espacio muestral (Gramatica, 2013). En una validación clásica de *Bootstrapping* se realiza la remoción al azar de  $K$  grupos compuestos por  $n$  elementos, en el conjunto de entrenamiento y se reemplazan los elementos removidos duplicando al azar algunos de los no removidos, de manera tal que la cantidad de elementos totales sea igual a la cantidad de elementos que conformaban el conjunto original. A continuación, se realizan modelos estadísticos, se evalúa el desempeño de estos sobre el subconjunto de elementos removidos y se calcula la media del desempeño de todas las rondas de validación. Se recomienda que esta remoción se reitere al menos 1000 veces (Egbert & Plonsky, 2020; Gramatica, 2013). La principal diferencia de esta técnica con la metodología de validación cruzada es que en este caso la cantidad de elementos del conjunto de entrenamiento es siempre la misma, mientras que en la validación cruzada el número de elementos del conjunto es  $N - n$ .

En la presente campaña se realizaron 10000 rondas, removiendo subgrupos de 32 proteínas en cada ronda de manera estratificada, es decir, tomando al azar 16 proteínas de cada categoría para conformar cada subgrupo.

### 5.2.6.5 Validación externa

Dado que un modelo computacional resultará útil solo si permite predecir la clasificación de casos (en este caso, proteínas) no incluidas en el conjunto de entrenamiento, diversos autores sostienen que la manera más rigurosa de evaluar un modelo clasificatorio es desafiarlo a un conjunto de validación externa (Gramatica, 2013; Ravichandran Veerasamy et al., 2011).

En esta oportunidad se evaluó la capacidad predictiva de los modelos en el conjunto de validación externa y se estableció como punto de corte un valor de 0,5 para discernir entre proteína Drogable y No Drogable. Dicho *score* surge del promedio de valores asignados a cada categoría, donde a las proteínas Drogables se le asignó un valor de 1, mientras que para las No Drogables se asignó 0 (Toropova & Toropov, 2017).

Para examinar el desempeño de los modelos se ha calculado la Exactitud (**Acc**, del inglés *accuracy*) tanto para los modelos originales como para los modelos que se obtuvieron mediante las distintas técnicas de validación interna. Adicionalmente frente al conjunto externo se evaluaron los parámetros de recuperación y precisión (**Pr**) definido por la **ecuación 5.3**.

$$Pr = \frac{VP}{VP+FP} \quad \text{Ec 5.3}$$

### 5.2.7 Dominio de aplicación

Existen varios métodos que permiten establecer si un caso pertenece al DA de los modelos y, por ende, si la predicción que el modelo hace de tal caso es confiable. En el presente trabajo se eligió la regla de la palanca. La técnica consiste en obtener un valor denominado  $h_i$  (**ecuación 5.4**) para cada elemento analizado (proteína en este caso) y compararlo con el valor crítico conocido como  $h^*$  (**ecuación 5.5**). En consecuencia, si  $h_i < h^*$ , entonces se considera que el elemento pertenece un espacio químico donde los modelos resultarán útiles, mientras que si  $h_i > h^*$  el elemento está fuera del alcance de los modelos (Eslamimanesh et al., 2012; Roy et al., 2015).

$$h_i = X(X^T X)^{-1} * X^T \quad \text{Ec. 5.4}$$

$$h^* = \frac{3k}{n} \quad \text{Ec. 5.5}$$

## 5.3 Resultados

### 5.3.1 Conformación y partición del set de datos

Se compiló un set de 222 proteínas que contaban con su correspondiente estructura tridimensional experimental. La estructura de cada proteína se descargó en formato .pdb a partir de la plataforma Protein Data Bank<sup>6</sup>.

Luego de aplicar los criterios de drogabilidad establecidos, el set de proteínas se dividió en dos conjuntos con 111 proteínas de cada uno. Tal como se describió en la metodología, se evaluó la variación del coeficiente Silhouette en función del número de agrupamientos generados ( $k$ ), generándose así dos grupos definidos por PC1 para el conjunto de proteínas Drogables, y tres agrupamientos definidos por la combinación de las componentes PC1 y PC2 para el conjunto de proteínas no drogables. A continuación, se presenta la **tabla 5.2**, donde se aprecia la conformación final para el conjunto de entrenamiento y el conjunto de validación.

Tabla 5.2: Conformación final de los conjuntos

| Clase          | Conjunto de entrenamiento | Conjunto de validación |
|----------------|---------------------------|------------------------|
| Drogables      | 78                        | 33                     |
| No Drogables   | 78                        | 33                     |
| <b>Totales</b> | <b>156</b>                | <b>66</b>              |

### 5.3.2 Generación y validación de los modelos

#### 5.3.2.1 Modelos basados en herramientas online

A partir de los 14 descriptores obtenidos mediante las herramientas *online* se generaron, mediante búsqueda exhaustiva, 3458 modelos diferentes. La cantidad de modelos obtenidos según la cantidad de descriptores incorporados se detalla en la **tabla 5.3**

Tabla 5.3: Cantidad de modelos según los descriptores involucrados

| Número de descriptores involucrados | Cantidad de modelos generados |
|-------------------------------------|-------------------------------|
| 2                                   | 92                            |
| 3                                   | 364                           |
| 4                                   | 1001                          |
| 5                                   | 2002                          |
| <b>Total</b>                        | <b>3458</b>                   |

A continuación, se muestra la ecuación del mejor modelo individual.

#### Modelo 1226

$$Fd = -0,170 + 0,475 * D3_1 + 0,035 * D1_2 + 0,537 * D2_2 + 0,022 * D2_4 - 0,036 * D3_4$$

Como se puede observar de la ecuación, el mejor modelo individual presenta el número máximo de descriptores permitidos, y en consecuencia el modelo podría presentar algún grado de sobreajuste frente a las evaluaciones de las métricas de validación interna. En la **tabla 5.4** se presenta una breve interpretación de cada descriptor involucrado en el modelo 1226.

**Tabla 5.4.** Descripción breve de los descriptores involucrados en el modelo 1226  
**Información / Interpretación**

|                       |  |
|-----------------------|--|
| <b>D3<sub>1</sub></b> | Promedio de <i>score</i> de los tres mejores bolsillos según PockDrug      |
| <b>D1<sub>2</sub></b> | Cantidad de bolsillos drogables según DoGSiteScore                         |
| <b>D2<sub>2</sub></b> | <i>Score</i> del mejor bolsillo drogable según DogSiteScore                |
| <b>D2<sub>4</sub></b> | Máxima cantidad de ligandos en el bolsillo más drogable según FTMap        |
| <b>D3<sub>4</sub></b> | Promedio de ligandos reconocidos en los tres mejores bolsillos según FTMap |

### 5.3.2.2 Modelos Basados en subespacios y pasos hacia adelante

Se generaron 1000 modelos computacionales con el conjunto de entrenamiento que contemplaba 161 descriptores (147 CTD + 14 herramientas *online*).

Seguidamente se describe la ecuación del mejor modelo individual para esta estrategia.

#### Modelo 778

$$Fd = 1,865 + 0,327 * D2_3 - 0,0405 * NormalizedVDWVD3075 + 0,0194 * ChargeD1050 + 0,416 * D2_1 + 0,169 * D1_3$$

Como puede apreciarse la ecuación del mejor modelo, el modelo 778, se obtiene mediante la combinación de descriptores provenientes de diferentes tipos de cálculos. A continuación, en la **tabla 5.5**, se presenta la información que brinda cada uno de los descriptores incorporados en el modelo 778.

**Tabla 5.5** Información de cada descriptor

|                           | <b>Información / interpretación</b>  |
|---------------------------|--|
| <b>D2<sub>3</sub></b>     | <i>Score</i> del mejor bolsillo drogable según COACH-D.                                |
| <b>NormalizedVDWD3075</b> | Radio de vdW normalizado para cada aa del grupo 3 presentes en el 75% de la secuencia. |
| <b>ChargeD1050</b>        | Carga (positiva) en los aa del grupo 1 presentes en el 50% de la secuencia.            |
| <b>D2<sub>1</sub></b>     | <i>Score</i> del mejor bolsillo drogable según PockDrug.                               |
| <b>D1<sub>3</sub></b>     | Cantidad de bolsillos drogables según COACH-D.   |

En la **tabla 5.6** se informan los resultados obtenidos para cada modelo según la metodología empleada. Los valores corresponden a la evaluación del modelo frente al conjunto de entrenamiento y al conjunto de validación. Adicionalmente se presenta en la **figura 5.2** la frecuencia de aparición de los valores obtenidos para cada métrica cuando se aplicó la técnica *bootstrapping* frente al conjunto de validación.

Tabla 5.6 Resultados de cada modelo frente a los respectivos conjuntos

| Modelo            | Conjunto de entrenamiento |              |           | Conjunto de validación |              |           |
|-------------------|---------------------------|--------------|-----------|------------------------|--------------|-----------|
|                   | Exactitud                 | Recuperación | Precisión | Exactitud              | Recuperación | Precisión |
| 1226 <sup>a</sup> | 0,718                     | 0,693        | 0,788     | 0,818                  | 0,818        | 0,818     |
| 778 <sup>b</sup>  | 0,872                     | 0,846        | 0,892     | 0,871                  | 0,848        | 0,800     |

<sup>a</sup> modelo que contempla solo descriptores de herramientas *online*. <sup>b</sup> modelo correspondiente a la combinación de descriptores. p-valor < 0,05.

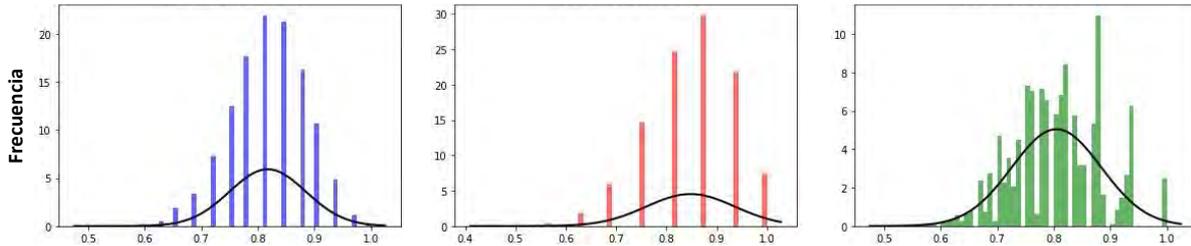


Figura 5.2. Frecuencia de aparición de los valores de cada métrica cuando se aplicó *bootstrapping* en el modelo 778 frente al conjunto de validación.

### 5.3.2.3 Validación de los modelos

Las estrategias de validación aplicadas fueron las mismas independientemente de la metodología utilizada para generar los modelos (búsqueda exhaustiva o subespacios aleatorizados)

En la **tabla 5.7** pueden observarse los resultados obtenidos para algunos de los modelos originales que reportaron mejores valores de exactitud, recuperación y precisión, como así también los resultados promedios de *Acc*, con su respectiva desviación estándar, obtenidos para cada una de las técnicas de validación interna que se aplicó.

Tabla 5.7 Valores de exactitud (*Acc*) para los diferentes modelos en el conjunto de entrenamiento y de validación

| Modelo            | <i>Acc</i> original | <i>Acc</i> Validación de Fisher | <i>Acc</i> Validación Cruzada |
|-------------------|---------------------|---------------------------------|-------------------------------|
| 1226 <sup>a</sup> | 0,782               | 0,500 (0,040)                   | 0,538 (0,092)                 |
| 778 <sup>b</sup>  | 0,872               | 0,503 (0,143)                   | 0,850 (0,086)                 |
| 457 <sup>b</sup>  | 0,852               | 0,499 (0,143)                   | 0,786 (0,109)                 |
| 309 <sup>b</sup>  | 0,846               | 0,500 (0,129)                   | 0,801 (0,103)                 |

<sup>a</sup> modelos que contemplan solo descriptores de herramientas *online*. <sup>b</sup> modelos correspondientes a la combinación de descriptores.

La validación de Fisher fue superada exitosamente por la mayoría de los modelos generados independientemente del conjunto de descriptores empleados. Esto se desprende de comparar los valores de *Acc* obtenidos para dicha técnica con los resultados que se

observan para los modelos originales. Se puede apreciar que en todos los casos el valor de *Acc* del modelo original supera ampliamente el valor obtenido mediante la aleatorización de Fisher (que, por otro modo, en promedio, corresponde a 0,5; el valor esperado para una clasificación aleatoria de conjuntos balanceados) y esto permite descartar que los protocolos de modelado implementados conduzcan a correlaciones espurias.

Sin embargo, también puede observarse que el modelo generado a partir de los descriptores obtenidos mediante las herramientas *online* (modelo 1226) no superó con éxito la etapa de validación cruzada, ya que la notoria disminución del valor de *Acc* para esta estrategia de validación, respecto al valor de *Acc* del modelo original, nos indica que el mismo presenta un fuerte sobreajuste. Es decir, al remover algunas proteínas del conjunto de entrenamiento, el modelo es incapaz de mantener su capacidad predictiva. Si bien aquí se presenta solo el modelo 1226, es interesante destacar que este comportamiento se observó en la mayoría de los modelos generados a partir de descriptores obtenidos exclusivamente de aplicaciones *online*.

Para los modelos generados a partir de la combinación de descriptores, se puede observar que el desempeño de los modelos se mantiene estable en la etapa de validación cruzada, ya que al remover diversos subgrupos de proteínas del conjunto de entrenamiento, los modelos son capaces de mantener aproximadamente estable su valor de *Acc*, es decir, que los subgrupos removidos no presentan características excesivamente influyentes frente al resto del conjunto; los modelos, en este caso, parecen reflejar características generales relacionadas con la propiedad modelada, y no características particulares (“accidentes” no generalizables) del conjunto de entrenamiento. En conclusión, los modelos obtenidos del pool total de descriptores considerados exhiben un comportamiento más robusto.

### 5.3.2.4 Bootstrapping

Dado que se desconoce la distribución de los datos, y que tampoco es posible asumir una distribución normal, se ha utilizado la técnica *bootstrapping* para estimar el intervalo de confianza de los valores de exactitud para los modelos que superaron la etapa de validación cruzada. En la **tabla 5.8** se indica para cada modelo el valor promedio de *Acc* y los límites del intervalo de confianza al 95% (IC<sub>95</sub>).

**Tabla 5.8.** Valores de la media y límites del intervalo de confianza para cada métrica

| Modelo                 | Exactitud |        |        | Recuperación |        |        | Precisión |        |        |
|------------------------|-----------|--------|--------|--------------|--------|--------|-----------|--------|--------|
|                        | Media     | L. inf | L. sup | Media        | L. inf | L. sup | Media     | L. inf | L. sup |
| <b>778<sup>b</sup></b> | 0,872     | 0,871  | 0,873  | 0,846        | 0,844  | 0,848  | 0,897     | 0,896  | 0,898  |
| <b>457<sup>b</sup></b> | 0,852     | 0,851  | 0,853  | 0,871        | 0,869  | 0,872  | 0,845     | 0,843  | 0,846  |
| <b>309<sup>b</sup></b> | 0,846     | 0,845  | 0,848  | 0,923        | 0,922  | 0,924  | 0,806     | 0,805  | 0,808  |

<sup>b</sup> Modelos correspondientes a la combinación de descriptores.

Los valores de los límites observados para las diferentes métricas permiten inferir que los modelos son robustos, dado que dichos límites están próximos a la media. Adicionalmente en todos los casos la desviación estándar (no informada aquí) se mantiene por debajo de 0,1; más aún, en el 75 % de modelos este parámetro está por debajo de 0,07.

### 5.3.2.5 Validación Externa

La capacidad predictiva de los modelos se estimó frente al conjunto de validación, evaluando la capacidad predictiva global y la tasa de recuperación de proteínas drogables para los modelos que superaron las instancias de validación interna.

El conjunto de validación contiene 66 proteínas distribuidas equitativamente entre las categorías Drogable (33) y No Drogable (33). En la **tabla 5.9** se presentan los datos recabados frente al conjunto de validación.

**Tabla 5.9.** Métricas del modelo 788 frente al conjunto de validación  
**Conjunto de validación**

| Modelo                 | Exactitud | Recuperación | Precisión |
|------------------------|-----------|--------------|-----------|
| <b>778<sup>b</sup></b> | 0,871     | 0,848        | 0,800     |

<sup>b</sup> Modelo correspondientes a descriptores a la combinación de descriptores.

Como puede observarse de la **tabla 5.9** el modelo ha mantenido, frente al conjunto de validación, valores similares a los valores de métricas obtenidos frente al conjunto de entrenamiento. Esto sugiere que el modelo es, además de robusto, útil para predecir o estimar el grado de drogabilidad de una proteína externa al conjunto de entrenamiento: el modelo ha capturado una relación generalizable entre la propiedad modelada y la estructura proteica.

### 5.3.3 Evaluación y resultados frente a la enzima TryS

El mejor modelo para el análisis de drogabilidad fue aplicado en el marco de la presente tesis. El modelo denominado 778 se utilizó para evaluar el grado de drogabilidad de la enzima **TryS** de *Leishmania major* (Lm**TryS**) (PDB Id: 2VOB) depositada en la plataforma PDB desde el año 2008 (<https://www.rcsb.org/>). Dado que varias de las herramientas *online* aquí utilizadas para el cálculo de descriptores requieren la estructura 3D de la proteína, no se ha podido estudiar el grado de drogabilidad de **TryS** para otras especies de tripanosomátidos, ya que no se cuenta hasta el momento con la estructura 3D de estas proteínas.

El modelo 778 arrojó un *score* de drogabilidad para la Lm**TryS** de 0,370, esto indica que, según el modelo, la proteína es No Drogable o difícil de modular farmacológicamente. Adicionalmente se evaluó el DA, y se obtuvo que  $h_i < h^*$ , por lo tanto la predicción realizada por este modelo se considera válida para la proteína.

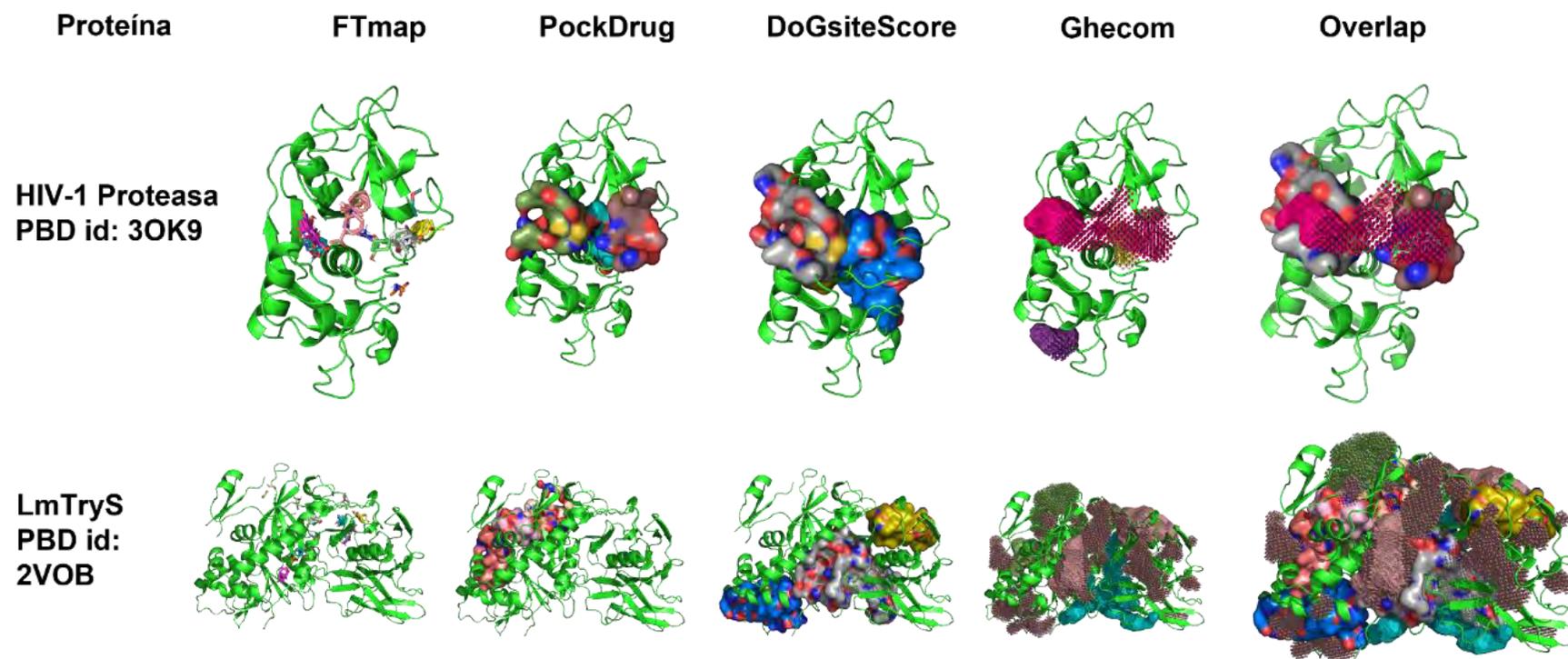
### 5.3.4. Análisis visual complementario frente a Lm TryS

Aprovechando los archivos pdb que ofrecen como salida las diferentes herramientas *online* se realizó un análisis visual de un grupo de catorce proteínas, entre las que se encuentran proteínas pertenecientes a las categorías Drogable y No Drogable. Posteriormente se analizó de igual manera la enzima Lm**TryS** (PDB Id: 2VOB).

A partir de los archivos pdb generados por las herramientas *online* PockDrug, DoGSiteScore, FTMap y Ghecom, se visualizó, mediante el software PyMOL 2.3.2, la ubicación

de tres bolsillos predichos por cada una de las herramientas, considerando para PockDrug y DoGSiteScore los bolsillos con mayores *scores* de drogabilidad, para FTMap se exhibieron los tres bolsillos con mayor cantidad de ligando, mientras que para la herramienta Ghecom se seleccionaron los bolsillos de mayor volúmenes. Finalmente se generó la superposición de los mejores tres bolsillos con el fin de observar el grado de consenso o no entre las diferentes herramientas.

A modo comparativo se presenta el análisis realizado para Lm **TryS** junto con una proteína altamente drogable como es la HIV-1 Proteasa (PDB Id: 3OK9). Como se desprende de la **figura 5.3**, la superposición de las diferentes herramientas muestra que existe un alto grado de consenso frente a la identificación de los mejores bolsillos para una proteína drogable; sin embargo, el análisis visual realizado frente a la enzima Lm **TryS** revela que las herramientas no presentan consenso sobre cuáles son los bolsillos drogables, tal como ocurre en el caso de la HIV-1 Proteasa. Este análisis parece sugerir, como nuestro mejor modelo, que la Lm **TryS** es una enzima difícil de modular farmacológicamente, y que dada la similitud de secuencia, estructura y función que presenta con Tc**TryS** podría asumirse un escenario similar para esta última.



**Figura 5.3.** Predicción de bolsillos de unión por diferentes herramientas de detección de bolsillos de unión y drogabilidad para la HIV 1-Proteasa (superior) y para LmTryS (inferior). La superposición de todos los bolsillos putativos se muestra a la derecha, lo que revela un alto consenso en las predicciones para la proteasa del VIH y un bajo consenso en el caso de la *Tripanotión sintetasa*.

### 5.4 Discusión de resultados

En el presente capítulo desarrollamos y validamos modelos que permiten establecer, a priori, una escala de drogabilidad para diversas proteínas.

Mediante búsqueda exhaustiva, se obtuvieron 3458 modelos utilizando descriptores provenientes solo de herramientas *online*; sin embargo, este conjunto de modelos no superó con éxito las etapas de validación interna, y en consecuencia los mismos no fueron aplicados para el análisis prospectivo de proteínas externas al conjunto de entrenamiento. Adicionalmente, y mediante la metodología de subespacios aleatorios combinada con selección de variables por pasos hacia adelante se generaron 1000 modelos clasificatorios combinando descriptores obtenidos por diversas aproximaciones; estos modelos sí superaron exitosamente todas las etapas de validación, incluyendo validación cruzada, aleatorización de Fisher, *bootstrapping* y validación externa. Por tanto, se puede asumir que los modelos provenientes a partir de la combinación de descriptores presentan mejor desempeño y robustez que los modelos generados únicamente a partir de descriptores obtenidos de herramientas *online*.

La metodología aquí desarrollada permitió establecer el grado de drogabilidad y clarificar el panorama actual sobre la enzima **TryS**, un blanco molecular interesante desde el punto de vista del análisis metabólico, pero que presenta, y así lo sugieren nuestros modelos y los métodos de predicción *online*, una alta resistencia a la modulación farmacológica. El comportamiento predicho podría explicar el bajo número de inhibidores reportados a la fecha y el mal desempeño en nuestras campañas de cribado *in silico* con validación experimental. Es posible aventurar, a manera de hipótesis, que el control de flujo que la **TryS** ejerce sobre la vía de biosíntesis de tripanotión podría explicar por qué, evolutivamente, esta proteína se ha mantenido difícil de modular.

Es importante mencionar que tanto el modelo desarrollado, como el análisis mediante superposición de bolsillos presentaron consenso en los resultados obtenidos; este hecho sugiere el acierto del algoritmo desarrollado, lo que nos alienta a seguir ampliando el conjunto de entrenamiento y el conjunto de validación con el fin de refinar el algoritmo obtenido.

Como perspectiva esta metodología permite seleccionar de manera rápida aquellas macromoléculas que presenten un *score* de drogabilidad mayor a 0,5, y en consecuencia discernir rápidamente qué macromoléculas poseen mayor probabilidad de éxitos frente a la búsqueda y descubrimiento de nuevas terapias farmacológicas.

**5.5 Referencias Bibliográficas**

- Agoni, C., Olotu, F. A., Ramharack, P., & Soliman, M. E. (2020). Druggability and drug-likeness concepts in drug design: Are biomodelling and predictive tools having their say? *Journal of Molecular Modeling*, *26*(6), 120. <https://doi.org/10.1007/s00894-020-04385-6>
- Alberca, L. N., Chuguransky, S. R., Álvarez, C. L., Talevi, A., & Salas-Sarduy, E. (2019). In silico Guided Drug Repurposing: Discovery of New Competitive and Non-competitive Inhibitors of Falcipain-2. *Frontiers in Chemistry*, *7*, 534. <https://doi.org/10.3389/fchem.2019.00534>
- Alberca, L. N., Sbaraglini, M. L., Balcazar, D., Fraccaroli, L., Carrillo, C., Medeiros, A., Benitez, D., Comini, M., & Talevi, A. (2016). Discovery of novel polyamine analogs with anti-protozoal activity by computer guided drug repositioning. *Journal of Computer-Aided Molecular Design*, *30*(4), 305-321. <https://doi.org/10.1007/s10822-016-9903-6>
- Alberca, L. N., Sbaraglini, M. L., Morales, J. F., Dietrich, R., Ruiz, M. D., Pino Martínez, A. M., Miranda, C. G., Fraccaroli, L., Alba Soto, C. D., Carrillo, C., Palestro, P. H., & Talevi, A. (2018). Cascade Ligand- and Structure-Based Virtual Screening to Identify New Trypanocidal Compounds Inhibiting Putrescine Uptake. *Frontiers in Cellular and Infection Microbiology*, *8*, 173. <https://doi.org/10.3389/fcimb.2018.00173>
- Barril, X. (2013). Druggability predictions: Methods, limitations, and applications. *WIREs Computational Molecular Science*, *3*(4), 327-338. <https://doi.org/10.1002/wcms.1134>
- Bélgamo, J. A., Alberca, L. N., Pórfido, J. L., Romero, F. N. C., Rodriguez, S., Talevi, A., Córscico, B., & Franchini, G. R. (2020). Application of target repositioning and in silico screening to exploit fatty acid binding proteins (FABPs) from *Echinococcus multilocularis* as possible drug targets. *Journal of Computer-Aided Molecular Design*, *34*(12), 1275-1288. <https://doi.org/10.1007/s10822-020-00352-8>
- Clark, J. J., Orban, Z. J., & Carlson, H. A. (2020). Predicting binding sites from unbound versus bound protein structures. *Scientific Reports*, *10*(1), 15856. <https://doi.org/10.1038/s41598-020-72906-7>
- Kawabata, T., & Go, N. (2007). Detection of pockets on protein surfaces using small and large probe spheres to find putative ligand binding sites. *Proteins: Structure, Function, and Bioinformatics*, *68*(2), 516-529. <https://doi.org/10.1002/prot.21283>
- Krasowski, A., Muthas, D., Sarkar, A., Schmitt, S., & Brenk, R. (2011). DrugPred: A Structure-Based Approach to Predict Protein Druggability Developed Using an Extensive Nonredundant Data Set. *Journal of Chemical Information and Modeling*, *51*(11), 2829-2842. <https://doi.org/10.1021/ci200266d>
- Laskowski, R. A., Luscombe, N. M., Swindells, M. B., & Thornton, J. M. (1996). Protein clefts in molecular recognition and function. *Protein Science: A Publication of the Protein Society*, *5*(12), 2438-2452.
- Volkamer, A., Kuhn, D., Grombacher, T., Rippmann, F., & Rarey, M. (2012). Combining Global and Local Measures for Structure-Based Druggability Predictions. *Journal of Chemical Information and Modeling*, *52*(2), 360-372. <https://doi.org/10.1021/ci200454v>

- Wu, Q., Peng, Z., Zhang, Y., & Yang, J. (2018). COACH-D: Improved protein–ligand binding sites prediction with refined ligand-binding poses through molecular docking. *Nucleic Acids Research*, *46*(W1), W438-W442. <https://doi.org/10.1093/nar/gky439>
- Yang, J., Roy, A., & Zhang, Y. (2013). Protein–ligand binding site recognition using complementary binding-specific substructure comparison and sequence profile alignment. *Bioinformatics*, *29*(20), 2588-2595. <https://doi.org/10.1093/bioinformatics/btt447>
- Celeux, G., Maugis-Rabusseau, C., & Sedki, M. (2019). Variable selection in model-based clustering and discriminant analysis with a regularization approach. *Advances in Data Analysis and Classification*, *13*(1), 259-278. <https://doi.org/10.1007/s11634-018-0322-5>
- Daho, M. E. H., & Chikh, M. A. (2015). Combining Bootstrapping Samples, Random Subspaces and Random Forests to Build Classifiers. *Journal of Medical Imaging and Health Informatics*, *5*(3), 539-544. <https://doi.org/10.1166/jmihi.2015.1423>
- Daoud, J. I. (2017). Multicollinearity and Regression Analysis. *Journal of Physics: Conference Series*, *949*, 012009. <https://doi.org/10.1088/1742-6596/949/1/012009>
- Dinh, D.-T., Fujinami, T., & Huynh, V.-N. (2019). Estimating the Optimal Number of Clusters in Categorical Data Clustering by Silhouette Coefficient. En J. Chen, V. N. Huynh, G.-N. Nguyen, & X. Tang (Eds.), *Knowledge and Systems Sciences* (pp. 1-17). Springer. [https://doi.org/10.1007/978-981-15-1209-4\\_1](https://doi.org/10.1007/978-981-15-1209-4_1)
- Dong, J., Yao, Z.-J., Zhang, L., Luo, F., Lin, Q., Lu, A.-P., Chen, A. F., & Cao, D.-S. (2018). PyBioMed: A python library for various molecular representations of chemicals, proteins and DNAs and their interactions. *Journal of Cheminformatics*, *10*(1), 16. <https://doi.org/10.1186/s13321-018-0270-2>
- Egbert, J., & Plonsky, L. (2020). Bootstrapping Techniques. En M. Paquot & S. Th. Gries (Eds.), *A Practical Handbook of Corpus Linguistics* (pp. 593-610). Springer International Publishing. [https://doi.org/10.1007/978-3-030-46216-1\\_24](https://doi.org/10.1007/978-3-030-46216-1_24)
- Eslamimanesh, A., Gharagheizi, F., Mohammadi, A. H., & Richon, D. (2012). A statistical method for evaluation of the experimental phase equilibrium data of simple clathrate hydrates. *Chemical Engineering Science*, *80*, 402-408. <https://doi.org/10.1016/j.ces.2012.06.019>
- Geladi, P., & Linderholm, J. (2020). Principal Component Analysis. En *Comprehensive Chemometrics* (pp. 17-37). Elsevier. <https://doi.org/10.1016/B978-0-12-409547-2.14892-9>
- Gramatica, P. (2013). On the Development and Validation of QSAR Models. En B. Reisfeld & A. N. Mayeno (Eds.), *Computational Toxicology* (Vol. 930, pp. 499-526). Humana Press. [https://doi.org/10.1007/978-1-62703-059-5\\_21](https://doi.org/10.1007/978-1-62703-059-5_21)
- Hajduk, P. J., Huth, J. R., & Tse, C. (2005). Predicting protein druggability. *Drug Discovery Today*, *10*(23-24), 1675-1682. [https://doi.org/10.1016/S1359-6446\(05\)03624-X](https://doi.org/10.1016/S1359-6446(05)03624-X)
- Hopkins, A. L., & Groom, C. R. (2003). Target Analysis: A Priori Assessment of Druggability. En H. Waldmann & M. Koppitz (Eds.), *Small Molecule—Protein Interactions* (pp. 11-17). Springer. [https://doi.org/10.1007/978-3-662-05314-0\\_2](https://doi.org/10.1007/978-3-662-05314-0_2)

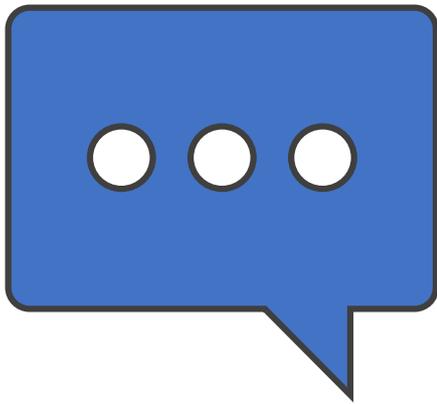
- Hussein, H. A., Borrel, A., Geneix, C., Petitjean, M., Regad, L., & Camproux, A.-C. (2015). PockDrug-Server: A new web server for predicting pocket druggability on holo and apo proteins. *Nucleic Acids Research*, *43*(W1), W436-W442. <https://doi.org/10.1093/nar/gkv462>
- Kawabata, T. (2010). Detection of multiscale pockets on protein surfaces using mathematical morphology. *Proteins: Structure, Function, and Bioinformatics*, *78*(5), 1195-1211. <https://doi.org/10.1002/prot.22639>
- Kozakov, D., Grove, L. E., Hall, D. R., Bohnuud, T., Mottarella, S., Luo, L., Xia, B., Beglov, D., & Vajda, S. (2015). The FTMap family of web servers for determining and characterizing ligand binding hot spots of proteins. *Nature protocols*, *10*(5), 733-755. <https://doi.org/10.1038/nprot.2015.043>
- Ravichandran Veerasamy, Rajak, H., Jain, A., Sivadasan, S., Christopher, P. V., & Agrawal, R. (2011). Validation of QSAR Models—Strategies and Importance. *Int J Drug Design and Discov*, *2*, 511-519.
- Roy, K., & Ambure, P. (2016). The “double cross-validation” software tool for MLR QSAR model development. *Chemometrics and Intelligent Laboratory Systems*, *159*, 108-126. <https://doi.org/10.1016/j.chemolab.2016.10.009>
- Roy, K., Kar, S., & Ambure, P. (2015). On a simple approach for determining applicability domain of QSAR models. *Chemometrics and Intelligent Laboratory Systems*, *145*, 22-29. <https://doi.org/10.1016/j.chemolab.2015.04.013>
- Sael, L., Li, B., La, D., Fang, Y., Ramani, K., Rustamov, R., & Kihara, D. (2008). Fast protein tertiary structure retrieval based on global surface shape similarity. *Proteins: Structure, Function, and Bioinformatics*, *72*(4), 1259-1273. <https://doi.org/10.1002/prot.22030>
- The PyBioMed Documentation—PyBioMed 1 documentation.* (s. f.). Recuperado 6 de octubre de 2021, de <https://pybiomed.readthedocs.io/en/latest/>
- Toropova, A. P., & Toropov, A. A. (2017). CORAL: Binary classifications (active/inactive) for drug-induced liver injury. *Toxicology Letters*, *268*, 51-57. <https://doi.org/10.1016/j.toxlet.2017.01.011>
- Truchon, J.-F., & Bayly, C. I. (2007). Evaluating Virtual Screening Methods: Good and Bad Metrics for the “Early Recognition” Problem. *Journal of Chemical Information and Modeling*, *47*(2), 488-508. <https://doi.org/10.1021/ci600426e>
- Varoquaux, G., Raamana, P. R., Engemann, D. A., Hoyos-Idrobo, A., Schwartz, Y., & Thirion, B. (2017). Assessing and tuning brain decoders: Cross-validation, caveats, and guidelines. *NeuroImage*, *145*, 166-179. <https://doi.org/10.1016/j.neuroimage.2016.10.038>
- Volkamer, A., Kuhn, D., Grombacher, T., Rippmann, F., & Rarey, M. (2012). Combining Global and Local Measures for Structure-Based Druggability Predictions. *Journal of Chemical Information and Modeling*, *52*(2), 360-372. <https://doi.org/10.1021/ci200454v>
- Volkamer, A., Kuhn, D., Rippmann, F., & Rarey, M. (2012). DoGSiteScorer: A web server for automatic binding site prediction, analysis and druggability assessment. *Bioinformatics*, *28*(15), 2074-2075. <https://doi.org/10.1093/bioinformatics/bts310>

Vyškovský, R., Schwarz, D., Janoušová, E., & Kašpárek, T. (2016). *Random Subspace Ensemble Artificial Neural Networks for First-episode Schizophrenia Classification*. 317-321. <https://doi.org/10.15439/2016F333>

Wang, X., Wang, R., Zhang, Y., & Zhang, H. (2013). Evolutionary Survey of Druggable Protein Targets with Respect to Their Subcellular Localizations. *Genome Biology and Evolution*, 5(7), 1291-1297. <https://doi.org/10.1093/gbe/evt092>

Wu, Q., Peng, Z., Zhang, Y., & Yang, J. (2018). COACH-D: Improved protein–ligand binding sites prediction with refined ligand-binding poses through molecular docking. *Nucleic Acids Research*, 46(W1), W438-W442. <https://doi.org/10.1093/nar/gky439>

# Capítulo 6 | Conclusiones



### 6.1 Conclusiones de la primera campaña

La primera campaña de CV se llevó adelante con un pequeño set de datos (109 compuestos totales), sin embargo, la diversidad estructural de moléculas que conforman dicho set, junto con una adecuada partición mediante la utilización y combinación de técnicas de agrupamiento y la aplicación de la metodología de subespacios aleatorios han permitido obtener 1000 modelos clasificadores con una buena relación descriptores/moléculas, y una tasa de buenas clasificaciones globales, en el orden del 75%, que se mantienen estables tanto en el conjunto de entrenamiento como en el conjunto de prueba. Estos resultados sugieren un bajo grado de sobreajuste por parte de los modelos y por lo tanto son aptos para el CV de bibliotecas externas. Sin embargo, con el objetivo de mejorar el desempeño de los modelos y garantizar la robustez de los mismos hemos aplicado metodologías de aprendizaje por ensamblado, donde se obtuvieron cuatro meta-clasificadores para cada operador de ensamblado (MÍNIMO; PROMEDIO, PROMEDIO del *RANKING* y PROMEDIO del *VOTO*). A partir del análisis de AUCROC y la desviación estándar que presentó cada ensamble frente al conjunto de prueba, se seleccionó el meta-clasificador conformado por la combinación diez modelos mediante el operador mínimo (E-10). Para este meta-clasificador se construyeron superficies PPV frente al conjunto de prueba, con el análisis de la superficie se logró establecer un valor de corte de 0,3026 donde la *Se* y la *Sp* fueron de 84,2% y 94,7% respectivamente. Este *score* permitió seleccionar un total de 169 compuestos durante la fase de CV de las bibliotecas DrugBank (Knox et al., 2011) y SweetLead (Novick et al., 2013). Teniendo en cuenta el potencial clínico que presentan los fármacos reposicionados el primer paso fue eliminar del conjunto de compuestos que superaron los filtros de CV aquellas moléculas para las que aún no se hayan realizado estudios clínicos o que no se encuentren aprobadas para el uso en medicina humana/veterinaria. El pool de 169 compuestos quedó entonces reducido a un conjunto de 44 moléculas, todas con aprobación para uso clínico.

Experimentalmente se evaluaron frente a la enzima *TryS* de tres tripanosomátidos (*T. cruzi*, *T. brucei* y *L. Infantum*) diez compuestos de los 44 fármacos seleccionados; sin embargo, ninguno de los candidatos evaluados produjo una inhibición próxima al 50% de la actividad enzimática a concentraciones 20  $\mu$ M. Estos resultados podrían ser explicados en base al bajo número de compuestos con el que se conformó el conjunto de entrenamiento, algunos autores sostienen que el número de casos para entrenar un modelo no debe ser menor a 150 (Tropsha, 2010), y en esta oportunidad dicho conjunto contaba solo con 90 compuestos.

La evaluación biológica realizada frente a *T. cruzi* demostró que Terbinafina presenta efecto inhibitorio frente a la proliferación de epimastigotes, y en la evaluación de la curva concentración-respuesta se obtuvo que para este estadio del parásito el fármaco presentó un valor de  $EC_{50}$  de 7,5  $\mu$ M ( $\pm 1,8$ ), en contraposición con lo que ocurrió con la forma tripomastigote donde a 20  $\mu$ M, Terbinafina no mantuvo el efecto observado frente a epimastigotes. En otras palabras, Terbinafina no mostró efecto tripanocida estadísticamente diferente del control negativo (DMSO 1%) frente a tripomastigotes.

El fármaco Diclomina fue reportado previamente por nuestro grupo como inhibidor de la proliferación de epimastigotes con un valor de  $EC_{50}$  de 4,22  $\mu$ M pero con un escenario

similar al fármaco Terbinafina, ya que tampoco presentó efecto tripanocida frente al estadio tripomastigote (L. Alberca et al., 2018).

Diciclomina fue predicho como un modulador negativo de la enzima N-miristoil transferasa (L. Alberca et al., 2018); mientras que Terbinafina inhibe la síntesis de Ergosterol a partir de la inhibición de la enzima Escualeno sintetasa (Lazardi et al., 1990). Estos resultados podrían ser aplicados para la búsqueda de nuevas terapias ya sea mediante optimizaciones estructurales de los núcleos farmacológicos o mediante la combinación de fármacos que interaccionen con diferentes blancos moleculares de la vía metabólica potenciando la actividad individual de cada uno de ellos (sinergismo farmacológico).

El estudio de los candidatos frente a *T. brucei* (cepa 427) permitió establecer que el fármaco Diciclomina presenta actividad tripanocida, y se determinó que dicho compuesto exhibe un valor de  $EC_{50}$  de 13,6  $\mu$ M frente al estadio sanguíneo. Si bien este fármaco no es capaz de inhibir la enzima *TryS*, el resultado exhibido frente a *T. brucei* resulta alentador para la búsqueda de una nueva terapia enfocando dicha pesquisa en estructuras análogas a la de este fármaco o derivados químicos del mismo, y de esta manera dilucidar nuevas terapias frente a la enfermedad del sueño, sobre todo en los casos de resistencia a los tratamientos convencionales o de pacientes alérgicos a las drogas utilizadas hasta el momento.

Resulta de interés proseguir las evaluaciones biológicas de los candidatos seleccionados frente a los parásitos, ya que aunque no se haya podido validar los modelos computacionales frente al blanco terapéutico elegido cualquier tratamiento que, como monoterapia o en combinación, pudiera mejorar los perfiles de seguridad y eficacia de BZN y/o NFX constituirá un avance importante para mejorar la calidad de vida de los pacientes que sufren esta patología.

### 6.2 Conclusiones de la segunda campaña

La colaboración establecida con el Dr. Marcelo Comini, nos permitió acceder a una base de datos perteneciente al Laboratorio de Biología Redox de Tripanosomátidos del Instituto Pasteur Montevideo, lo cual permitió aumentar el número de compuestos del set de datos original y a su vez ampliar la diversidad estructural, garantizando la exploración de un mayor espacio químico por parte de los modelos generados a partir del nuevo set. Con esta incorporación el conjunto de entrenamiento creció en más de un 60% respecto al mismo conjunto de la primera campaña, ya que el número de ejemplos de entrenamiento se incrementó de 90 a 140 compuestos. El aumento de la cantidad de compuestos habilitó elevar de nueve a catorce la cantidad de descriptores que se permitía incorporar en cada uno de los modelos, manteniendo fija la relación de un descriptor por cada diez casos de entrenamiento (moléculas). Adicionalmente, y debido a la cantidad de nuevos compuestos, se lograron generar dos conjuntos de prueba (DUD-e A y DUD-e B), de manera tal que la validación externa de los modelos individuales se realizó con uno de ellos (DUD-e A), mientras que la validación externa del meta-clasificador se realizó sobre el conjunto DUD-e B.

En la presente campaña se desarrollaron 1000 modelos computacionales independientes de la conformación, y con capacidad de identificar compuestos con promisoría

actividad inhibitoria frente a la enzima **TryS**. Estos modelos mantuvieron una relación de descriptores incorporados respecto a la cantidad de compuestos en el conjunto de entrenamiento siempre inferior a la permitida, lo que minimiza notablemente la probabilidad de sobreajuste, tal como se desprende del análisis de la **tabla 4.10 del capítulo 4**, donde a partir de la validación cruzada puede apreciarse que el desempeño de los modelos es robusto.

En cuando a la validación mediante la técnica de aleatorización de Fisher se puede comprobar que la comparación de resultados denota la bajísima probabilidad de generar buenos modelos clasificatorios a partir del azar, ya que como se discutió en la sección de resultados los modelos generados mediante aleatorización manifiestan una importante caída en los %BC, con relación a los modelos verdaderos. Estos resultados en combinación con los valores de buenas clasificaciones obtenidas en la validación cruzada sugieren que la utilización de las técnicas de agrupamiento que se implementaron en la presente tesis han permitido realizar una correcta partición del set de datos, evitando enriquecer el conjunto de entrenamiento en moléculas que compartan una misma región del espacio químico, es decir, que la partición generada ha permitido la exploración representativa de todo el espacio químico contemplado en el set de datos.

Los resultados obtenidos para los modelos individuales frente al conjunto DUD-e A permitieron estimar la capacidad predictiva de los mismos frente a un conjunto externo. Adicionalmente la construcción de curvas AUCROC versus combinaciones lineales de dos a cien modelos para el conjunto DUD-e A permitió identificar cuatro meta-clasificadores, que respecto al modelo 319, reflejan mejores valores de AUCROC frente al conjunto DUD-e A. La generación de gráficas PPV, junto con la evaluación de diversas métricas de enriquecimientos (EF y BEDROC) (las cuales permiten estimar la probabilidad de encontrar compuestos **ACTIVOS** en las primeras posiciones del Ranking de una lista ordenada según el score asignado por el meta-modelo), han permitido establecer que la combinación lineal de 90 modelos obtenida mediante el operador mínimo (meta-clasificador (E-90)) reporta valores de métricas que superan a los resultados alcanzados por los ensambles de 4, 5 y 50 modelos que generaron mediante el mismo operador.

Para evaluar el desempeño del meta-clasificador E-90 se realizó un CV retrospectivo frente al conjunto DUD-e B; esto permite establecer si existe algún sesgo involuntario durante la fase de combinación de modelos, ya que de existir sobreajuste o tendencia de los modelos combinados a reconocer un acotado espacio químico se vería reflejado en la disminución de los valores de métricas evaluadas. Como se discutió en la sección de resultados el ensamble mantuvo estable su desempeño frente al conjunto externo DUD-e B, y dichos parámetros sugieren que E-90 es un meta-clasificador confiable para realizar el CV de bibliotecas externas. Si bien es cierto que la combinación de modelos trae aparejada una pérdida de la capacidad explicativa, lo que se busca durante las etapas de CV es mejorar la capacidad predictiva, es decir, que los modelos sean capaces de reconocer compuestos que inhiban/modulen la actividad enzimática de la **TryS**, y no la explicación de interacciones entre los compuestos y la enzima, ya que esta información podrá ser extraída, mediante otras metodologías, evaluando aquellos compuestos que el meta-clasificador (E-90) seleccione como **ACTIVOS**.

El E-90 se aplicó en la fase de CV frente a las bibliotecas SweetLead (Novick et al., 2013) y DrugBank (Wishart et al., 2018) donde se recogieron 82 compuestos con promisoría

actividad inhibitoria de la **TryS**, de los cuales 30 corresponden a fármacos aprobados o compuestos que se encuentran cursando estudios clínicos, un dato no menor teniendo en cuenta el enfoque de reposicionamiento de fármacos que se planteó al inicio de esta tesis. Posteriormente se logró adquirir cinco de los 30 candidatos seleccionados, y evaluarlos frente a la enzima **TryS** de tres especies de tripanosomátidos (*T. cruzi*, *T. brucei* y *L. infantum*); sin embargo, ninguno de los cinco fármacos logró, a 25  $\mu\text{M}$ , inhibir significativamente a la enzima, ya que los análisis ANOVA de una vía y posterior prueba de Tukey que se realizaron para cada una de las enzimas (**TcTryS**, **TbTryS** y **LiTryS**) sugieren que no existen diferencias significativas entre los fármacos evaluados y el control negativo (DMSO 1%). Esto permite inferir que ninguno de los cinco fármacos examinados es capaz de inhibir la síntesis de Tripanotión en los tripanosomátidos.

Los ensayos biológicos realizados frente a *T. cruzi* no demostraron actividad inhibitoria frente a la proliferación de epimastigotes por parte de los candidatos evaluados, como así tampoco se logró identificar actividad tripanocida de algunos de los fármacos frente al estadio tripomastigote para esta misma especie. El fármaco Celecoxib mostro actividad tripanocida frente a *T. brucei*, y se logró determinar una  $EC_{50}$  de 17,54  $\mu\text{M}$  ( $IC_{95}$ : 11,28  $\mu\text{M}$ ; 27,26  $\mu\text{M}$ ) frente a la línea celular 427 cepa BB3. Dicho resultado sugiere que Celecoxib podría ser útil como punto de partida para el diseño de nuevos compuestos que, manteniendo el núcleo químico del fármaco, mejoren el perfil tripanocida frente a *T. brucei*, y permitan potencialmente obtener un nuevo tratamiento farmacológico para la enfermedad del sueño, sobre todo ante la posibilidad de aparición a futuro de resistencia parasitaria a los tratamientos vigentes, o para pacientes que no toleren adecuadamente los tratamientos actuales.

### 6.3 Un balance global de las dos campañas

Un análisis que podría desprenderse a partir de los resultados obtenidos hasta aquí es que los modelos desarrollados mediante la metodología ALD no son capaces de reconocer propiedades fisicoquímicas relevantes de diferentes moduladores. Sin embargo, existe una vasta experiencia sobre como el ALD y el aprendizaje por ensamblado, tal cual se emplearon aquí, son metodologías útiles para la búsqueda de moduladores de diferentes blancos moleculares, ya sean estos enzimas o transportadores, o para el cribado fenotípico de potenciales soluciones terapéuticas para diferentes patologías (L. Alberca et al., 2018; L. Alberca et al., 2016, 2019; Bélgamo et al., 2020; Bellera et al., 2018; Meneses-Marcel et al., 2005). Por lo tanto, una mirada integral de los datos en su conjunto, como son el análisis de los resultados de CV y validación experimental, en combinación con la relación de inhibidores/no inhibidores reportados hasta la fecha, y la dificultad para establecer la cristalización y resolución de la estructura 3D de la proteína (debido a la flexibilidad/movilidad de algunas de sus regiones) han dado lugar al planteo de una nueva pregunta: “¿Podría ser que la enzima presente resistencia a la modulación farmacológica?”, este cuestionamiento derivó en el desarrollo de un algoritmo para el análisis inverso, es decir evaluar a partir de un pool de proteínas cuáles de ellas presentan características válidas para ser moduladas

farmacológicamente, y aplicar luego de este proceso técnicas de CV enfocadas en proteínas Drogables.

### 6.4 Conclusiones sobre drogabilidad

En el presente trabajo de tesis se aplicaron dos técnicas QSAR basadas en ALD para construir modelos con capacidad de discriminar proteínas Drogables y No Drogables.

La primera estrategia fue realizar la búsqueda exhaustiva de combinaciones de descriptores originados mediante el procesamiento manual de los *outputs* de diferentes herramientas *online* para calcular drogabilidad de bolsillos de unión a ligandos. A partir de tal combinación de descriptores, y limitando a cinco el número de variables independientes por modelo a fin de prevenir el sobreajuste, se obtuvieron 3458 modelos que incorporaban entre 2 y 5 descriptores cada uno. Sin embargo, ninguno de estos modelos fue capaz de superar aceptablemente las instancias de validación interna a las que fueron sometidos, y en consecuencia se decidió no proseguir con esta estrategia.

La segunda estrategia consistió en generar modelos basados en la combinación de 147 descriptores del tipo CTD y los 14 descriptores que se obtuvieron a partir de las mismas herramientas *online* mencionadas en el párrafo anterior. Aquí se aplicó la metodología de subespacios aleatorizados en conjunto con la técnica de pasos hacia adelante, lográndose obtener y validar un pool de 1000 modelos clasificadores, los cuales superaron con éxito las diferentes etapas de validación a las que fueron sometidos. En esta instancia el mejor modelo individual resultó ser el modelo que denominamos 778, para el cual, frente al conjunto de validación, se obtuvieron valores de exactitud mayores al 87%, mientras que la tasa de recuperación y la precisión fueron del 84,8% y 80% respectivamente.

El mejor modelo individual aplicado para el análisis de la enzima LmTryS arrojó un *score* de 0,370, es decir que dicha proteína se encuentra por debajo del *score* establecido para discernir entre proteínas Drogables y No Drogables, ya que el valor de corte definido para discriminar proteínas drogables de no drogables fue de 0,5 correspondiente a la media del valor asignado arbitrariamente a cada categoría como etiqueta de clase (Drogable =1 y No drogable =0).

Adicionalmente, y con el objetivo de desafiar los resultados obtenidos a partir del modelado QSAR, se realizó un análisis de consenso mediante la superposición de bolsillos drogables. Dicho análisis reforzó los resultados obtenidos durante la etapa de modelado, ya que se observó que no existe consenso entre las diferentes herramientas *online* frente a la identificación de al menos un bolsillo drogable para la enzima LmTryS. Este dato parece reafirmar la predicción realizada por nuestro modelo (ausencia de bolsillos drogables) y en consecuencia se puede inferir que la proteína presenta resistencia a la modulación farmacológica.

Los resultados son consistentes con la hipótesis surgida durante el desarrollo de la tesis, sobre la posibilidad de que la enzima resultara farmacológicamente difícil de modular. Si bien la enzima elegida como blanco es interesante a partir de los resultados obtenidos por otros autores considerando su peso en el control metabólico de la biosíntesis de Tripanotión,

nuestro análisis sugiere que no es fácil de modular y que posiblemente convenga reorientar la búsqueda de nuevos agentes tripanocidas a otros blancos moleculares más susceptibles de ser modulados. Otra opción posible sería el abordaje de la enzima no ya desde el cribado virtual sino desde la perspectiva del diseño de fármacos, tal vez con foco en inhibidores competitivos que puedan unirse al sitio catalítico.

Como perspectiva, el algoritmo para predecir drogabilidad será incluido en el corto plazo entre las herramientas bioinformáticas de código abierto que el Laboratorio de Investigación y Desarrollo de Bioactivos (LIDeB) de la UNLP ofrece de manera gratuita en su sitio web.

### 6.5 Perspectivas

Las predicciones de los modelos computacionales no han podido ser validados mediante ensayos húmedos (es decir, no se confirmó la actividad inhibitoria predicha por los modelos). Sin embargo, a partir de estos resultados desalentadores e intentando dar una explicación a los mismos, hemos desarrollado un nuevo algoritmo computacional que permite evaluar la drogabilidad de proteínas y contribuye a establecer la potencialidad de las mismas como blancos moleculares de fármacos. Este podrá ser aprovechado de aquí en más para estudiar de manera temprana a las proteínas que se postulan como blancos, evitando así encaminar una búsqueda de fármacos frente a blancos moleculares difíciles de modular. En este sentido creemos que es de suma importancia compartir los resultados con la comunidad científica, ya que la búsqueda de moduladores exógenos para proteínas de diversas especies animales, vegetales y patógenas resulta un campo de investigación en continuo crecimiento y desarrollo. Los modelos desarrollados para evaluar drogabilidad son una herramienta útil que permitirán a los investigadores predecir, *a priori*, cuáles son las proteínas con mayor probabilidad de ser moduladas farmacológicamente; sin necesidad de evaluaciones biológicas previas, es decir que el desarrollo de nuevas terapias podría comenzar con el estudio bioinformático de diversas proteínas y la selección de proteínas drogables y recién entonces implementar campañas de cribado virtual y la subsecuente confirmación experimental de las predicciones.

### 6.6 Detalle de publicaciones, congresos y becas

Durante el trabajo de tesis doctoral se realizaron 13 presentaciones a congresos y reuniones científicas (8 nacionales y 5 internacionales), 2 publicaciones en revistas científicas con referato y 2 capítulos de libros.

En el año 2017 se obtuvo por concurso 1 subsidio de viajes otorgado por la Universidad Nacional de la Plata para una estadía en el exterior y 1 beca otorgada por la Sociedad Chilena de Parasitología para asistir al “XXIV Congreso Latinoamericano de Parasitología” en la ciudad de Santiago de Chile, Chile. En el año 2019 se obtuvieron 3 becas más, una de ellas otorgada por el Colegio de Farmacéuticos de La Plata para financiar una estadía en el Instituto Pasteur de Montevideo, Uruguay; otra beca otorgada por la Sociedad Argentina de Protozoología para asistir a la “Reunión Anual de Biociencias 2019”, en la ciudad de Mar del Plata y una tercera

beca de inscripción para realizar el curso: “*Training School: Biological Meaning of Protein Disorder*”.

Durante el desarrollo de la tesis se realizaron dos estadias de investigación en el Instituto Pasteur de Montevideo -Uruguay en los años 2017 y 2019.

A continuación, se detallan las publicaciones realizadas en el transcurso de la beca doctoral:

### Artículos:

- 1- Morales JF, Chuguransky S, Alberca LN, **Alice JI**, Goicoechea S, Ruiz ME, Bellera CL, Talevi A. Positive Predictive Value Surfaces as a Complementary Tool to Assess the Performance of Virtual Screening Methods. *Mini Rev Med Chem*. 2020;20(14):1447-1460.
- 2- **Alice JI**, Bellera CL, Benítez D, Comini MA, Duchowicz PR, Talevi A. Ensemble learning application to discover new trypanothione synthetase inhibitors. *Mol Divers*. 2021 Aug;25(3):1361-1373.

### Capítulos de libros:

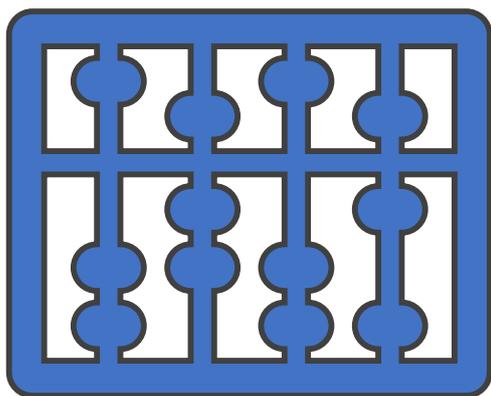
- 1- Bellera CL, Sbaraglini ML, Alberca LN, **Alice JI**, Talevi, A. In silico modeling of FDA-approved drugs for discovery of therapies against neglected diseases: A drug repurposing approach; Elsevier; 2019; 625-644.
- 2- Rodríguez S, **Alice JI**, Bellera CL, Talevi A. Structure-based binding pocket detection and druggability assessment. In: Drug Target Selection and Validation. Springer Nature. 2022. In press.

## 6.7 Referencias bibliográficas

- Alberca, L. N., Chuguransky, S. R., Álvarez, C. L., Talevi, A., & Salas-Sarduy, E. (2019). In silico Guided Drug Repurposing: Discovery of New Competitive and Non-competitive Inhibitors of Falcipain-2. *Frontiers in Chemistry*, 7, 534. <https://doi.org/10.3389/fchem.2019.00534>
- Alberca, L. N., Sbaraglini, M. L., Balcazar, D., Fraccaroli, L., Carrillo, C., Medeiros, A., Benitez, D., Comini, M., & Talevi, A. (2016). Discovery of novel polyamine analogs with anti-protozoal activity by computer guided drug repositioning. *Journal of Computer-Aided Molecular Design*, 30(4), 305-321. <https://doi.org/10.1007/s10822-016-9903-6>
- Alberca, L., Sbaraglini, M., Ruiz, D., Larrea, S. V., Villamil, S. F., Carrillo, C., & Talevi, A. (2018). Computer-aided search of novel inhibitors of n-myristoyl transferase with trypanocidal effects. *International Journal of Infectious Diseases*, 73, 311. <https://doi.org/10.1016/j.ijid.2018.04.4121>
- Bélgamo, J. A., Alberca, L. N., Pórfido, J. L., Romero, F. N. C., Rodríguez, S., Talevi, A., Córscico, B., & Franchini, G. R. (2020). Application of target repositioning and in silico screening to exploit fatty acid binding proteins (FABPs) from *Echinococcus multilocularis* as possible drug targets. *Journal of Computer-Aided Molecular Design*, 34(12), 1275-1288. <https://doi.org/10.1007/s10822-020-00352-8>
- Bellera, C. L., Sbaraglini, M. L., & Talevi, A. (2018). Modern Approaches for the Discovery of Anti-Infectious Drugs for the Treatment of Neglected Diseases. *Current Topics in Medicinal Chemistry*, 18(5), 369-381. <https://doi.org/10.2174/1568026618666180509151146>
- Knox, C., Law, V., Jewison, T., Liu, P., Ly, S., Frolkis, A., Pon, A., Banco, K., Mak, C., Neveu, V., Djoumbou, Y., Eisner, R., Guo, A. C., & Wishart, D. S. (2011). DrugBank 3.0: A comprehensive resource for «Omics» research on drugs. *Nucleic Acids Research*, 39(Database), D1035-D1041. <https://doi.org/10.1093/nar/gkq1126>
- Lazardi, K., Urbina, J. A., & Souza, W. de. (1990). Ultrastructural alterations induced by two ergosterol biosynthesis inhibitors, ketoconazole and terbinafine, on epimastigotes and amastigotes of *Trypanosoma (Schizotrypanum) cruzi*. *Antimicrobial Agents and Chemotherapy*. <https://doi.org/10.1128/AAC.34.11.2097>
- Meneses-Marcel, A., Marrero-Ponce, Y., Machado-Tugores, Y., Montero-Torres, A., Pereira, D. M., Escario, J. A., Nogal-Ruiz, J. J., Ochoa, C., Arán, V. J., Martínez-Fernández, A. R., & García Sánchez, R. N. (2005). A linear discrimination analysis based virtual screening of trichomonacidal lead-like compounds: Outcomes of in silico studies supported by experimental results. *Bioorganic & Medicinal Chemistry Letters*, 15(17), 3838-3843. <https://doi.org/10.1016/j.bmcl.2005.05.124>
- Novick, P. A., Ortiz, O. F., Poelman, J., Abdulhay, A. Y., & Pande, V. S. (2013). SWEETLEAD: An In Silico Database of Approved Drugs, Regulated Chemicals, and Herbal Isolates for Computer-Aided Drug Discovery. *PLoS ONE*, 8(11), e79568. <https://doi.org/10.1371/journal.pone.0079568>
- Tropsha, A. (2010). Best Practices for QSAR Model Development, Validation, and Exploitation. *Molecular Informatics*, 29(6-7), 476-488. <https://doi.org/10.1002/minf.201000061>

Wishart, D. S., Feunang, Y. D., Guo, A. C., Lo, E. J., Marcu, A., Grant, J. R., Sajed, T., Johnson, D., Li, C., Sayeeda, Z., Assempour, N., Iynkkaran, I., Liu, Y., Maciejewski, A., Gale, N., Wilson, A., Chin, L., Cummings, R., Le, D., ... Wilson, M. (2018). DrugBank 5.0: A major update to the DrugBank database for 2018. *Nucleic Acids Research*, *46*(D1), D1074-D1082.  
<https://doi.org/10.1093/nar/gkx1037>

# Anexo 1 | Descriptores



A continuación, se presenta una breve descripción de los descriptores incorporados en los mejores modelos individuales de las dos campañas de cribado virtual.

### Descriptores involucrados en los mejores 5 modelos individuales de la primera campaña CV

#### Modelo 321

$$\text{Score} = 0.193 - 0.2199 * NssO + 0.8548 * MATS5m + 3.6289 * RBF$$

#### Modelo 744

$$\text{Score} = -0.12786 - 0.6876 * B10[C - O] + 0.72926 * S-110 + 0.34613 * B09[N - F] + 0.37486 * C-004 + 0.10336 * SpMAD_Dz(e) - 0.04838 * CATS2D_07_AL$$

#### Modelo 628

$$\text{Score} = 0.30265 - 0.01026 * P_VSA_LogP_5 + 37.08550 * JGI8 + 0.22687 * B01[C - F] - 0.12146 * NsssCH$$

#### Modelo 429

$$\text{Score} = -2.0646 - 0.6556 * Eig04_EA(dm) - 0.4912 * SpDiam_EA + 0.3155 * B10[N - N]$$

#### Modelo 470

$$\text{Score} = -0.3005 - 0.0510 * Eig13_EA(ed) - 1.1307 * GATS7i - 0.4879 * B10[C - O] + 0.1952 * TI2_L + 0.5214 * B05[S - S] - 0.3847 * O-059 - 0.926 * GATS6p - 0.4252 * B03[N - S] - 0.6932 * nRSR$$

En la siguiente tabla se presenta una breve interpretación de los descriptores involucrados en los mejores 5 modelos individuales de la primera campaña de cribado virtual

| Descriptor      | Descripción/interpretación                                    | Bloque              |
|-----------------|---|---------------------|
| <b>B01[C-F]</b> | Presencia/ausencia del par C-F a una distancia topológica 1.  | Pares de átomos 2D. |
| <b>B03[N-S]</b> | Presencia/ausencia del par N-S a una distancia topológica 3.  | Pares de átomos 2D. |
| <b>B05[S-S]</b> | Presencia/ausencia del par S-S a una distancia topológica 5.  | Pares de átomos 2D. |
| <b>B09[N-F]</b> | Presencia/ausencia del par N-F a una distancia topológica 9.  | Pares de átomos 2D. |
| <b>B10[C-O]</b> | Presencia/ausencia del par C-O a una distancia topológica 10. | Pares de átomos 2D. |
| <b>B10[N-N]</b> | Presencia/ausencia del par N-N a una distancia topológica 10. | Pares de átomos 2D. |

|              |   |                                       |
|--------------|---|---------------------------------------|
| O-059        | Número de fragmentos Al-O-Al.   | Fragmentos centrados en el átomo.     |
| S-110        | Numero de fragmentos R-SO <sub>2</sub> -R   | Fragmentos centrados en el átomo.     |
| C-004        | Número de fragmentos CR <sub>4</sub> .  | Fragmentos centrados en el átomo.     |
| SpDiam_EA    | Diámetro de espectral de la matriz de adyacencia de bordes.   | Índice de adyacencia de bordes        |
| Eig04_EA(dm) | Autovalor de orden 4 de la matriz de adyacencia ponderada por el momento dipolar.                                 | Índice de adyacencia de bordes.       |
| Eig13_EA(ed) | Valor propio número 13 de la matriz de adyacencia de borde ponderada por el grado del borde.                      | Índice de adyacencia de bordes.       |
| GATS7i       | Autocorrelación de Geary de grado 7 ponderado por potencial de ionización.  | Autocorrelaciones 2D.                 |
| GATS6p       | Autocorrelación de Geary de grado 6 ponderado por polarización.   | Autocorrelaciones 2D.                 |
| JGI8         | Índice topológico de carga de orden 8.  | Autocorrelaciones 2D.                 |
| MATS5m       | Autocorrelación de Moran de grado 5 ponderada por la masa.  | Autocorrelaciones 2D.                 |
| nRSR         | Número de sulfuros.   | Recuento de grupos funcionales.       |
| NssO         | Numero de átomos tipo ssO.  | Índices de estado E de tipo de átomo. |
| NsssCH       | Numero de átomos tipo sssCH.  | Índices de estado E de tipo de átomo. |
| P_VSA_LogP_5 | Tipo P_VSA como LogP, compartimento 5.  | Descriptores tipo P_VSA.              |
| RBF          | Fracción de enlaces rotables.   | Índice constitucionales.              |
| SpMAD_Dz(e)  | Desviación absoluta de la media espectral de la matriz de Barysz ponderada por la electronegatividad de Sanderson | Descriptores basado en matrices 2D.   |
| T12_L        | Segundo índice de Mohar de la matriz de Laplace   | Descriptores basados en matrices 2D.  |
| CATS2D_07_AL | Par de átomos aceptor-lipofílico CATS2D a una distancia de 7.   | CAST 2D.                              |

### Descriptores basados en pares de átomos 2D

Los descriptores incluidos dentro del bloque pares de átomos 2D codifican la presencia o ausencia de los pares de átomos presentes en una molécula y la distancia topológica que conecta a dichos átomos. Un par de átomos se define como dos átomos diferentes al hidrogeno, y donde la distancia a considerar es la que resulte del recorrido más corto.

Los mejores cinco modelos de la primera campaña incluyeron los descriptores **B01[C-F]**, **B03[N-S]**, **B05[S-S]**, **B09[N-F]**, **B10[C-O]** y **B10[N-N]**. El descriptor **B01[C-F]** indica la presencia o ausencia del par Carbono-Flúor a una distancia topológica de 1, **B03[N-S]** corresponde al par de átomos Nitrógeno y Azufre separado por una distancia topología de tres, mientras que **B05[S-S]** indica una distancia topológica de cinco en la conectividad del par de átomos definido por dos Azufres. La presencia o no del par Nitrógeno-Flúor separado por una distancia

topológica de 9 está dado por el descriptor **B09[N-F]**. **B10 [C-O]** y **B10[N-N]** señalan la presencia/ ausencia del par Carbono- Oxígeno y del par Nitrógeno-Nitrógeno, respetivamente, separado por una distancia topología de diez.

### Descriptores de fragmentos centrados en el átomo

Este conjunto de descriptores, correspondiente a los descriptores 1D, permite conocer propiedades de la molécula a partir la subestructura de la misma, también se los suelen mencionar como descriptores del tipo indicadores, ya que los mismos permiten conocer la presencia/ ausencia (o frecuencia de aparición) de determinado fragmento en una molécula. El fragmento se define en términos del átomo y todos los átomos, distintos al Hidrogeno, que se enlazan con el mismo. Es decir que el fragmento más pequeño puede estar comprendido de dos átomos y un enlace, mientras que el mayor tamaño de fragmento se compone de cinco átomos y cuatro enlaces químicos.

**O-059** Este descriptor indica la presencia del fragmento Al-O-Al, donde el átomo central es Oxígeno y esta enlazado químicamente a dos cadenas alifáticas. **S-110** indica la presencia del fragmento R-SO<sub>2</sub>-R, donde el Azufre ocupa la posición central, mientras que R representa cualquier grupo funcional enlazados al átomo centra mediante un Carbono, a su vez el S encuentra químicamente enlazado a dos átomos de Oxígeno. El descriptor **C-004** denota el fragmento CR<sub>4</sub>, el cual se compone de un átomo central de Carbono enlazado a cuatro ligandos/grupos mediante otro átomo de Carbono.

### Descriptores basados en índice de adyacencia de bordes

La matriz de adyacencia de bordes, también conocida como matriz G, parte de la premisa que G es una función definida por los vértices (átomos) y los bordes (enlaces). Esta será una matriz cuadra y simétrica definida como  $A = [a_{ij}]_n$  donde los  $n$  representa el número de vértices, mientras que los elementos  $[a_{ij}]$  se encuentran definido de la siguiente manera:

$$a_{ij} = \begin{cases} 1; & \text{si } e_k \in E/e_{k \sim v_i, v_j} \\ 0; & \text{si no cumple} \end{cases}$$

Es decir que existe un borde  $e_k$  (enlace) entre  $V_i$  y  $V_j$  entonces ambos vértices son adyacentes y por lo tanto  $a_{ij} = 1$ , en otras palabras, el átomo 1 y el átomo 2 están conectados mediante un enlace químico.

La adyacencia no es una propiedad exclusiva del conjunto de vértices, ya que dos enlaces pueden ser adyacentes cuando existe un vértice  $V^*$  que incide con bordes a ambos lados. Y a partir de operaciones algebraicas pude expresarse la matriz mediante la siguiente ecuación:

$$\epsilon(\rho) = \sum_l [\delta(e_i)\delta(e_j)]_l^{-1/2}$$

Donde la suma se realiza sobre todo los bordes adyacentes ( $l$ ) del gráfico que representa la topología molecular del compuesto analizado.

En los modelos se incluyeron tres descriptores pertenecientes a este bloque, **SpDiam\_EA** que representa el diámetro de la esfera espectral de la matriz de adyacencia de borde,

**Eig04\_EA(dm)** el cual representa el cuarto autovalor de la matriz de adyacencia de bordes ponderado por el momento dipolar de la molecular, y **Eig13\_EA(ed)** que denota el decimotercer autovalor de la matriz de adyacencia de borde ponderada por el grado del borde.

### Descriptores de autocorrelación 2D

El bloque de descriptores de autocorrelación 2D codifican las propiedades estructurales y fisicoquímicas de las moléculas. Esta información esta codificada mediante vectores, los cuales están definido en términos de distancias, y dichas distancias quedan definidas por las uniones químicas que presenta un par de átomos. En términos generales la distancia puede calcularse con la siguiente expresión matemática:

$$\text{Autocorrelacion}(r_a; r_b) = \sum_i \sum_j \delta(r_a \leq r_{ij} \leq r_b) p_i p_j$$

Donde  $r_{ij}$  representa la distancia entre los átomos  $i$  y  $j$  de la molécula.  $P_i$  y  $P_j$  corresponde a las propiedades atómicas utilizadas para ponderar la autocorrelación de cada átomo.  $r_a$  y  $r_b$  definen los límites inferiores y superiores de cada distancia consecutiva. A su vez estas propiedades pueden describirse más allá de las características geométricas de una molécula, ponderando las distancias de pares atómicos respecto a diversas propiedades de los átomos como carga parcial, polarizabilidad, electronegatividad, etc.

### Autocorrelación 2D de Geary

**GATS7i** y **GATS6p** son descriptores de autocorrelación 2D de Gary de diversos grados (7 y 6) ponderados por el potencial de ionización y el grado de polarización respectivamente. Este tipo de descriptores pueden ser calculado a partir de la siguiente ecuación:

$$GATS(w) = \frac{1/2K \sum_i \sum_j (w_i - w_j) \delta(L, d_{ij})}{1/(n-1) \sum_i (w_i - \bar{w})^2}$$

Donde  $w_i$  y  $w_j$  representan la propiedad atómica de los átomos  $i$  y  $j$  respectivamente, mientras que  $\bar{w}$  corresponde al valor promedio de la propiedad en toda la molécula,  $n$  es el número de átomos que conforman la molécula,  $L$  es el orden de la autocorrelación, y  $\delta(L, d_{ij})$  representa la distancia entre  $L$  y  $d_{ij}$ , es decir que para el cálculo se tendrán en cuenta solo aquellos pares de átomos que se encuentran separados por una distancia  $L$ , mientras que  $K$  corresponde a la cantidad de pares de átomos que comparten una distancia  $L$ .

### Autocorrelación 2D de Moran

Los descriptores incluidos en el bloque de autocorrelación 2D de Moran pueden definirse con la siguiente expresión matemática.

$$MATS(w) = \frac{1/K \sum_i \sum_j (w_i - \bar{w})(w_j - \bar{w}) \delta(d_{ij}; k)}{1/n \sum_i (w_i - \bar{w})^2}$$

Donde  $w_i$  y  $w_j$  representan la propiedad atómica de los átomos  $i$  y  $j$  respectivamente, mientras que  $\bar{w}$  corresponde al valor promedio de la propiedad en toda la molécula.  $n$  es el número de átomos que conforman la molécula.  $d_{ij}$  corresponde a la distancia entre los átomos ( $i;j$ ) mientras que  $\delta(d_{ij};k)$  delta de Kronecker y vale 1 cuando  $d_{ij}=k$  y cero de lo contrario. El descriptor **MATS5m** corresponde a la autocorrelación de Moran de grado cinco cuando la propiedad ponderada es la masa atómica.

### Descriptores basados en recuento de grupos funcionales

Este grupo de descriptores se encuentra dentro de los descriptores 1D, y al igual que los descriptores de fragmentos centrados en el átomo, codifican la presencia/ausencia o frecuencia de aparición de determinados grupos funcionales. En esta campaña, el descriptor **nRSR**, incorporado en el modelo 427, codifica la frecuencia de aparición de grupos sulfuros, enlazados químicamente mediante átomos de Carbono a cualquier otro grupo funcional, cadenas alifáticas, o grupos aromáticos.

### Descriptores de Índices de estado E de tipo de átomo

Los descriptores basados en índices de estados electrotopológicos (E) proporcionan información sobre el estado electrónico y topológico de los átomos en una molécula, identificando a cada tipo de átomo según su número atómico (Z) y estado de valencia. Dicho estado se define según la siguiente ecuación.

$$S_i = I_i + \Delta I_i$$

Donde  $I_i$  es el estado intrínseco del átomo, mientras que  $\Delta I_i$  corresponde a la perturbación del campo sobre el átomo  $i$ -ésimo, es decir el efecto de todos los demás átomos de la molécula.

El estado E para un átomo está definido como la suma de los estados ( $I_i$ ) y todas las perturbaciones ( $\Delta I_i$ ) que sufre el mismo, este valor se puede determinar a partir de la ecuación expresada a continuación.

$$\Delta I_{ij} = \sum_j \frac{I_i + I_j}{r_{ij}^2}$$

Donde  $r_{ij}$  corresponde a la distancia entre los átomos  $i$  y  $j$ .

**NssO** y **NsssCH** corresponden al número de átomos con estado de oxidación ssO y estado de valencia sssCH respectivamente.

### Descriptores tipo P\_VSA.

Este conjunto de descriptores está definido como el área superficial de van der Waals (VSA) que presenta una propiedad  $P$  en cierto rango de área. Para generar los descriptores P\_VSA se debe estimar el área de van der Waals de cada átomo con la siguiente expresión:

$$VAS_i = 4\pi R^2 - \pi R_i \sum_i^A a_{ij} \left( \frac{R_j^2 - (R_i - g_{ij})^2}{g_{ij}} \right)$$

Donde  $R$  corresponde al radio atómico de van der Waals, la suma abarca todos los átomos, pero sólo tiene en cuenta las contribuciones de los átomos enlazados al  $i$ -ésimo átomo, siendo  $a_{ij}$  los elementos de la matriz de adyacencia. La cantidad  $g_{ij}$  se calcula como:

$$g_{ij} = \min\{\max\{|R_i - R_j|, b_{ij}\}, (R_i + R_j)\}$$

El término  $b_{ij}$  representa la longitud ideal de enlace entre 2 átomos ( $i, j$ ), y se calcula de la siguiente manera:

$$b_{ij} = r_{ij}^* - c_{ij}$$

Donde  $r_{ij}$  representa una longitud de enlace de referencia y  $c_{ij}$  es un término de corrección dependiente de la multiplicidad de enlaces: 0 para enlaces sencillos, 0,1 para enlaces aromáticos, 0,2 para enlaces dobles y 0,3 para enlaces triples.

Llamemos  $P_i$  a una propiedad  $P$  del  $i$ -ésimo átomo. Entonces, los descriptores de  $P\_VSA$  se definen como:

$$P\_VSA_k = \sum_{i=1}^A VSA_i \delta(P_i \in [a_{k-1}, a_k]); k = 1, 2, \dots, n$$

La suma comprende todos los átomos, pero  $VSA_i$  es el área de superficie de van der Waals del  $i$ -ésimo átomo,  $\delta(P_i \in [a_{k-1}, a_k])$  es la función delta de Dirac, a la cual es igual a 1 para átomos con valores de propiedades en el rango especificado y cero en caso contrario;  $a_0 \leq a_k \leq a_n$  son los límites del intervalo tales que  $[a_0, a_n]$  delimita todos los valores de la propiedad  $P$  en cualquier molécula del conjunto de datos.

### Descriptores de incidencias constitucionales

Este bloque contempla descriptores que se pueden obtener a partir del conocimiento de la fórmula química de la molécula y, por tanto, no requieren representar ni vectorizar la estructura molecular. El valor del descriptor **RBF** representan la fracción de enlaces rotables que se encuentran presentes en una molécula determinada.

### Descriptores basados en macices 2D

#### CATS2D – Chemically Advanced Template Search 2D

Los descriptores CATS2D examinan la distancia topológica (independiente de la conformación) entre dos tipos de puntos farmacofóricos de interés, de esta manera se determina la cantidad de veces que se repite un determinado par de puntos en la molécula incluyendo en su determinación la distancia entre dichos puntos.

Generalmente se suelen contemplar, en los modelos farmacofóricos, cinco puntos de interés que incluyen: punto donador de puentes de hidrógeno (**D**), aceptor de puentes de hidrógeno (**A**), ionizables o cargados positivamente (**P**), ionizables o cargados negativamente (**N**), y lipofílico (**L**). Un átomo puede no pertenecer a alguno de estos cinco tipos, por lo cual no se considera o puede ser asignado a más de uno. Bajo estas condiciones existen quince posibles

combinaciones de pares (DD, DA, DP, DN, DL, AA, AP, AN, AL, PP, PN, PL, NN, NL, LL) y cada par es luego determinado por el número de enlaces intervinientes entre los dos puntos.

CATS2D\_07\_AL representan la frecuencia de aparición de la combinación del par denotado por la combinación de un aceptor (A) y un punto lipofílico (L) a una distancia de siete enlaces.

### Descriptores involucrados en el mejor modelo individual de la segunda campaña de cribado virtual.

#### Modelo 319

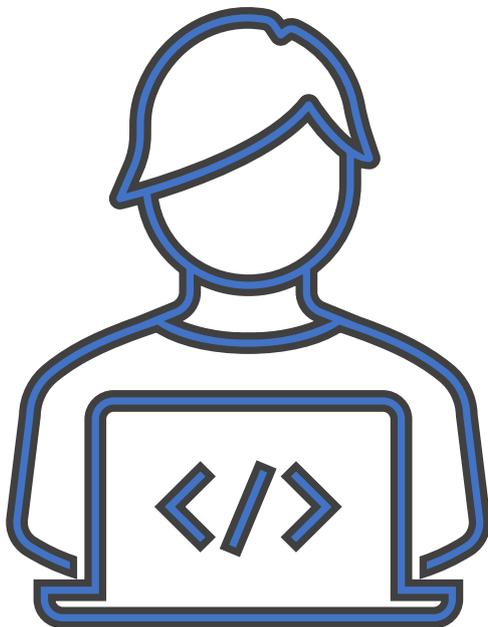
$$\begin{aligned} \text{Score} = & -1,6391 + 0,68769 * B10[Cl - Br] + 0,4902 * B01[N - N] + 2,92669 \\ & * X2Av + 0,26707 * B01[C - F] + 0,61351 * GATS3s - 0,36813 \\ & * CAST2D_09_DD + 0,18382 * N - 067 - 0,02424 * CATS2D_07_LL \\ & + 0,02869 * nCbH + 0,16267 * SpMax_AEA(dm) \end{aligned}$$

En la siguiente tabla se presenta una breve interpretación de los descriptores involucrados en mejor modelo individual de la segunda campaña de cribado virtual

| Descriptor    | Descripción/interpretación  | Bloque                              |
|---------------|---|-------------------------------------|
| B10[Cl-Br]    | Presencia/ausencia del par Cl-Br a una distancia topológica 10.                                     | Pares de átomos 2D.                 |
| B01[N-N]      | Presencia/ausencia del par N-N a una distancia topológica 1.  | Pares de átomos 2D.                 |
| B01[C-F]      | Presencia/ausencia del par C-F a una distancia topológica 1.  | Pares de átomos 2D.                 |
| GATS3s        | Autocorrelación de Geary de grado 3 ponderado por el estado I.                                      | Autocorrelaciones 2D.               |
| CASTS2D_09_DD | Par de átomos Donor - Donor CATS2D a una distancia de 9.  | CAST 2D.                            |
| CATSD_07_LL   | Par de átomos Lipofílico - Lipofílico CATS2D a una distancia de 7.                                  | CAST 2D.                            |
| N-069         | Numero de fragmentos Ar-NH <sub>2</sub> /X-NH <sub>2</sub>  | Fragmentos centrados en el átomo    |
| nCbH          | Número carbonos (sp <sup>2</sup> ) sin sustituyentes en el benceno.                                 | Recuento de grupos funcionales.     |
| SpMax_AEA(dm) | Autovalor principal de la matriz de adyacencia de bordes ampliada, ponderada por el momento dipolar | Descriptores basados en matrices 2D |

Como podrá observarse los descriptores incluidos en el mejor modelo individual de la segunda campaña, corresponden a los mismos bloques descriptos para la campaña previa.

# Anexo 2 | Drogabilidad (PDB id)



## Anexo 2 | Drogabilidad (PDB id)

En las siguientes tablas se presentan los códigos PDB para los conjuntos de proteínas que conforman el conjunto de entrenamiento y conjunto de validación utilizados en el desarrollo y validación del algoritmo descrito en el capítulo 5.

### Proteínas del conjunto de entrenamiento

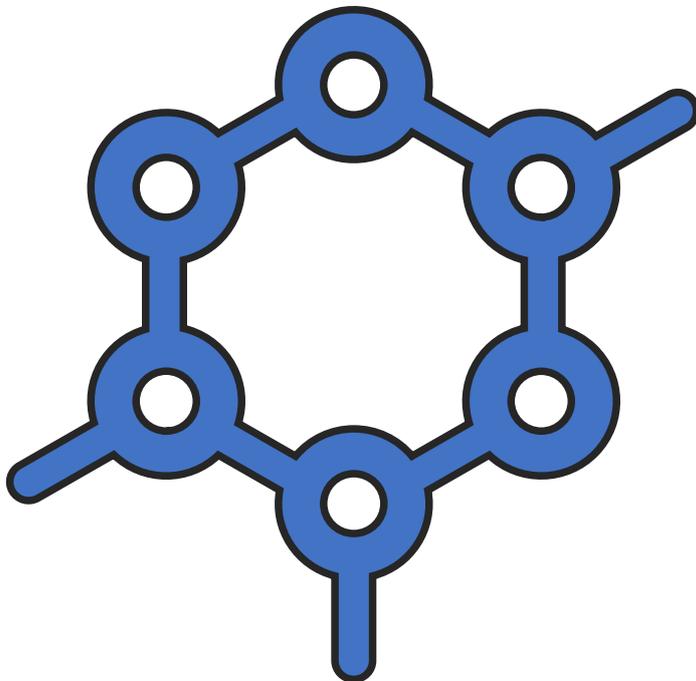
|      |      |      |      |      |      |
|------|------|------|------|------|------|
| 1CAH | 1TB7 | 1NHZ | 1AYU | 1Q1M | 4YWA |
| 1CAM | 1TBB | 1OYN | 1AYV | 1QS4 | 5EU5 |
| 1CAO | 1U2P | 1PTW | 1AYW | 1RNJ | 5MVX |
| 1CD2 | 1UXY | 1Q9M | 1BGO | 1SYL | 6EHT |
| 1DAJ | 1VJ3 | 1QBS | 1MEM | 1WM1 | 6GMN |
| 1DTL | 2AGT | 1S9E | 1NLJ | 1XIW | 1EUW |
| 1DYR | 2AXA | 1S9G | 1Q6K | 1Y32 | 1I9Y |
| 1E3G | 2CD2 | 1T63 | 1SNK | 2AST | 1I9Z |
| 1F5F | 2FNM | 1TV6 | 1TU6 | 2CI9 | 1RN8 |
| 1FM6 | 2H15 | 1ZGY | 1U9V | 2DVJ | 1RWK |
| 1IHI | 2H4N | 2A3I | 1U9X | 2HXM | 1RWN |
| 1J96 | 2HDJ | 2AA2 | 1YK7 | 2JGB | 1RWO |
| 1JSV | 2OAX | 2HND | 2AUZ | 2JV5 | 1RWP |
| 1KLK | 2OVH | 2J32 | 2R6N | 2L9I | 1RWV |
| 1LHN | 2OZ7 | 2OPP | 1B3J | 2L9R | 1RWX |
| 1LHU | 3KFY | 2R4R | 1B55 | 2LFW | 2H48 |
| 1LY3 | 3MDZ | 2ZD1 | 1E8C | 2MX4 | 2H4W |
| 1LY4 | 3V93 | 3BPF | 1EUP | 3POS | 2H4Y |
| 1ME4 | 4I15 | 3CLD | 1G1F | 3PTD | 2H51 |
| 1QKT | 4KN0 | 3D1X | 1GYM | 3X01 | 2H54 |
| 1R39 | 4PH9 | 3D4S | 1IAR | 4CNM | 2HBQ |
| 1RDT | 4YAA | 3OC2 | 1JF7 | 4J33 | 2HBR |
| 1S3Y | 6GL8 | 3OK9 | 1NNY | 4NUU | 2HBY |
| 1SQN | 6GU2 | 4APC | 1ONZ | 4PL8 | 2HBZ |
| 1SR7 | 1HXW | 5YC8 | 1PTY | 4PTD | 2HR6 |
| 1T5Z | 1KNU | 6O69 | 1PVB | 4UNL | 2HRM |

**Proteínas del conjunto de validación**

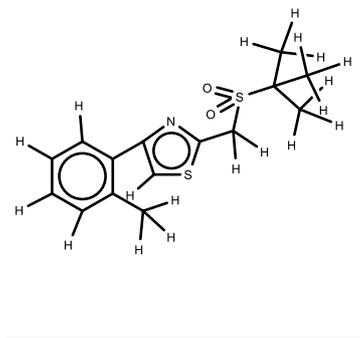
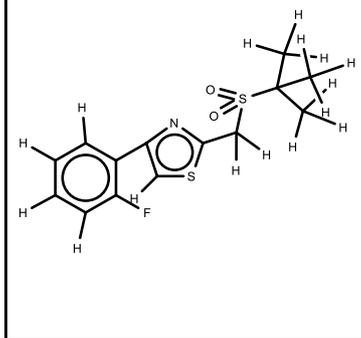
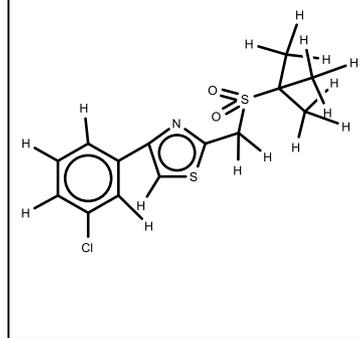
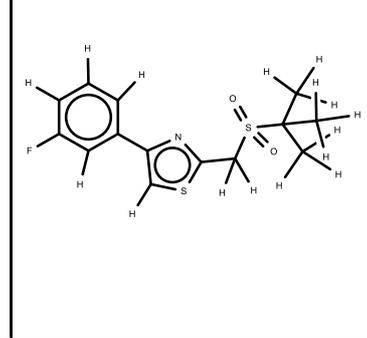
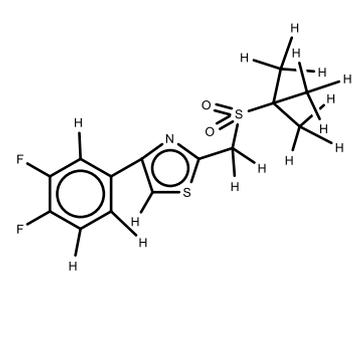
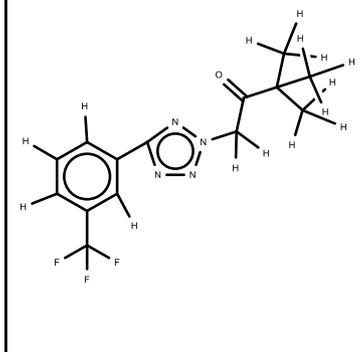
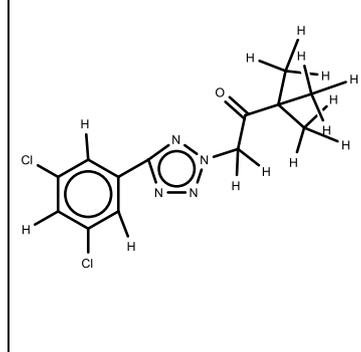
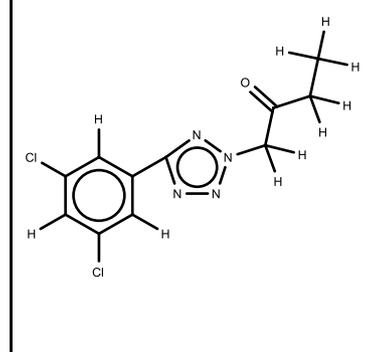
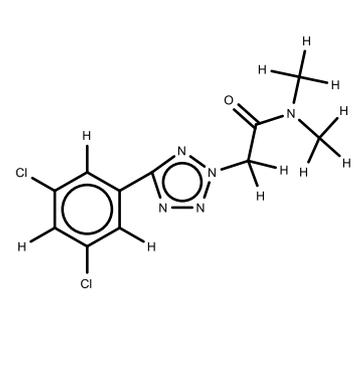
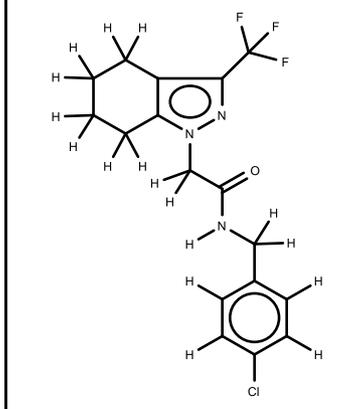
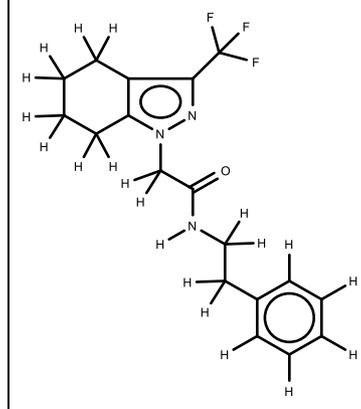
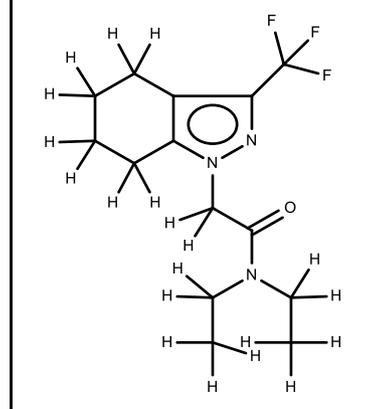
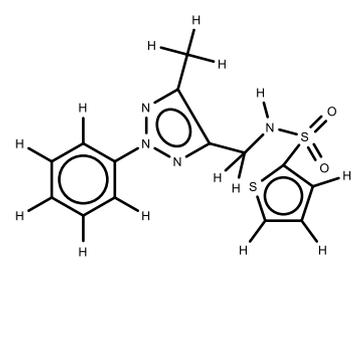
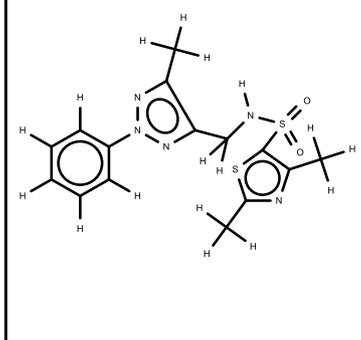
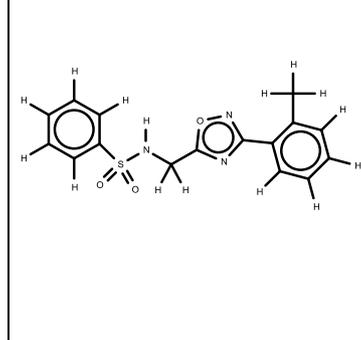
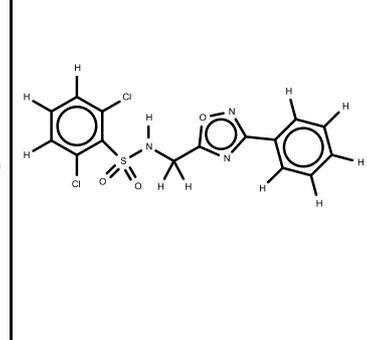
|      |      |      |      |      |      |
|------|------|------|------|------|------|
| 1A28 | 4XAR | 2PTD | 3ERT | 2JOX | 1CEN |
| 1A30 | 5UFW | 2YP5 | 3GOW | 1ATK | 1DMO |
| 1BBC | 1FK9 | 2YVQ | 3HNG | 1AU3 | 1BMQ |
| 4ZYO | 1G1T | 2Z15 | 1XOM | 1AU4 | 1DUD |
| 5HZ5 | 1HQU | 3CXD | 1XON | 1YT7 | 1DUP |
| 5X68 | 1HVR | 3HD7 | 1Y9R | 2ATO | 2LGQ |
| 6FYL | 1HWR | 3IXS | 1YA3 | 2AUX | 2LLH |
| 2PDU | 2AA5 | 3NFK | 1Z95 | 5XOJ | 2N77 |
| 2POU | 2B6A | 1BMD | 1ZEO | 6CEJ | 1SC3 |
| 2Q8G | 2BE2 | 1BTK | 1ZUC | 6CWS | 1SEH |
| 3CAJ | 2HBT | 1BWN | 4BSJ | 2ODC | 2D9W |

## **Anexo 3 |**

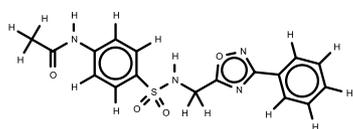
# **Estructuras de compuestos activos e inactivos del set de datos, segunda campaña de CV**



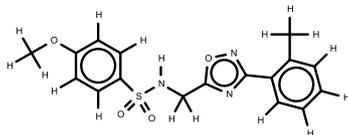
A continuación se presenta el conjunto de entrenamiento, el cual se encuentra conformado por 90 compuestos y comprende desde la estructura denominada **A003** hasta la molécula denominada **I290**. También se adjuntan los conjuntos de prueba DUD-e A y DUD-e B. En este anexo solo presentan las estructuras de los compuestos **ACTIVOS** e **INACTIVOS** reales que conforman cada uno de los conjuntos. El conjunto DUD-e A contempla desde la molécula denominada **A001** hasta la estructura nombrada como **I289**, mientras que las moléculas correspondiente al conjunto DUD-e B son aquellas comprendidas entre la estructura **A011** y la molécula **A286**

|  |   |  |   |
|--|---|--|---|
| <p style="text-align: center;"><b>1</b></p>  <p style="text-align: center;">A 003</p>    | <p style="text-align: center;"><b>2</b></p>  <p style="text-align: center;">A 004</p>    | <p style="text-align: center;"><b>3</b></p>  <p style="text-align: center;">A 007</p>    | <p style="text-align: center;"><b>4</b></p>  <p style="text-align: center;">A 008</p>    |
| <p style="text-align: center;"><b>5</b></p>  <p style="text-align: center;">A 010</p>   | <p style="text-align: center;"><b>6</b></p>  <p style="text-align: center;">A 013</p>   | <p style="text-align: center;"><b>7</b></p>  <p style="text-align: center;">A 015</p>   | <p style="text-align: center;"><b>8</b></p>  <p style="text-align: center;">A 017</p>   |
| <p style="text-align: center;"><b>9</b></p>  <p style="text-align: center;">A 018</p>  | <p style="text-align: center;"><b>10</b></p>  <p style="text-align: center;">A 020</p> | <p style="text-align: center;"><b>11</b></p>  <p style="text-align: center;">A 022</p> | <p style="text-align: center;"><b>12</b></p>  <p style="text-align: center;">A 025</p> |
| <p style="text-align: center;"><b>13</b></p>  <p style="text-align: center;">A 026</p> | <p style="text-align: center;"><b>14</b></p>  <p style="text-align: center;">A 027</p> | <p style="text-align: center;"><b>15</b></p>  <p style="text-align: center;">A 029</p> | <p style="text-align: center;"><b>16</b></p>  <p style="text-align: center;">A 030</p> |

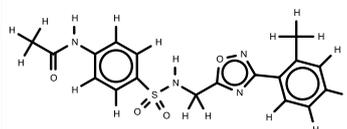
17



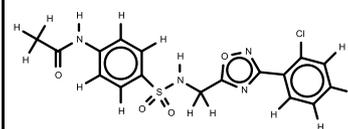
18



19



20



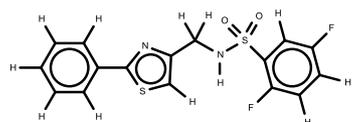
A 031

A 033

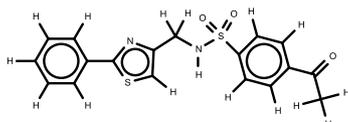
A 034

A 035

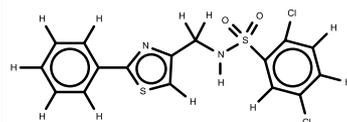
21



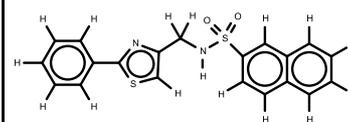
22



23



24



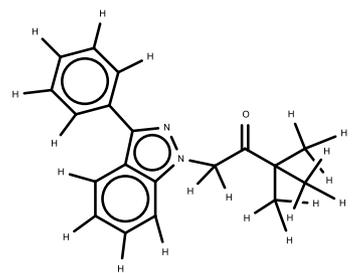
A 036

A 037

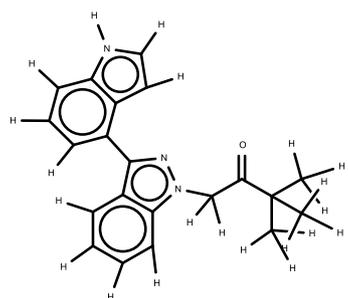
A 038

A 040

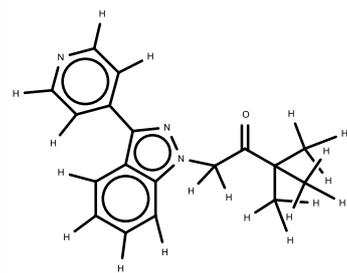
25



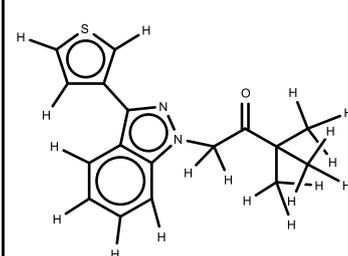
26



27



28



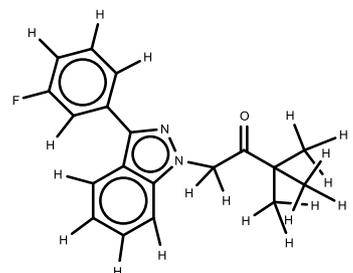
A 041

A 042

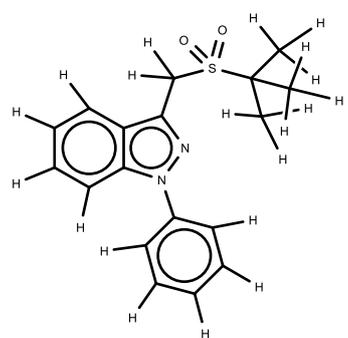
A 043

A 046

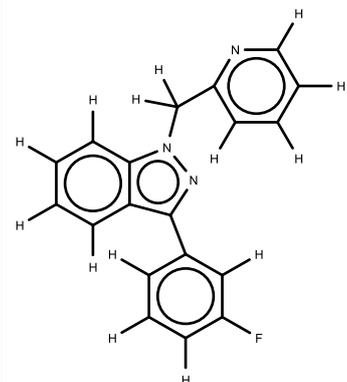
29



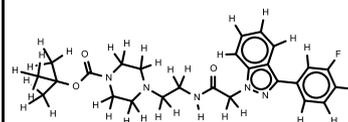
30



31



32



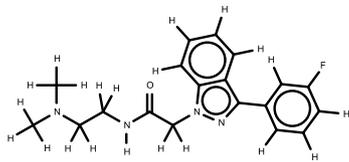
A 048

A 049

A 052

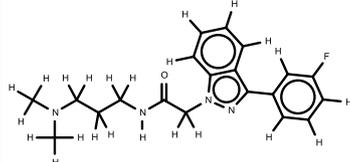
A 056

33



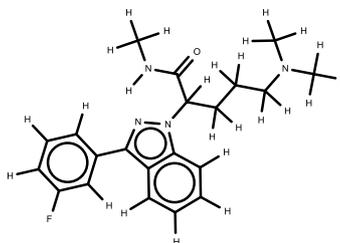
A 057

34



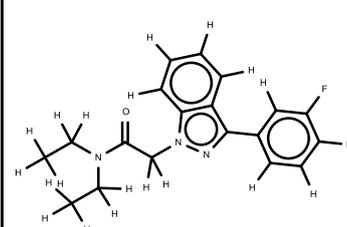
A 058

35



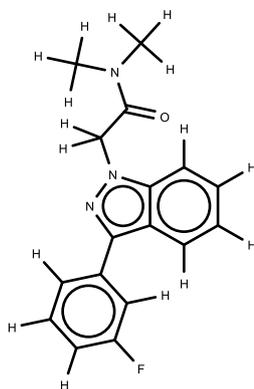
A 059

36



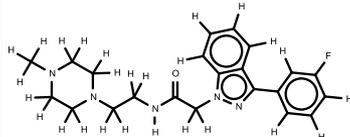
A 060

37



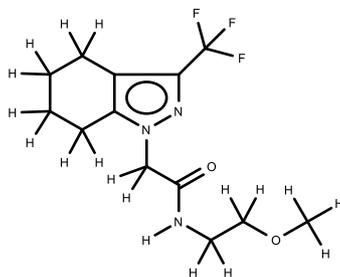
A 061

38



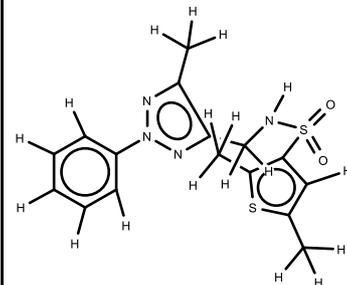
A 063

39



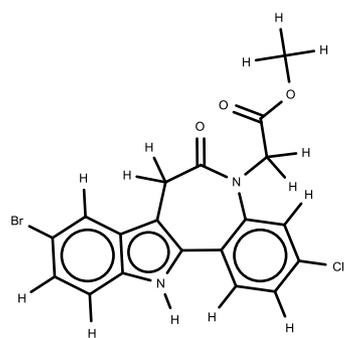
A 065

40



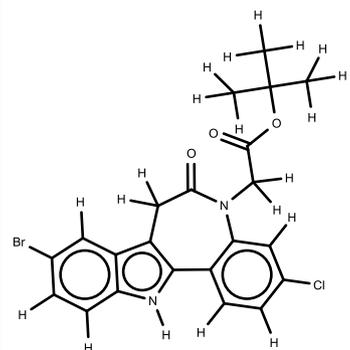
A 066

41



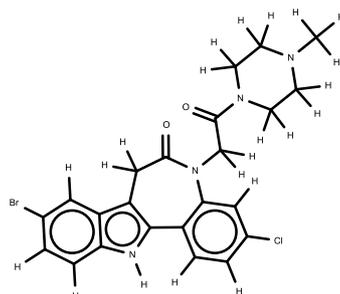
A 069

42



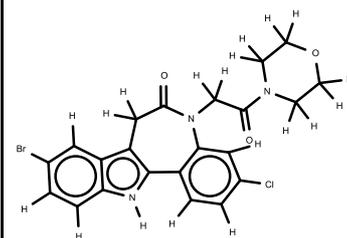
A 070

43



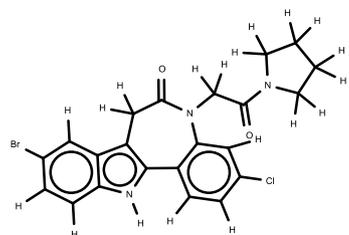
A 072

44



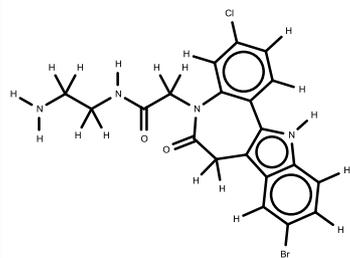
A 073

45



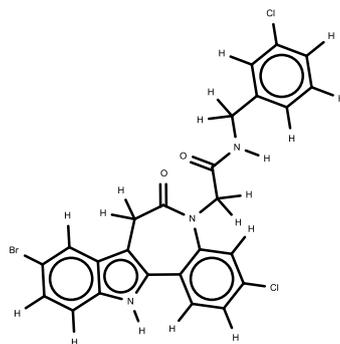
A 074

46



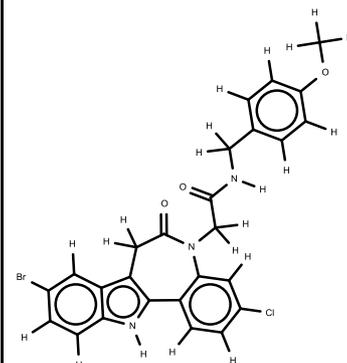
A 076

47



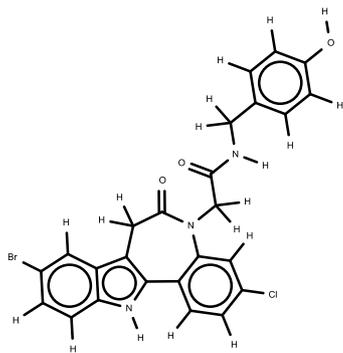
A 077

48



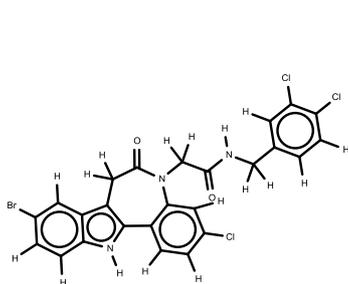
A 079

49



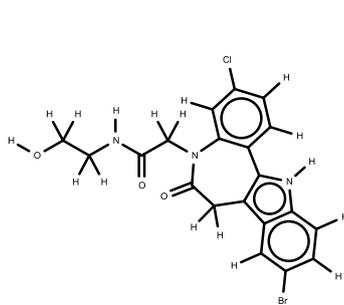
A 080

50



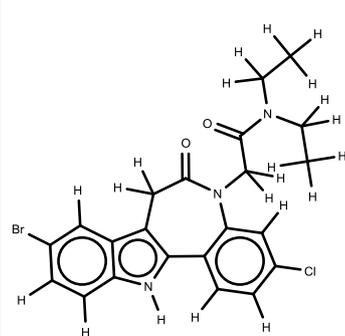
A 081

51



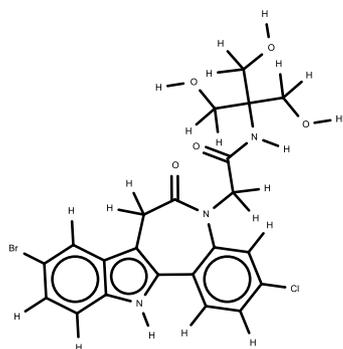
A 082

52



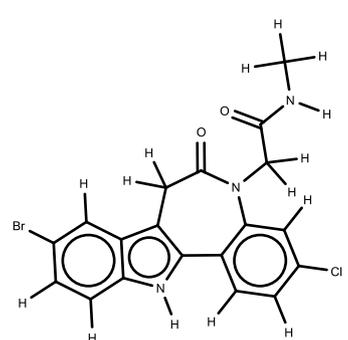
A 084

53



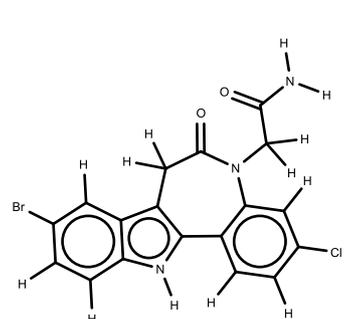
A 085

54



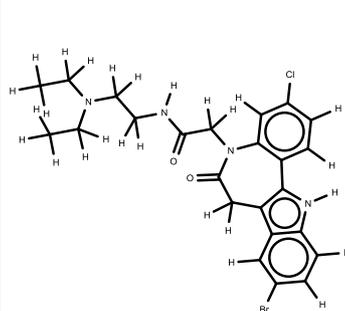
A 086

55



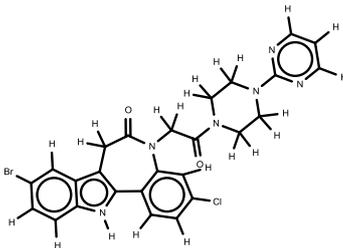
A 088

56



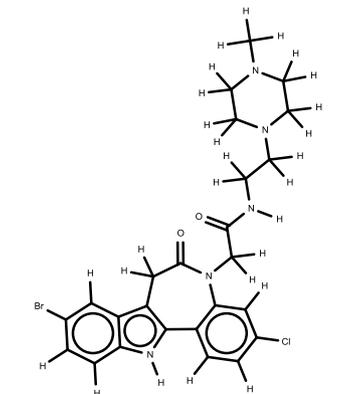
A 089

57



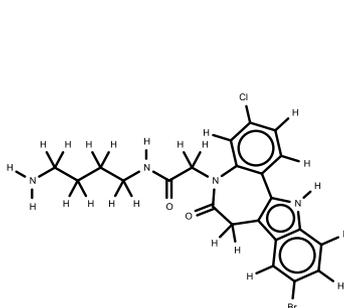
A 090

58



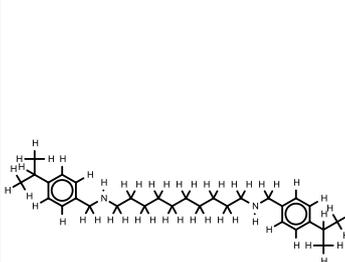
A 093

59



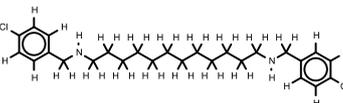
A 096

60



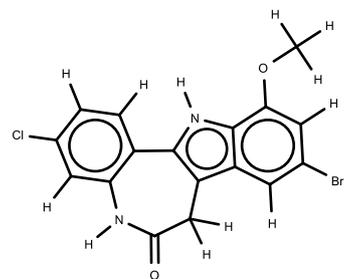
A 098

61



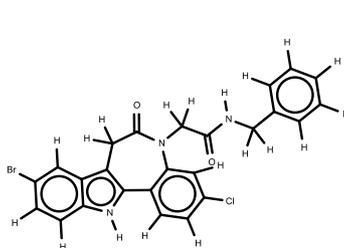
A 099

62



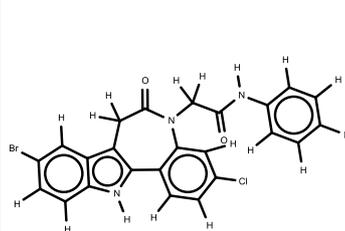
A 100

63

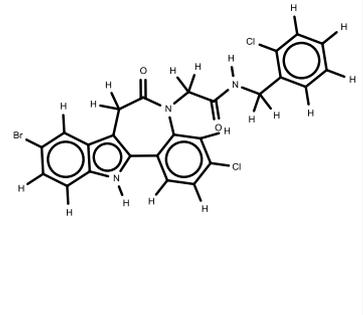
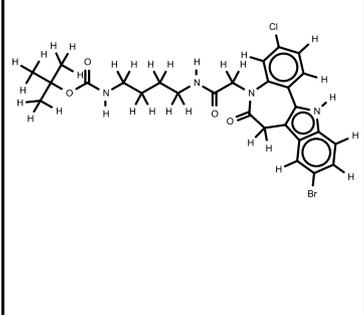
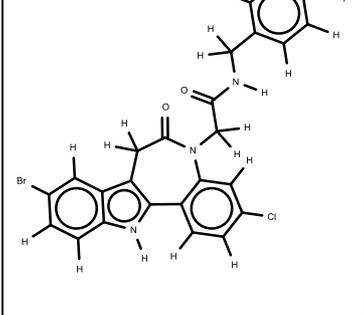
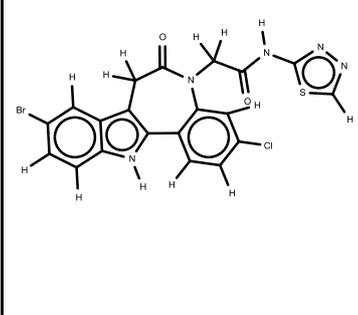
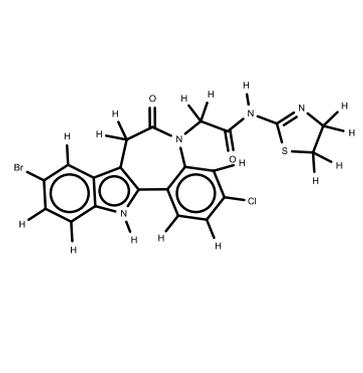
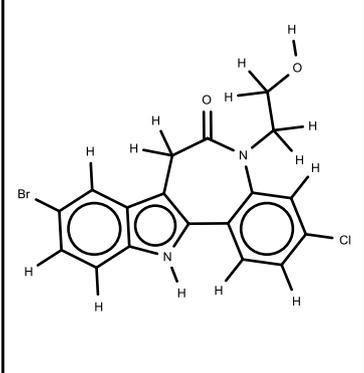
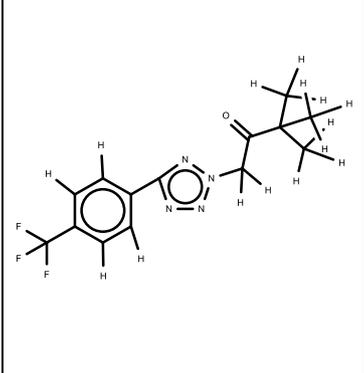
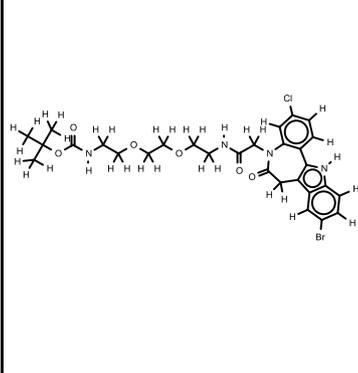
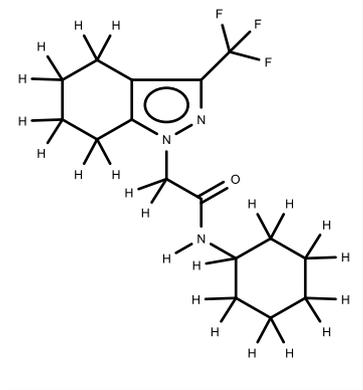
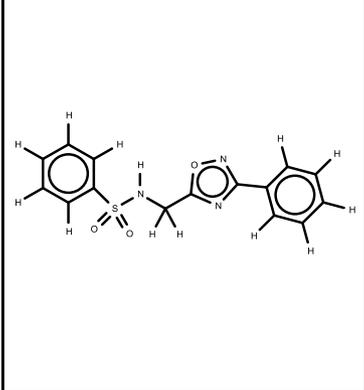
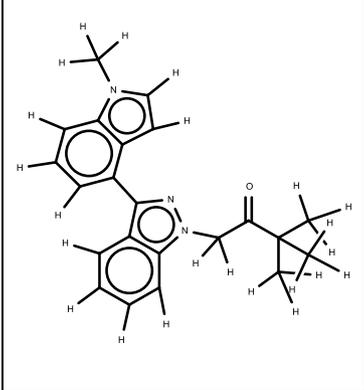
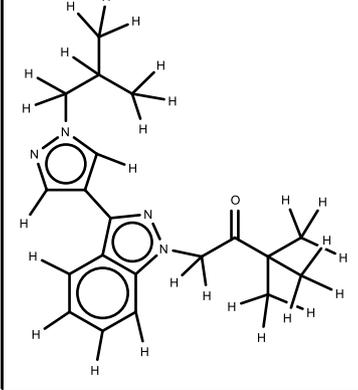
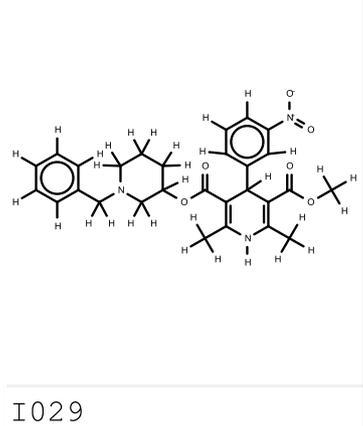
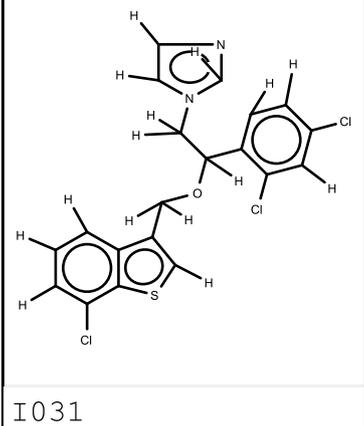
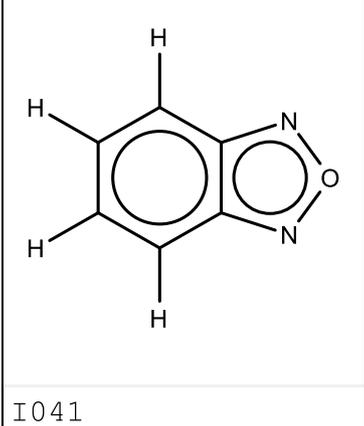
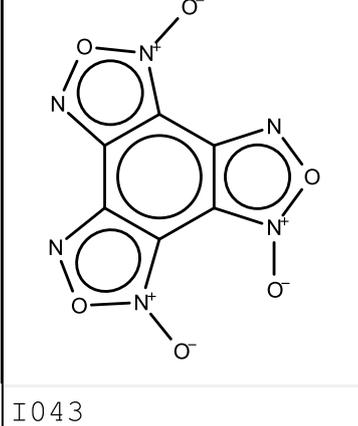


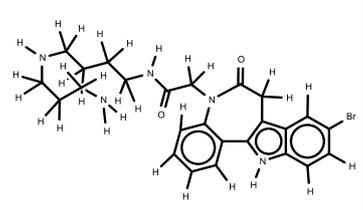
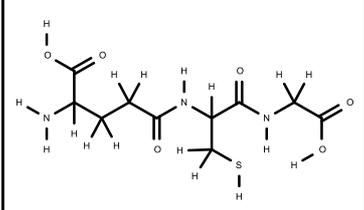
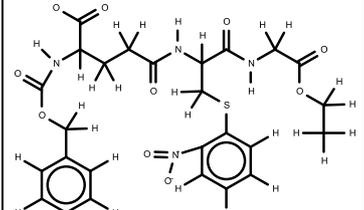
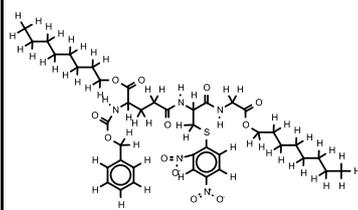
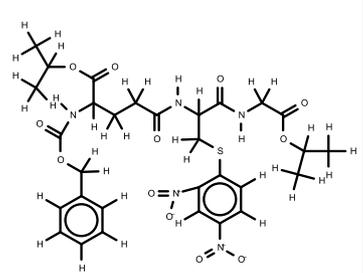
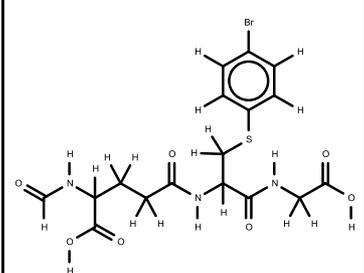
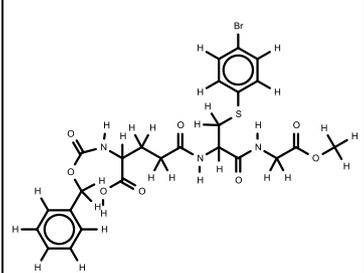
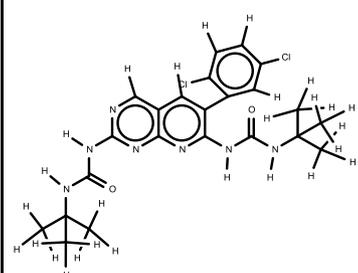
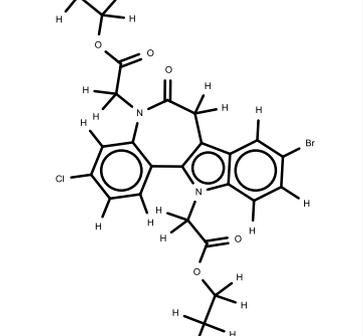
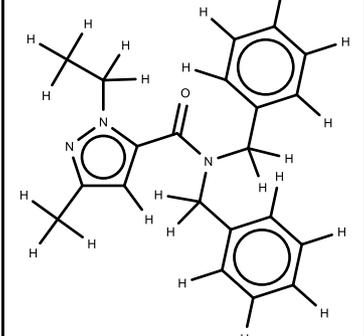
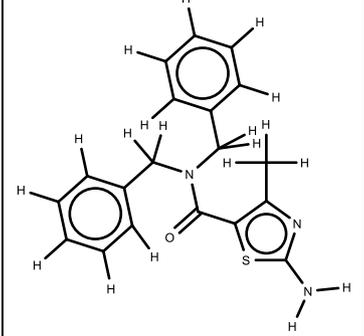
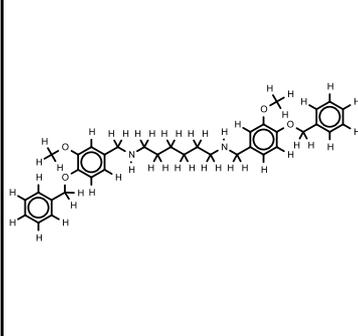
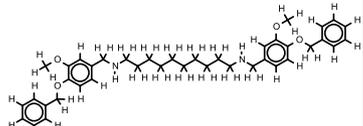
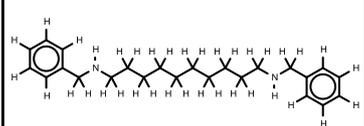
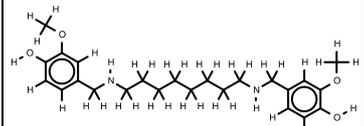
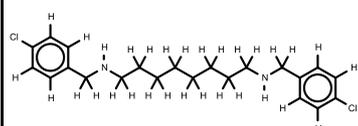
A 101

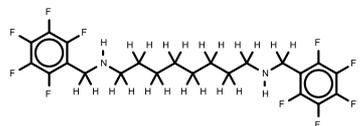
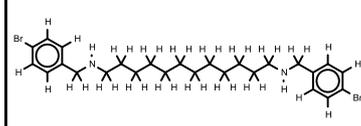
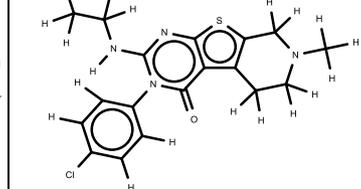
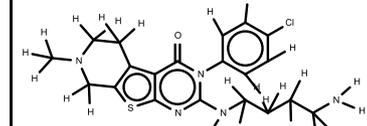
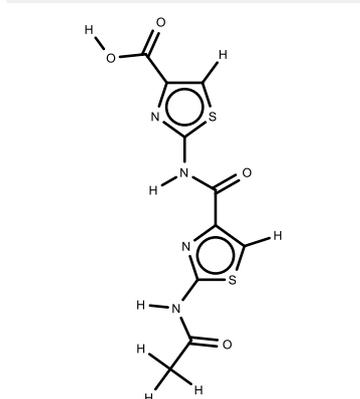
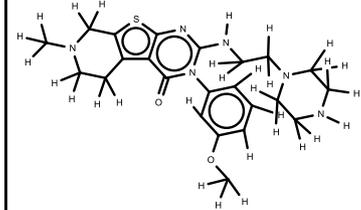
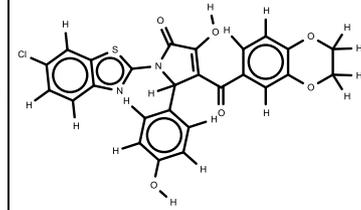
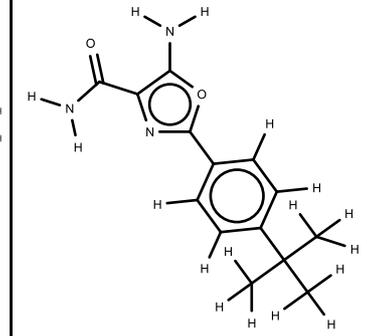
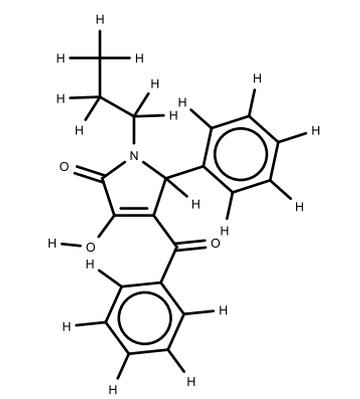
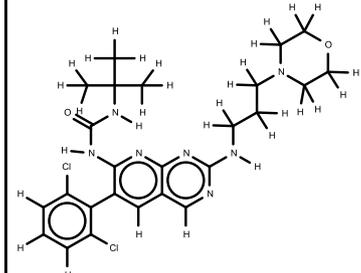
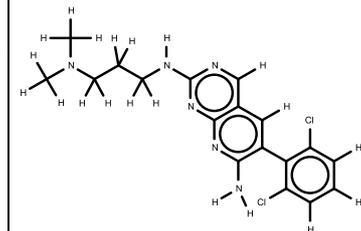
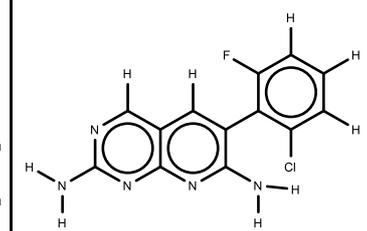
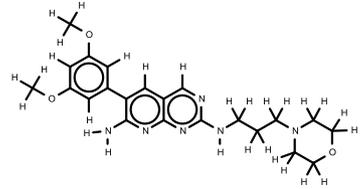
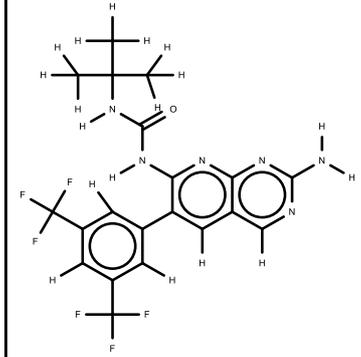
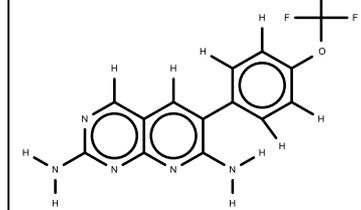
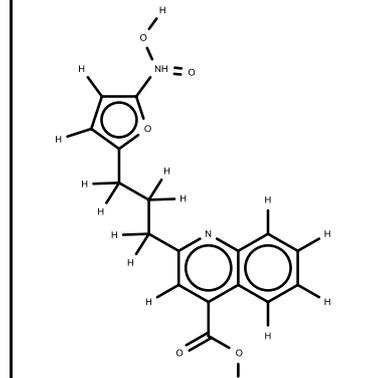
64



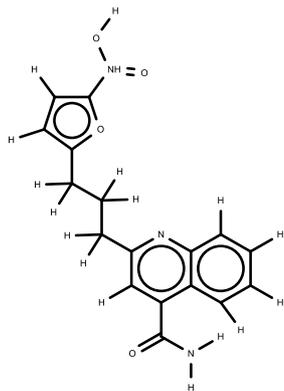
A 102

|   |  |   |  |
|---|--|---|--|
| <p style="text-align: center;"><b>65</b></p>  <p>A 104</p>  | <p style="text-align: center;"><b>66</b></p>  <p>A 106</p>  | <p style="text-align: center;"><b>67</b></p>  <p>A 107</p>  | <p style="text-align: center;"><b>68</b></p>  <p>A 108</p>  |
| <p style="text-align: center;"><b>69</b></p>  <p>A 109</p> | <p style="text-align: center;"><b>70</b></p>  <p>A 110</p> | <p style="text-align: center;"><b>71</b></p>  <p>I004</p>  | <p style="text-align: center;"><b>72</b></p>  <p>I006</p>  |
| <p style="text-align: center;"><b>73</b></p>  <p>I008</p> | <p style="text-align: center;"><b>74</b></p>  <p>I014</p> | <p style="text-align: center;"><b>75</b></p>  <p>I020</p> | <p style="text-align: center;"><b>76</b></p>  <p>I022</p> |
| <p style="text-align: center;"><b>77</b></p>  <p>I029</p> | <p style="text-align: center;"><b>78</b></p>  <p>I031</p> | <p style="text-align: center;"><b>79</b></p>  <p>I041</p> | <p style="text-align: center;"><b>80</b></p>  <p>I043</p> |

|   |  |   |  |
|---|--|---|--|
| <p style="text-align: center;"><b>81</b></p>  <p style="text-align: center;">I050</p>   | <p style="text-align: center;"><b>82</b></p>  <p style="text-align: center;">I053</p>   | <p style="text-align: center;"><b>83</b></p>  <p style="text-align: center;">I055</p>   | <p style="text-align: center;"><b>84</b></p>  <p style="text-align: center;">I060</p>   |
| <p style="text-align: center;"><b>85</b></p>  <p style="text-align: center;">I061</p>   | <p style="text-align: center;"><b>86</b></p>  <p style="text-align: center;">I066</p>   | <p style="text-align: center;"><b>87</b></p>  <p style="text-align: center;">I068</p>   | <p style="text-align: center;"><b>88</b></p>  <p style="text-align: center;">I071</p>   |
| <p style="text-align: center;"><b>89</b></p>  <p style="text-align: center;">I079</p> | <p style="text-align: center;"><b>90</b></p>  <p style="text-align: center;">I082</p> | <p style="text-align: center;"><b>91</b></p>  <p style="text-align: center;">I083</p> | <p style="text-align: center;"><b>92</b></p>  <p style="text-align: center;">I093</p> |
| <p style="text-align: center;"><b>93</b></p>  <p style="text-align: center;">I096</p> | <p style="text-align: center;"><b>94</b></p>  <p style="text-align: center;">I101</p> | <p style="text-align: center;"><b>95</b></p>  <p style="text-align: center;">I106</p> | <p style="text-align: center;"><b>96</b></p>  <p style="text-align: center;">I112</p> |

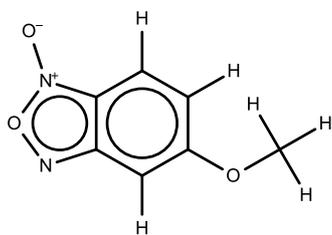
|  |   |  |   |
|--|---|--|---|
| <p style="text-align: center;"><b>97</b></p>     | <p style="text-align: center;"><b>98</b></p>     | <p style="text-align: center;"><b>99</b></p>     | <p style="text-align: center;"><b>100</b></p>    |
| <p>I115</p>  | <p>I121</p>   | <p>I127</p>  | <p>I128</p>   |
| <p style="text-align: center;"><b>101</b></p>    | <p style="text-align: center;"><b>102</b></p>    | <p style="text-align: center;"><b>103</b></p>    | <p style="text-align: center;"><b>104</b></p>    |
| <p>I130</p>  | <p>I134</p>   | <p>I140</p>  | <p>I141</p>   |
| <p style="text-align: center;"><b>105</b></p>  | <p style="text-align: center;"><b>106</b></p>  | <p style="text-align: center;"><b>107</b></p>  | <p style="text-align: center;"><b>108</b></p>  |
| <p>I142</p>  | <p>I146</p>   | <p>I147</p>  | <p>I151</p>   |
| <p style="text-align: center;"><b>109</b></p>  | <p style="text-align: center;"><b>110</b></p>  | <p style="text-align: center;"><b>111</b></p>  | <p style="text-align: center;"><b>112</b></p>  |
| <p>I153</p>  | <p>I155</p>   | <p>I160</p>  | <p>I165</p>   |

113



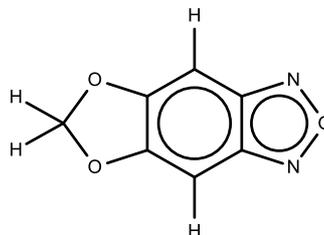
I166

114



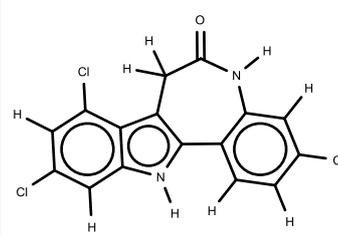
I177

115



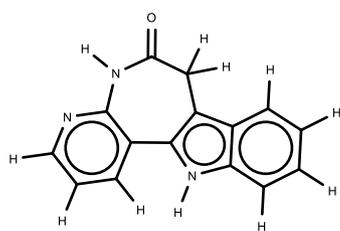
I183

116



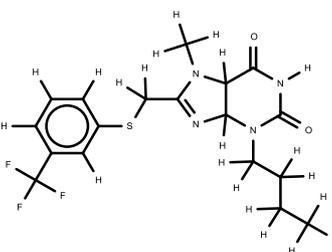
I190

117



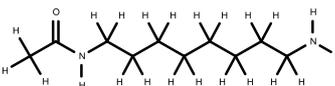
I191

118



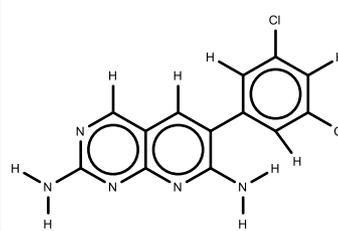
I195

119



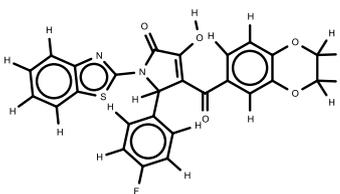
I200

120



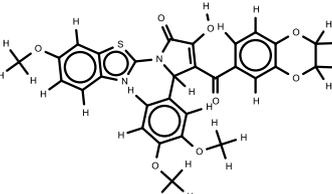
I204

121



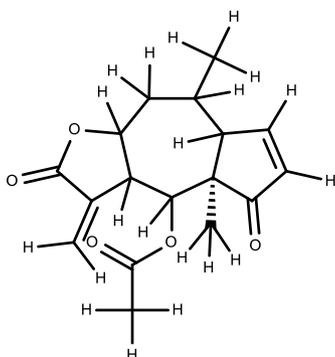
I207

122



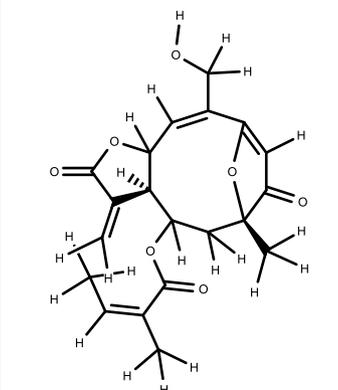
I211

123



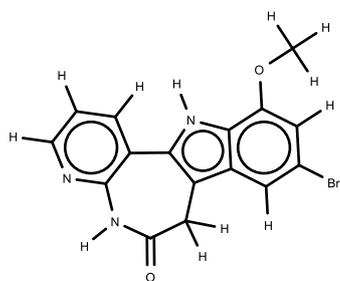
I223

124



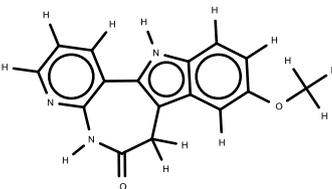
I224

125



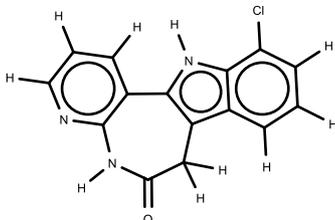
I230

126



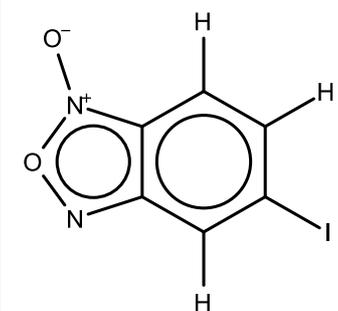
I234

127



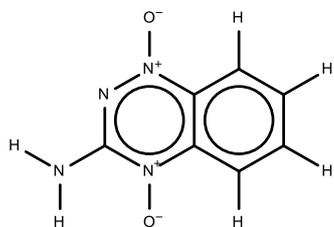
I235

128



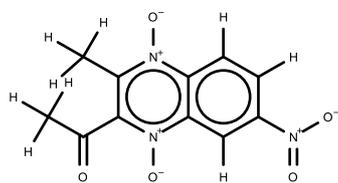
I239

129



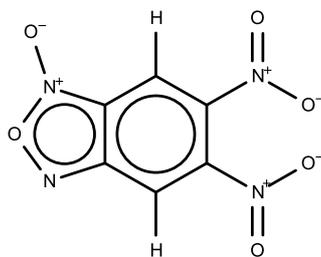
I240

130



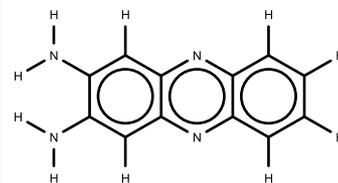
I242

131



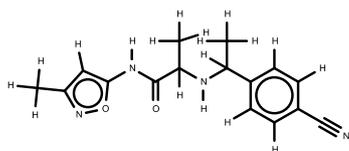
I247

132



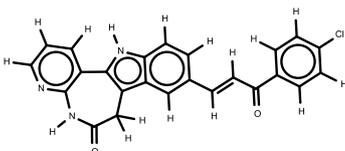
I254

133



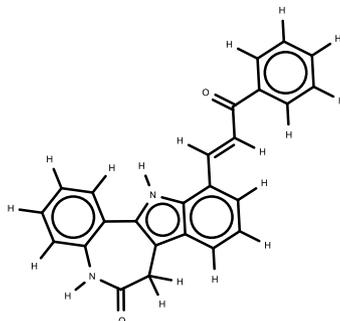
I258

134



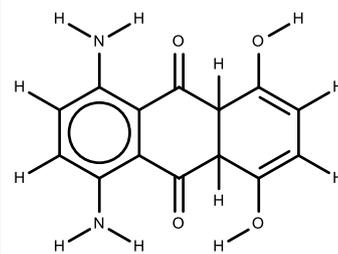
I270

135



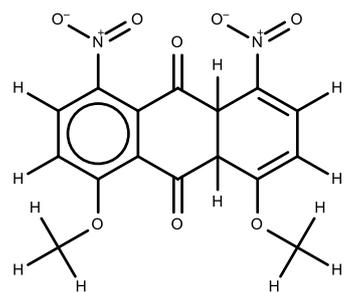
I275

136



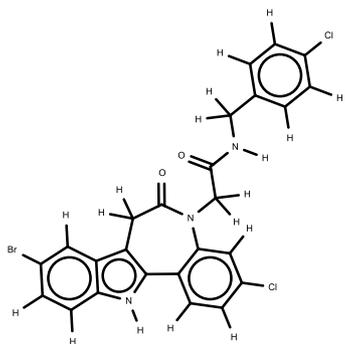
I276

137



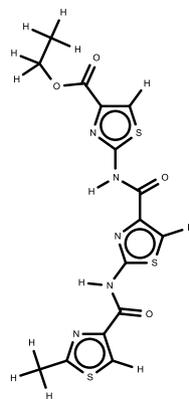
I281

138



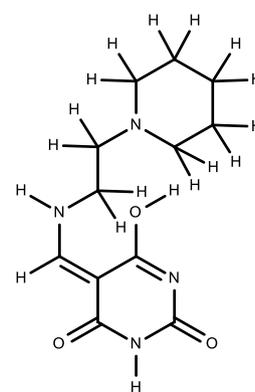
I282

139

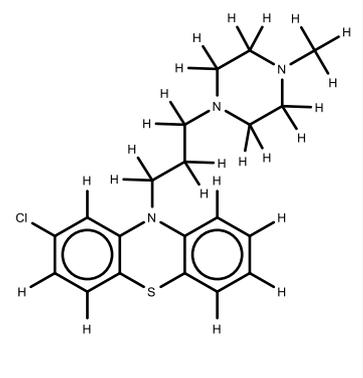
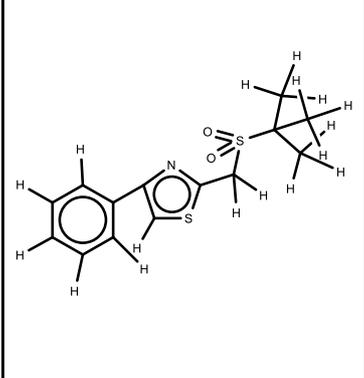
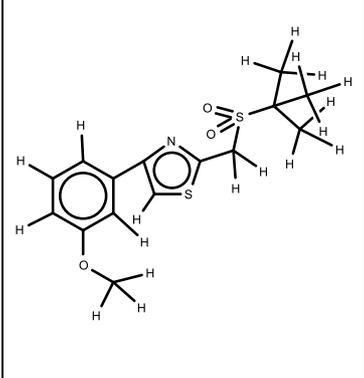
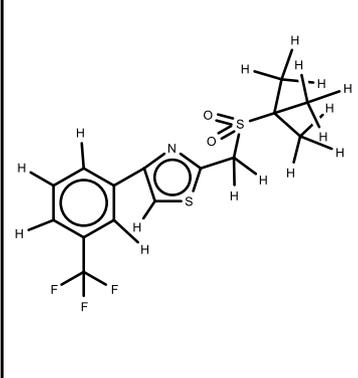
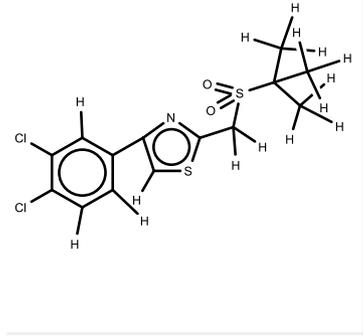
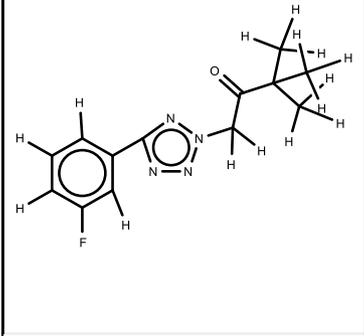
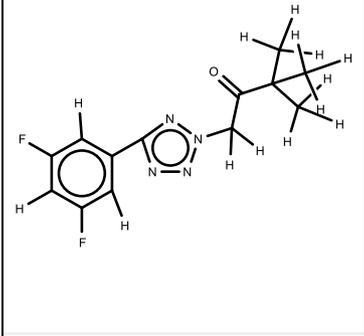
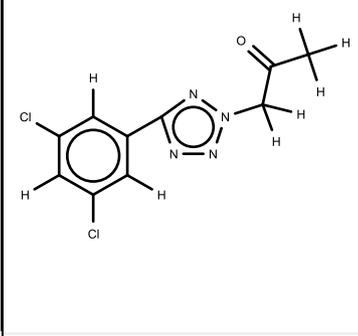
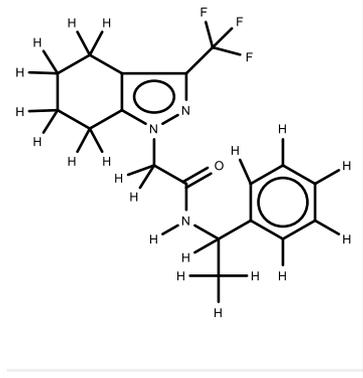
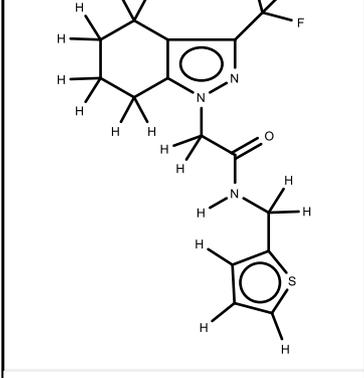
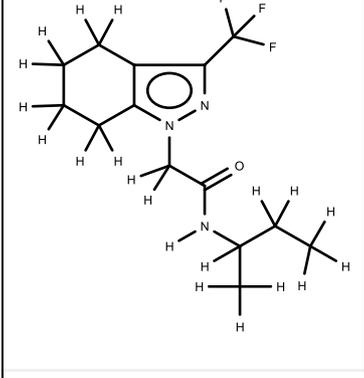
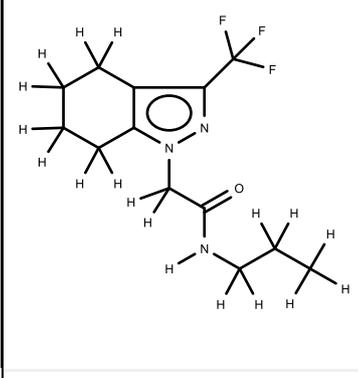
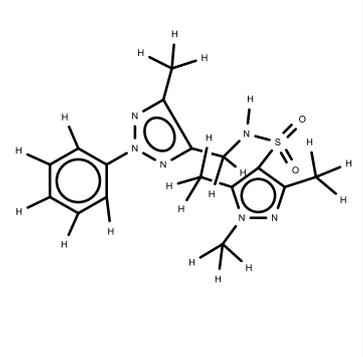
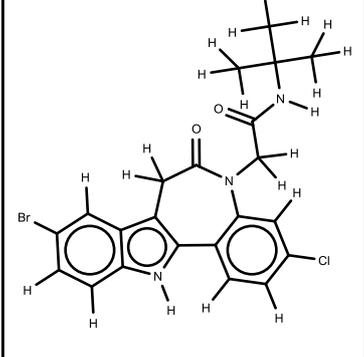
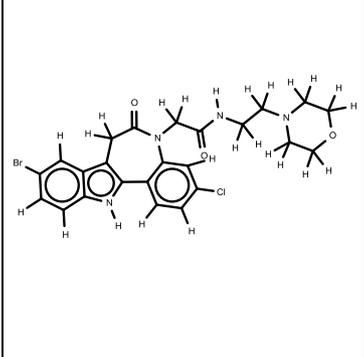
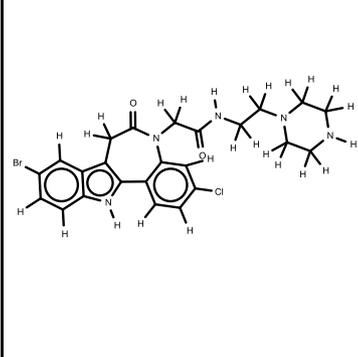


I287

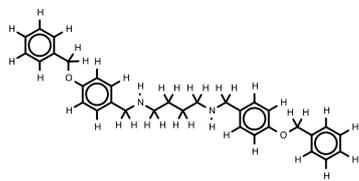
140



I290

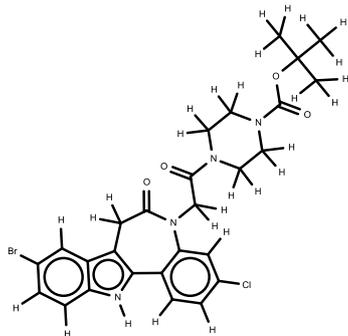
|   |  |   |  |
|---|--|---|--|
| <p style="text-align: center;"><b>1</b></p>     | <p style="text-align: center;"><b>2</b></p>     | <p style="text-align: center;"><b>3</b></p>     | <p style="text-align: center;"><b>4</b></p>     |
| A 001   | A 002  | A 005   | A 006  |
| <p style="text-align: center;"><b>5</b></p>    | <p style="text-align: center;"><b>6</b></p>    | <p style="text-align: center;"><b>7</b></p>    | <p style="text-align: center;"><b>8</b></p>    |
| A 009   | A 012  | A 014   | A 016  |
| <p style="text-align: center;"><b>9</b></p>   | <p style="text-align: center;"><b>10</b></p>  | <p style="text-align: center;"><b>11</b></p>  | <p style="text-align: center;"><b>12</b></p>  |
| A 019   | A 021  | A 023   | A 024  |
| <p style="text-align: center;"><b>13</b></p>  | <p style="text-align: center;"><b>14</b></p>  | <p style="text-align: center;"><b>15</b></p>  | <p style="text-align: center;"><b>16</b></p>  |
| A 028   | A 083  | A 091   | A 095  |

17



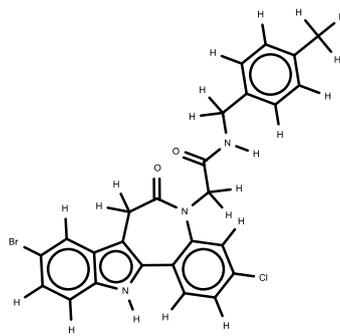
A 097

18



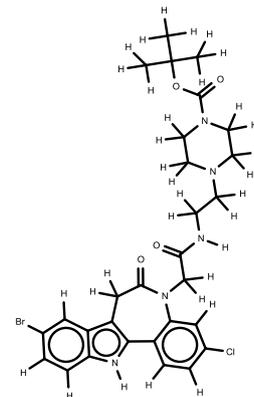
A 103

19



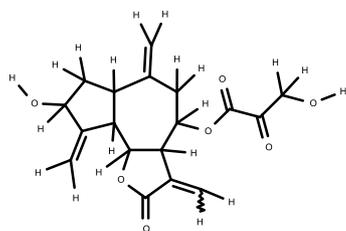
A 105

20



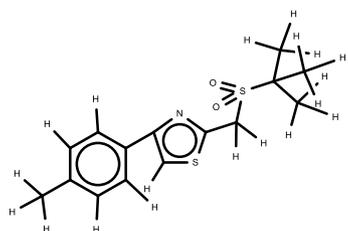
A 111

21



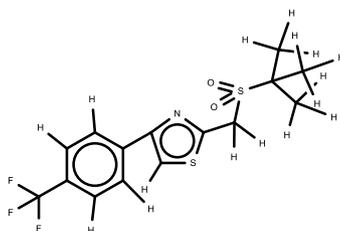
I001

22



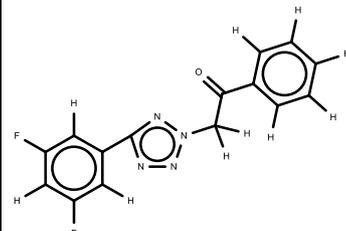
I002

23



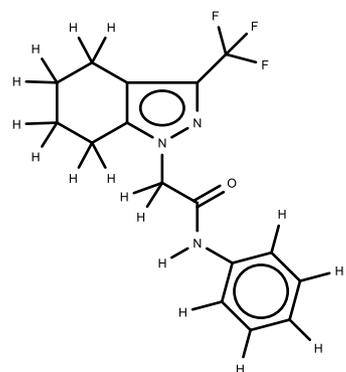
I003

24



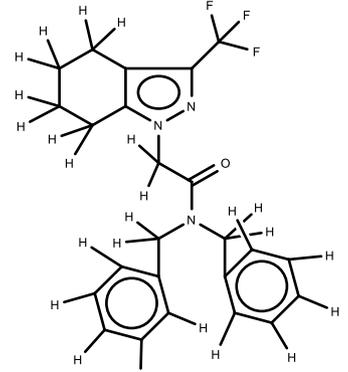
I005

25



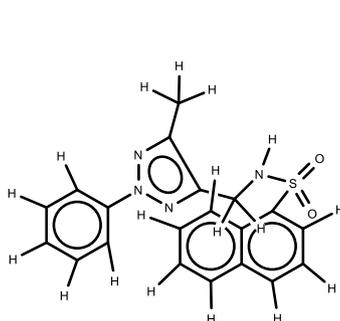
I007

26



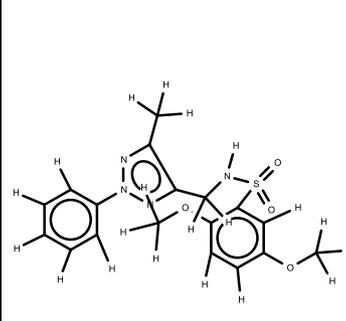
I009

27



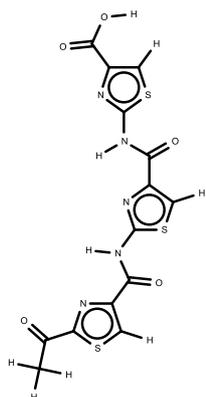
I010

28



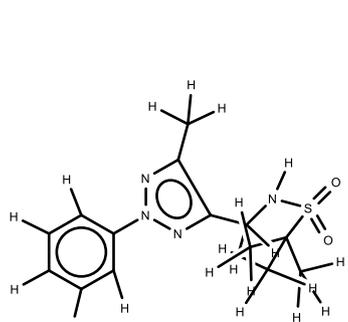
I011

29



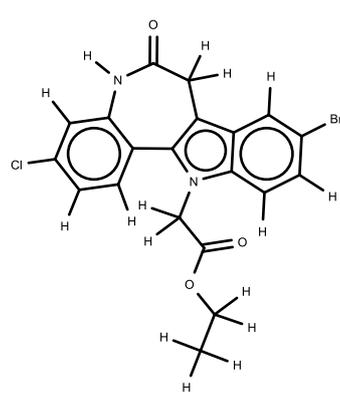
I012

30



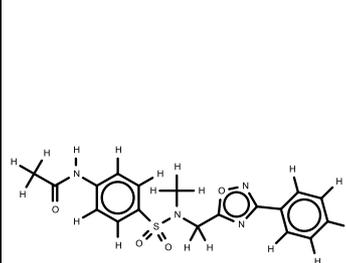
I013

31



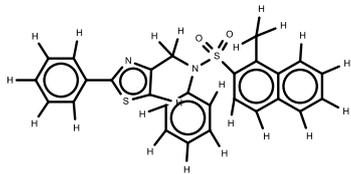
I015

32



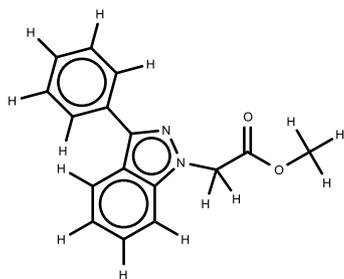
I016

33



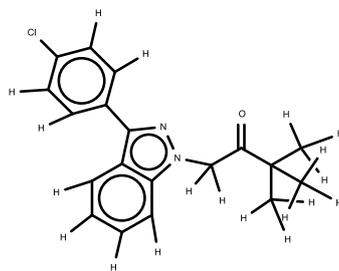
I017

34



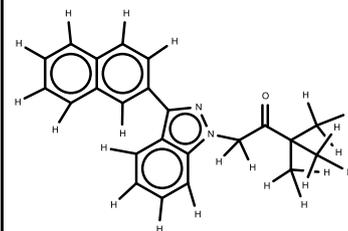
I018

35



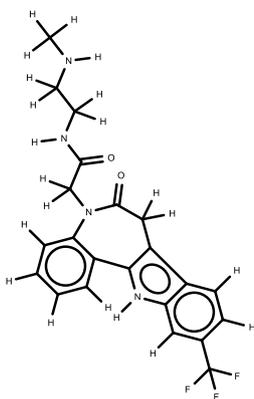
I019

36



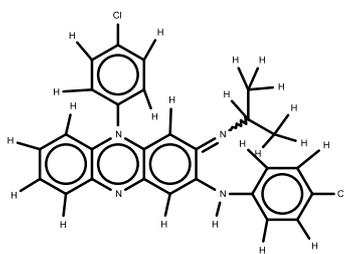
I021

37



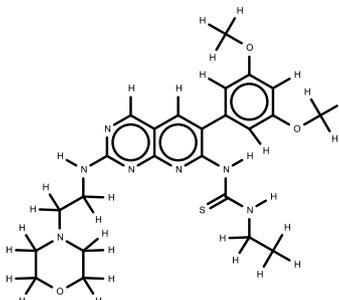
I027

38



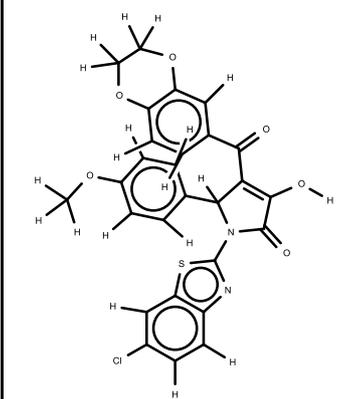
I030

39



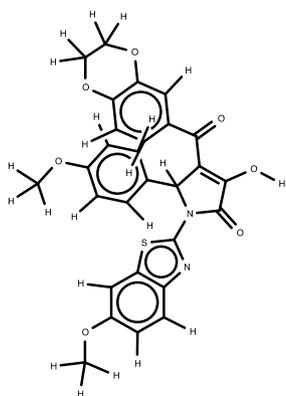
I035

40



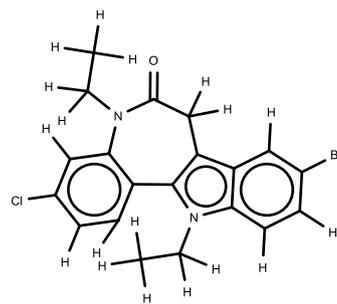
I037

41



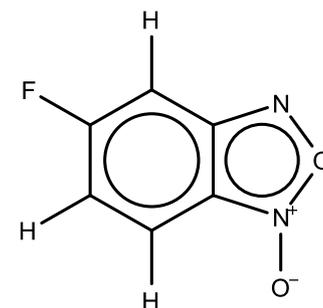
I039

42



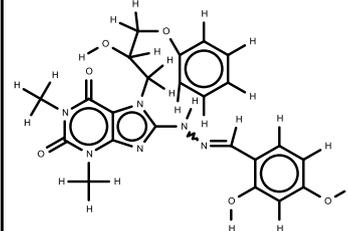
I040

43



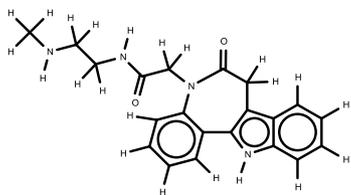
I044

44



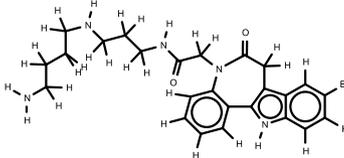
I045

45



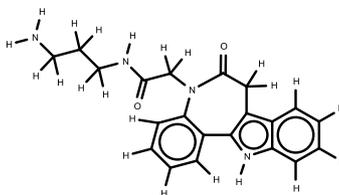
I047

46



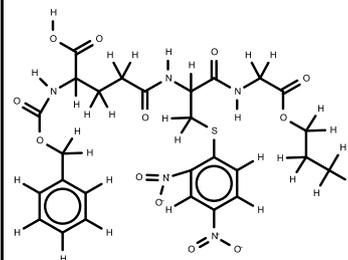
I051

47



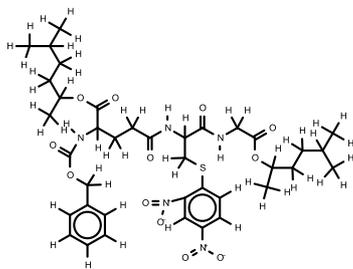
I052

48



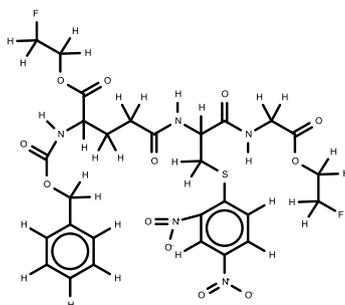
I056

49



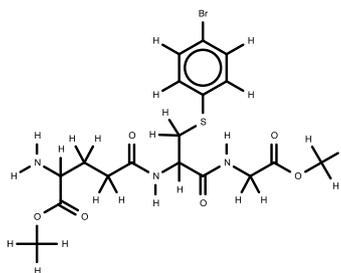
I062

50



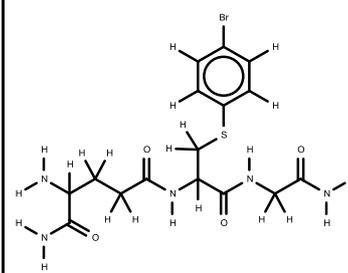
I063

51



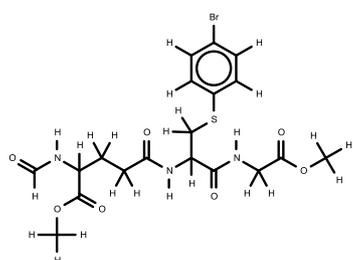
I064

52



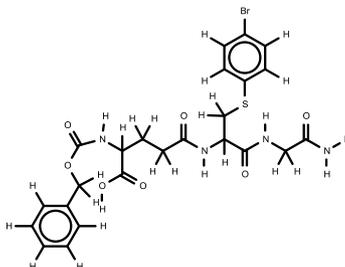
I065

53



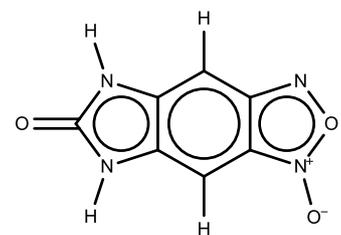
I067

54



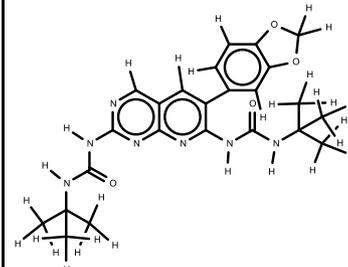
I069

55



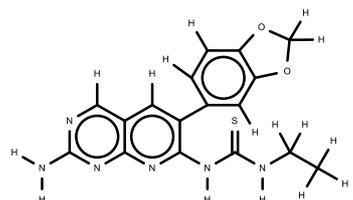
I070

56



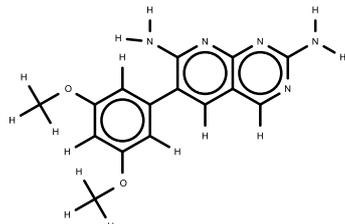
I072

57



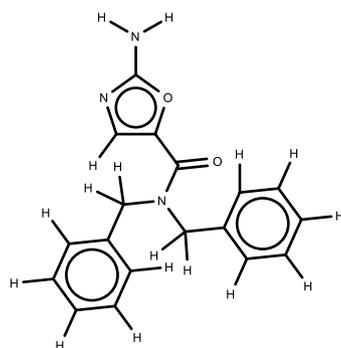
I073

58



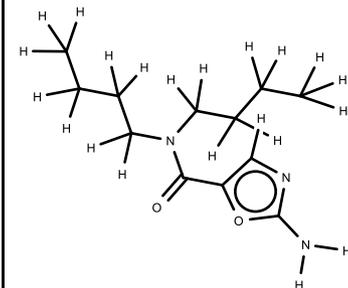
I074

59



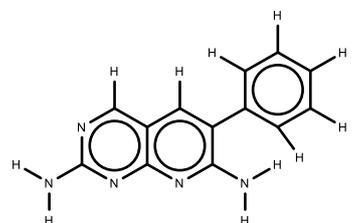
I075

60



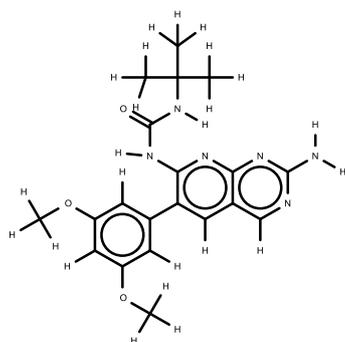
I076

61



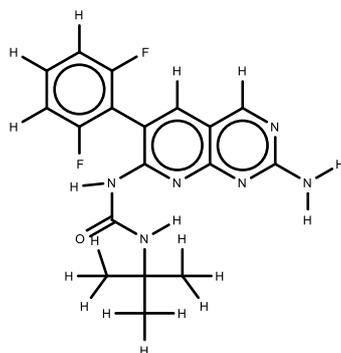
I077

62



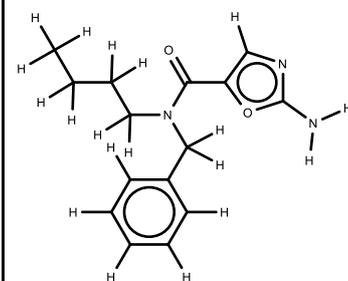
I078

63

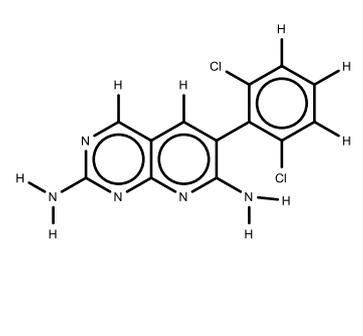
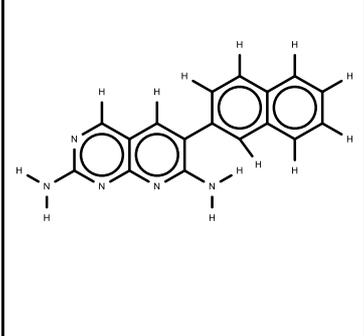
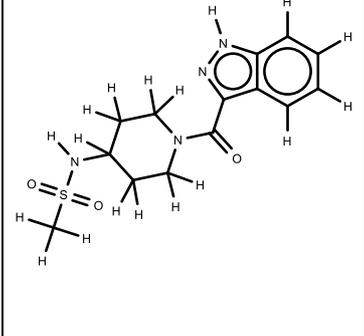
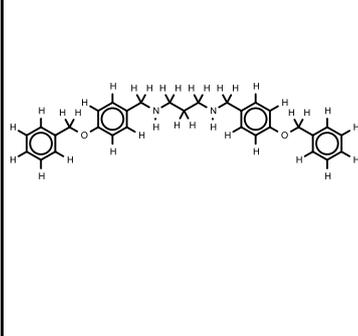
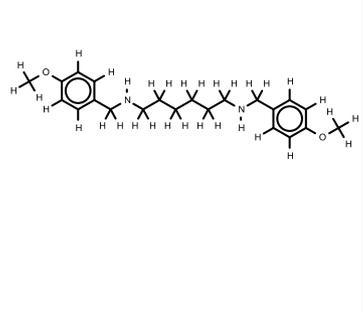
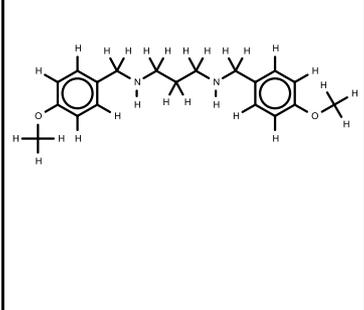
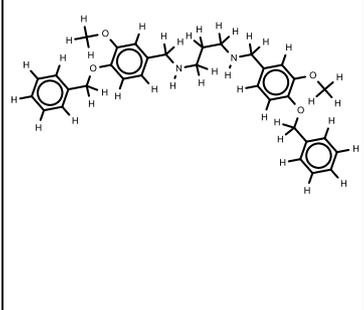
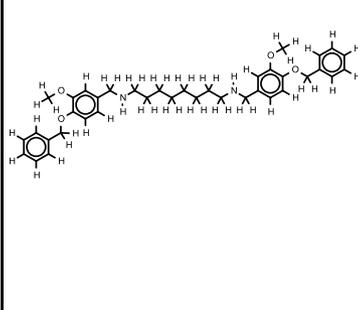
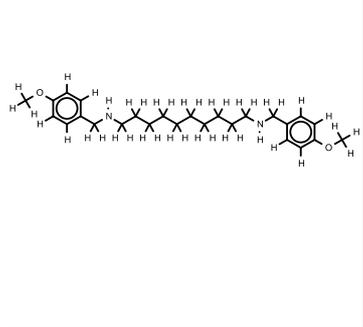
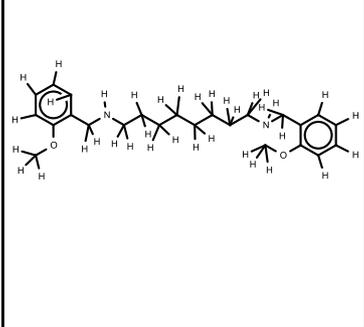
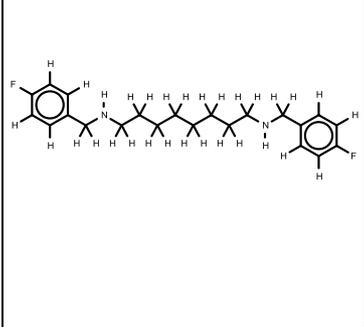
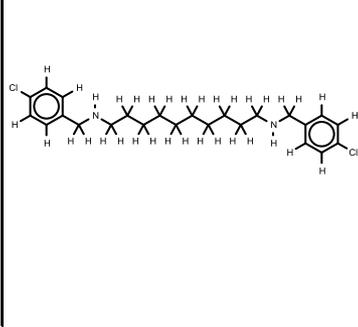
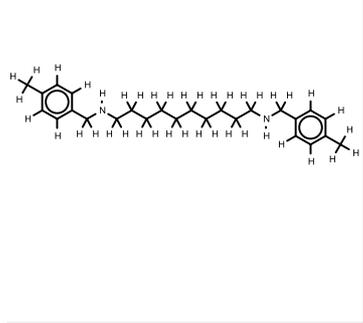
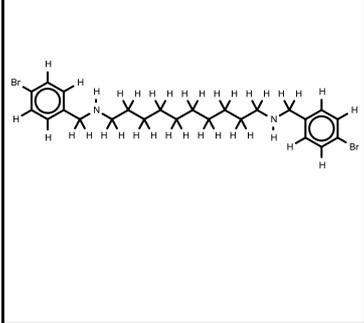
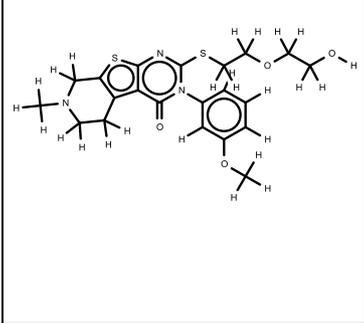
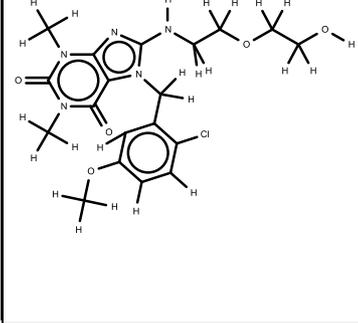


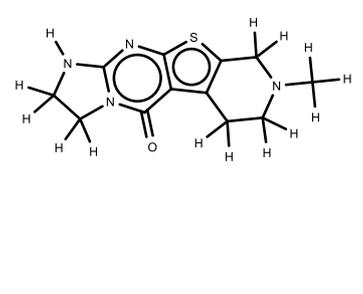
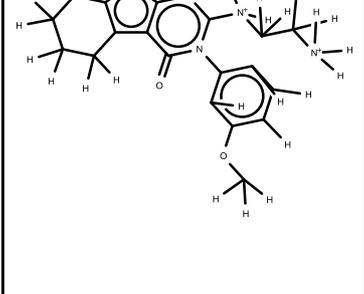
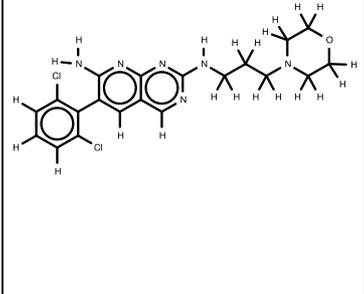
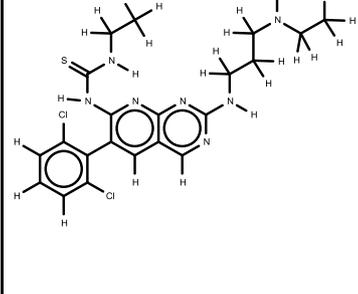
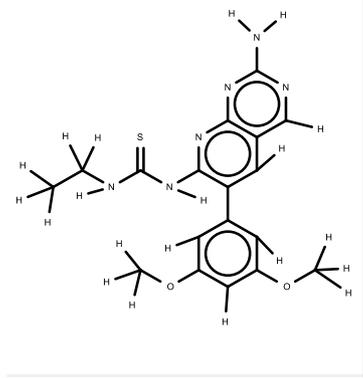
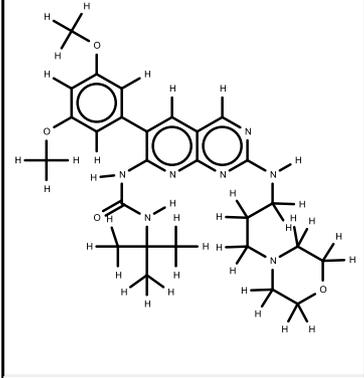
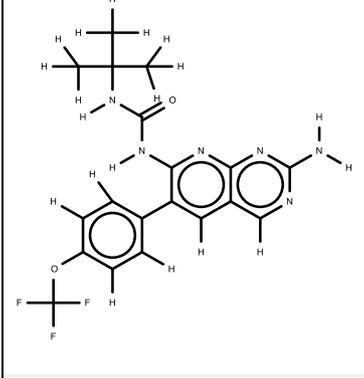
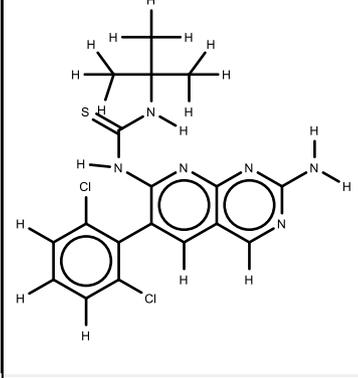
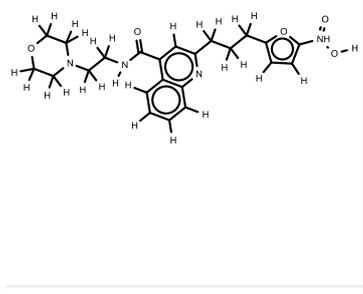
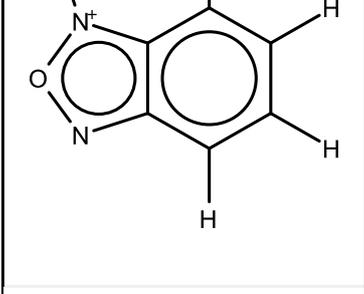
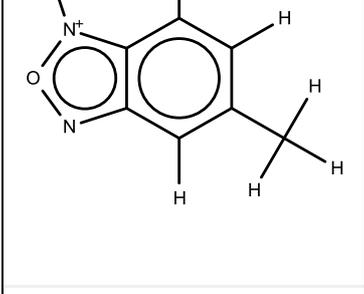
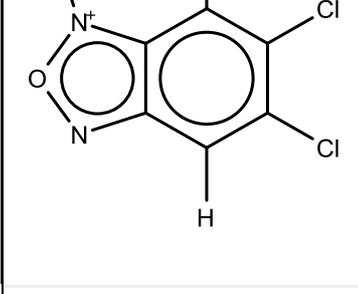
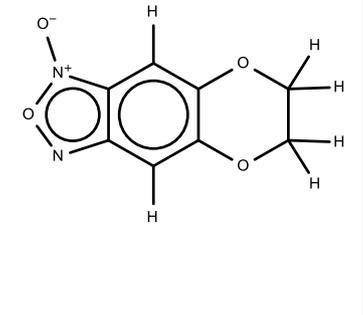
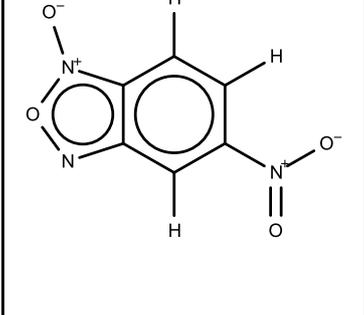
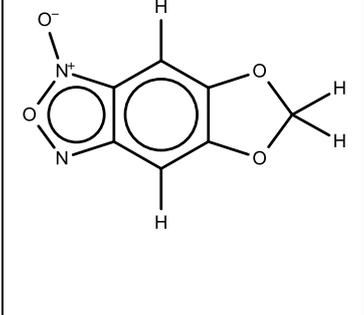
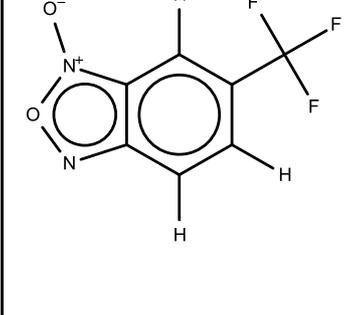
I080

64

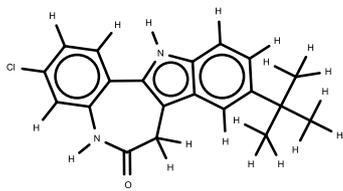


I081

|   |  |   |  |
|---|--|---|--|
| <p style="text-align: center;"><b>65</b></p>  <p style="text-align: center;">I084</p>   | <p style="text-align: center;"><b>66</b></p>  <p style="text-align: center;">I085</p>   | <p style="text-align: center;"><b>67</b></p>  <p style="text-align: center;">I088</p>   | <p style="text-align: center;"><b>68</b></p>  <p style="text-align: center;">I089</p>   |
| <p style="text-align: center;"><b>69</b></p>  <p style="text-align: center;">I091</p>  | <p style="text-align: center;"><b>70</b></p>  <p style="text-align: center;">I097</p>  | <p style="text-align: center;"><b>71</b></p>  <p style="text-align: center;">I098</p>  | <p style="text-align: center;"><b>72</b></p>  <p style="text-align: center;">I102</p>  |
| <p style="text-align: center;"><b>73</b></p>  <p style="text-align: center;">I103</p> | <p style="text-align: center;"><b>74</b></p>  <p style="text-align: center;">I107</p> | <p style="text-align: center;"><b>75</b></p>  <p style="text-align: center;">I111</p> | <p style="text-align: center;"><b>76</b></p>  <p style="text-align: center;">I118</p> |
| <p style="text-align: center;"><b>77</b></p>  <p style="text-align: center;">I119</p> | <p style="text-align: center;"><b>78</b></p>  <p style="text-align: center;">I120</p> | <p style="text-align: center;"><b>79</b></p>  <p style="text-align: center;">I124</p> | <p style="text-align: center;"><b>80</b></p>  <p style="text-align: center;">I129</p> |

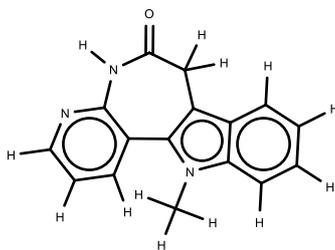
|   |  |   |  |
|---|--|---|--|
| <p style="text-align: center;"><b>81</b></p>    | <p style="text-align: center;"><b>82</b></p>    | <p style="text-align: center;"><b>83</b></p>    | <p style="text-align: center;"><b>84</b></p>    |
| I132  | I135   | I145  | I148   |
| <p style="text-align: center;"><b>85</b></p>   | <p style="text-align: center;"><b>86</b></p>   | <p style="text-align: center;"><b>87</b></p>   | <p style="text-align: center;"><b>88</b></p>   |
| I150  | I156   | I161  | I162   |
| <p style="text-align: center;"><b>89</b></p>  | <p style="text-align: center;"><b>90</b></p>  | <p style="text-align: center;"><b>91</b></p>  | <p style="text-align: center;"><b>92</b></p>  |
| I169  | I172   | I173  | I176   |
| <p style="text-align: center;"><b>93</b></p>  | <p style="text-align: center;"><b>94</b></p>  | <p style="text-align: center;"><b>95</b></p>  | <p style="text-align: center;"><b>96</b></p>  |
| I178  | I179   | I184  | I186   |

97



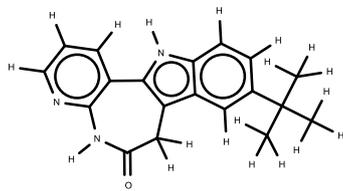
I188

98



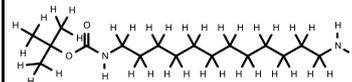
I192

99



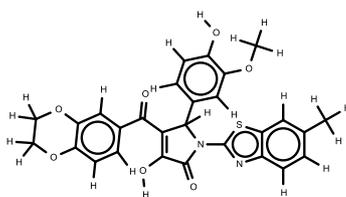
I193

100



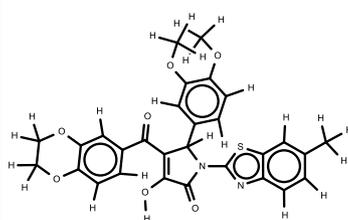
I201

101



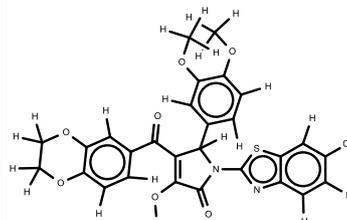
I208

102



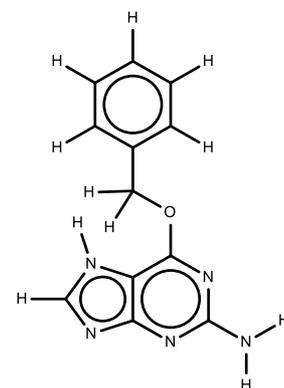
I209

103



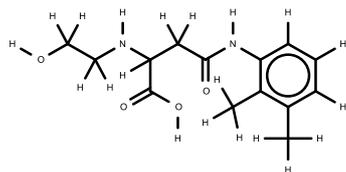
I213

104



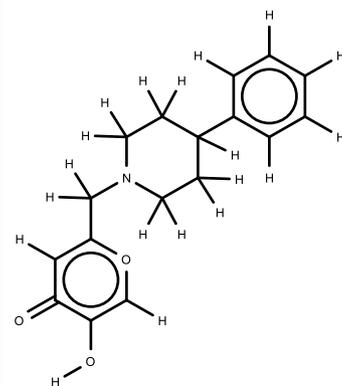
I214

105



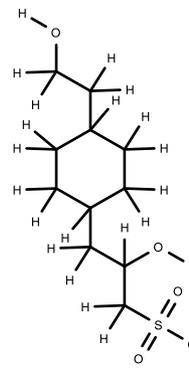
I217

106



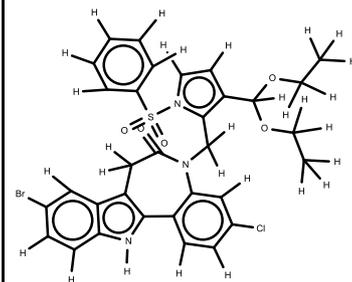
I221

107



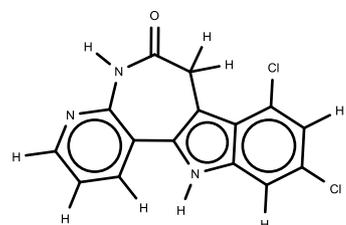
I222

108



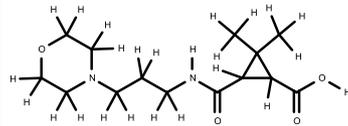
I228

109



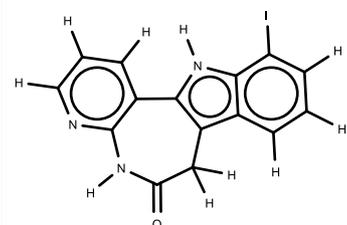
I229

110



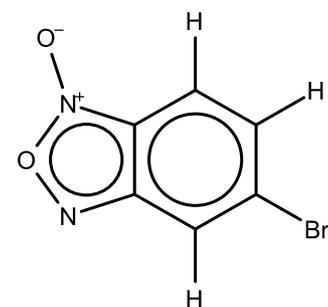
I231

111



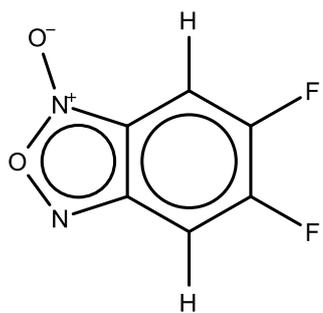
I237

112



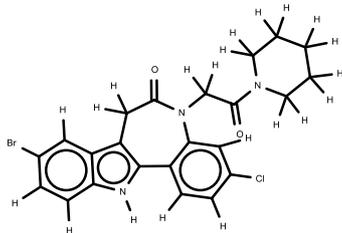
I245

113



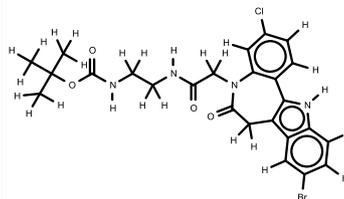
I251

114



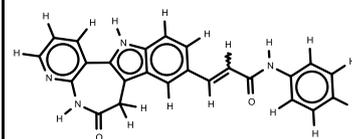
I255

115



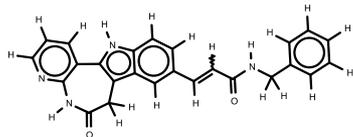
I256

116



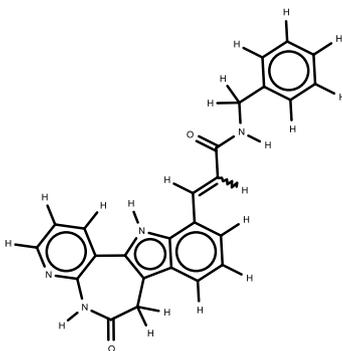
I259

117



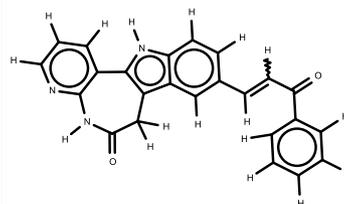
I261

118



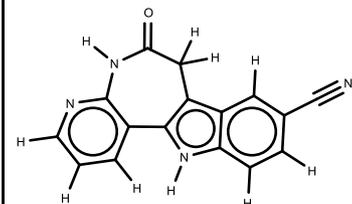
I262

119



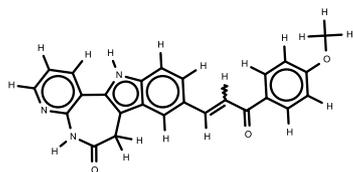
I263

120



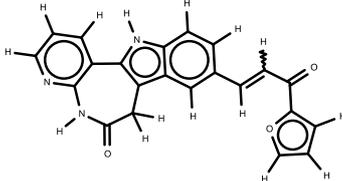
I266

121



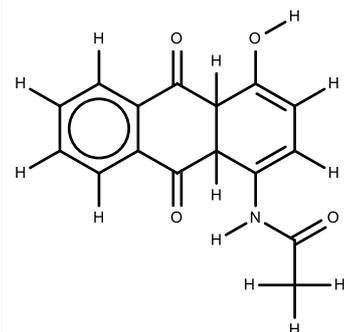
I269

122



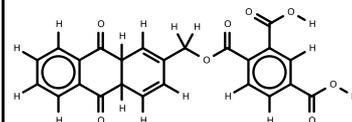
I271

123



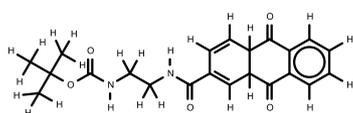
I277

124



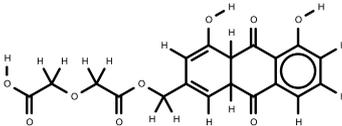
I278

125



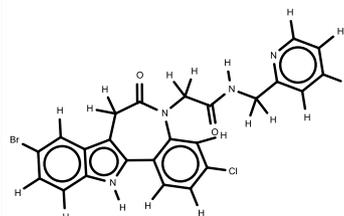
I279

126



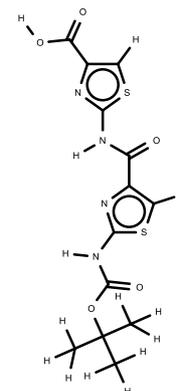
I280

127



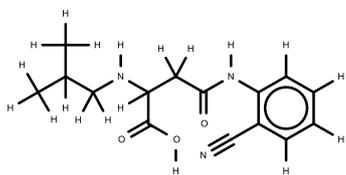
I283

128

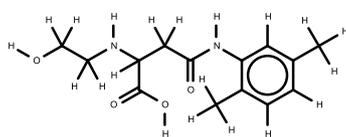


I284

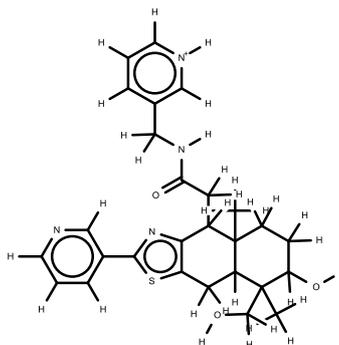
129



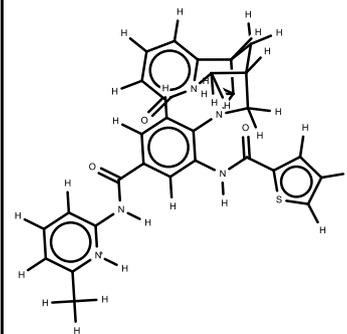
130



131

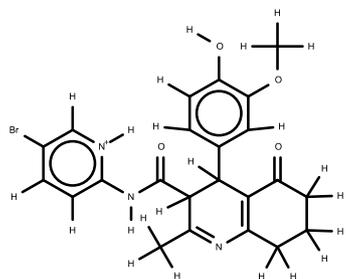


132



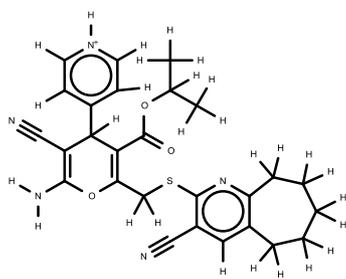
I288

133

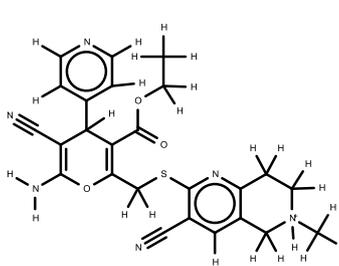


I289

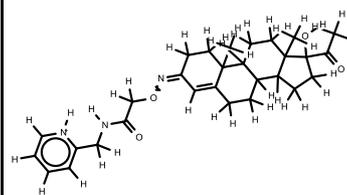
134



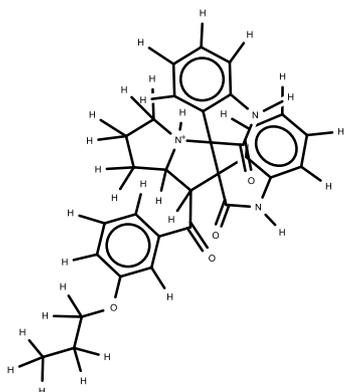
135



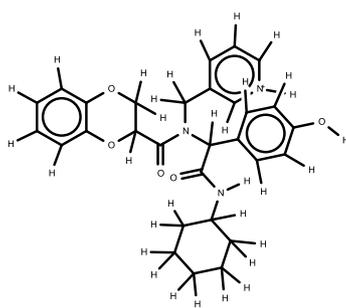
136



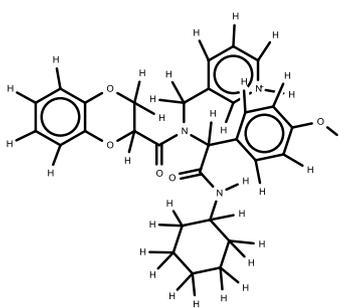
137



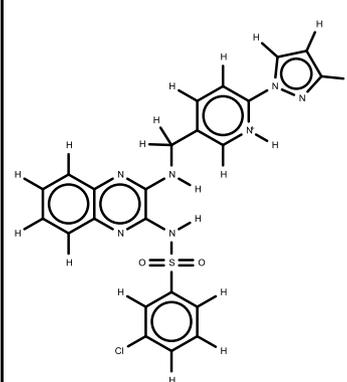
138



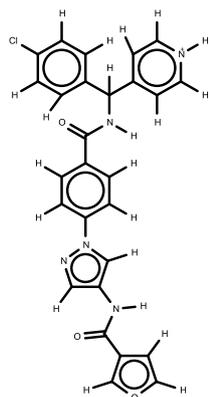
139



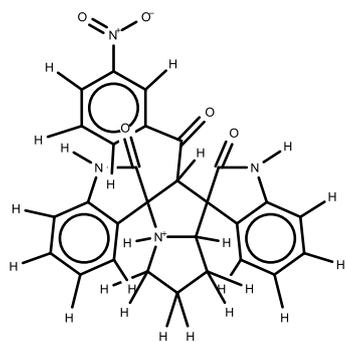
140



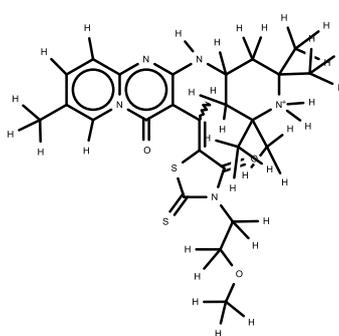
141



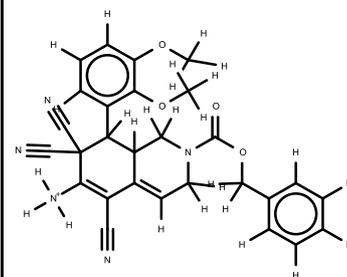
142



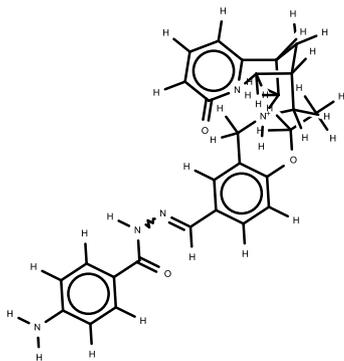
143



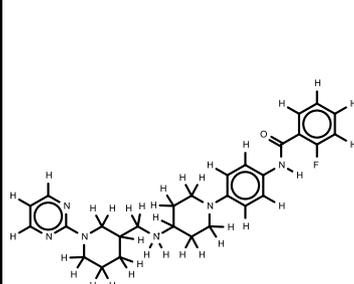
144



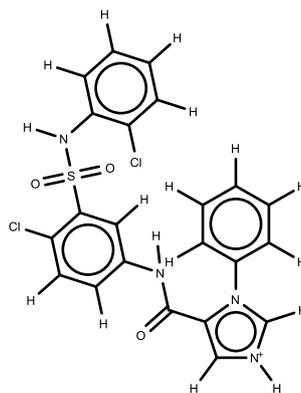
145



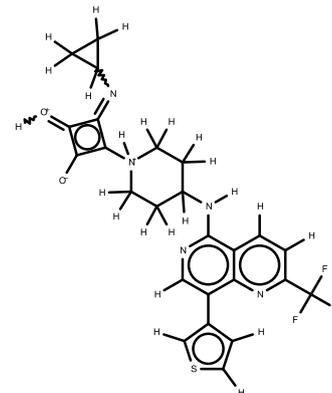
146



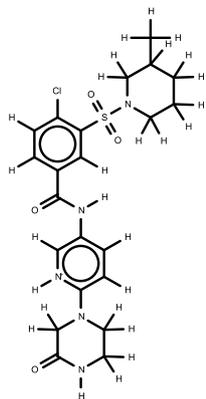
147



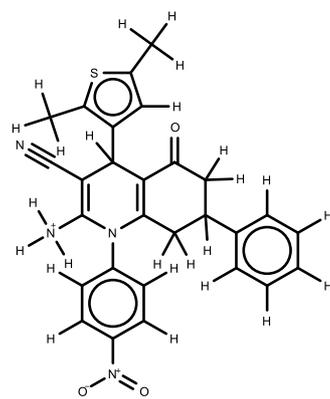
148



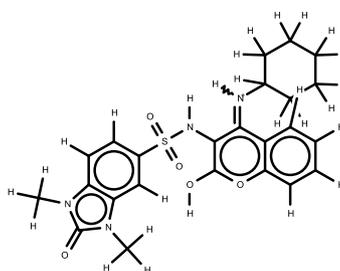
149



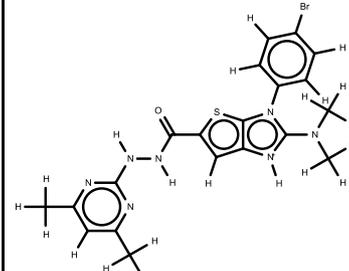
150



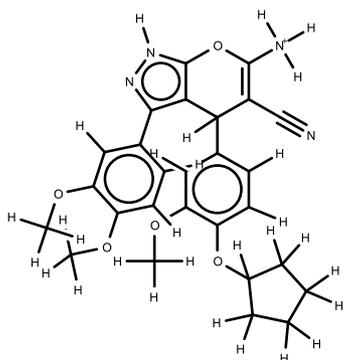
151



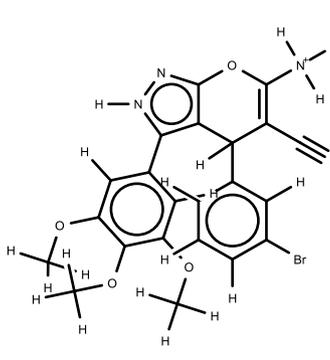
152



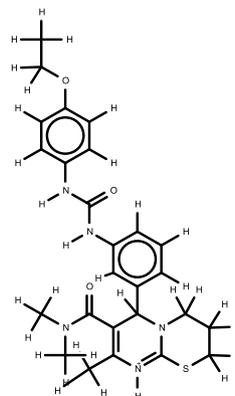
153



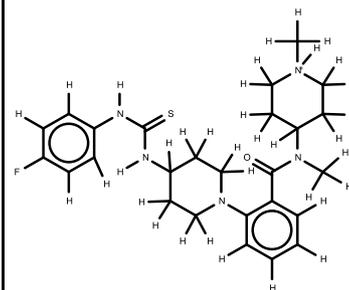
154



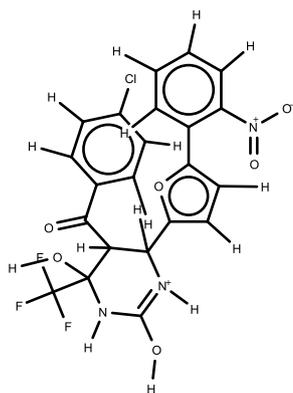
155



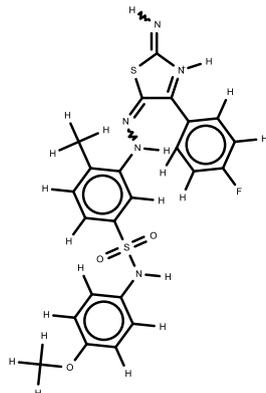
156



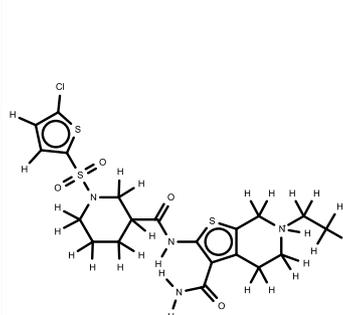
157



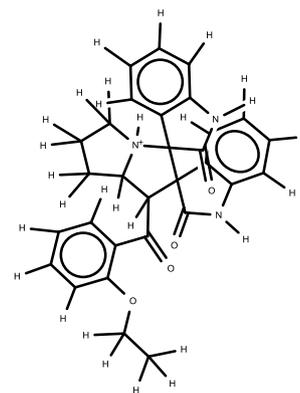
158



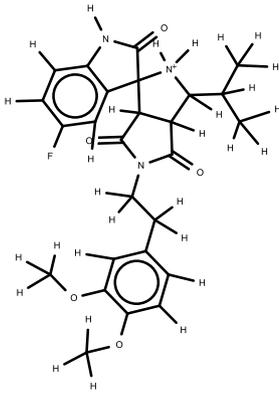
159



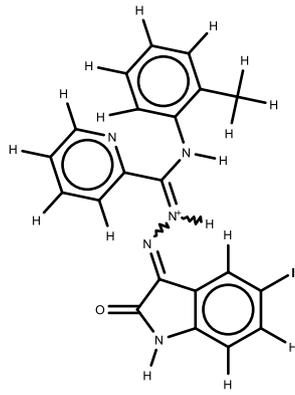
160



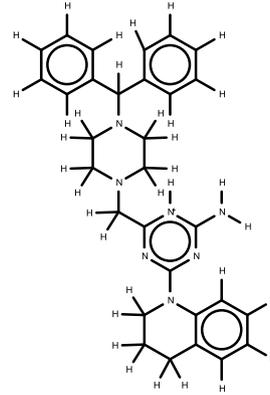
161



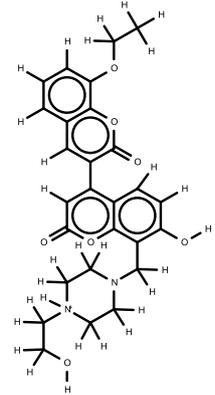
162



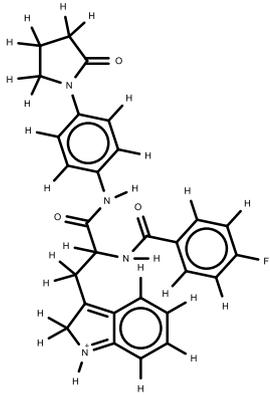
163



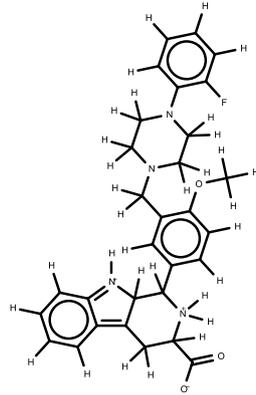
164



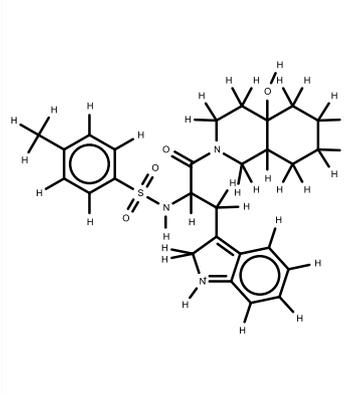
165



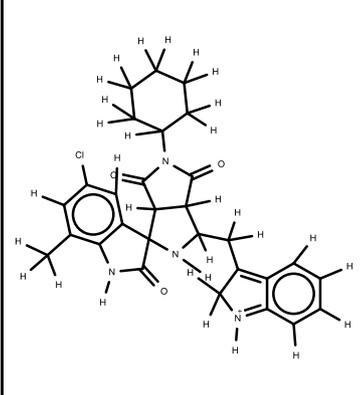
166



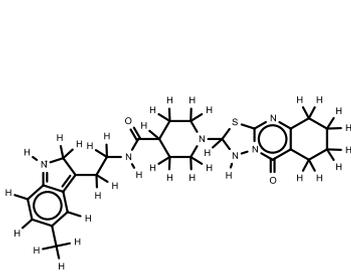
167



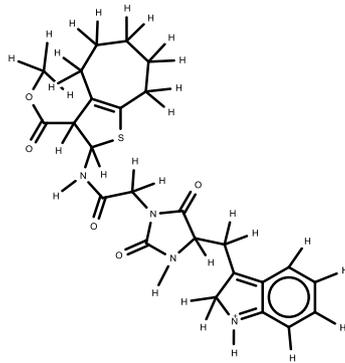
168



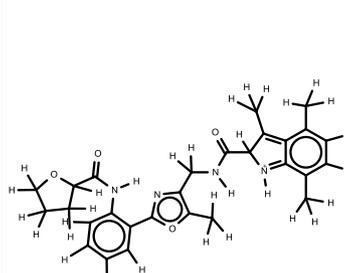
169



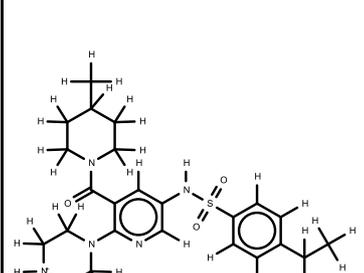
170



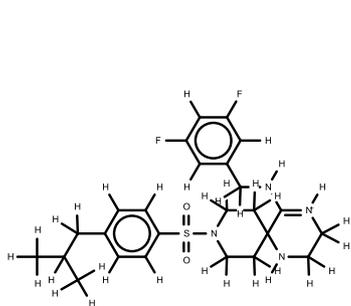
171



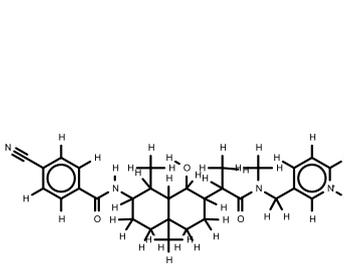
172



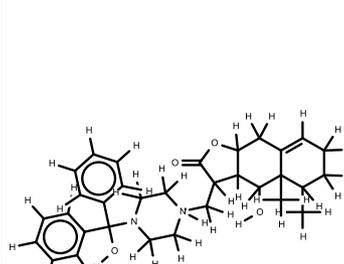
173



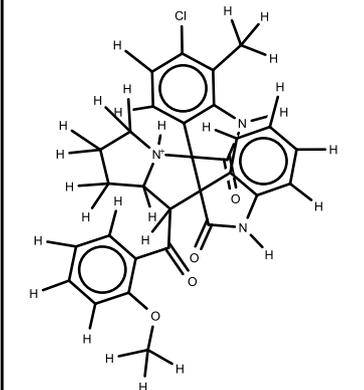
174



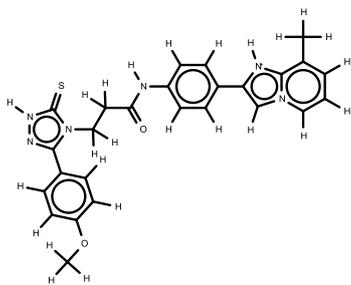
175



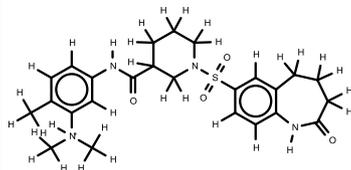
176



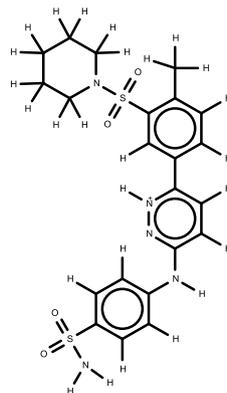
177



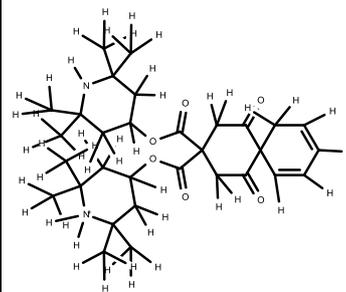
178



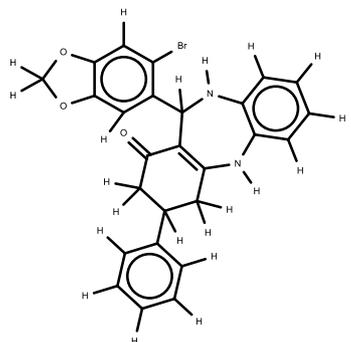
179



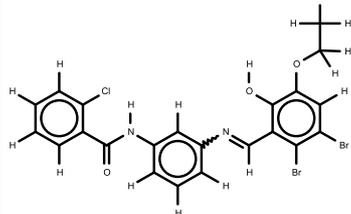
180



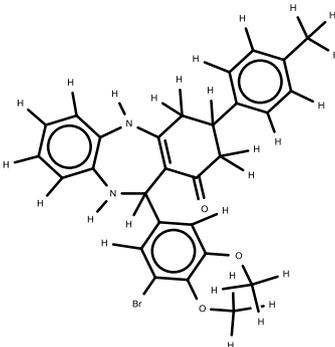
181



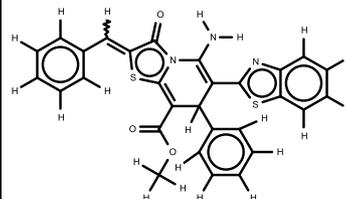
182



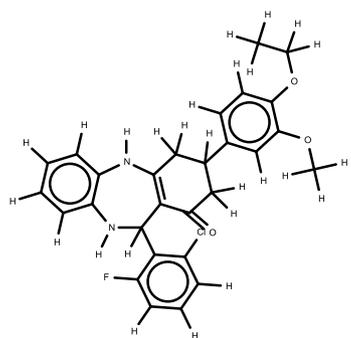
183



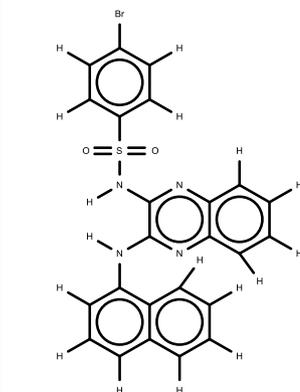
184



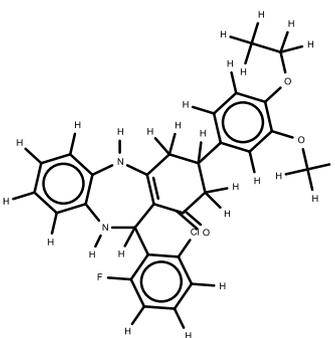
185



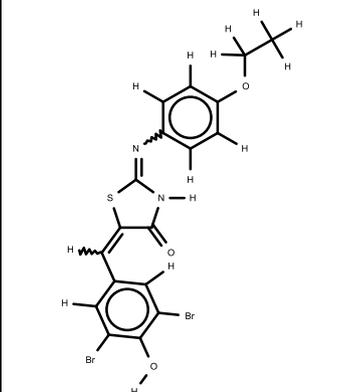
186



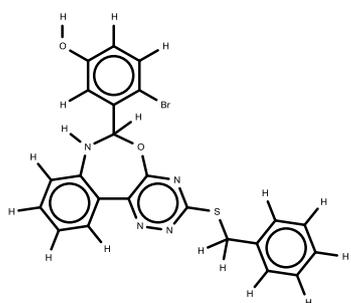
187



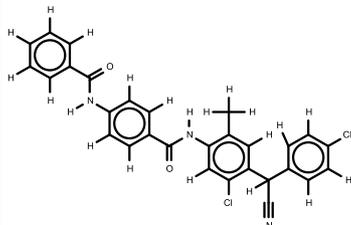
188



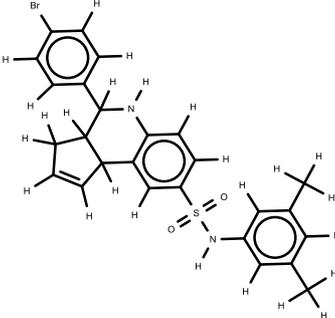
189



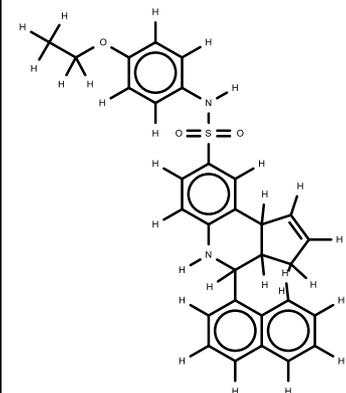
190



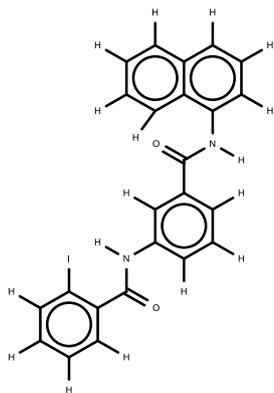
191



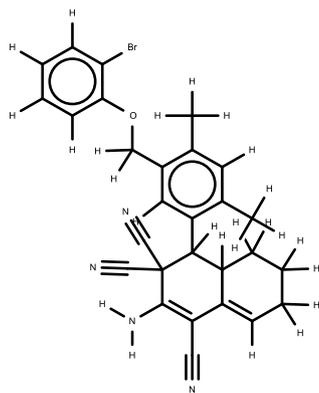
192



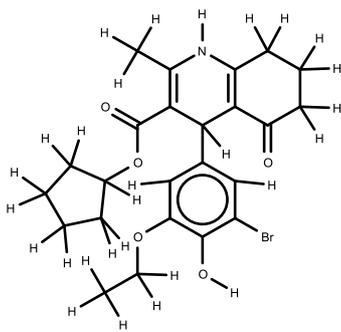
193



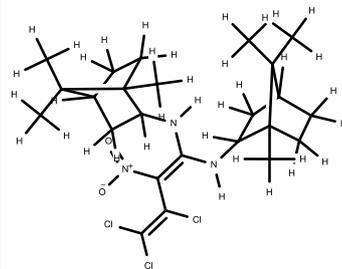
194



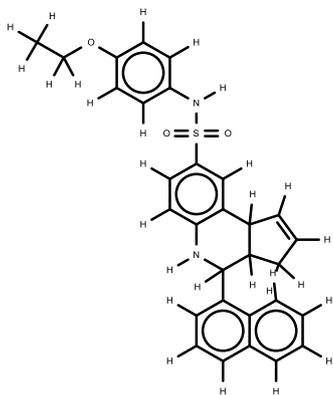
195



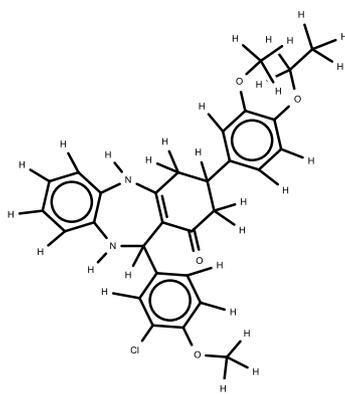
196



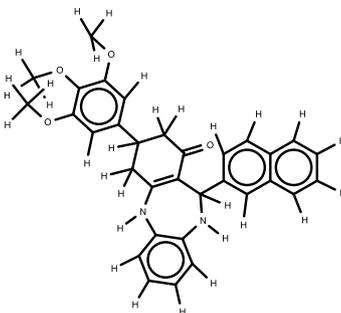
197



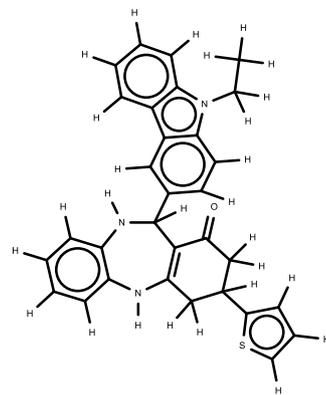
198



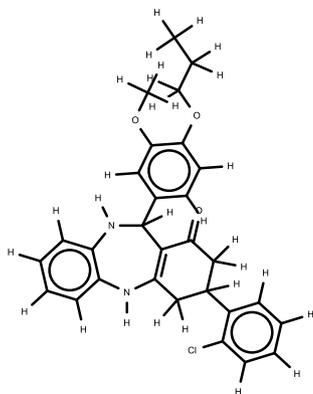
199



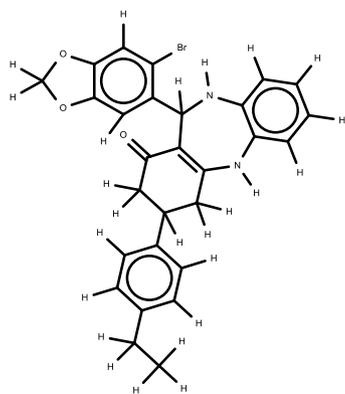
200



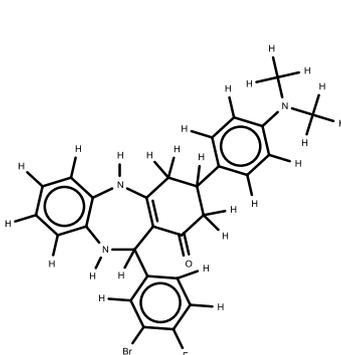
201



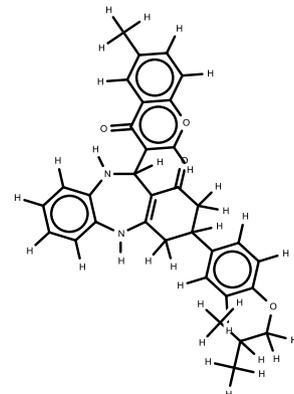
202



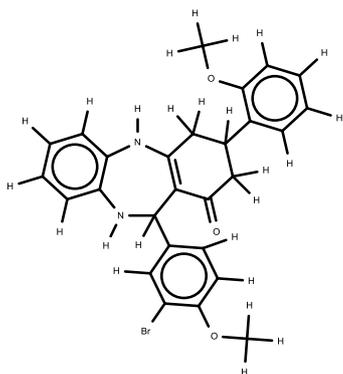
203



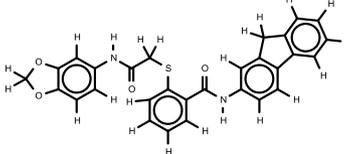
204



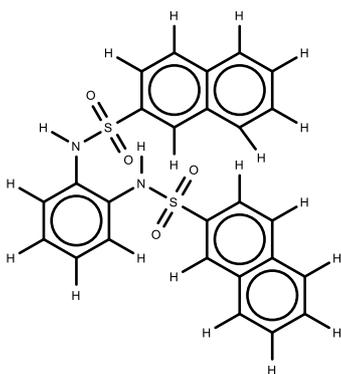
205



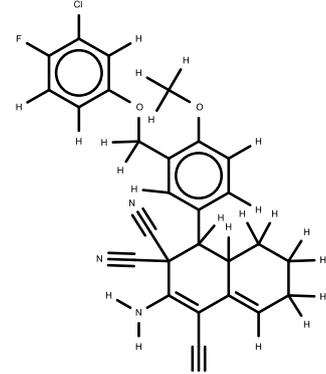
206



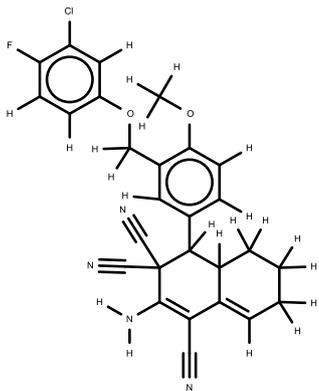
207



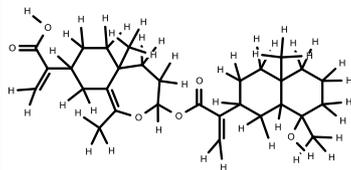
208



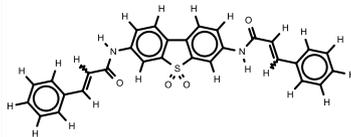
209



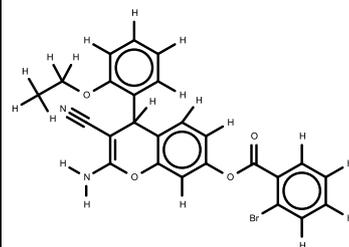
210



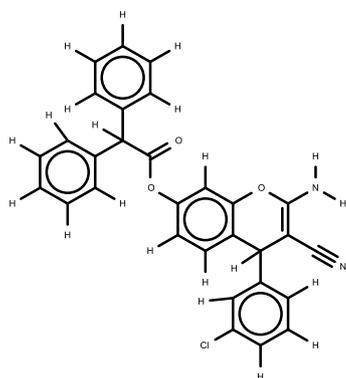
211



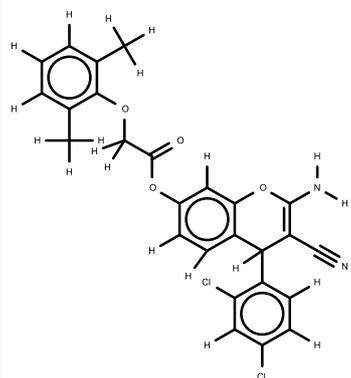
212



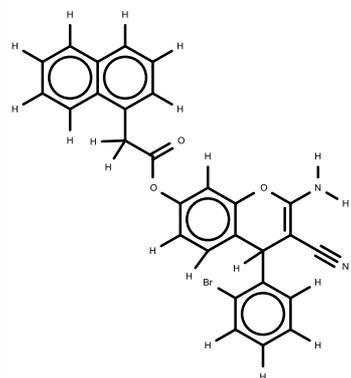
213



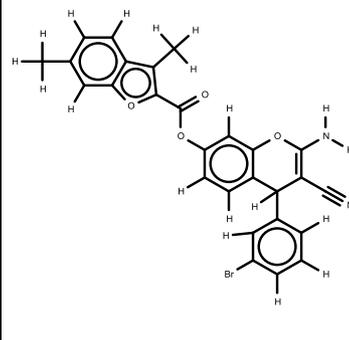
214



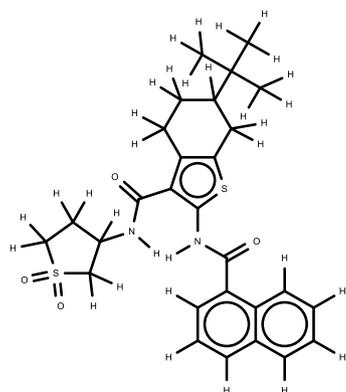
215



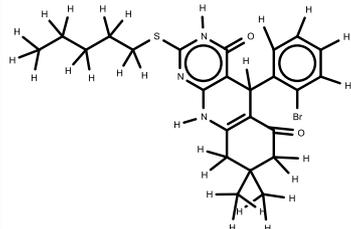
216



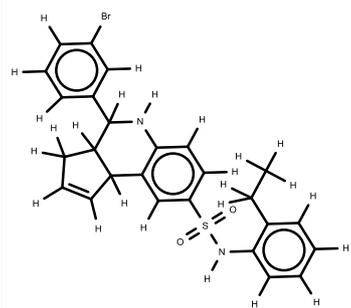
217



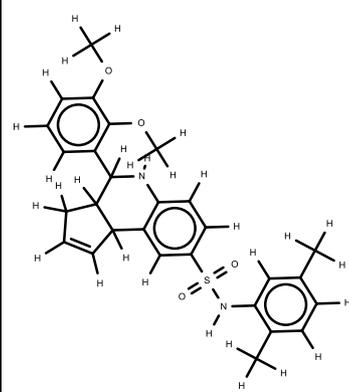
218



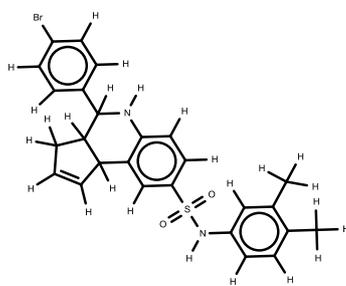
219



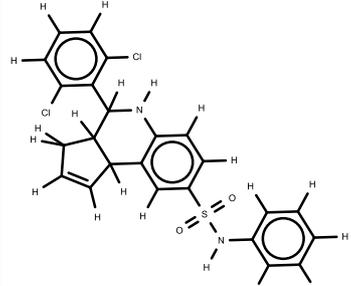
220



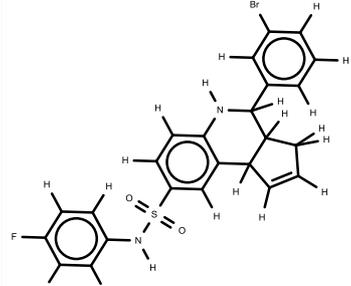
221



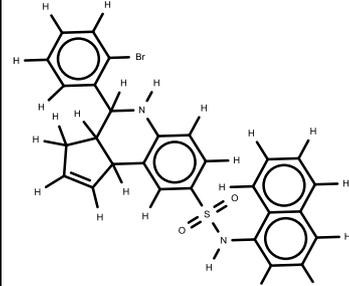
222



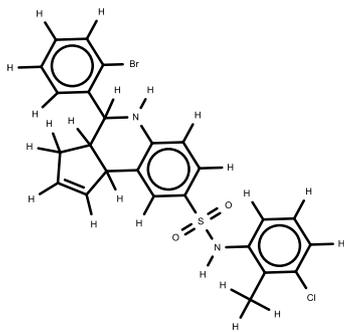
223



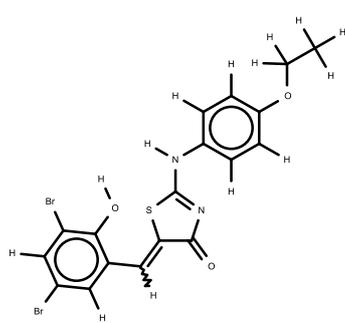
224



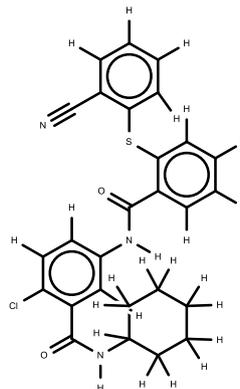
225



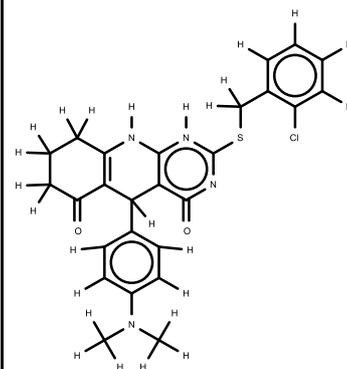
226



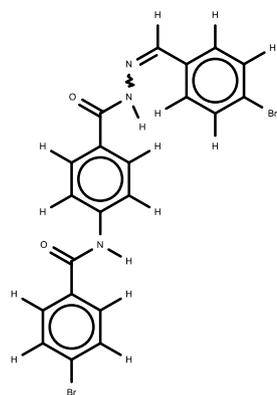
227



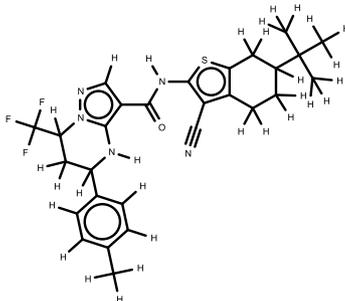
228



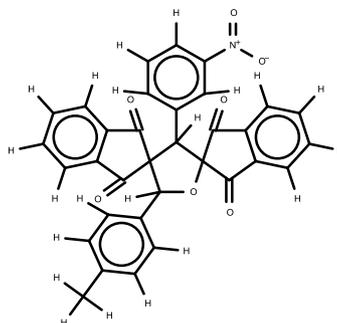
229



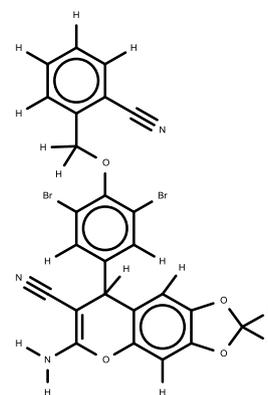
230



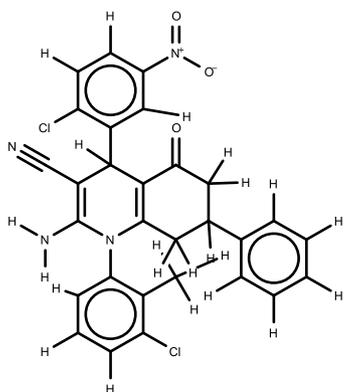
231



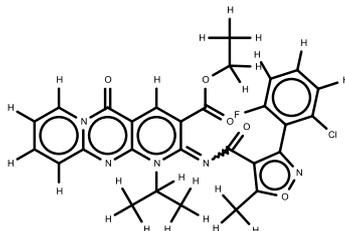
232



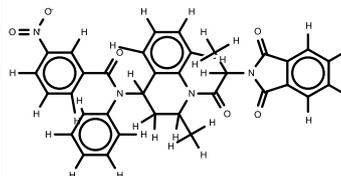
233



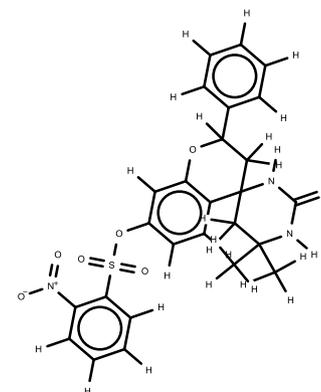
234



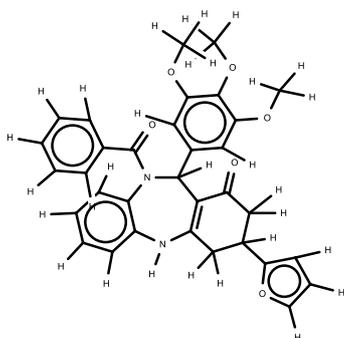
235



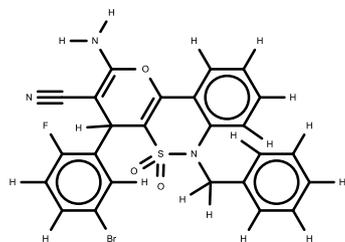
236



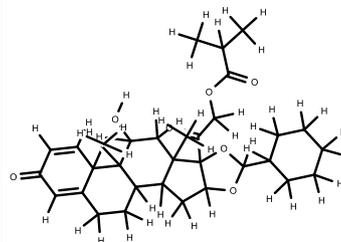
237



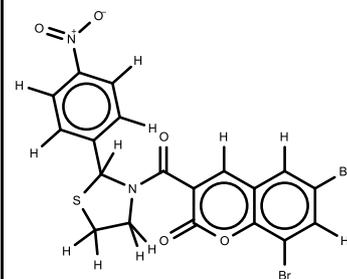
238



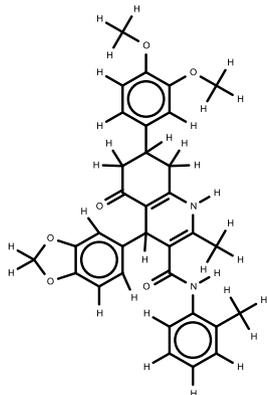
239



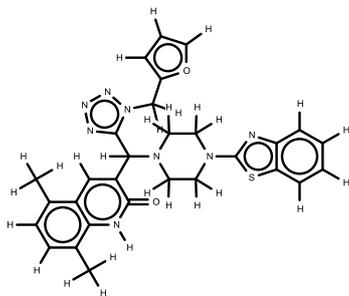
240



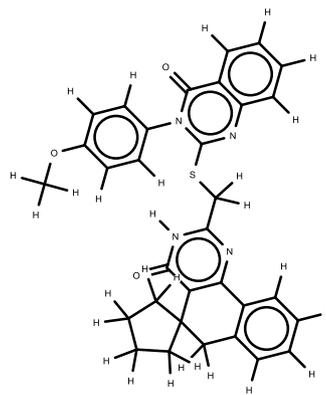
241



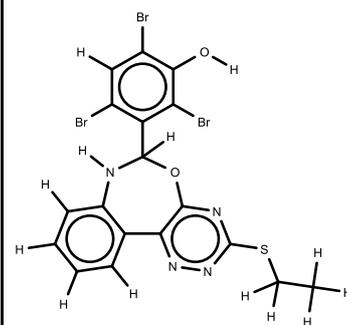
242



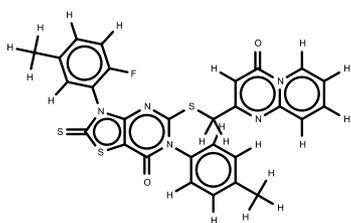
243



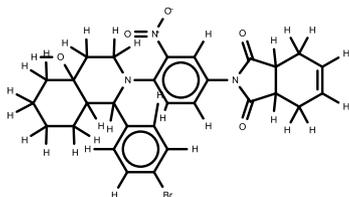
244



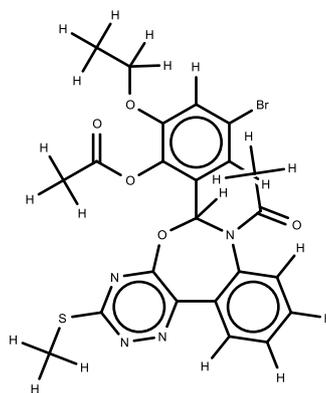
245



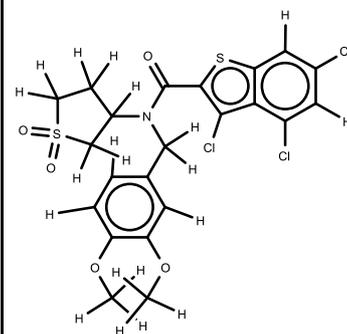
246



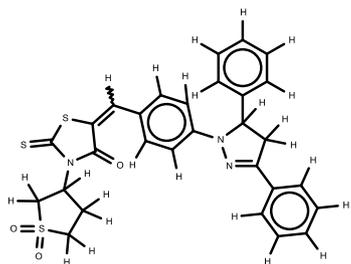
247



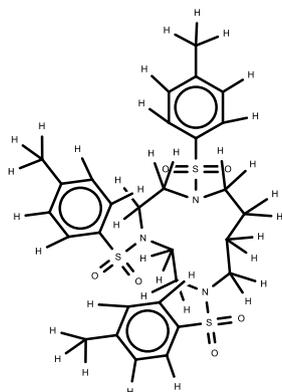
248



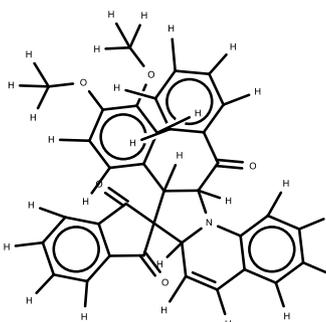
249



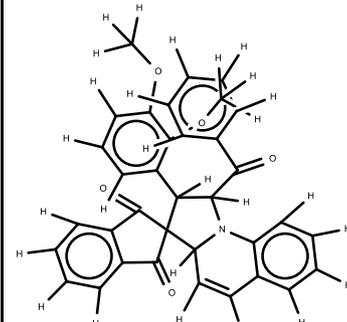
250



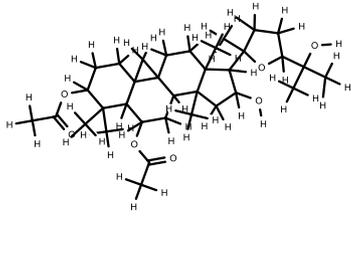
251



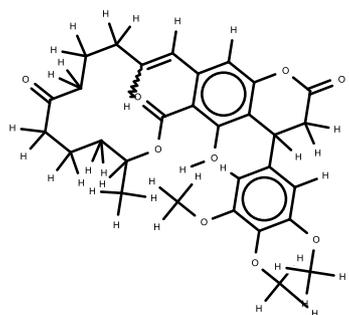
252



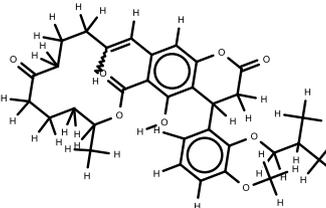
253



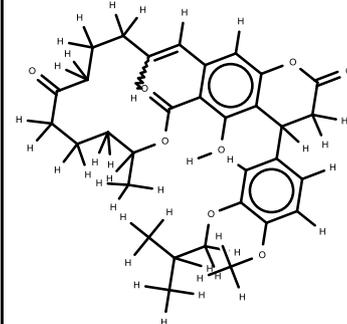
254



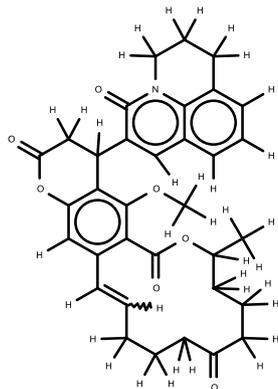
255



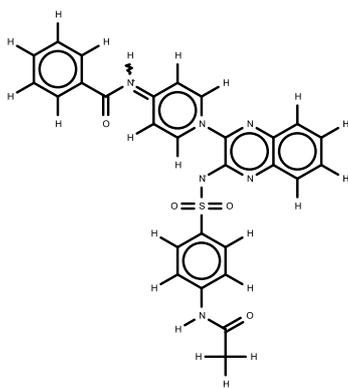
256



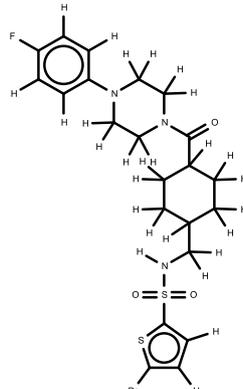
257



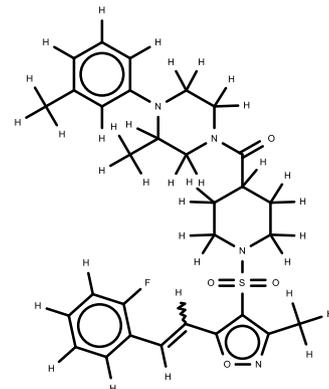
258



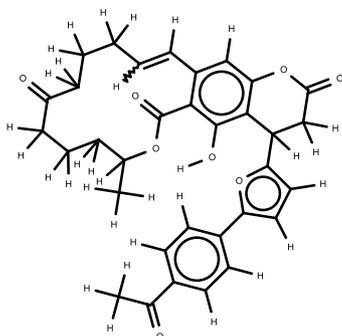
259



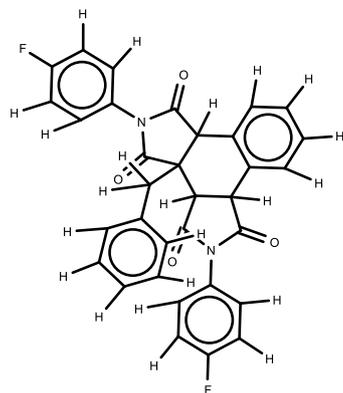
260



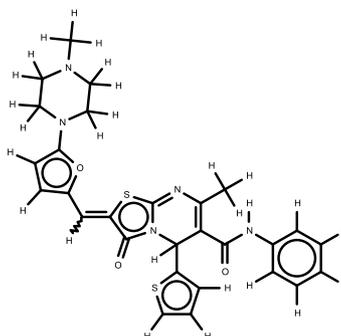
261



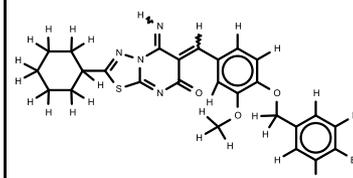
262



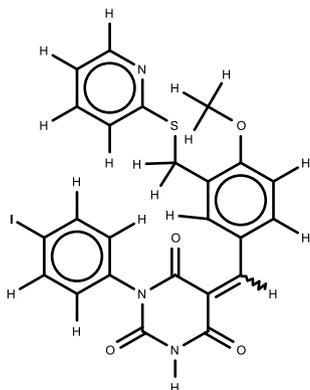
263



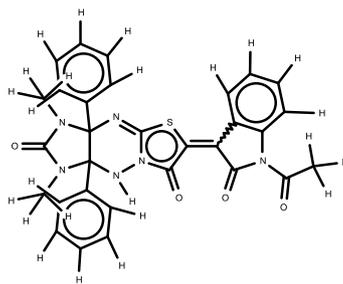
264



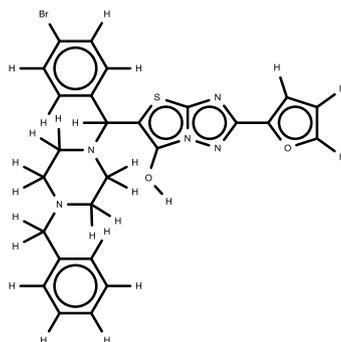
265



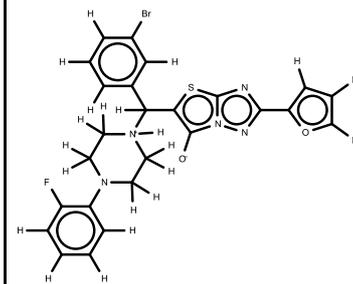
266



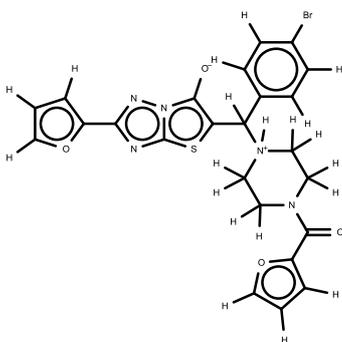
267



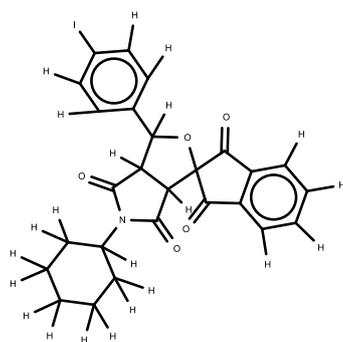
268



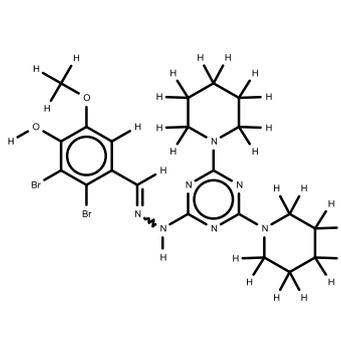
269



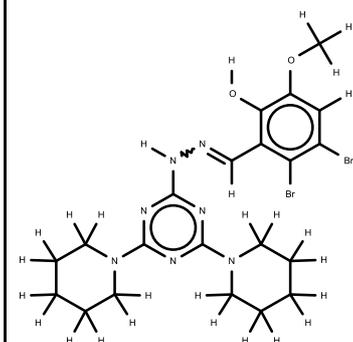
270



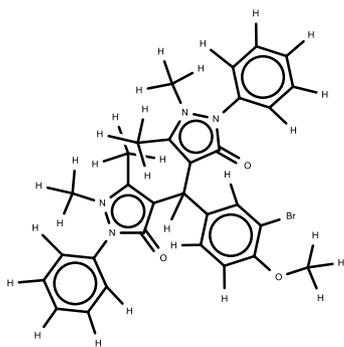
271



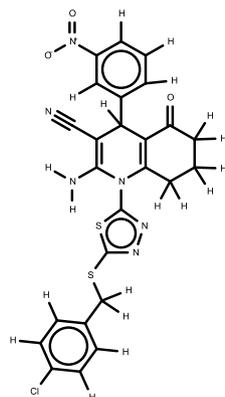
272



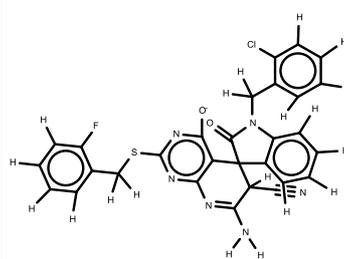
273



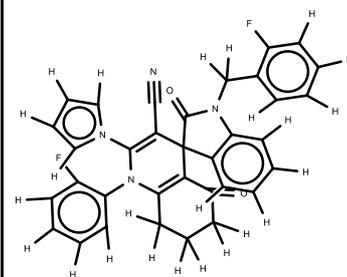
274



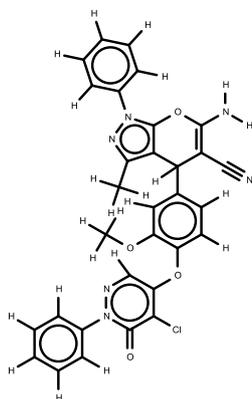
275



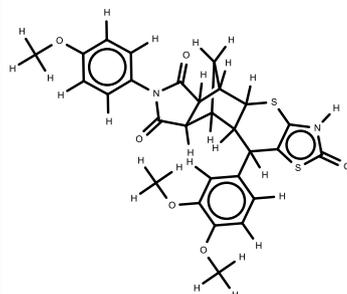
276



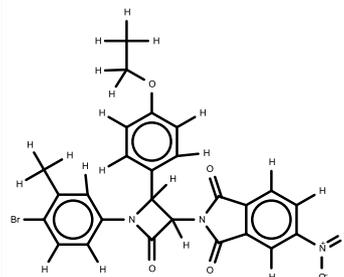
277



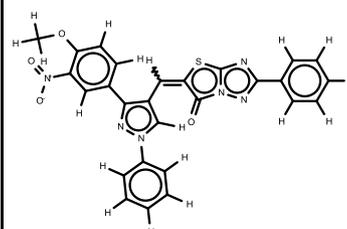
278



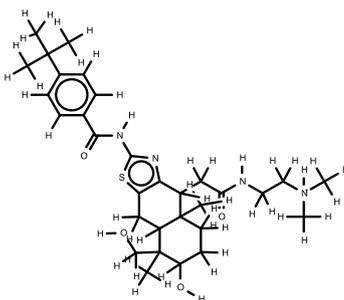
279



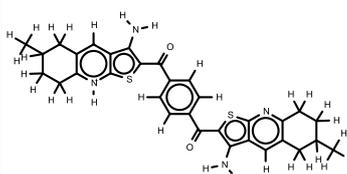
280



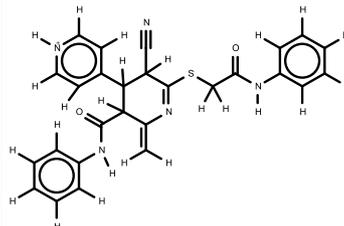
281



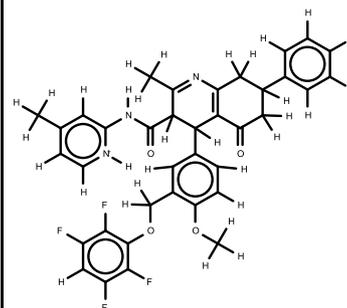
282



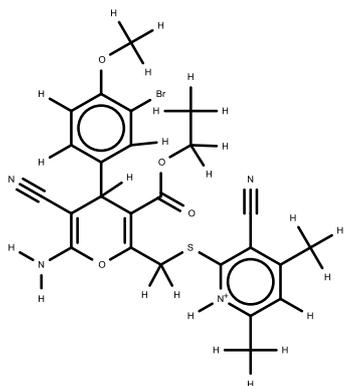
283



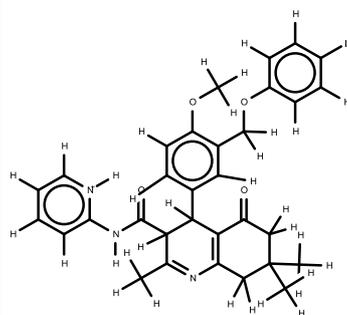
284



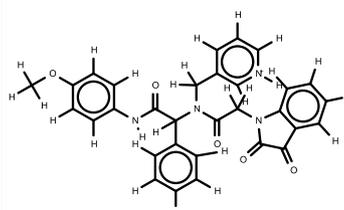
285



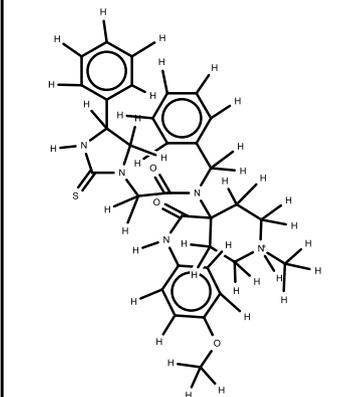
286



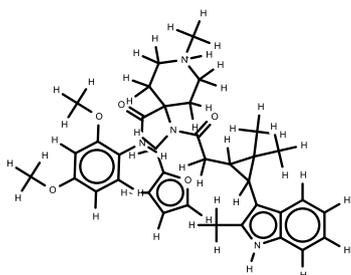
287



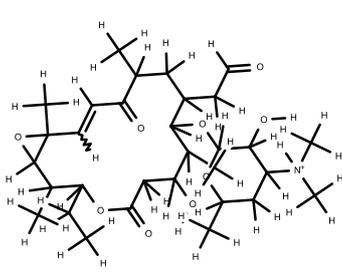
288



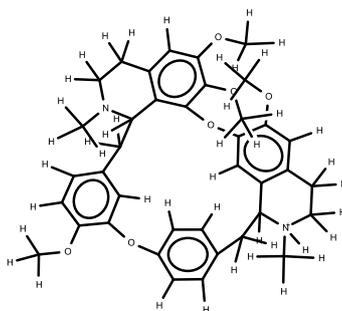
289



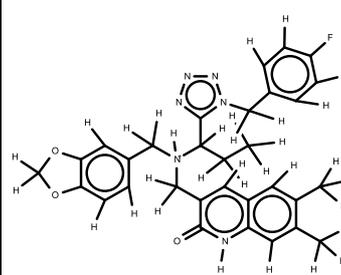
290



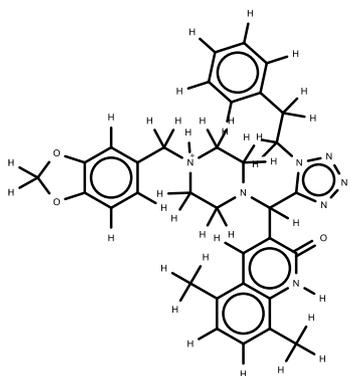
291



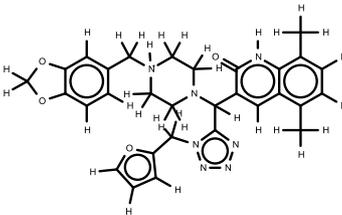
292



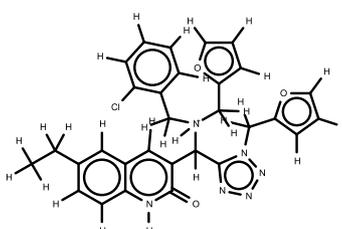
293



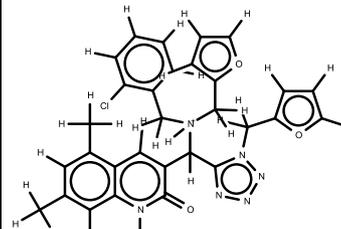
294



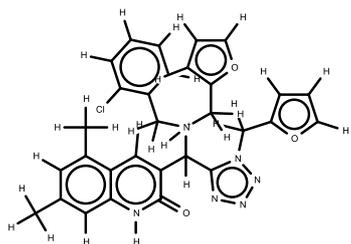
295



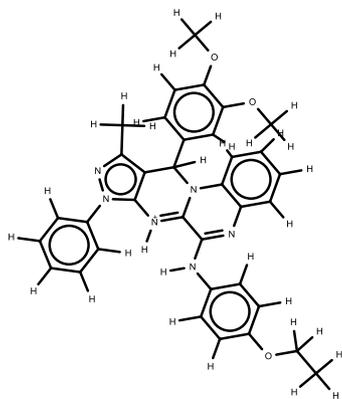
296



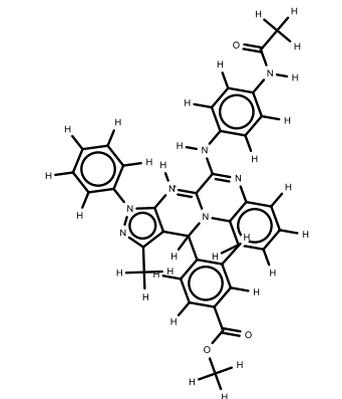
297



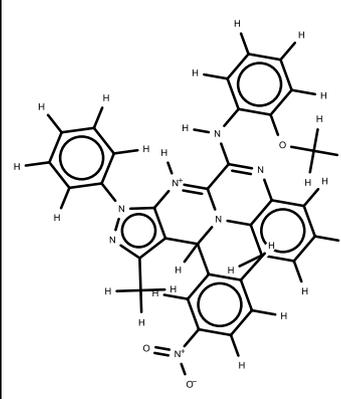
298



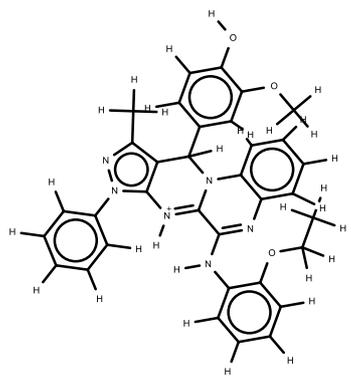
299



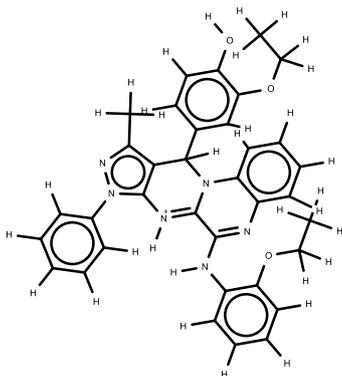
300



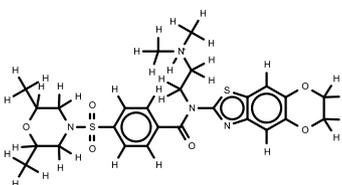
301



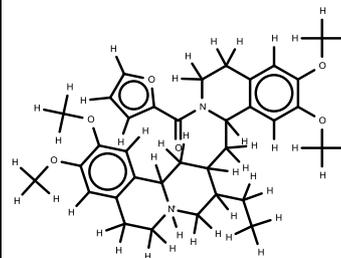
302



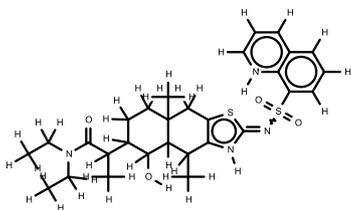
303



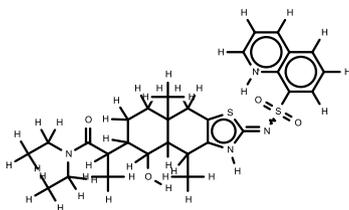
304



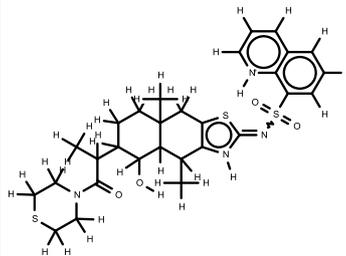
305



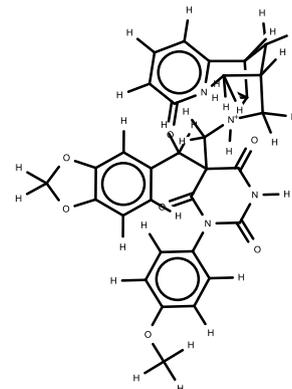
306



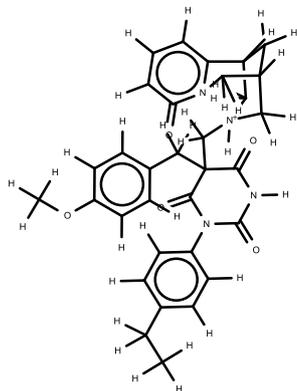
307



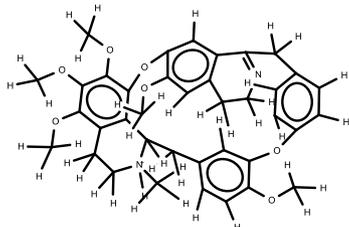
308



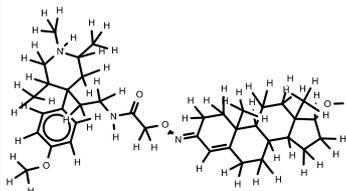
309



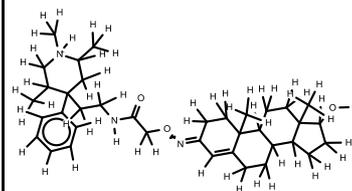
310



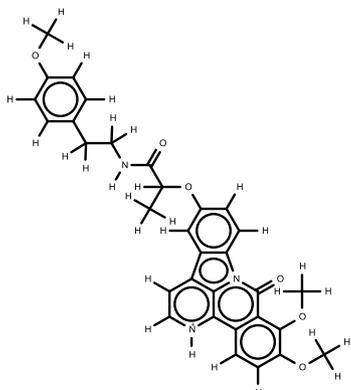
311



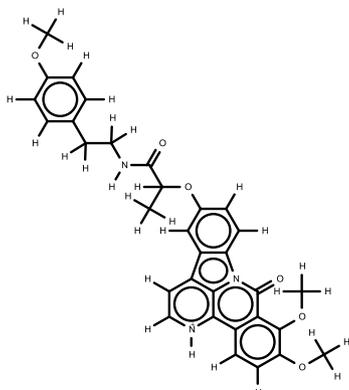
312



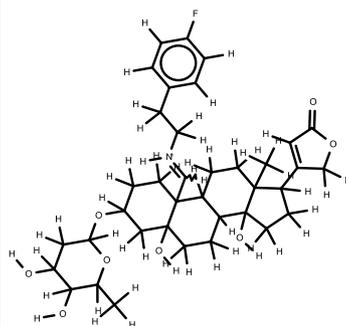
313



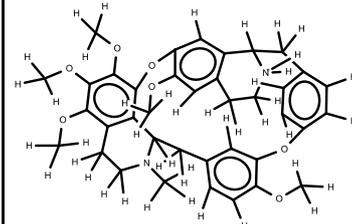
314



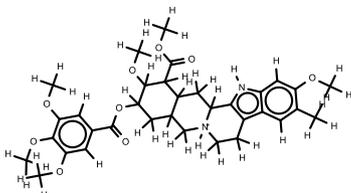
315



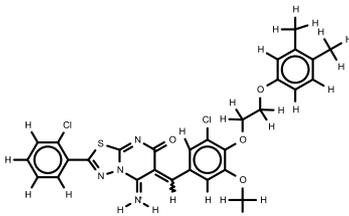
316



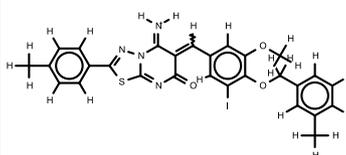
317



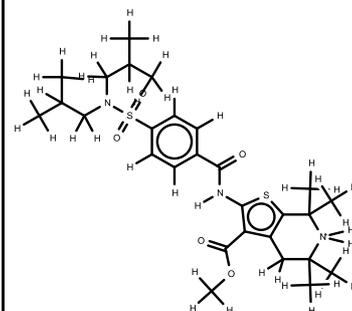
318



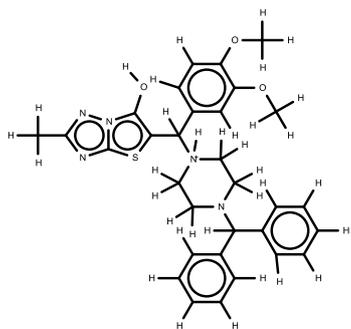
319



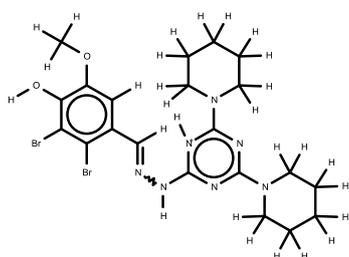
320



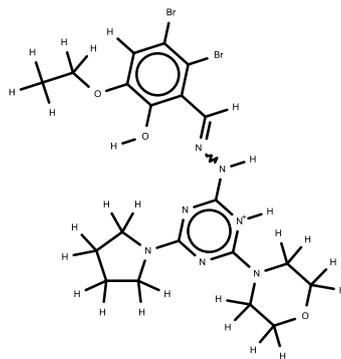
321



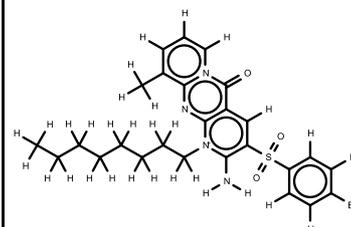
322



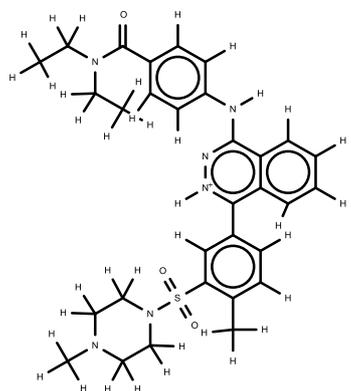
323



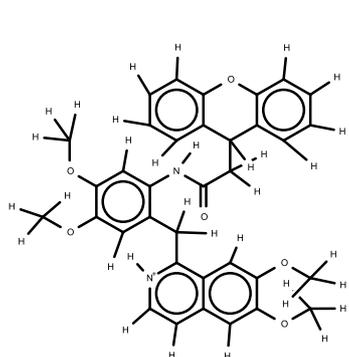
324



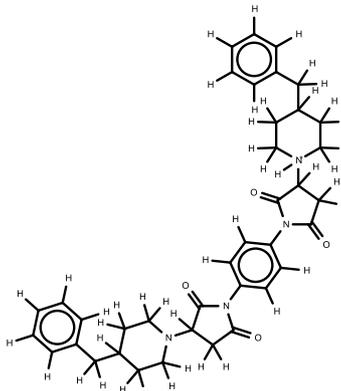
325



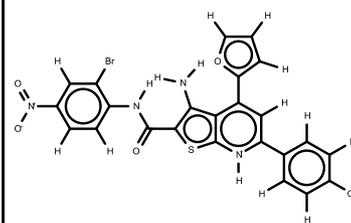
326



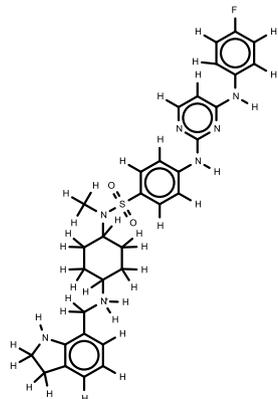
327



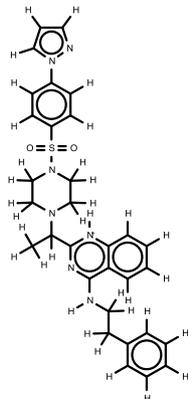
328



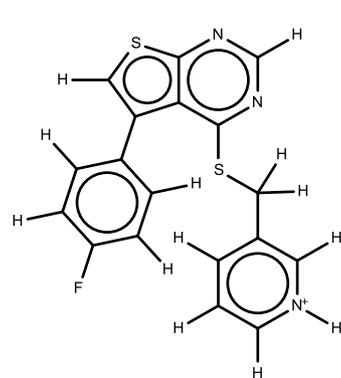
329



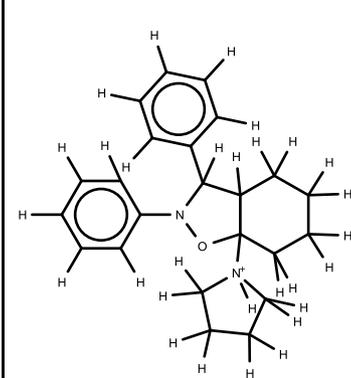
330



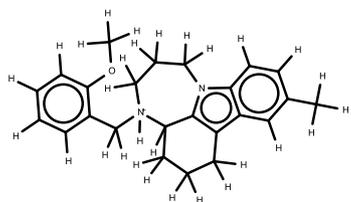
331



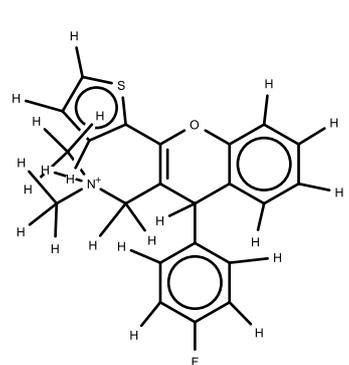
332



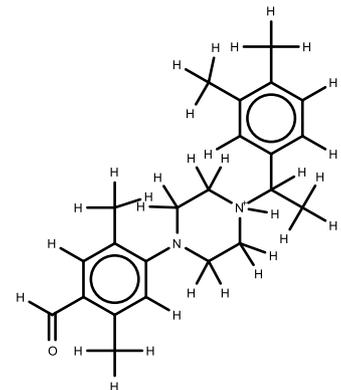
333



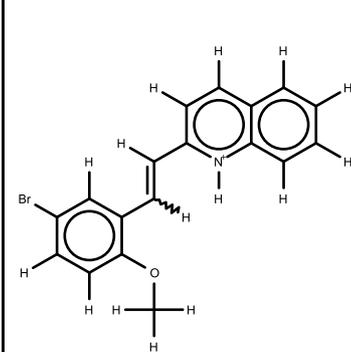
334



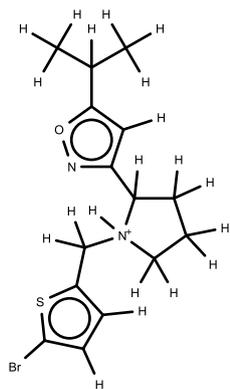
335



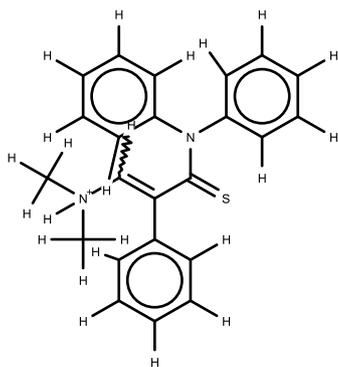
336



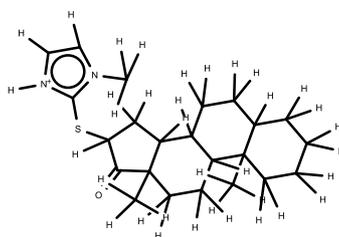
337



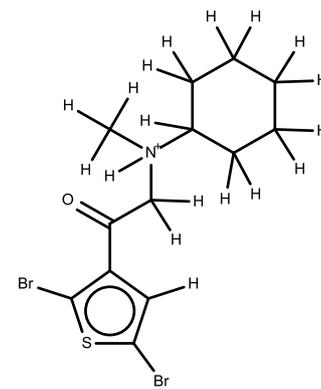
338



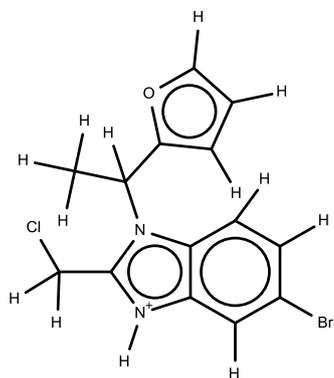
339



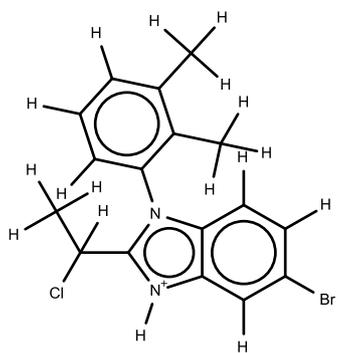
340



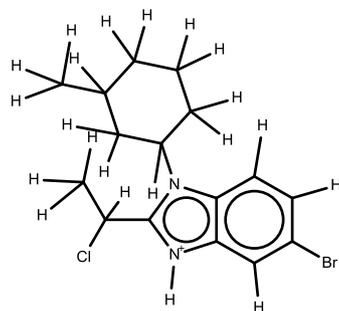
341



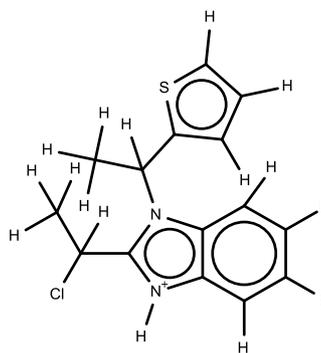
342



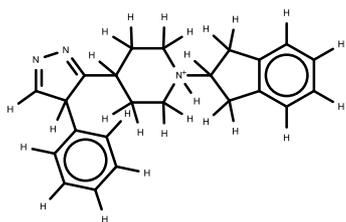
343



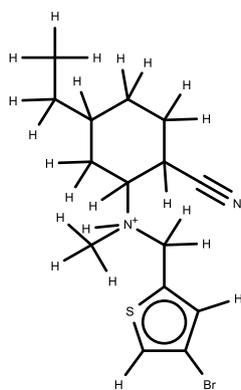
344



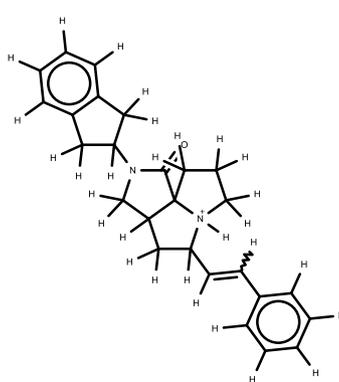
345



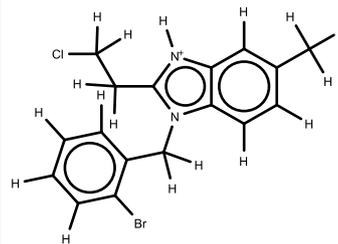
346



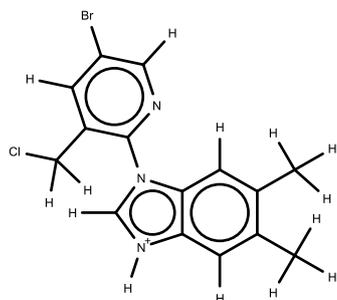
347



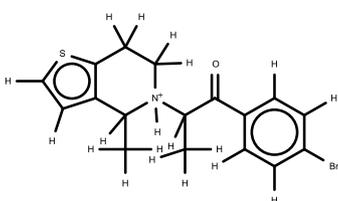
348



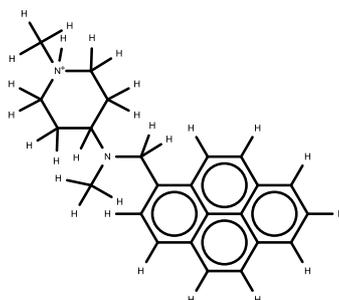
349



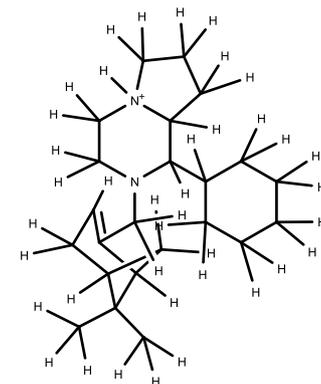
350



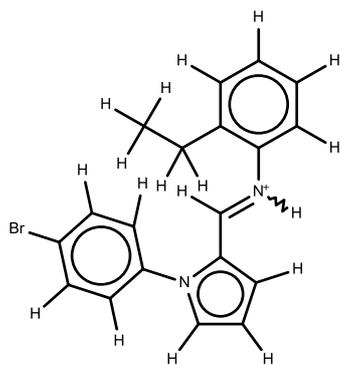
351



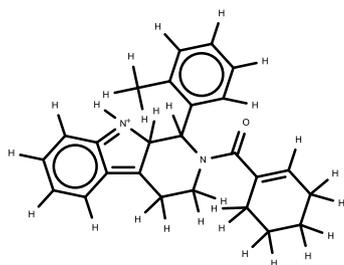
352



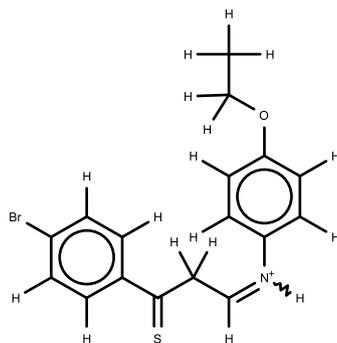
353



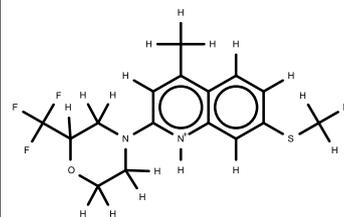
354



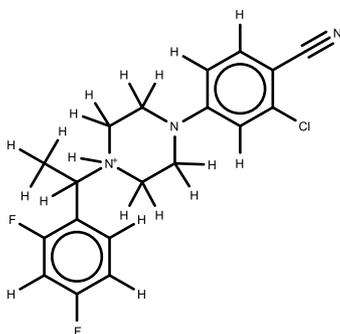
355



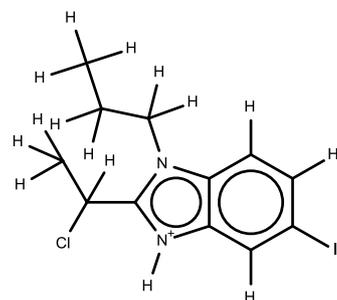
356



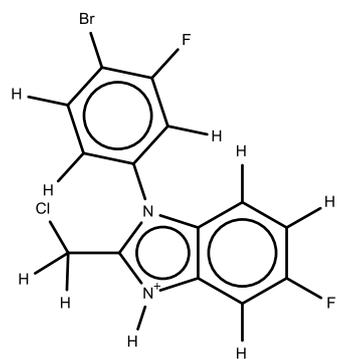
357



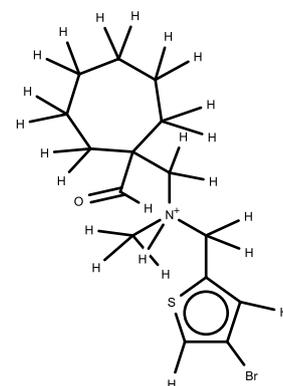
358



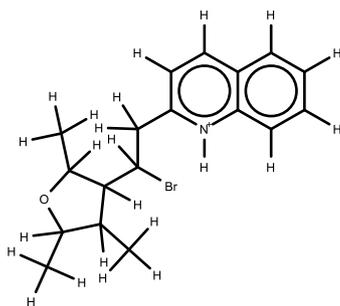
359



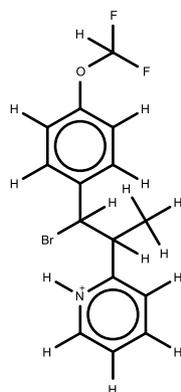
360



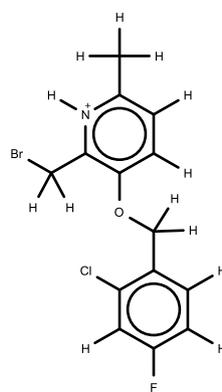
361



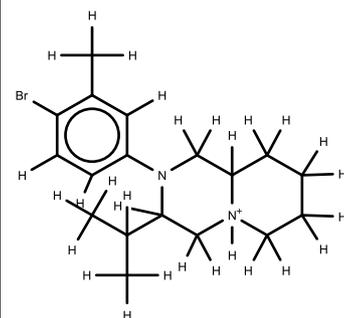
362



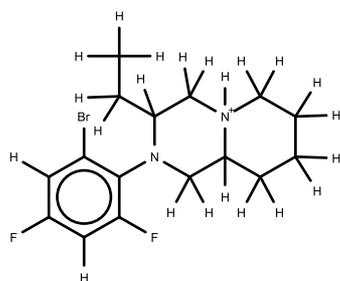
363



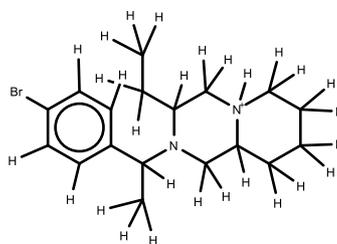
364



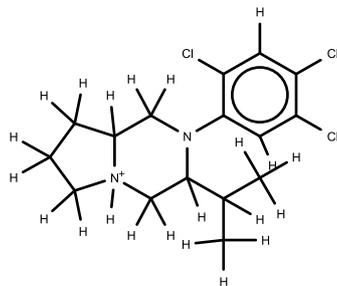
365



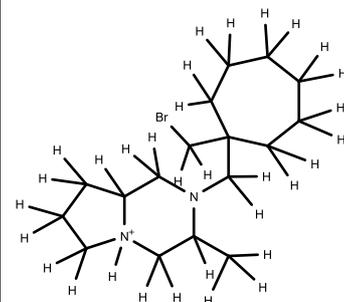
366



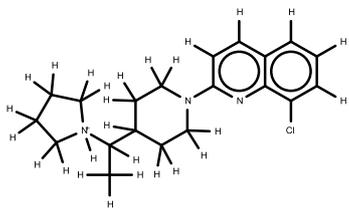
367



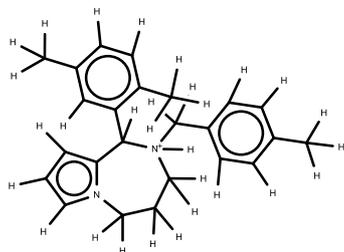
368



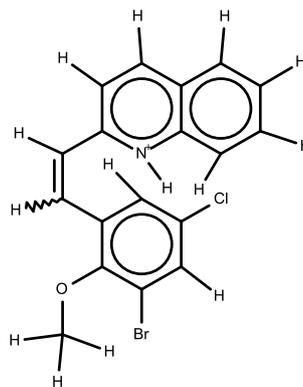
369



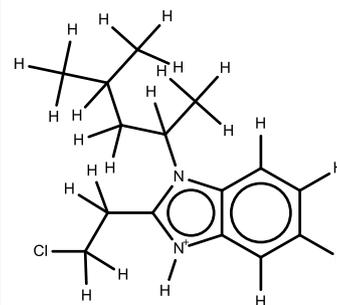
370



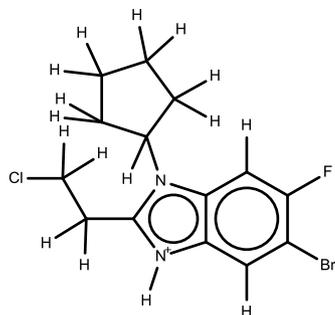
371



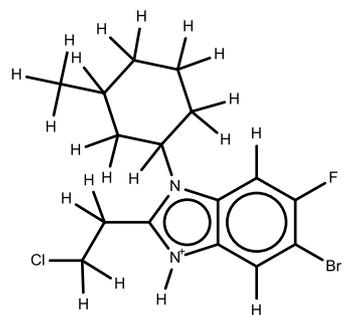
372



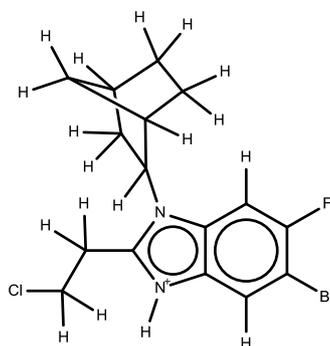
373



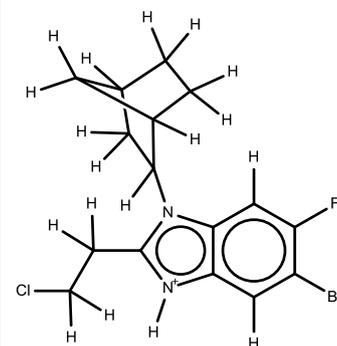
374



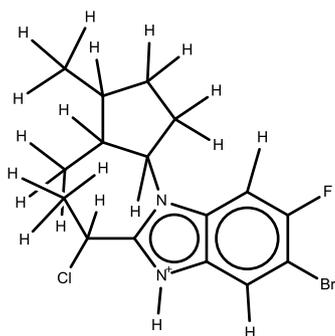
375



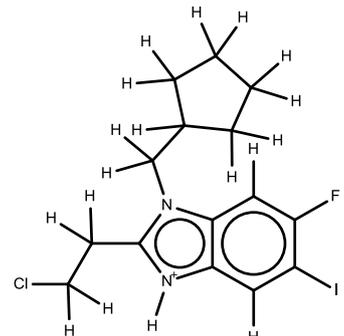
376



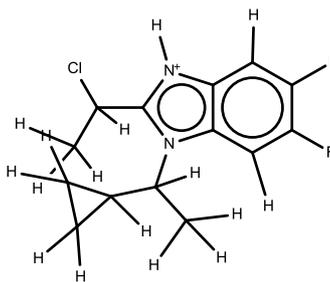
377



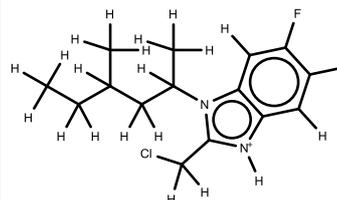
378



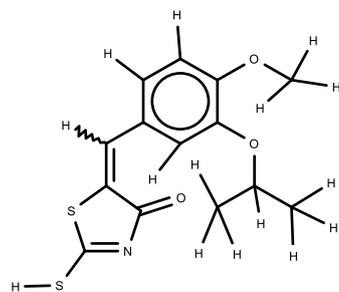
379



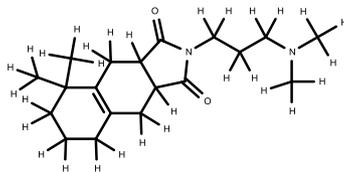
380



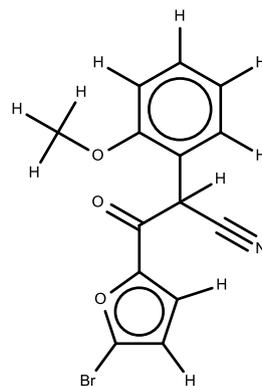
381



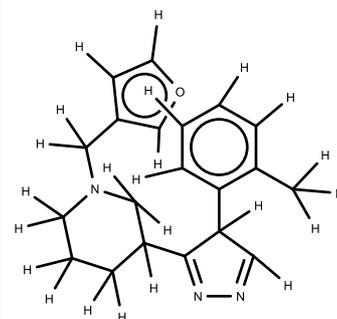
382



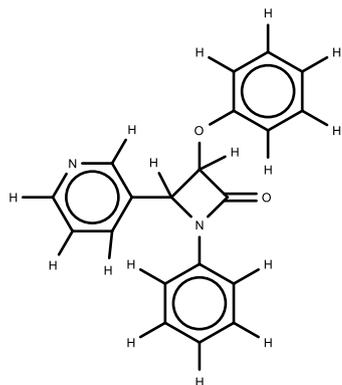
383



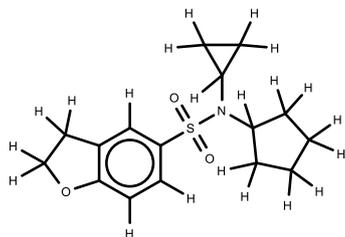
384



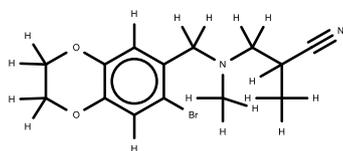
385



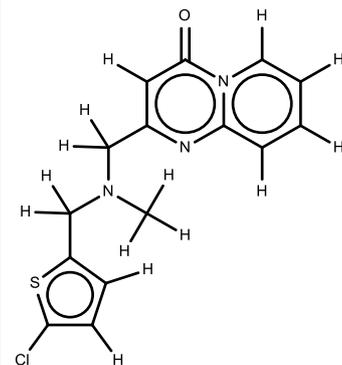
386



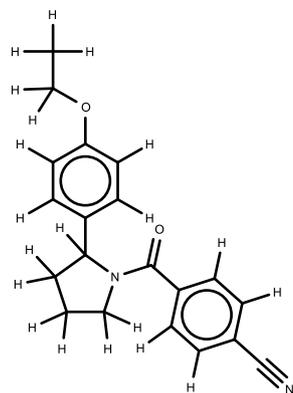
387



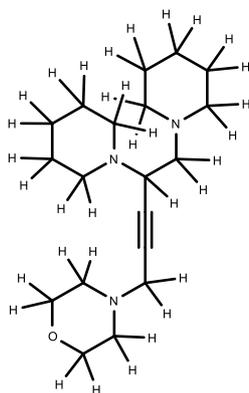
388



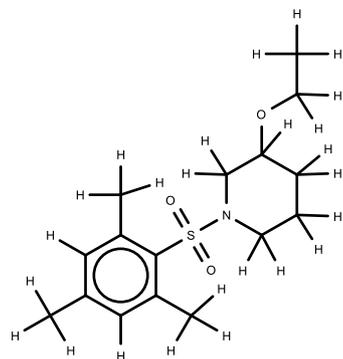
389



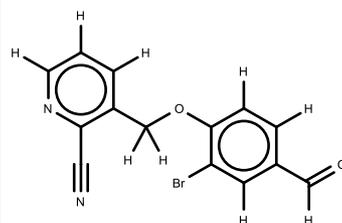
390



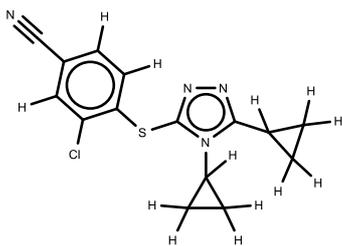
391



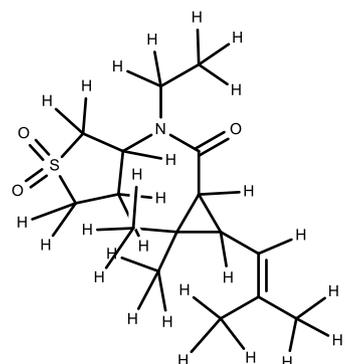
392



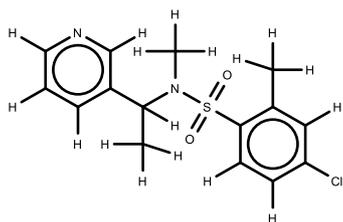
393



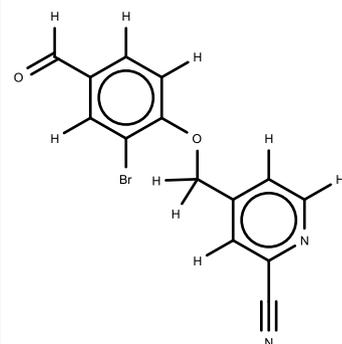
394



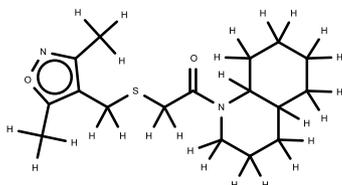
395



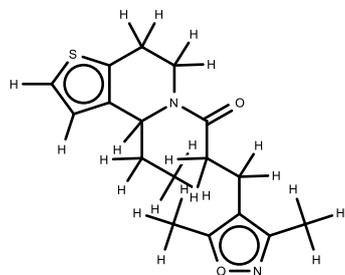
396



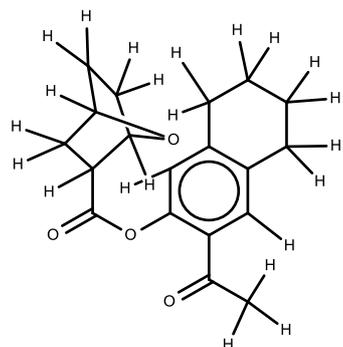
397



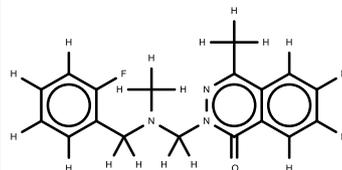
398



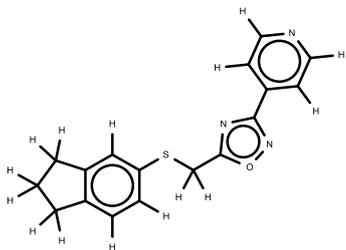
399



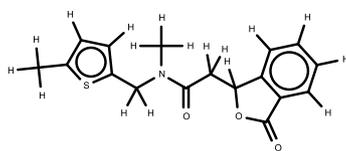
400



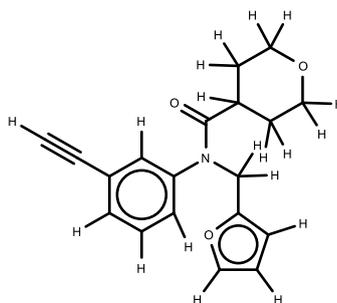
401



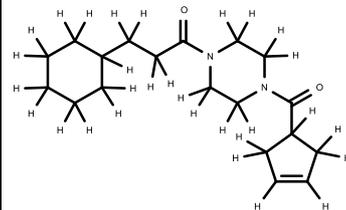
402



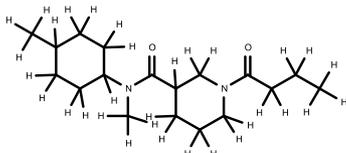
403



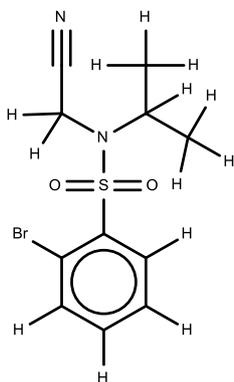
404



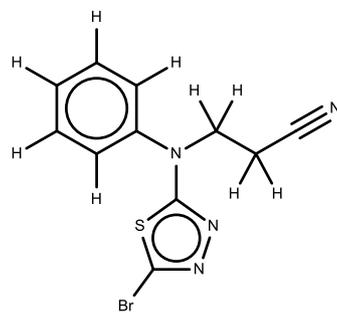
405



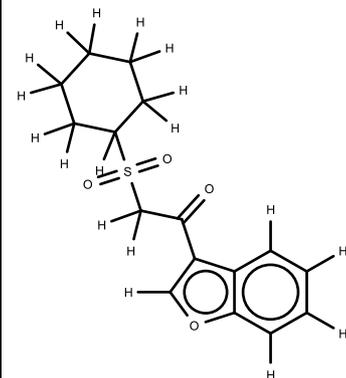
406



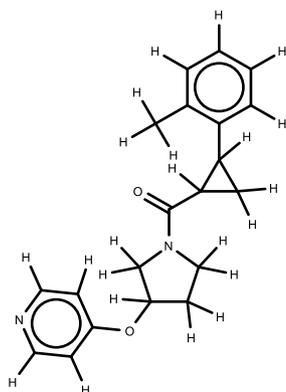
407



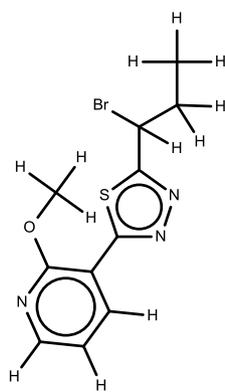
408



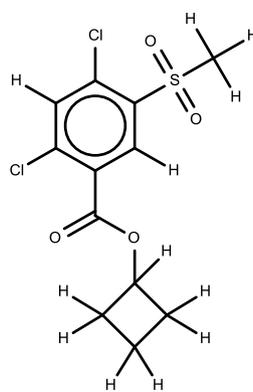
409



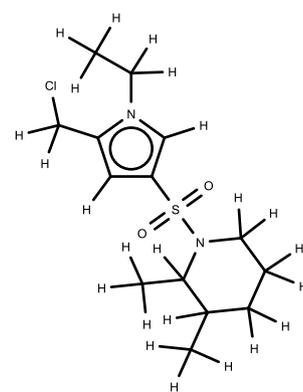
410



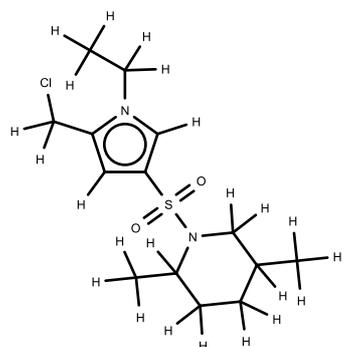
411



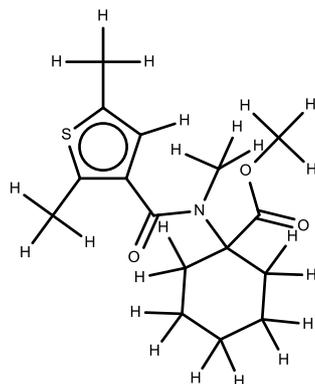
412



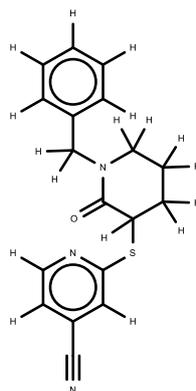
413



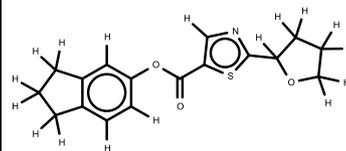
414



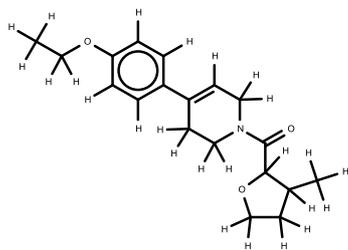
415



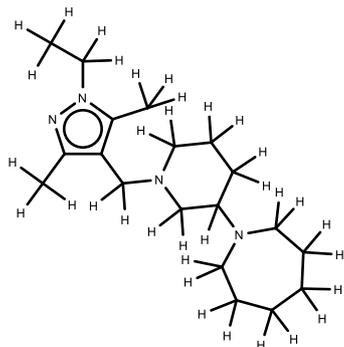
416



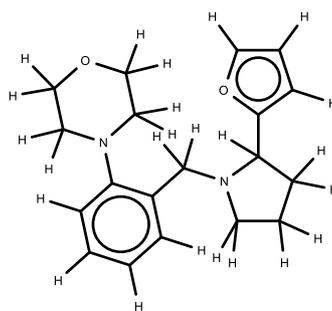
417



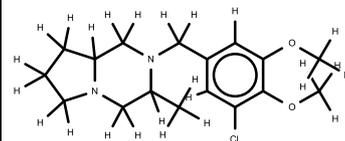
418



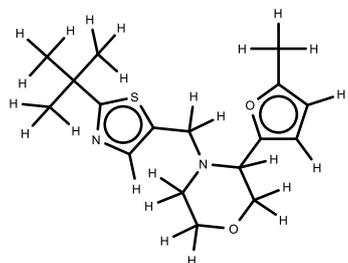
419



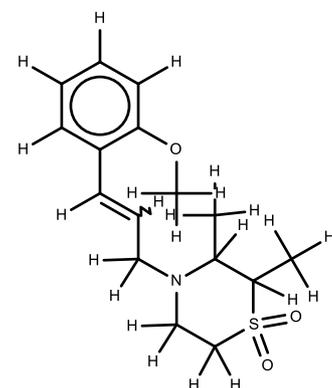
420



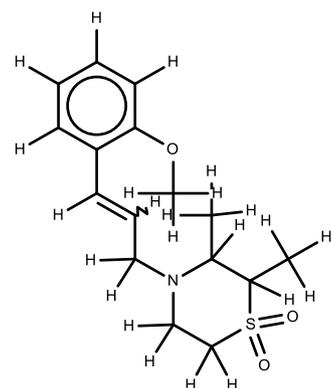
421



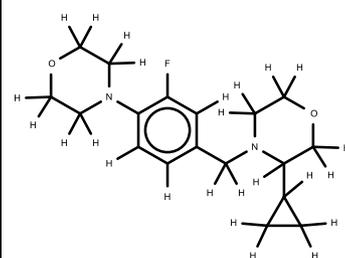
422



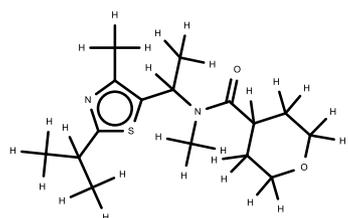
423



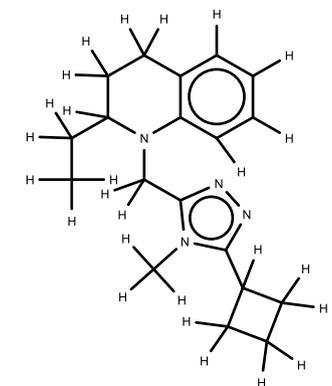
424



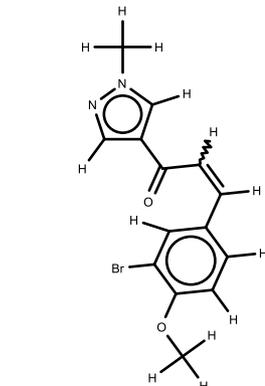
425



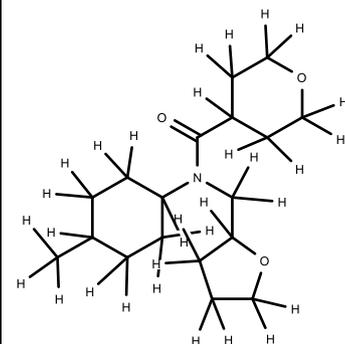
426



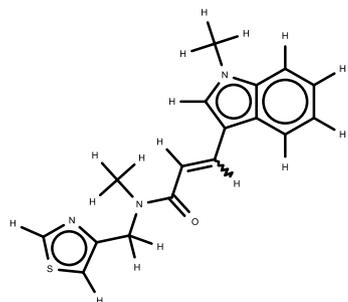
427



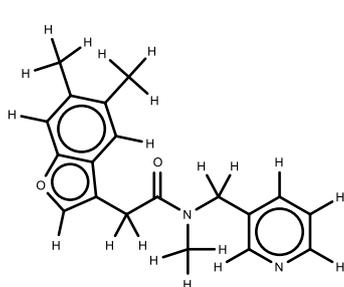
428



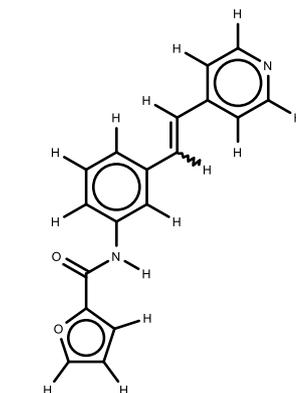
429



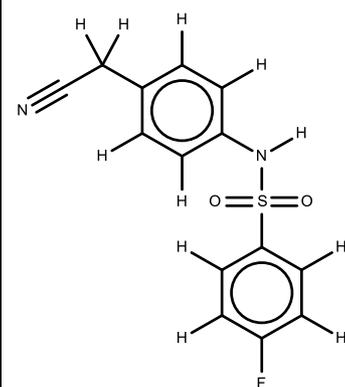
430



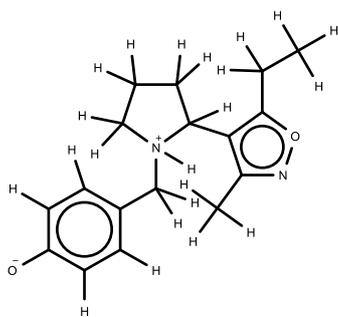
431



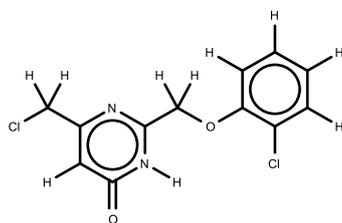
432



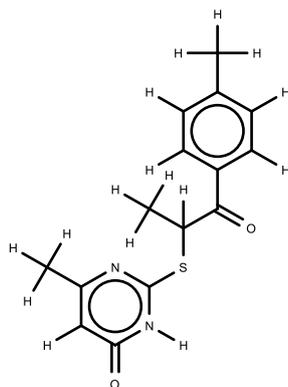
433



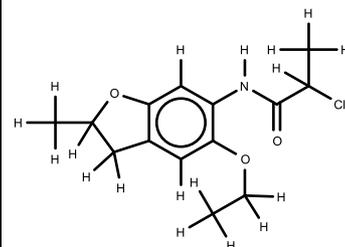
434



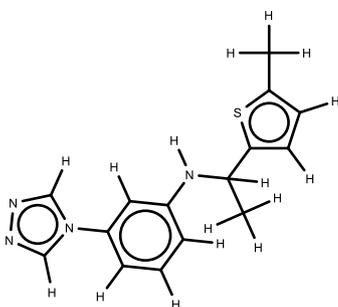
435



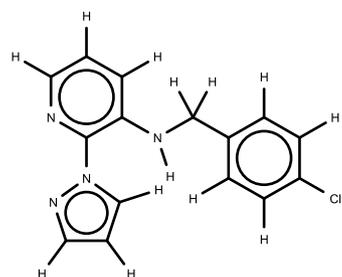
436



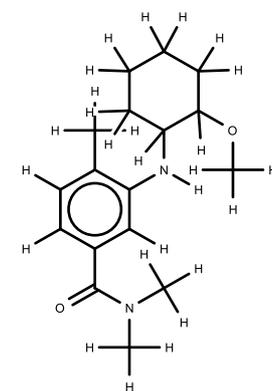
437



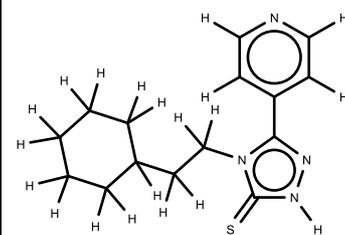
438



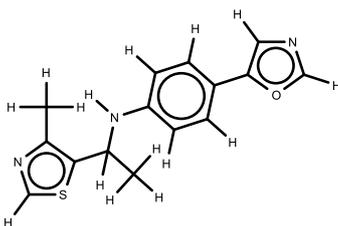
439



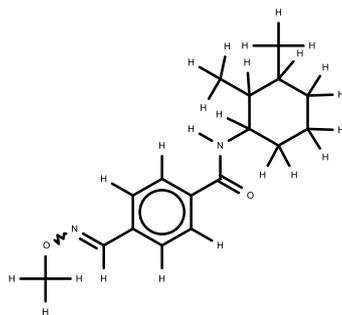
440



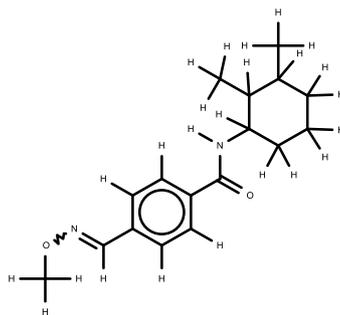
441



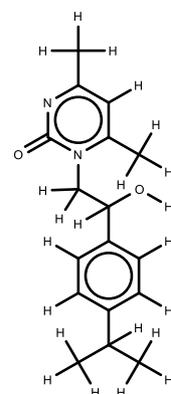
442



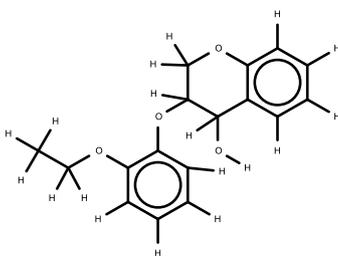
443



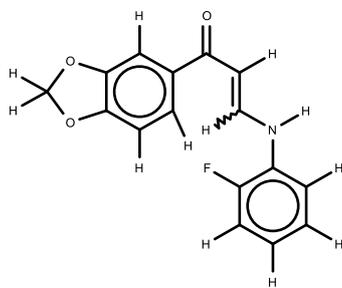
444



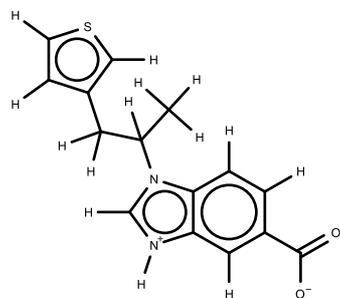
445



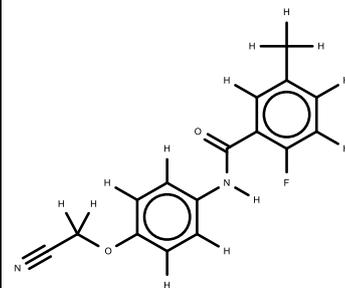
446



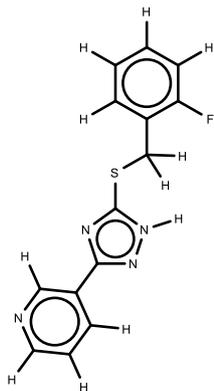
447



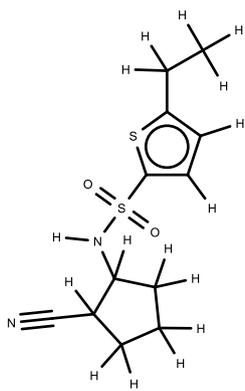
448



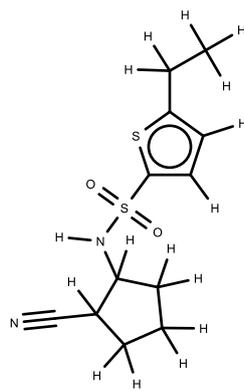
449



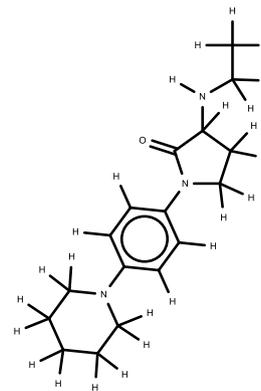
450



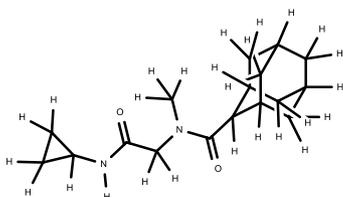
451



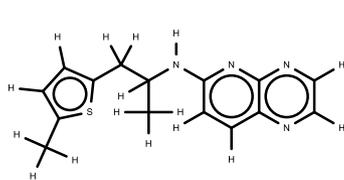
452



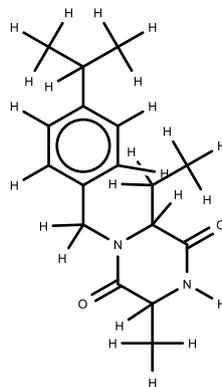
453



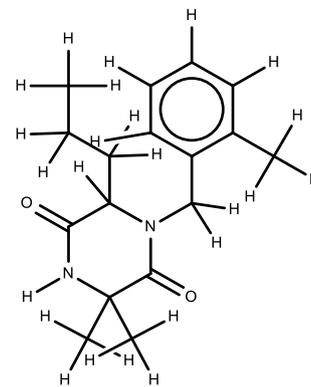
454



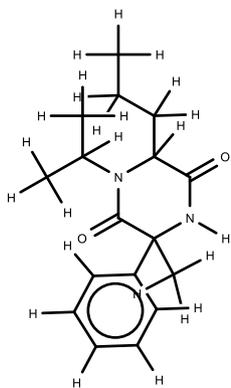
455



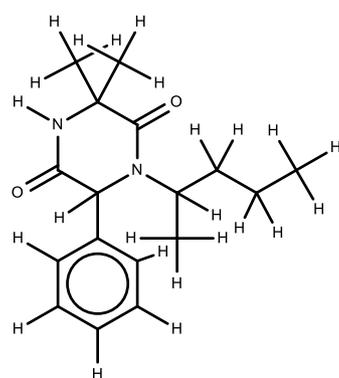
456



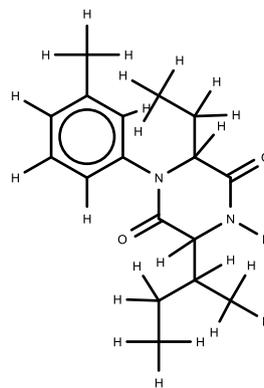
457



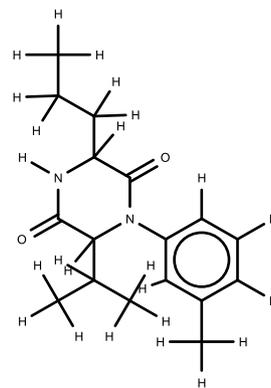
458



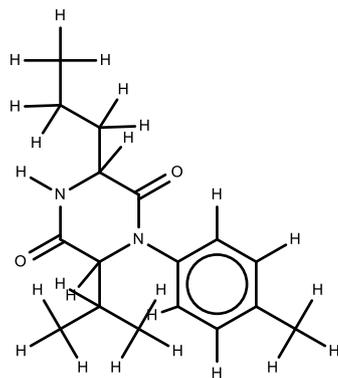
459



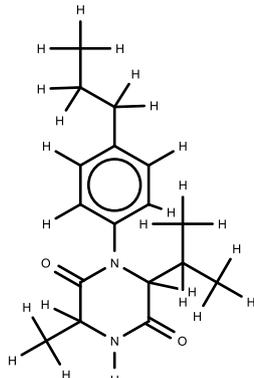
460



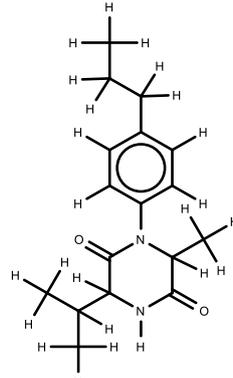
461



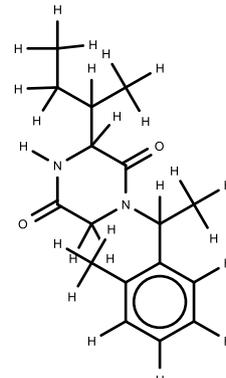
462



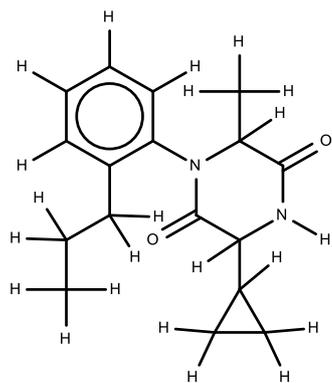
463



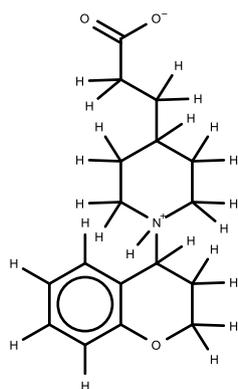
464



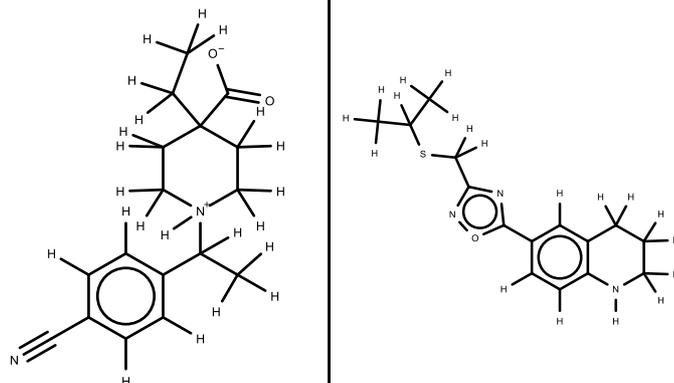
465



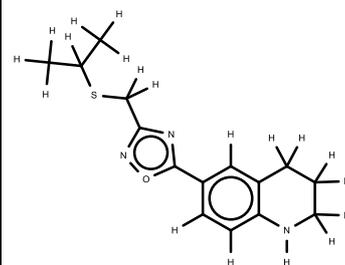
466



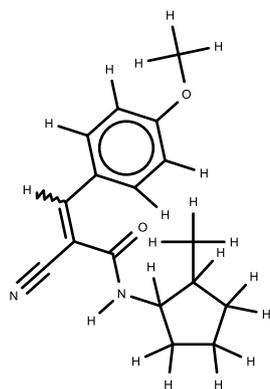
467



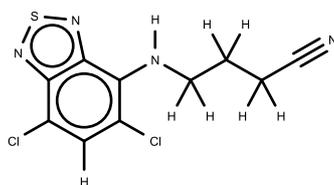
468



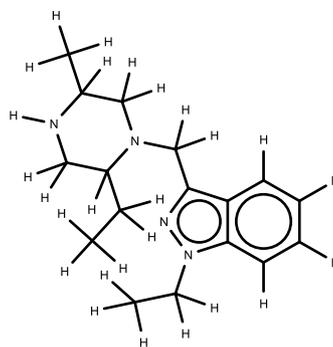
469



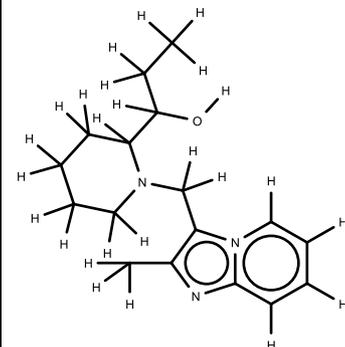
470



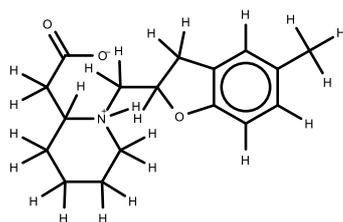
471



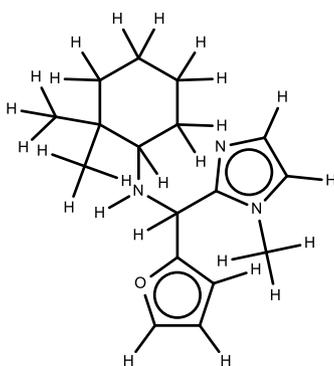
472



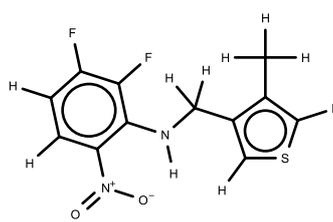
473



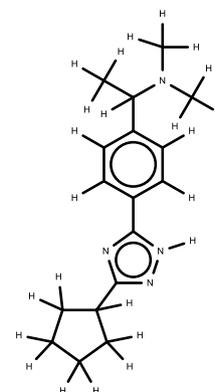
474



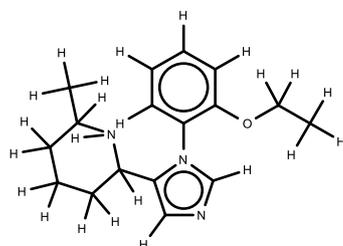
475



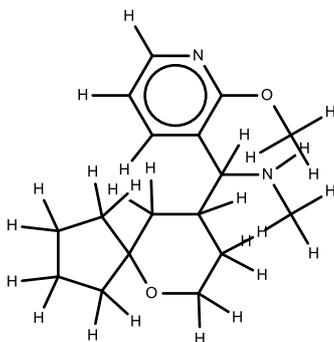
476



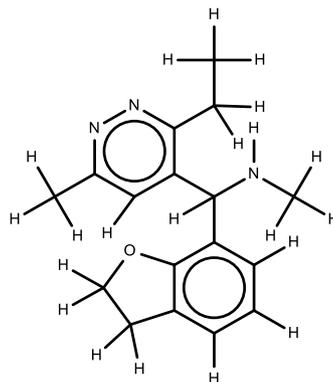
477



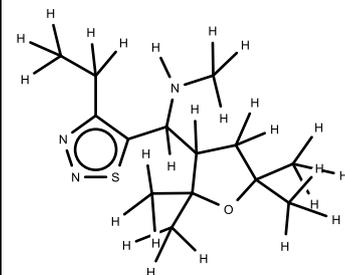
478



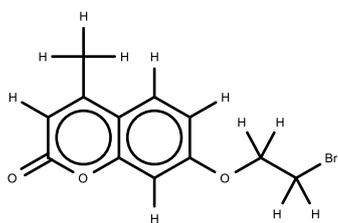
479



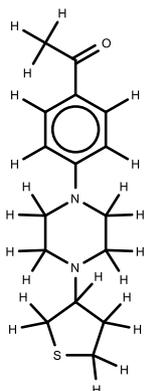
480



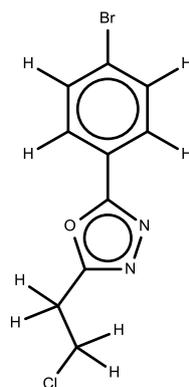
481



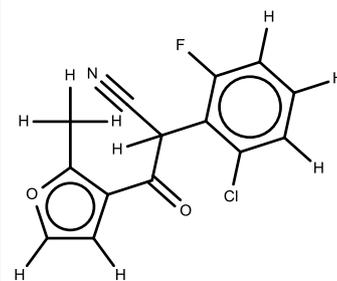
482



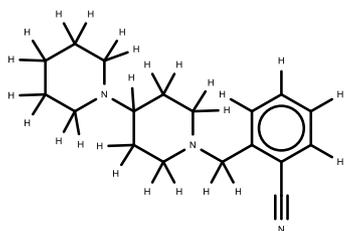
483



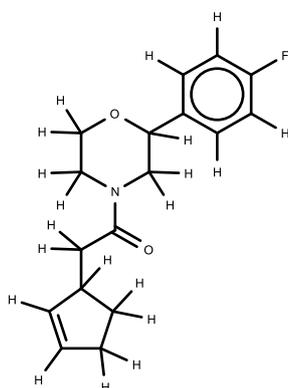
484



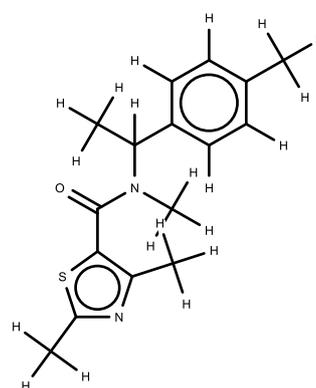
485



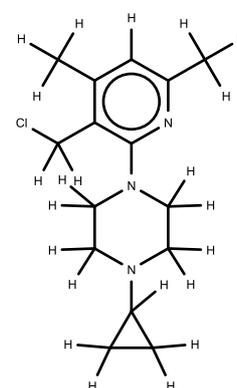
486



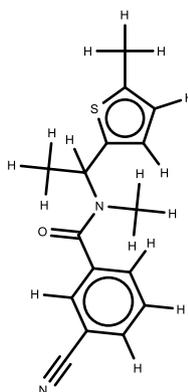
487



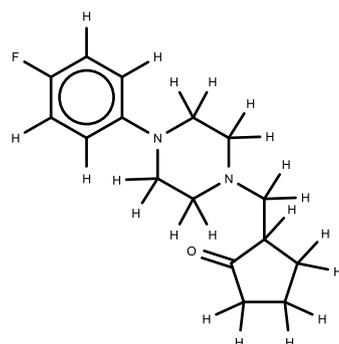
488



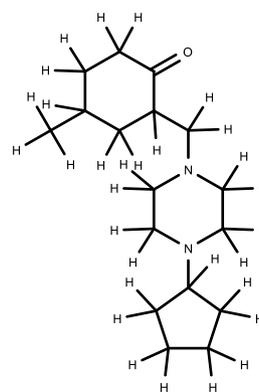
489



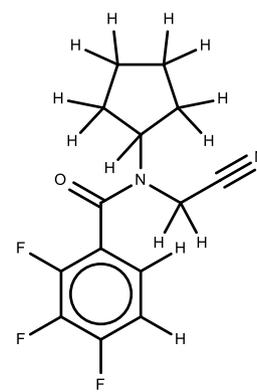
490



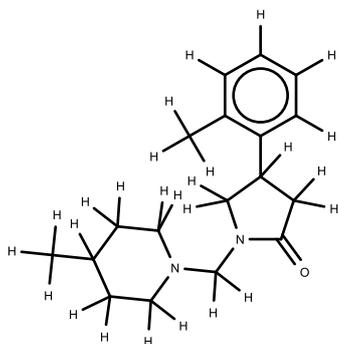
491



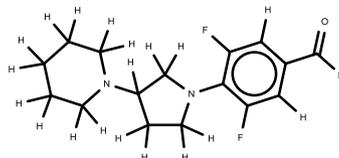
492



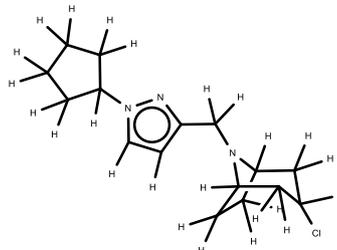
493



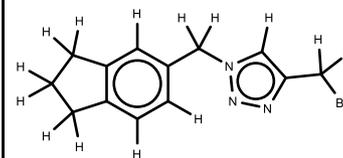
494



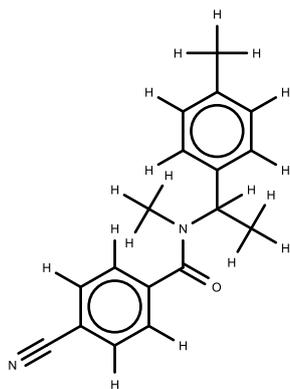
495



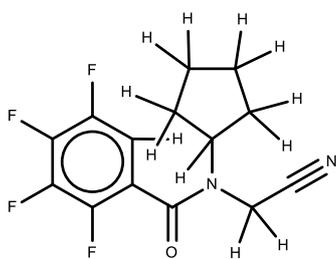
496



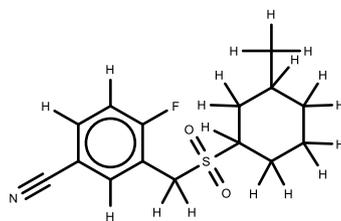
497



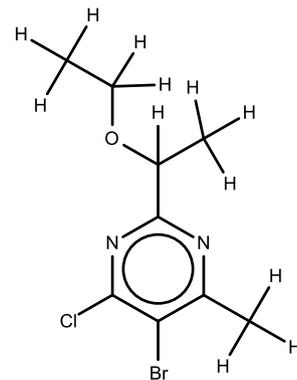
498



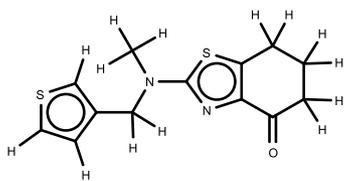
499



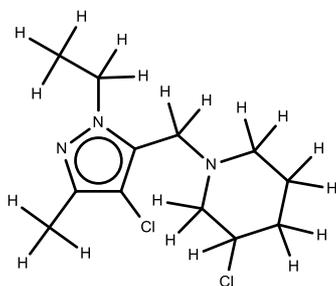
500



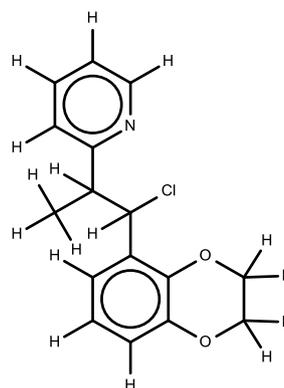
501



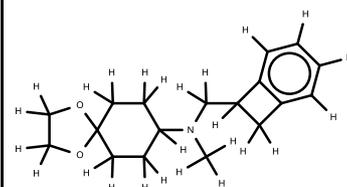
502



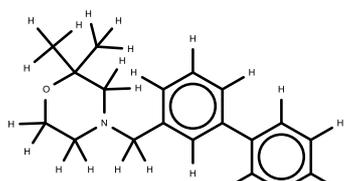
503



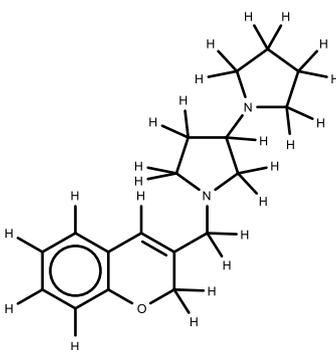
504



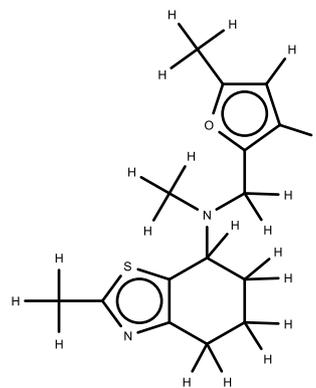
505



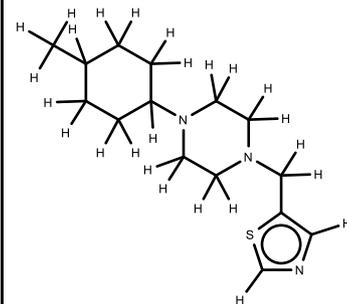
506



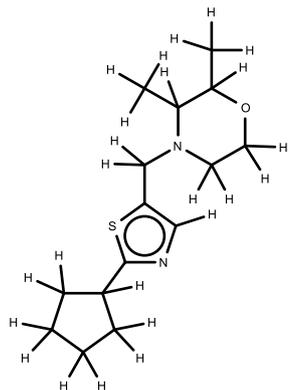
507



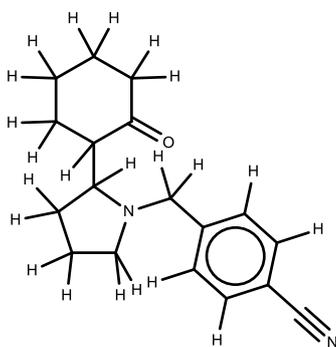
508



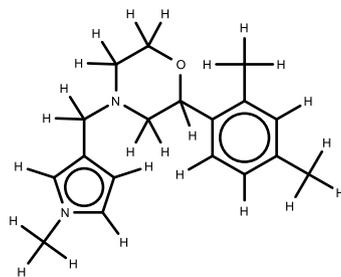
509



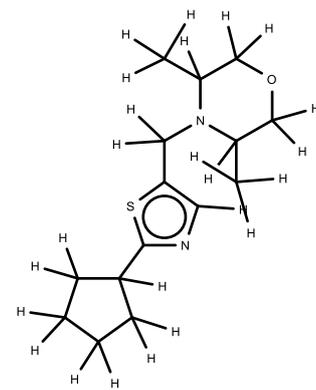
510



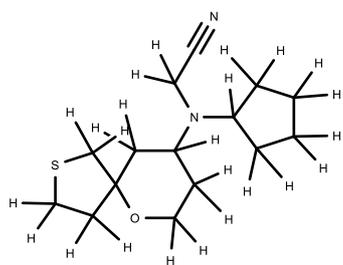
511



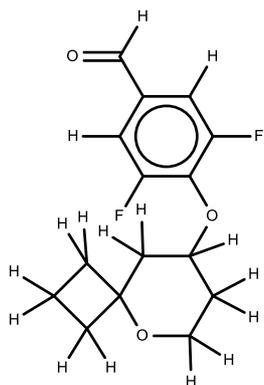
512



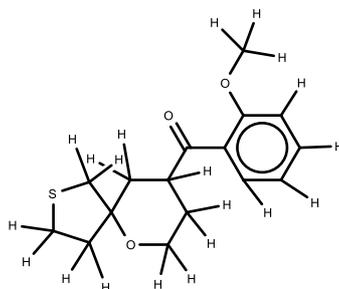
513



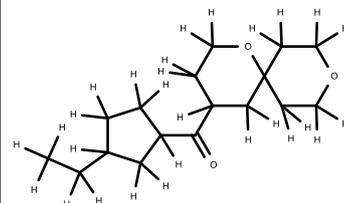
514



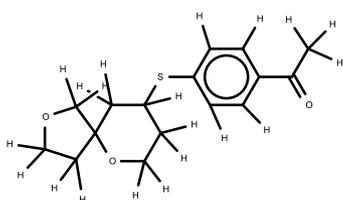
515



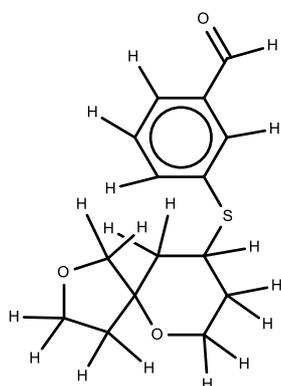
516



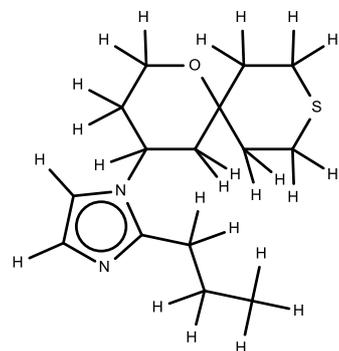
517



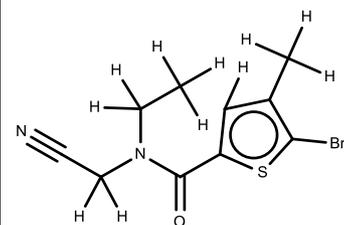
518



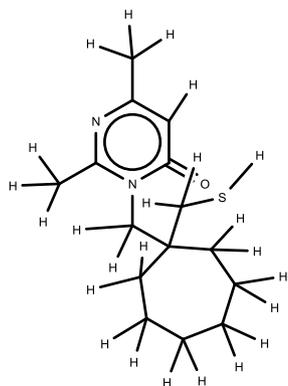
519



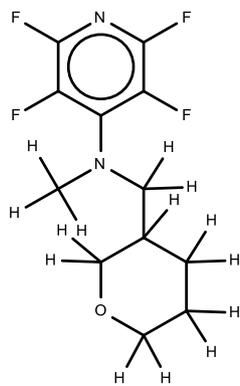
520



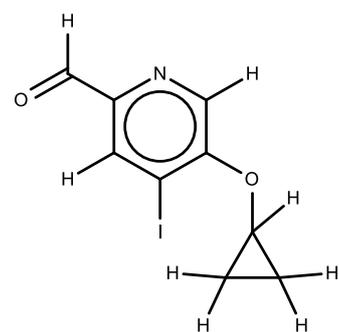
521



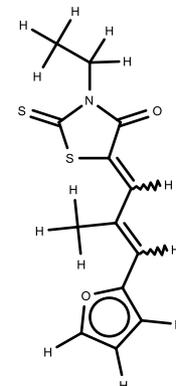
522



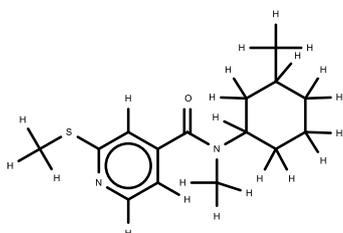
523



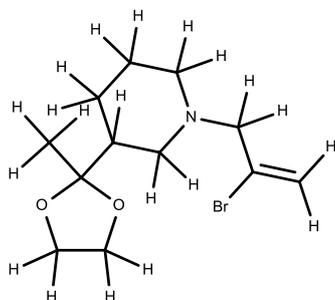
524



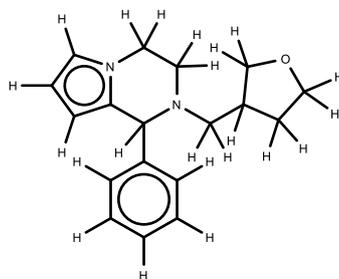
525



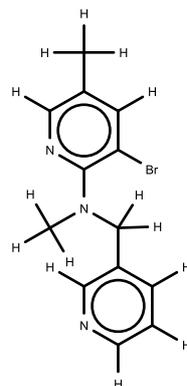
526



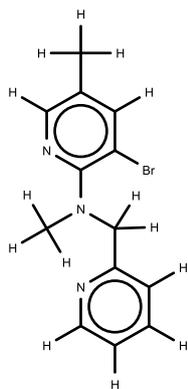
527



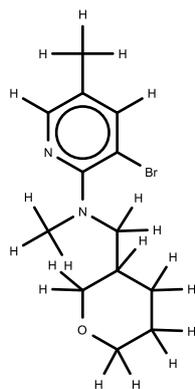
528



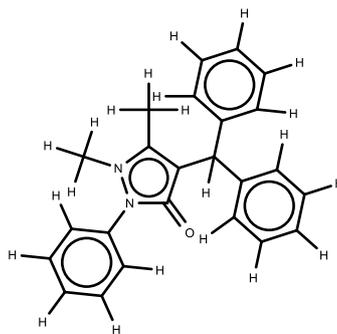
529



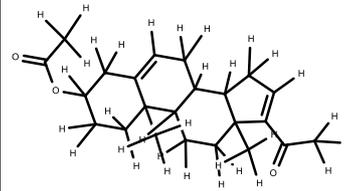
530



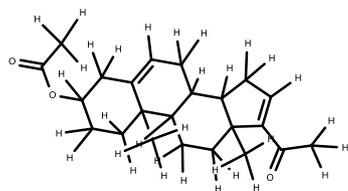
531



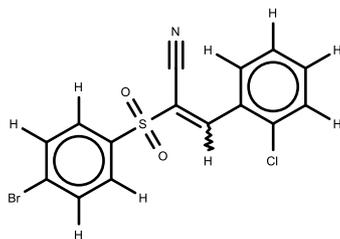
532



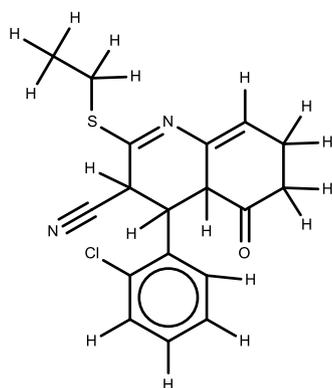
533



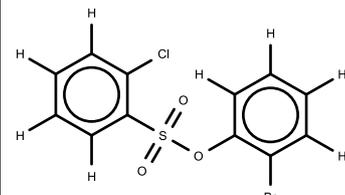
534



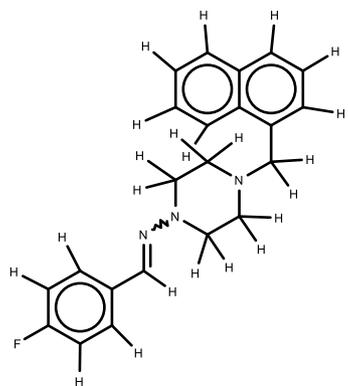
535



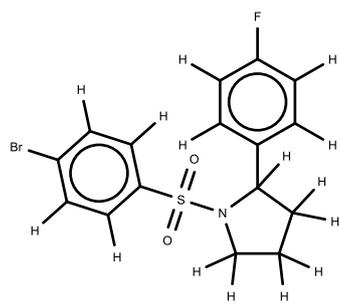
536



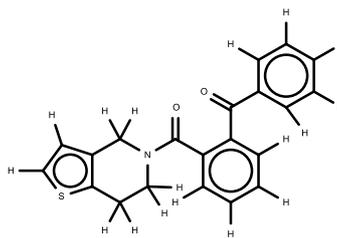
537



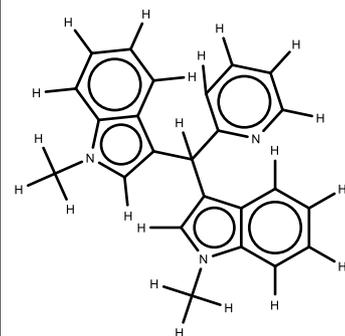
538



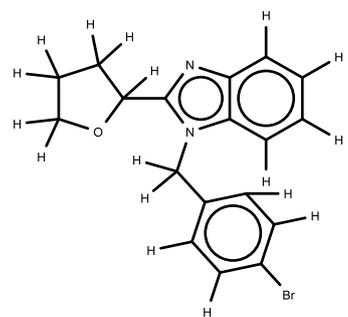
539



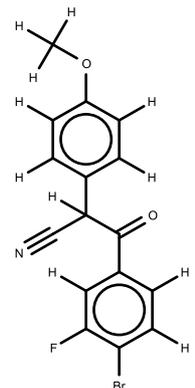
540



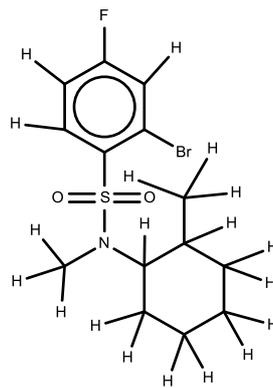
541



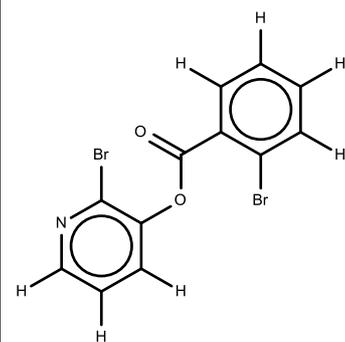
542



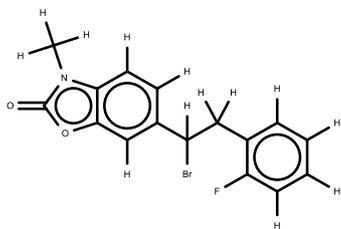
543



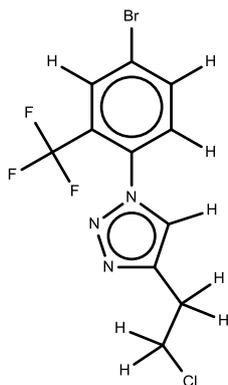
544



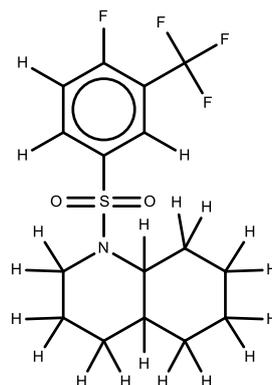
545



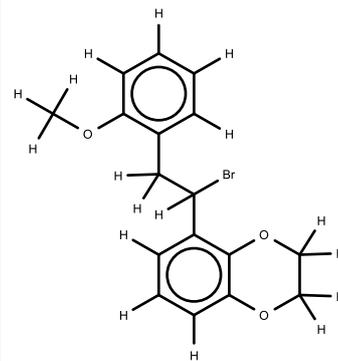
546



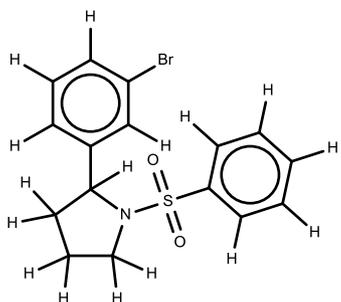
547



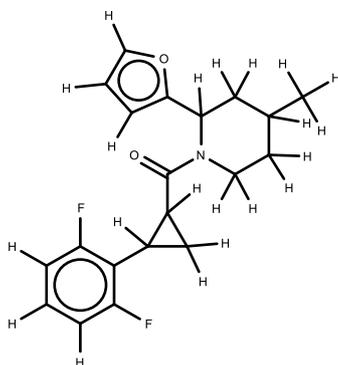
548



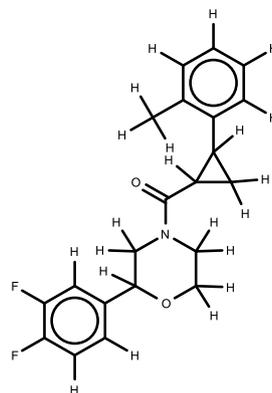
549



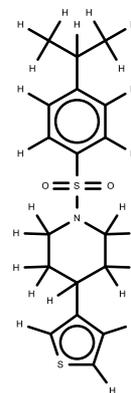
550



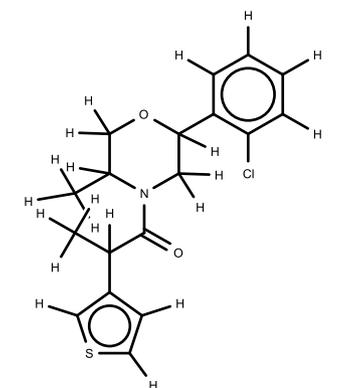
551



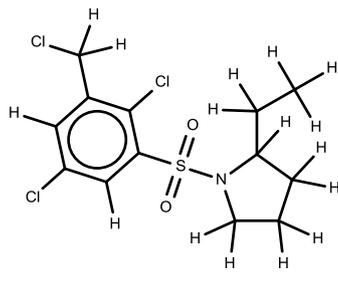
552



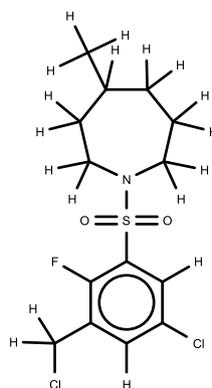
553



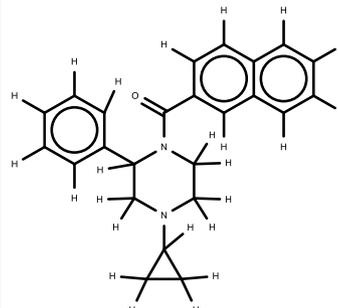
554



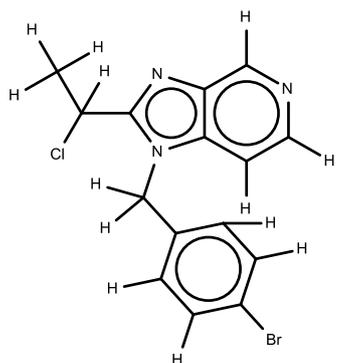
555



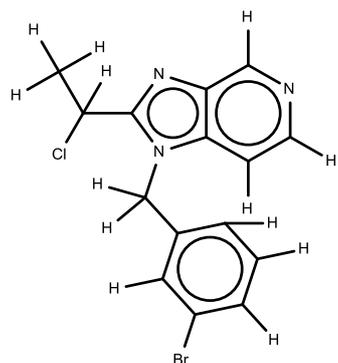
556



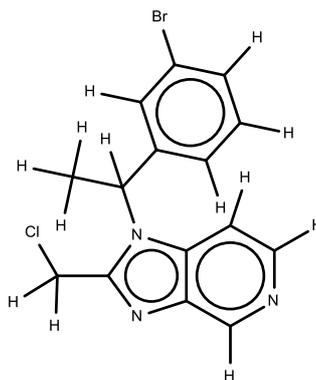
557



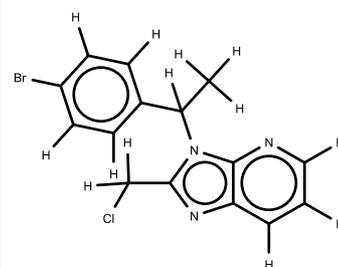
558



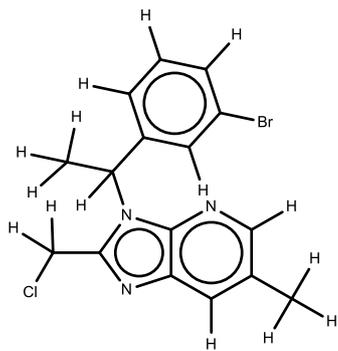
559



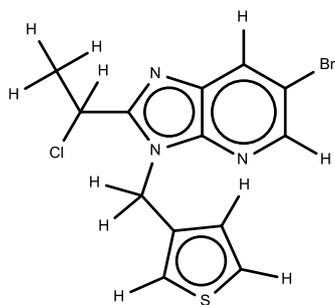
560



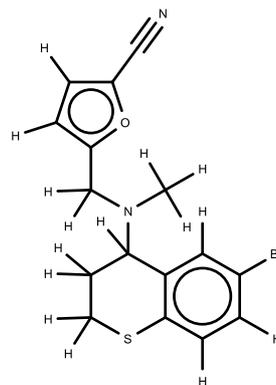
561



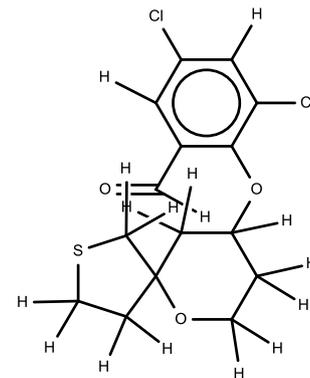
562



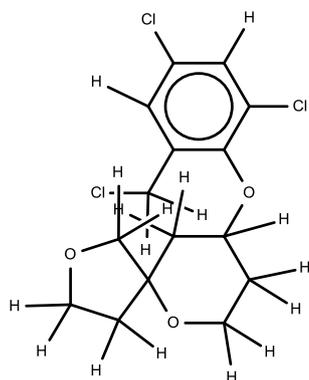
563



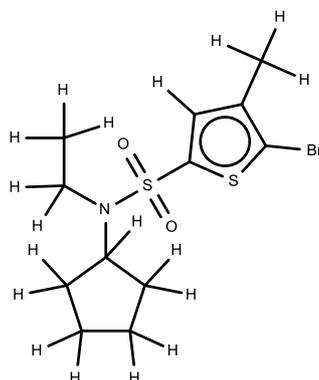
564



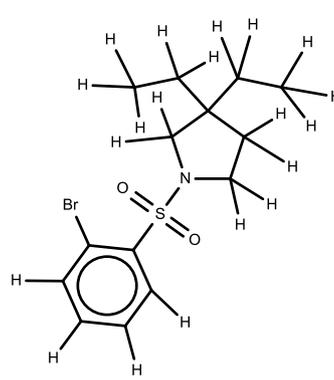
565



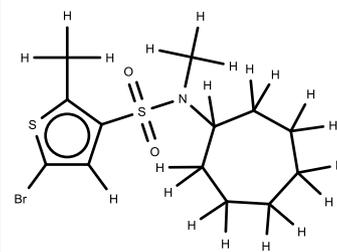
566



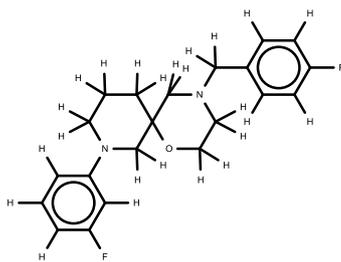
567



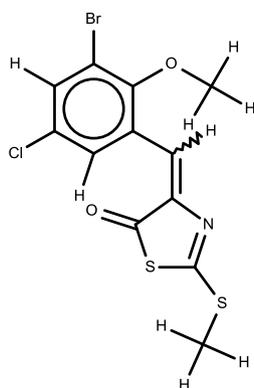
568



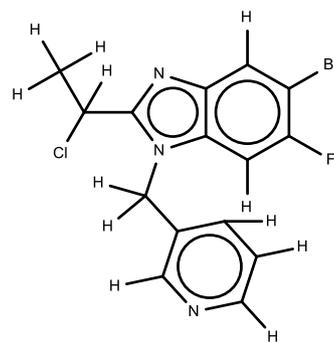
569



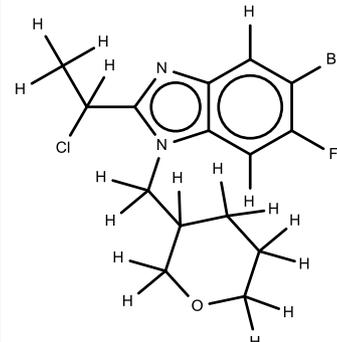
570



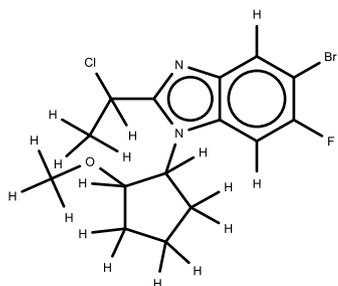
571



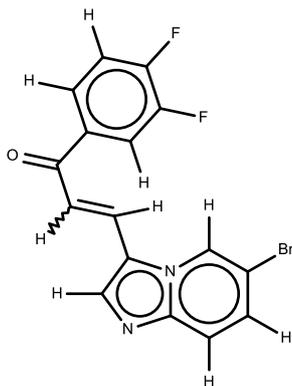
572



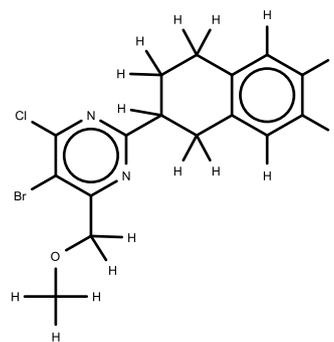
573



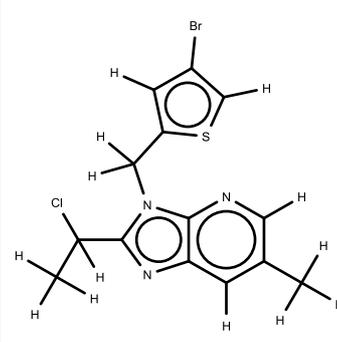
574



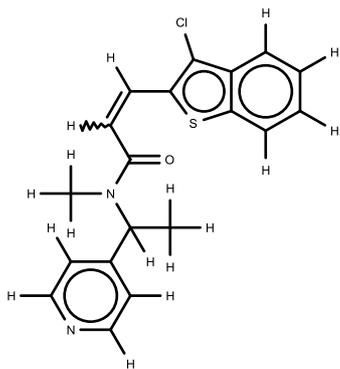
575



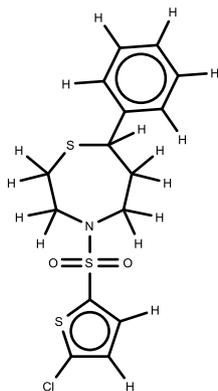
576



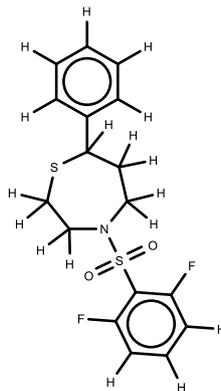
577



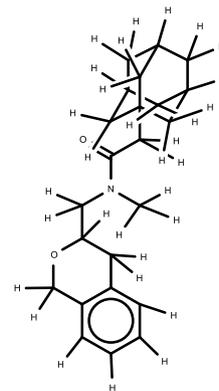
578



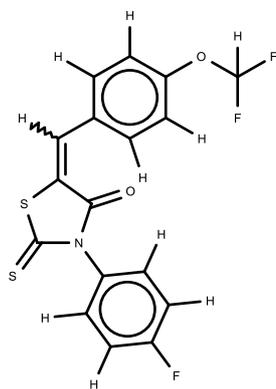
579



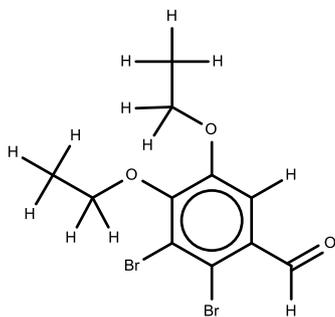
580



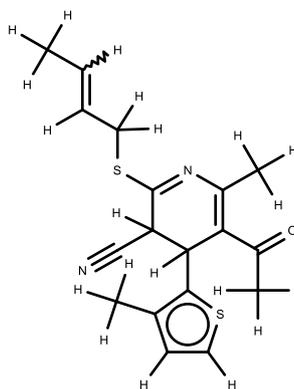
581



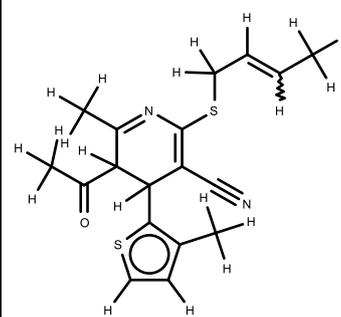
582



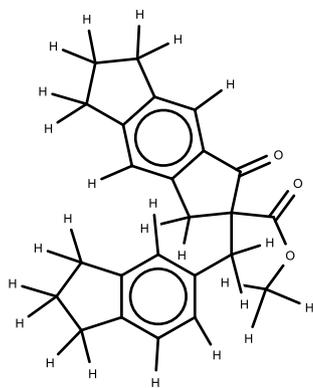
583



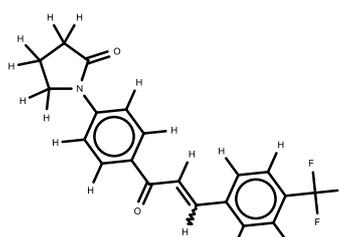
584



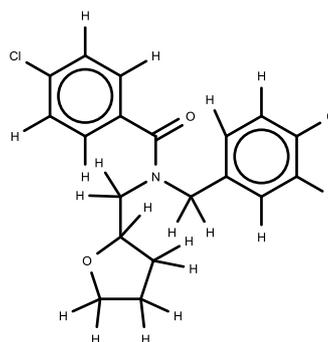
585



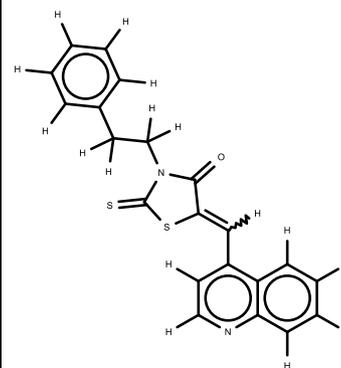
586



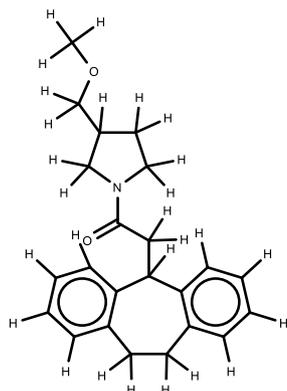
587



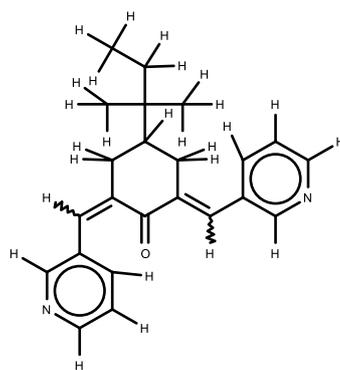
588



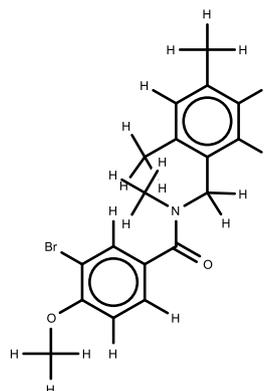
589



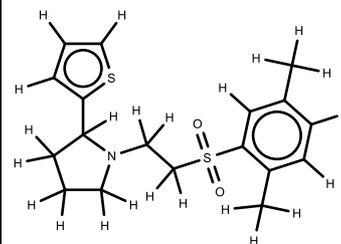
590



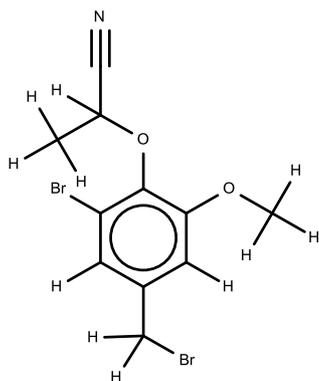
591



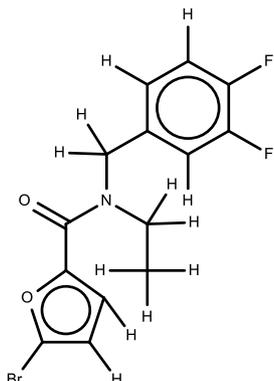
592



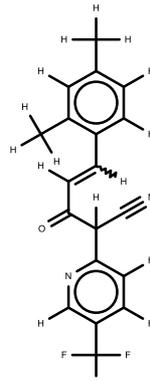
593



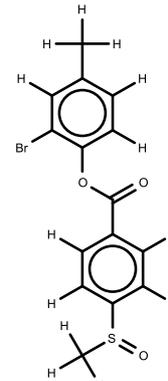
594



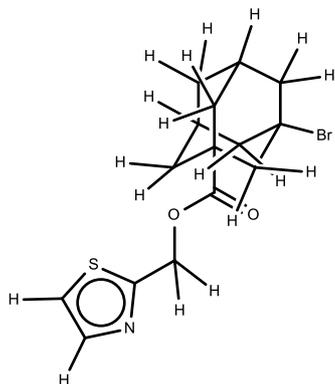
595



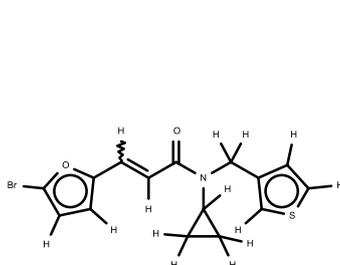
596



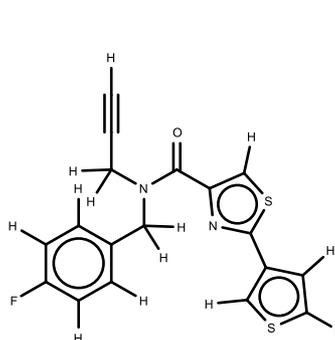
597



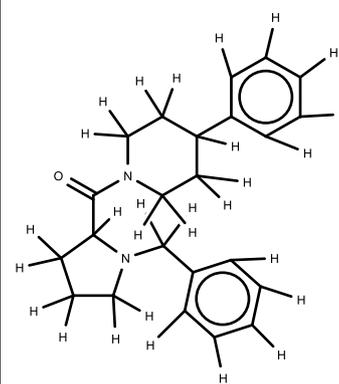
598



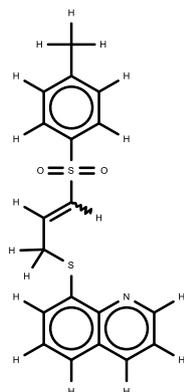
599



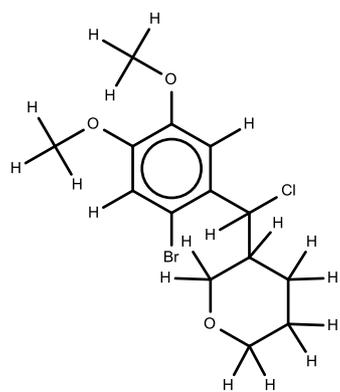
600



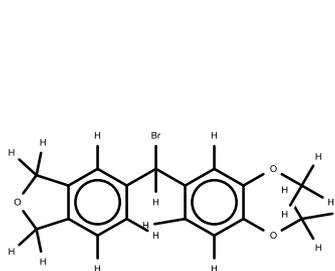
601



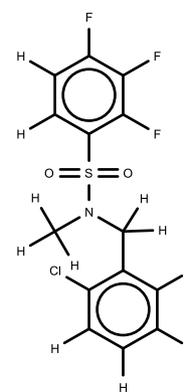
602



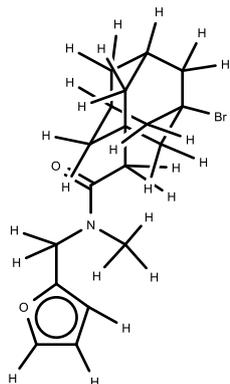
603



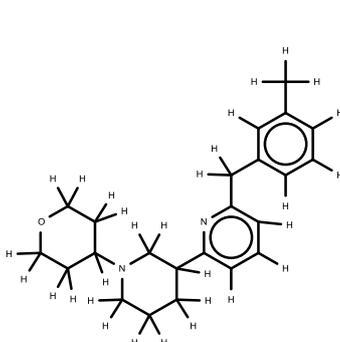
604



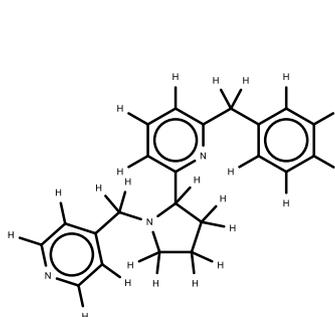
605



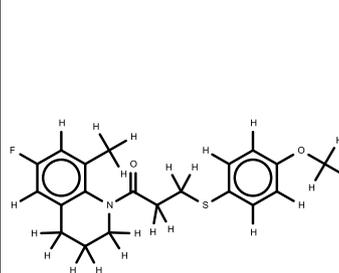
606



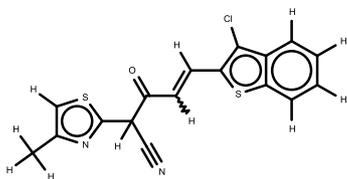
607



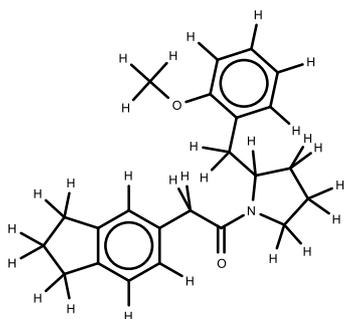
608



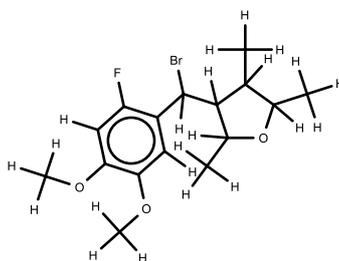
609



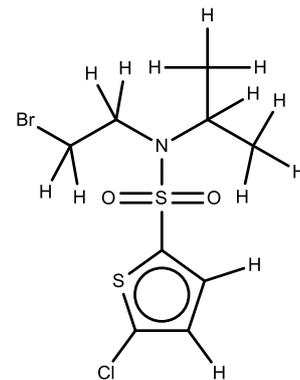
610



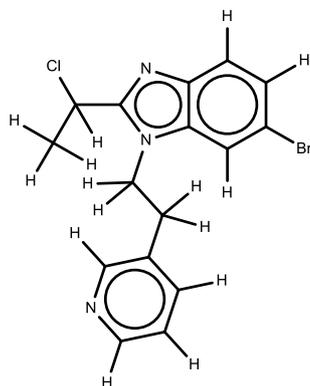
611



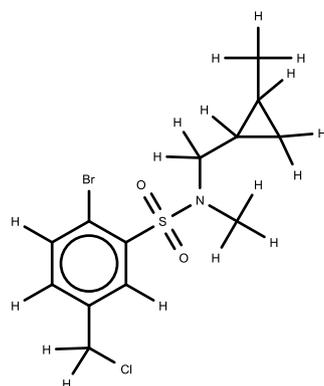
612



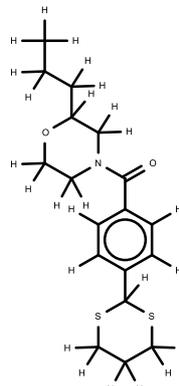
613



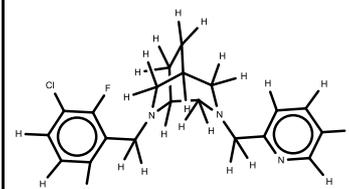
614



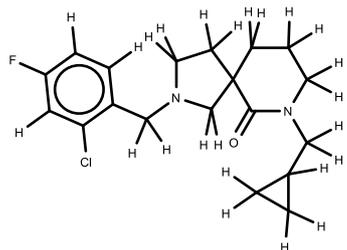
615



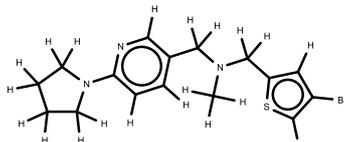
616



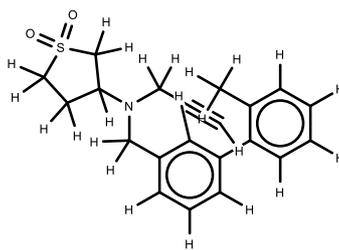
617



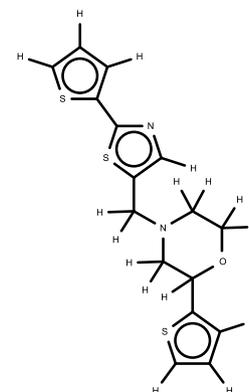
618



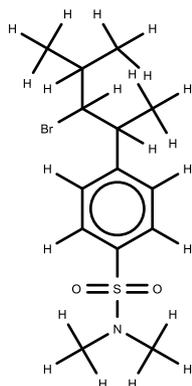
619



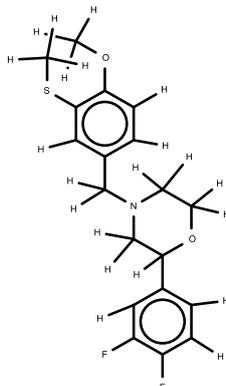
620



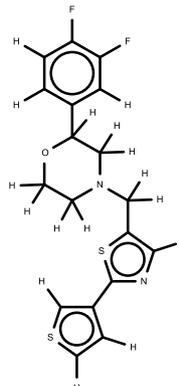
621



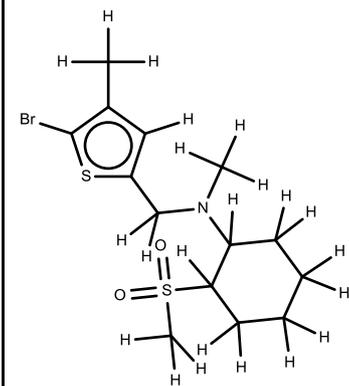
622



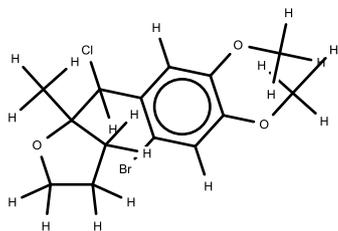
623



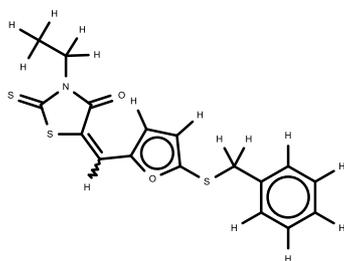
624



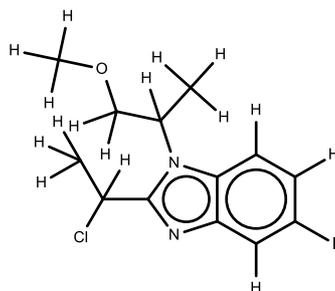
625



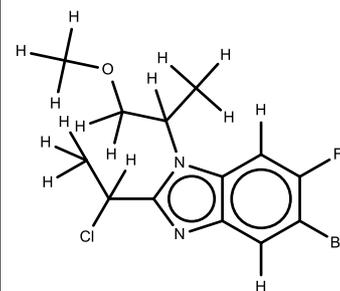
626



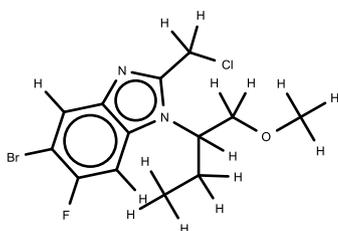
627



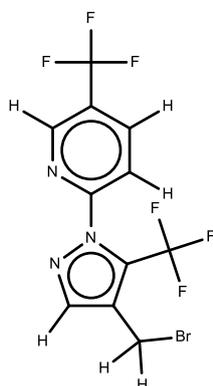
628



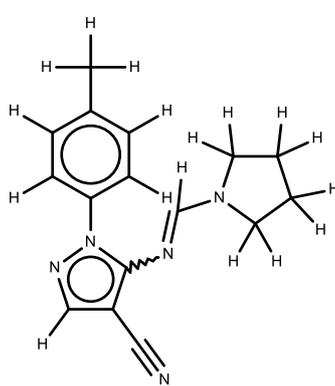
629



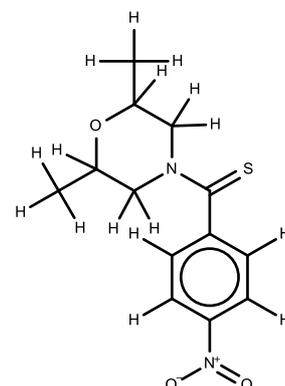
630



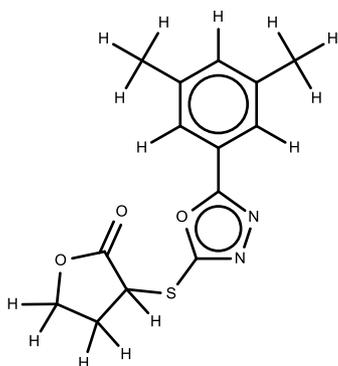
631



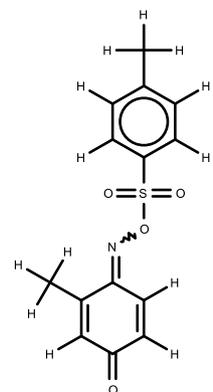
632



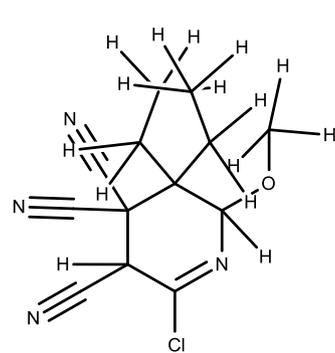
633



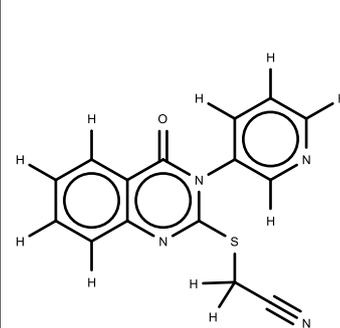
634



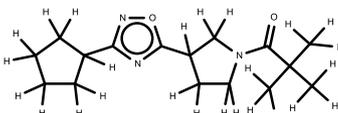
635



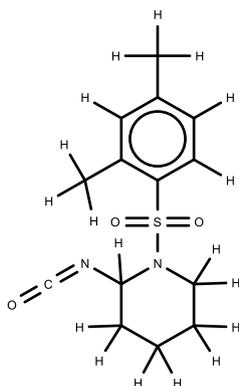
636



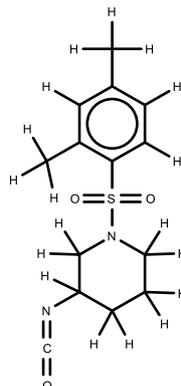
637



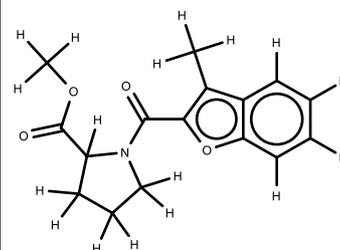
638



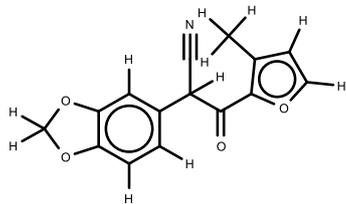
639



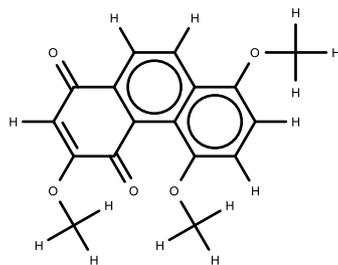
640



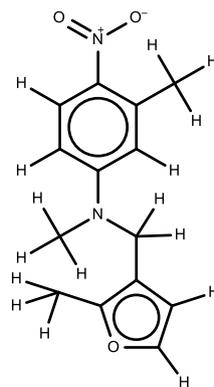
641



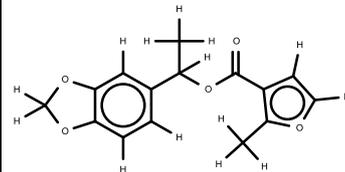
642



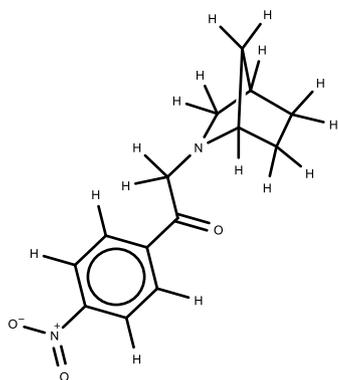
643



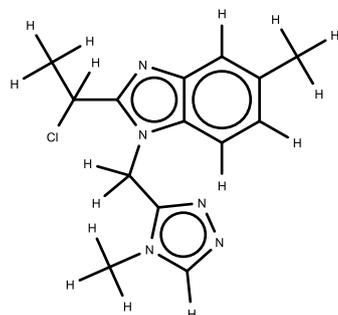
644



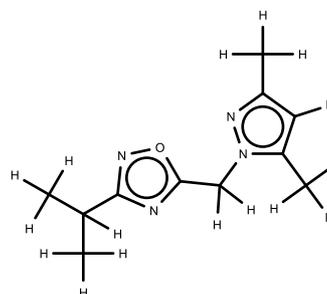
645



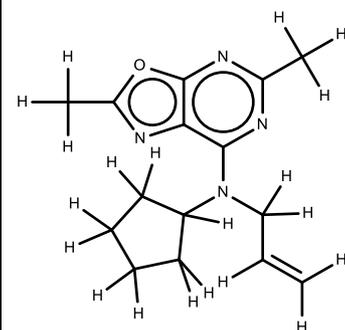
646



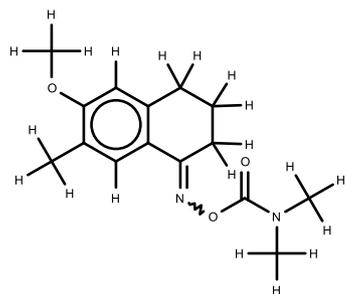
647



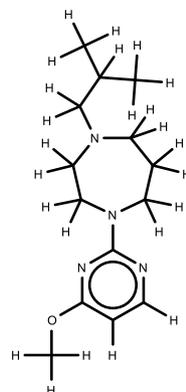
648



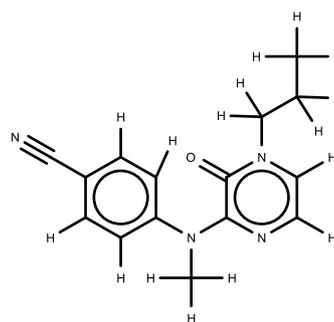
649



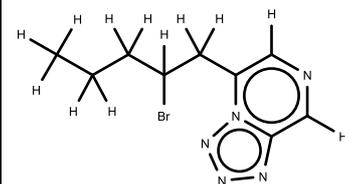
650



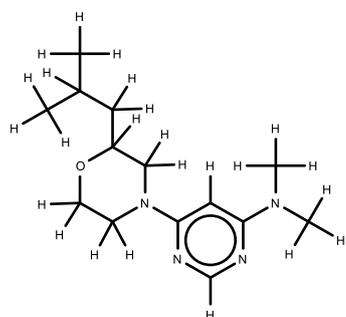
651



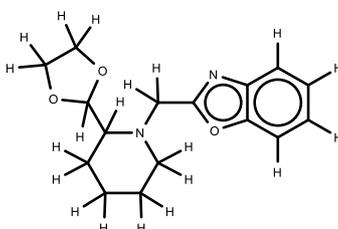
652



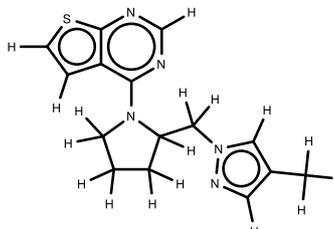
653



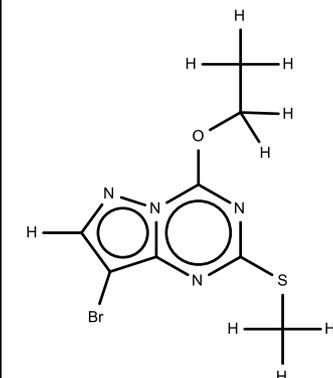
654



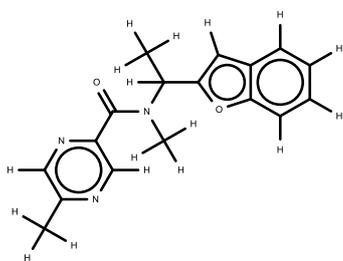
655



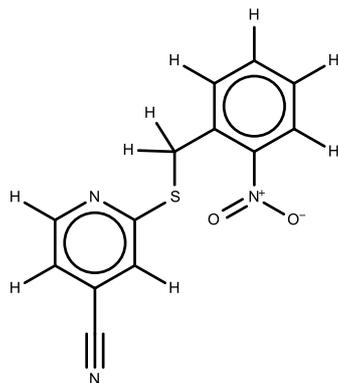
656



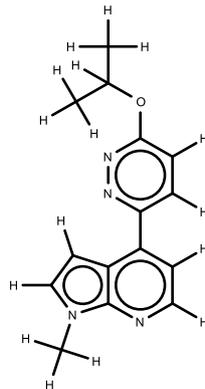
657



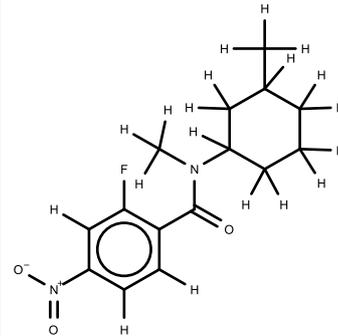
658



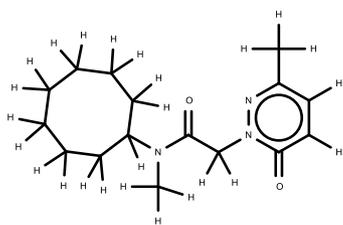
659



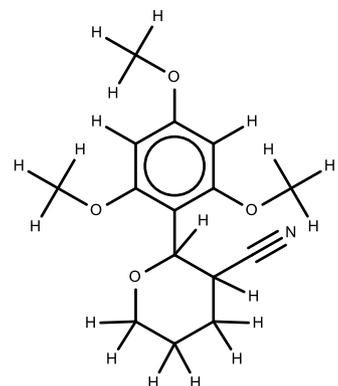
660



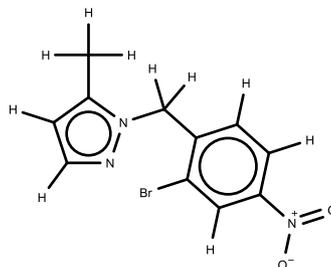
661



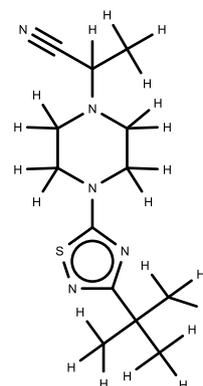
662



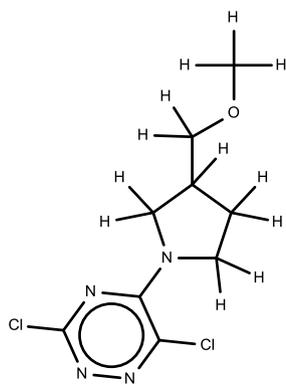
663



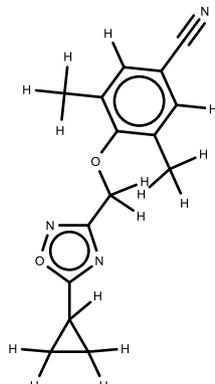
664



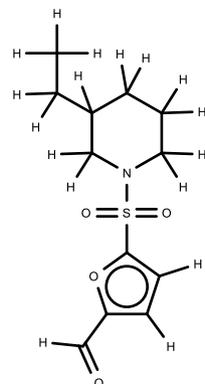
665



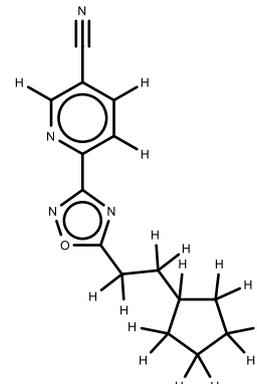
666



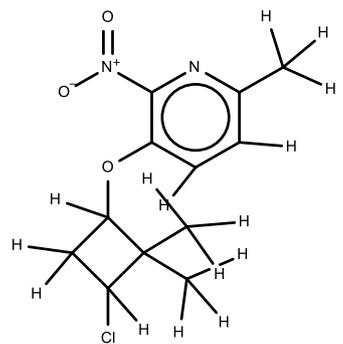
667



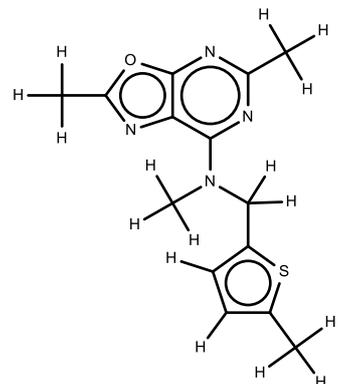
668



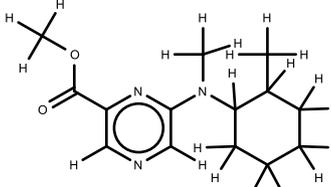
669



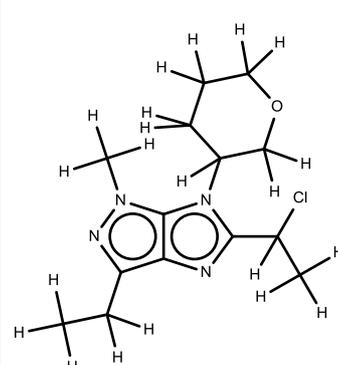
670



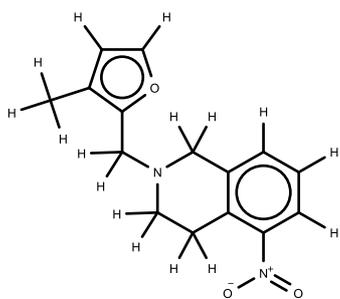
671



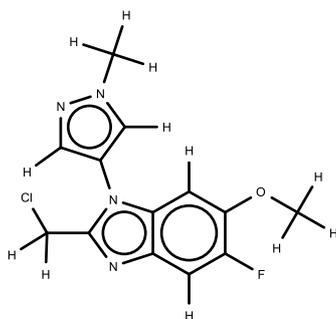
672



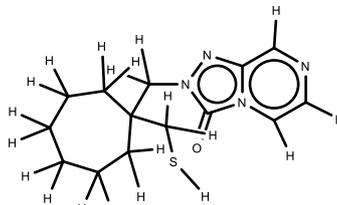
673



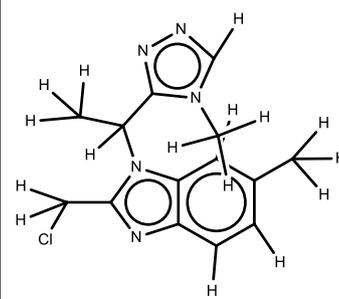
674



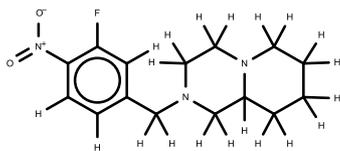
675



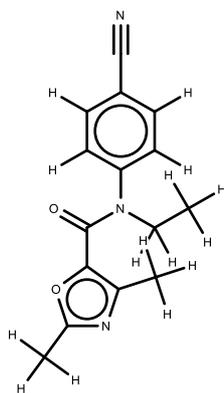
676



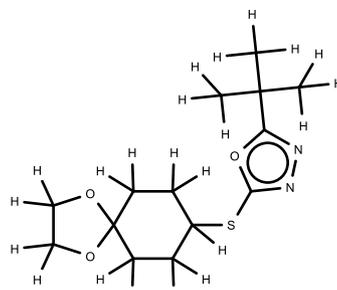
677



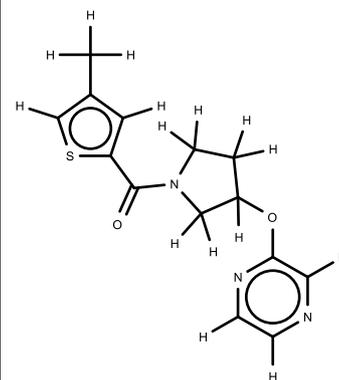
678



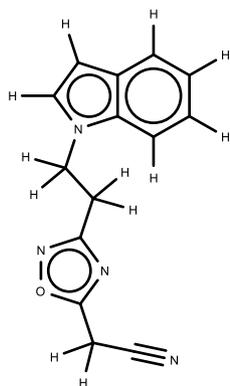
679



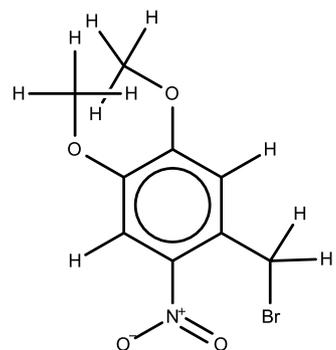
680



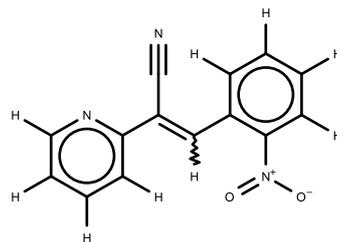
681



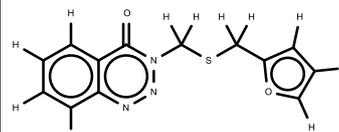
682



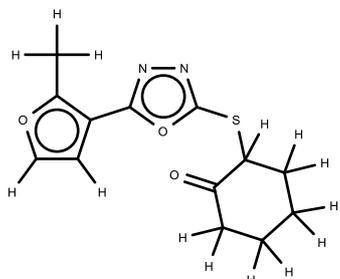
683



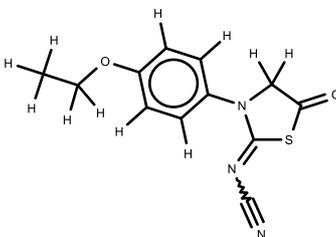
684



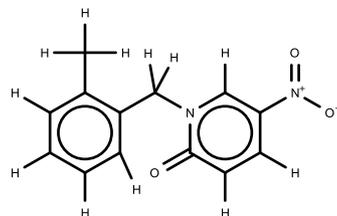
685



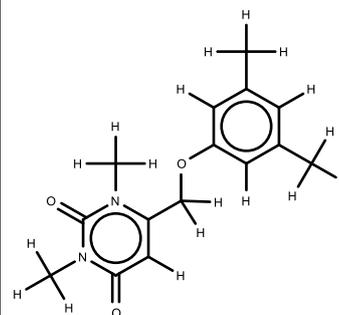
686



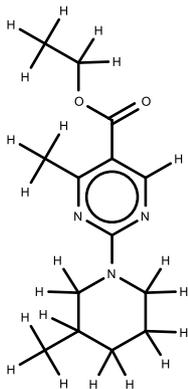
687



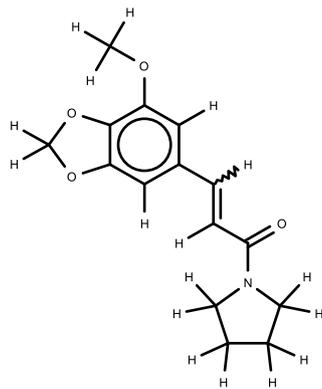
688



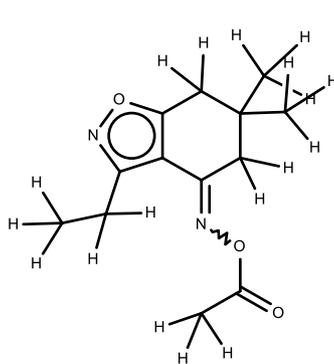
689



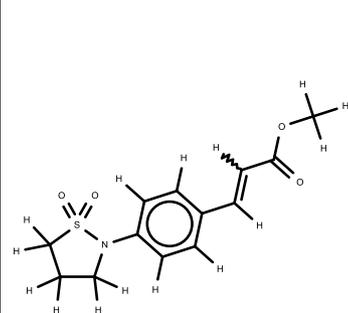
690



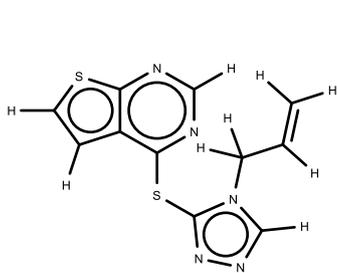
691



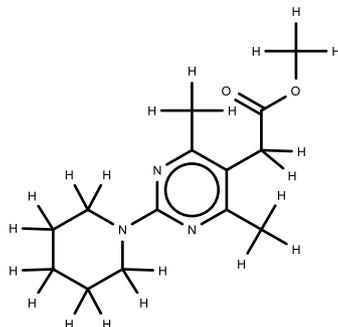
692



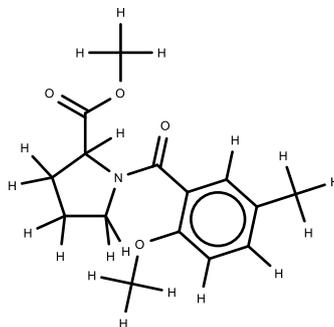
693



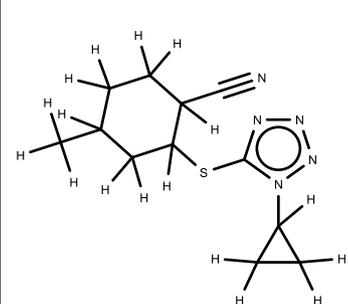
694



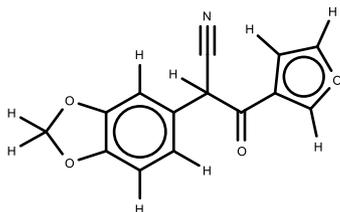
695



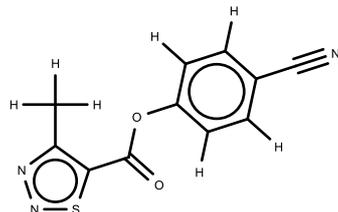
696



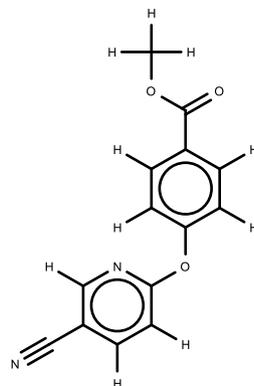
697



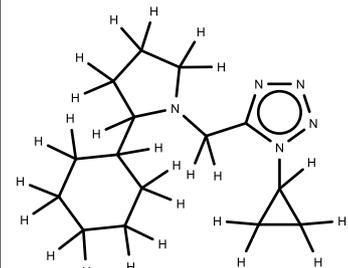
698



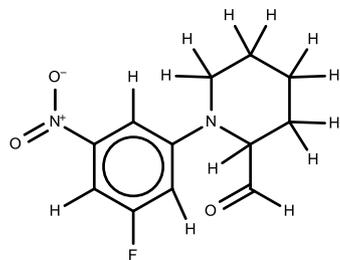
699



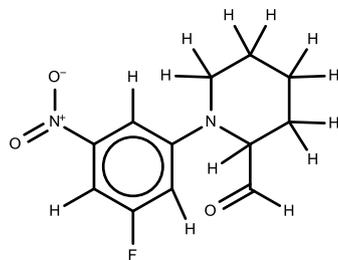
700



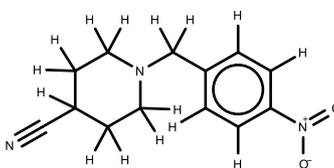
701



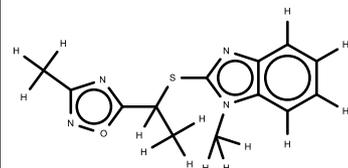
702



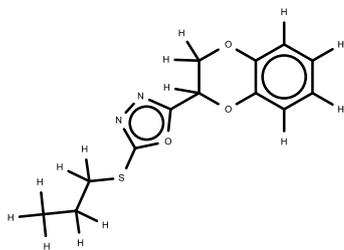
703



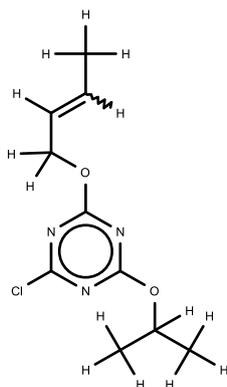
704



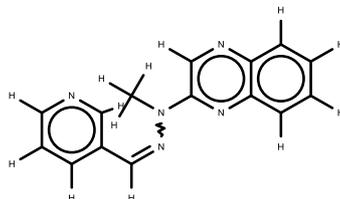
705



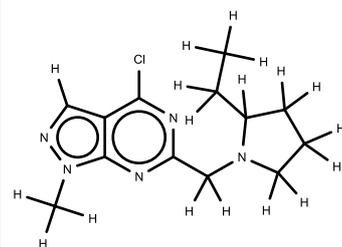
706



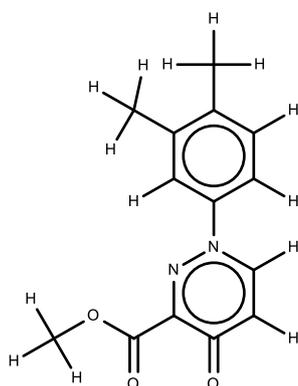
707



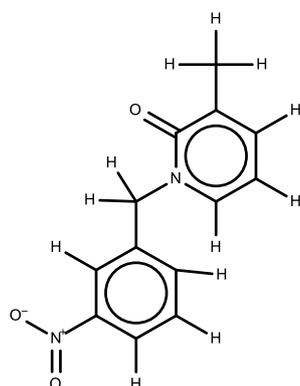
708



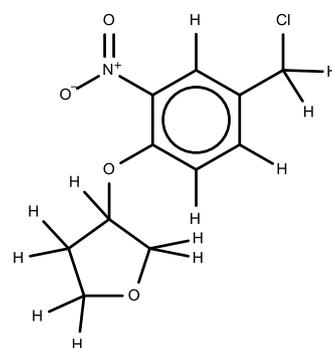
709



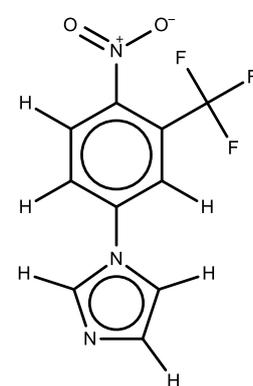
710



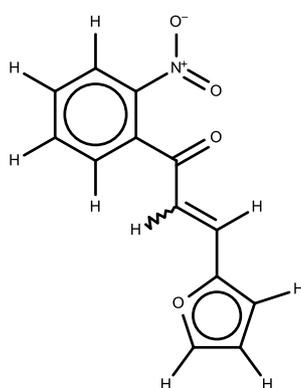
711



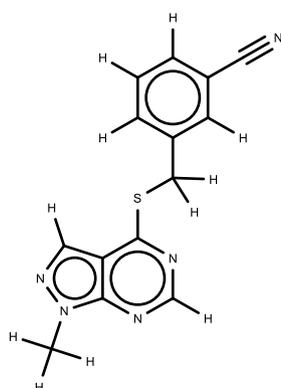
712



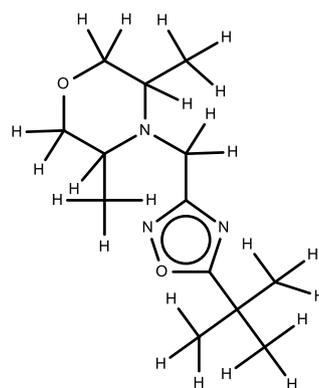
713



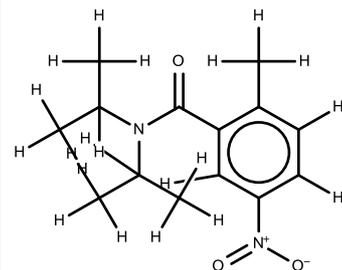
714



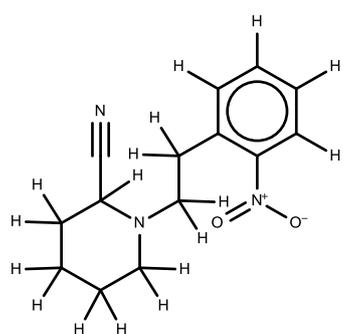
715



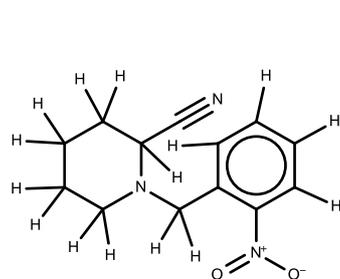
716



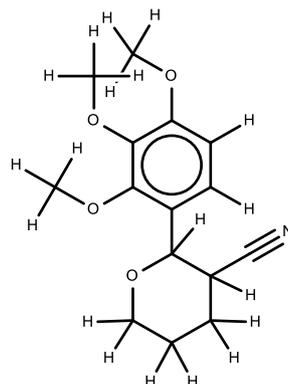
717



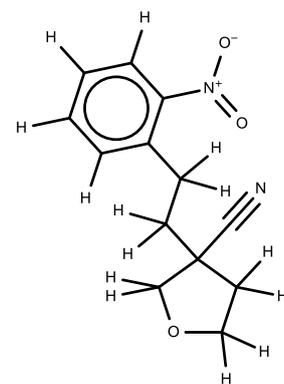
718



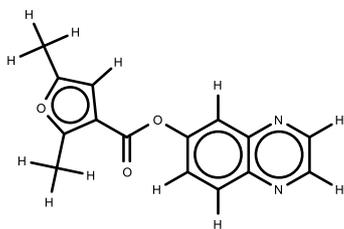
719



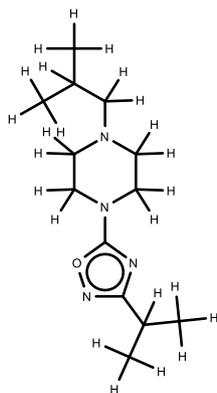
720



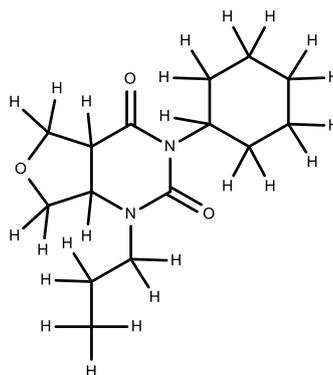
721



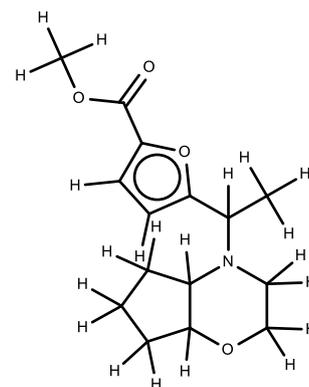
722



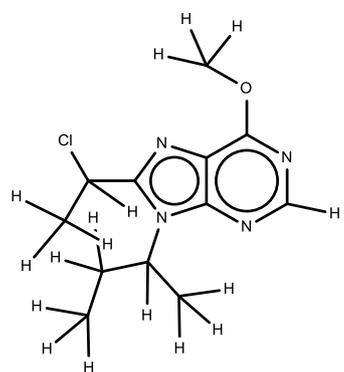
723



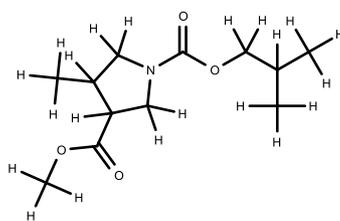
724



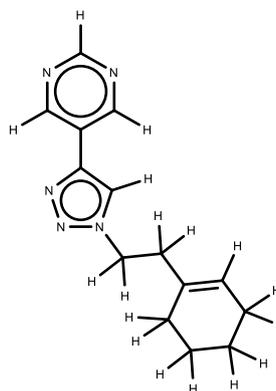
725



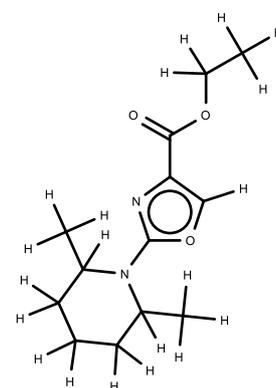
726



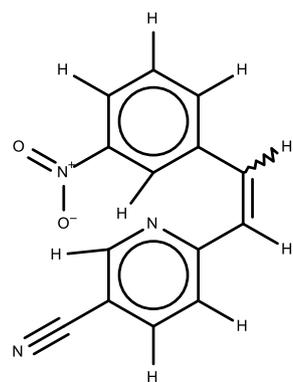
727



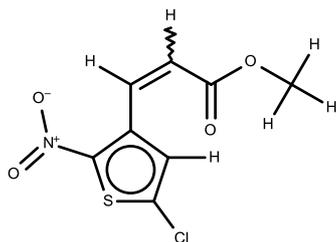
728



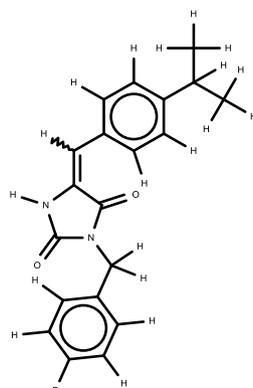
729



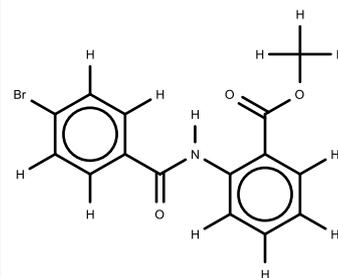
730



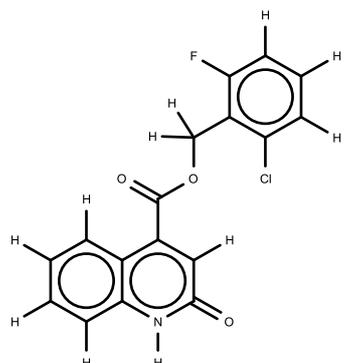
731



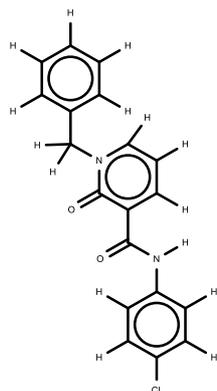
732



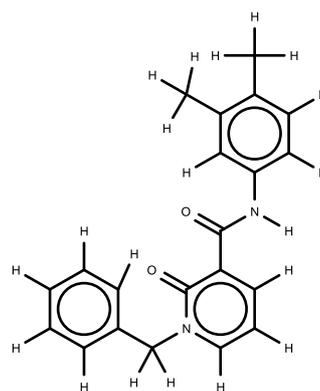
733



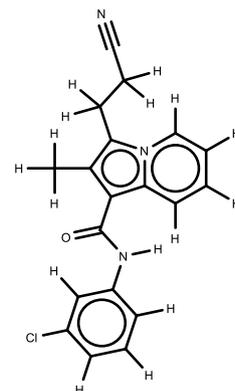
734



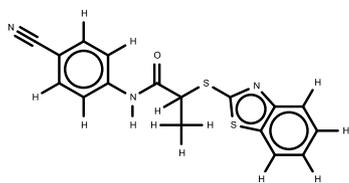
735



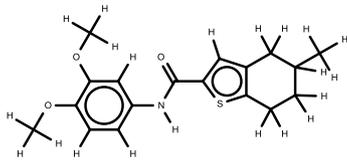
736



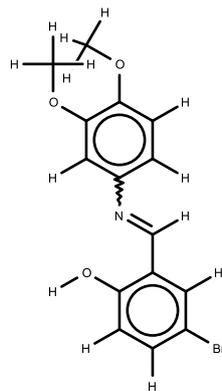
737



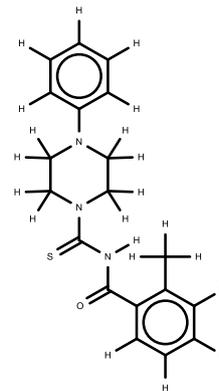
738



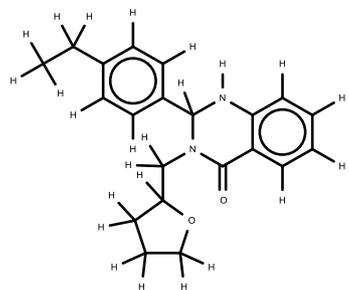
739



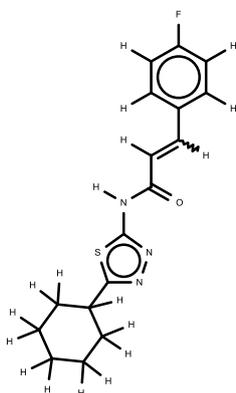
740



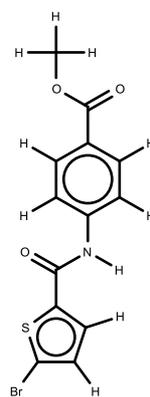
741



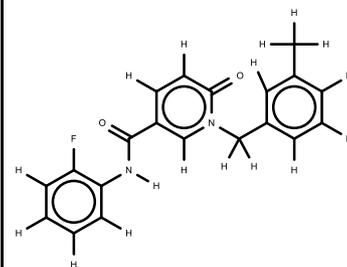
742



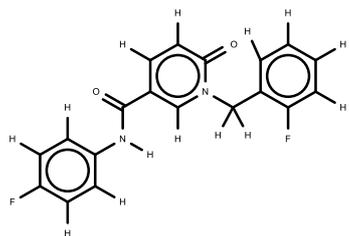
743



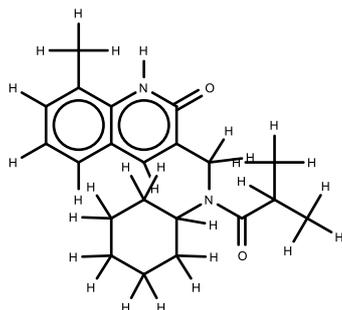
744



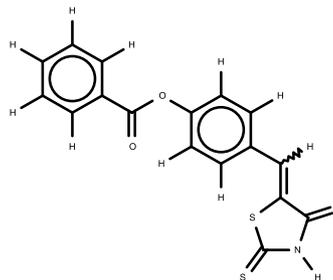
745



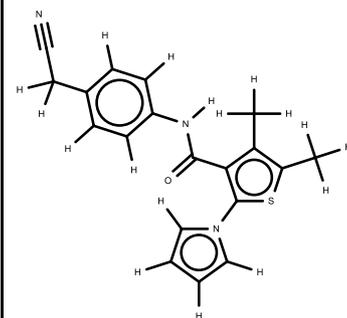
746



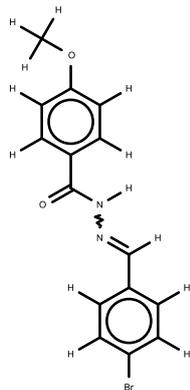
747



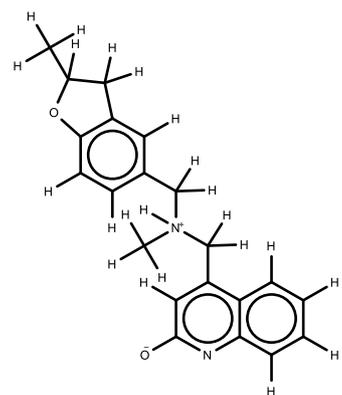
748



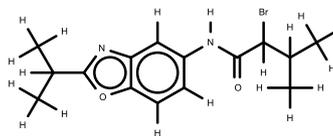
749



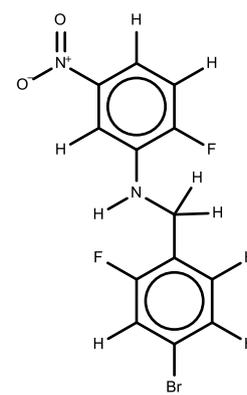
750



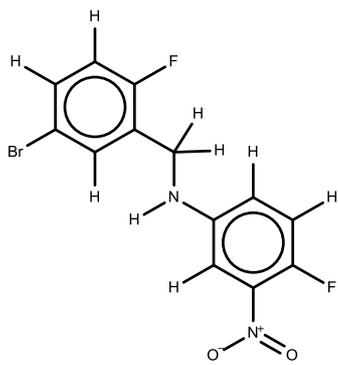
751



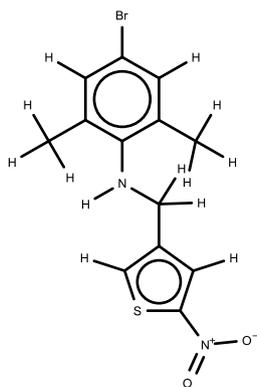
752



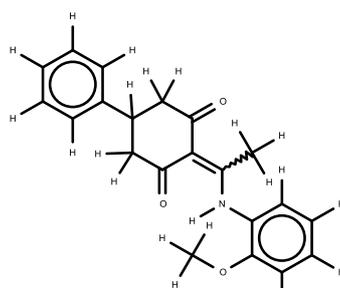
753



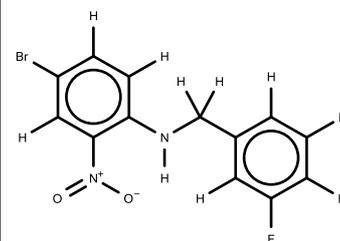
754



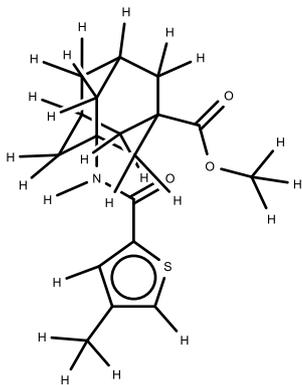
755



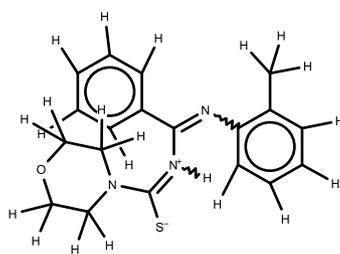
756



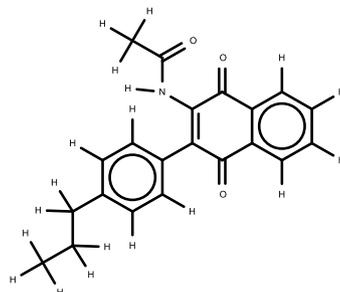
757



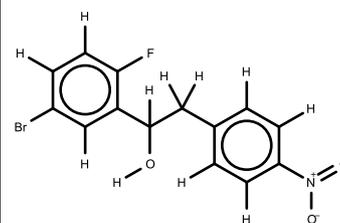
758



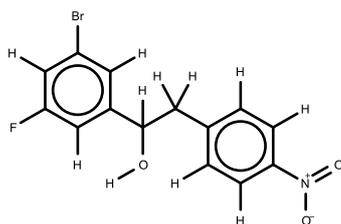
759



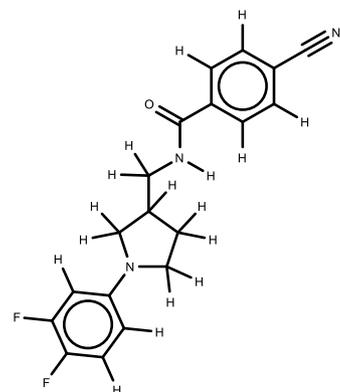
760



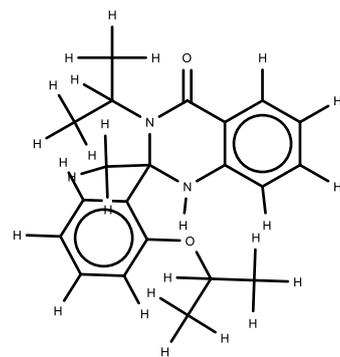
761



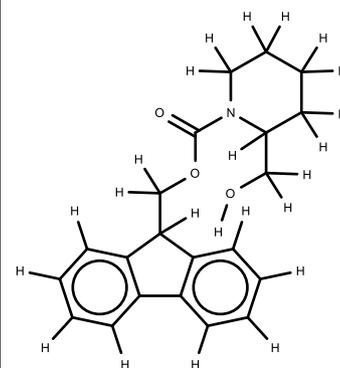
762



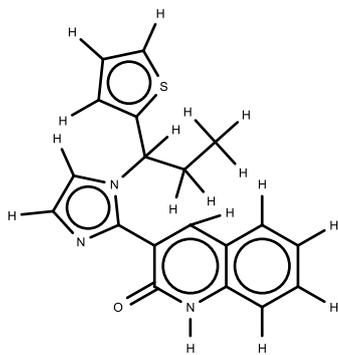
763



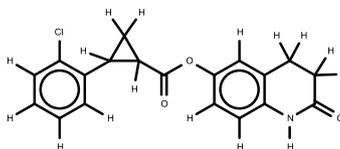
764



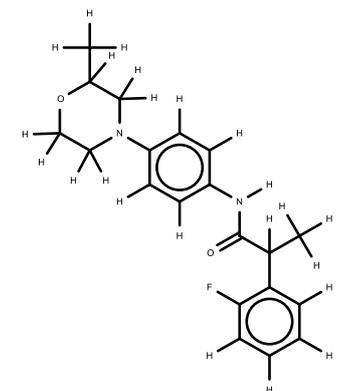
765



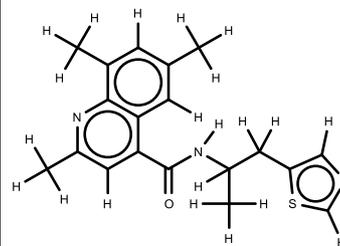
766



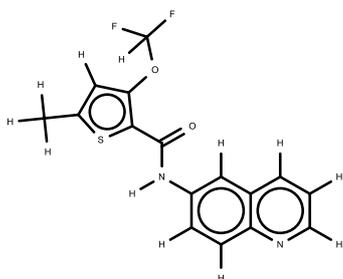
767



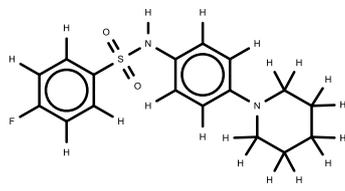
768



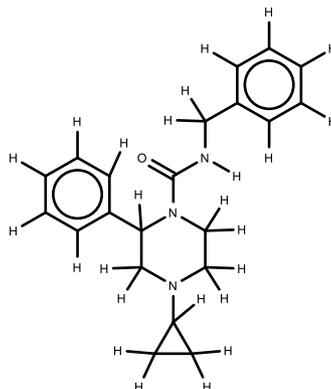
769



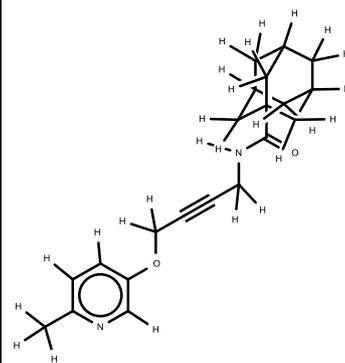
770



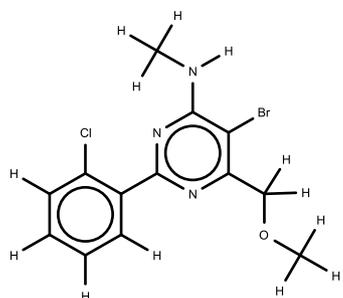
771



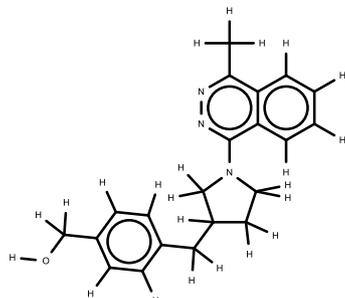
772



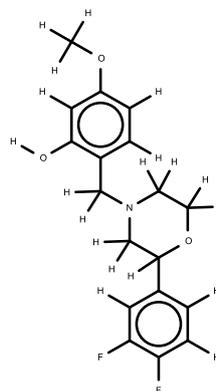
773



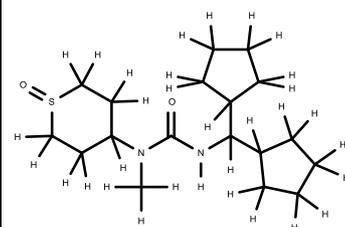
774



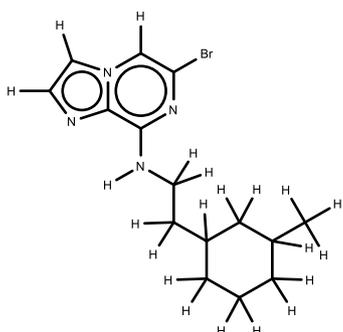
775



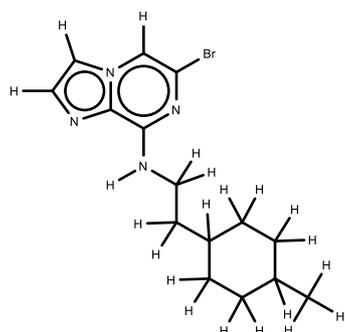
776



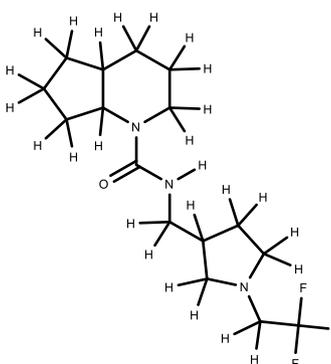
777



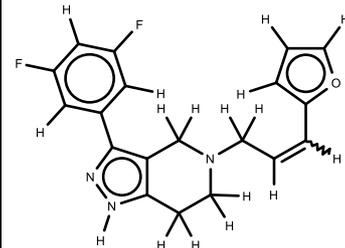
778



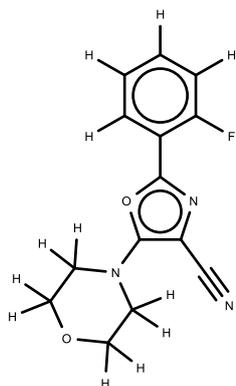
779



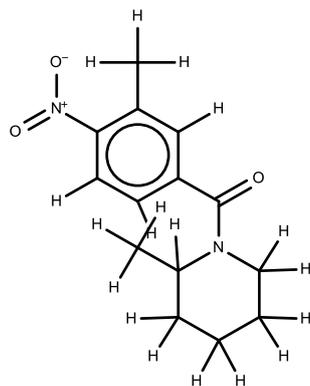
780



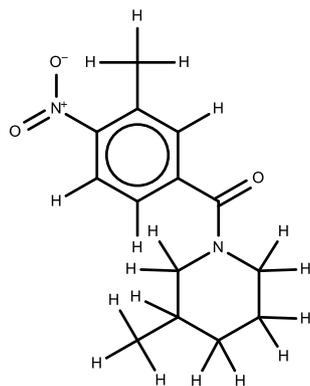
781



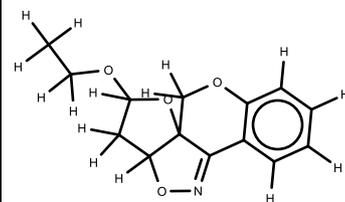
782



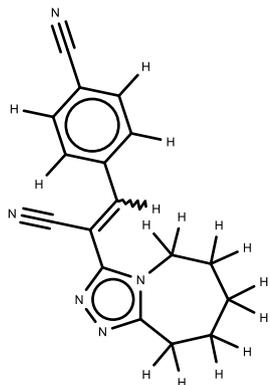
783



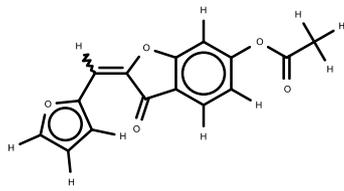
784



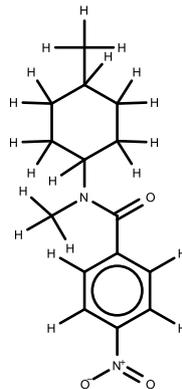
785



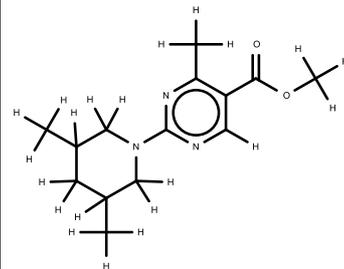
786



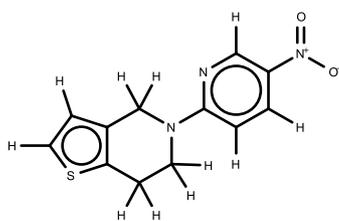
787



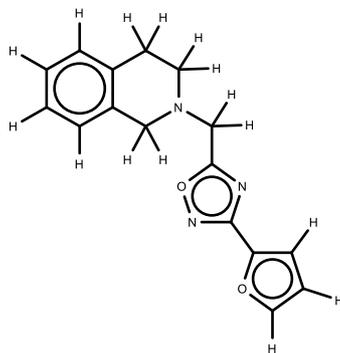
788



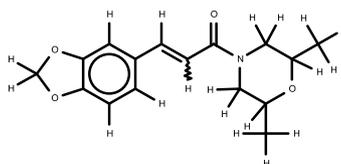
789



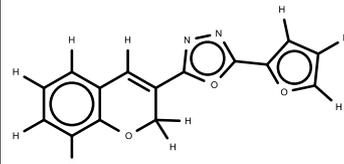
790



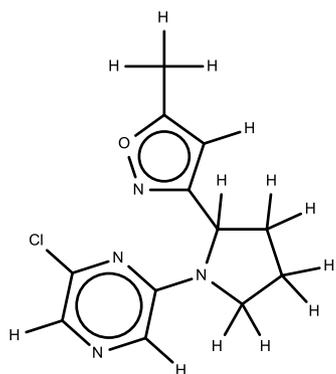
791



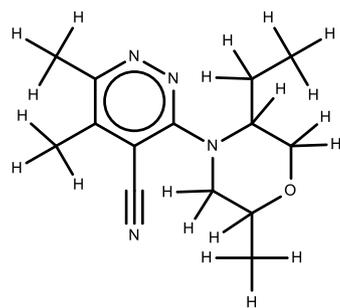
792



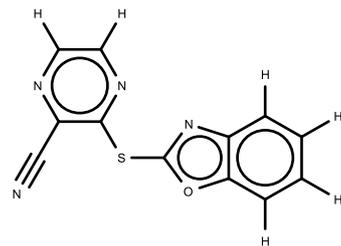
793



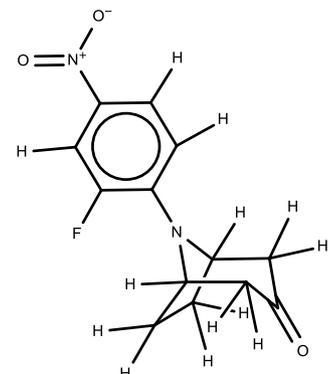
794



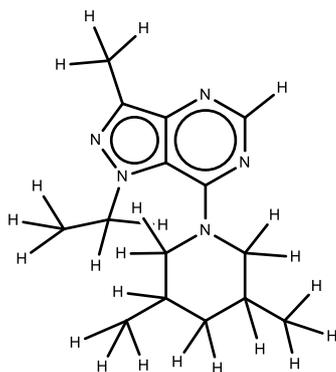
795



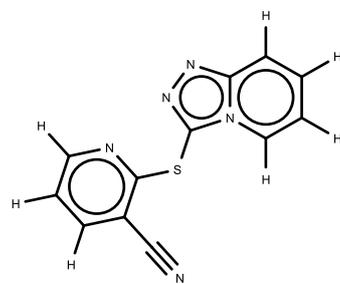
796



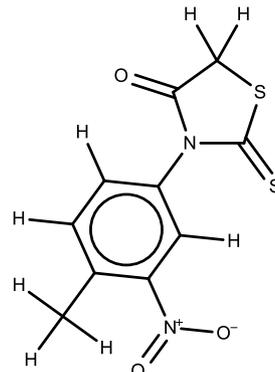
797



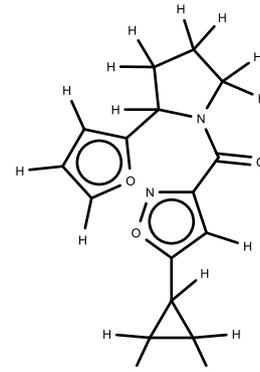
798



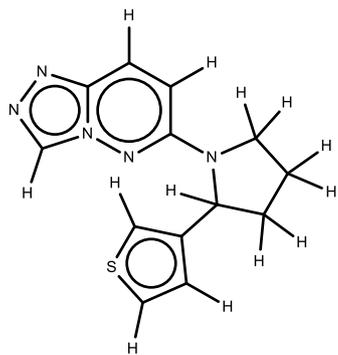
799



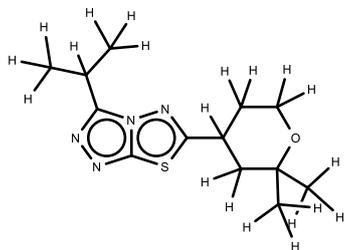
800



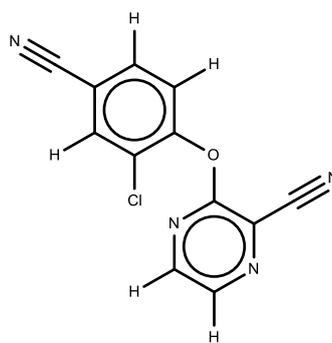
801



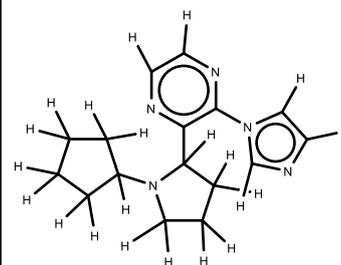
802



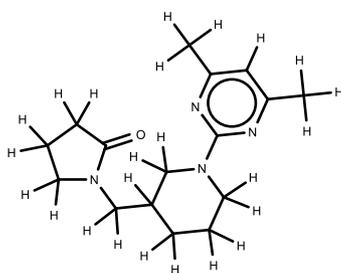
803



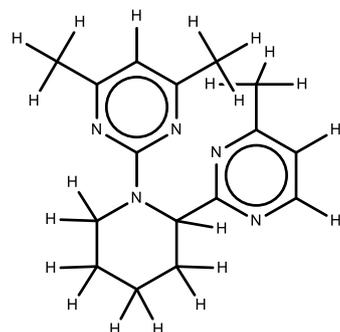
804



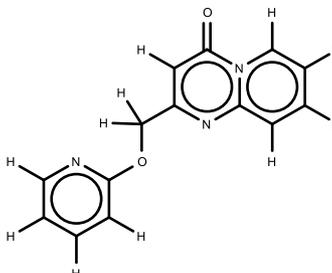
805



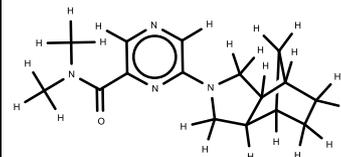
806



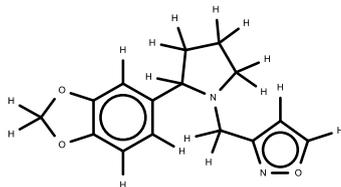
807



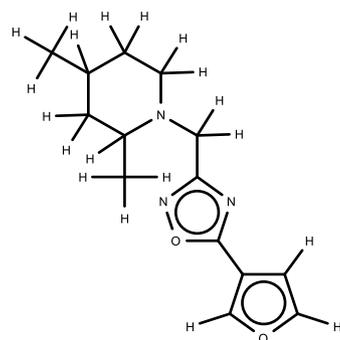
808



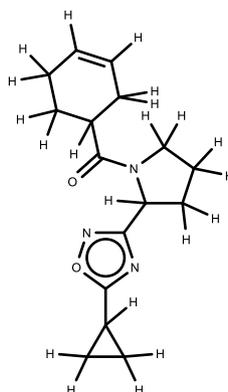
809



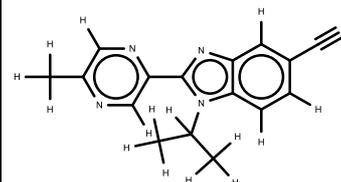
810



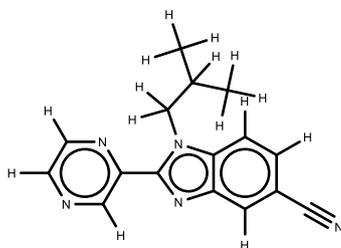
811



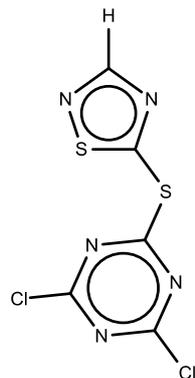
812



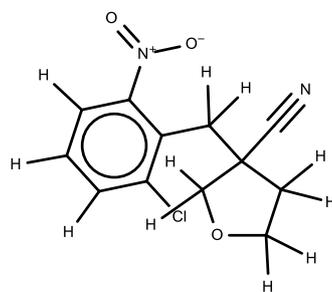
813



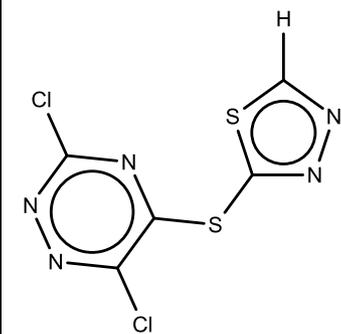
814



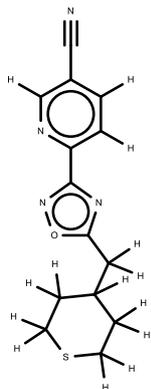
815



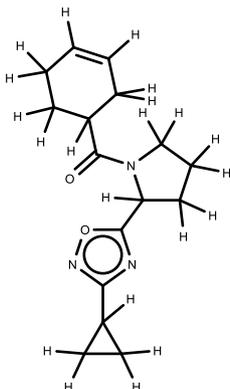
816



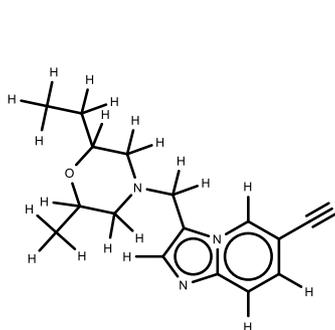
817



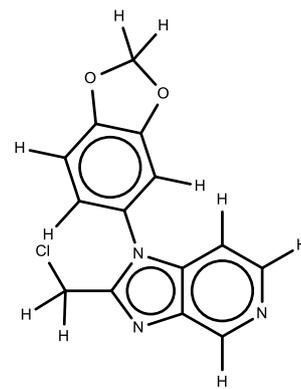
818



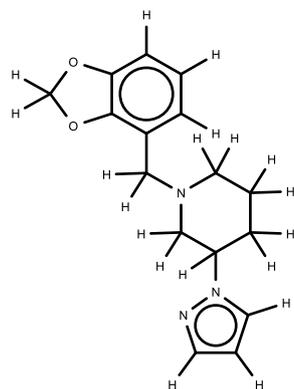
819



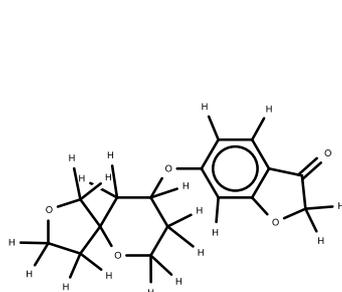
820



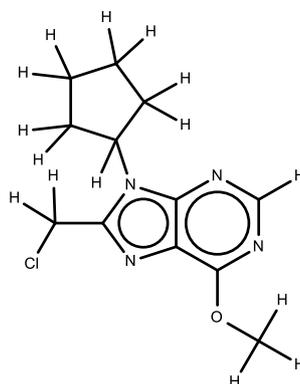
821



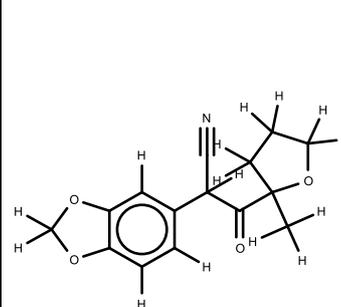
822



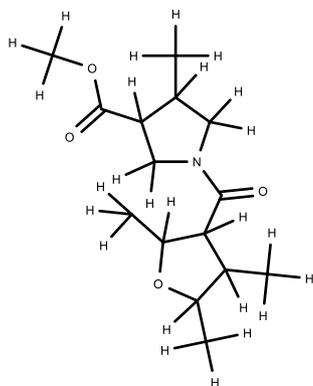
823



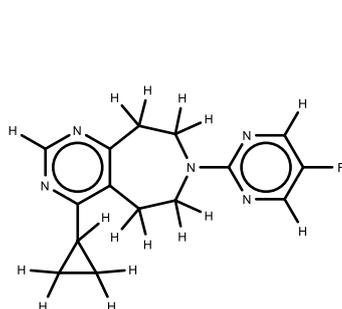
824



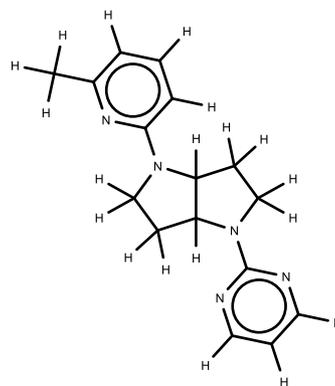
825



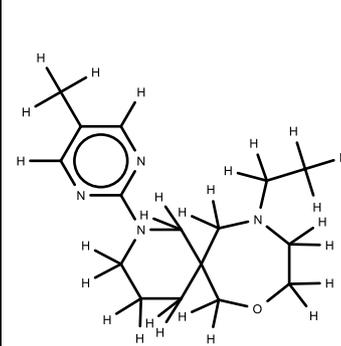
826



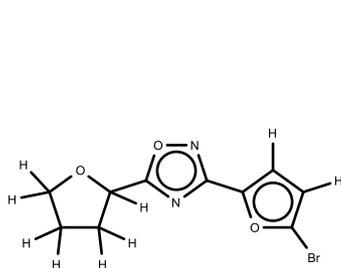
827



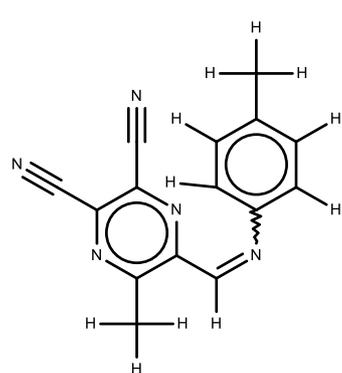
828



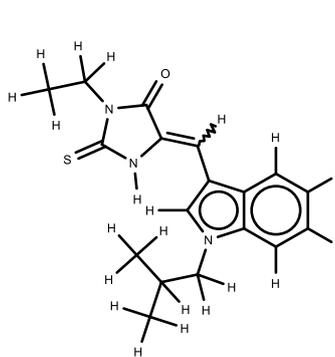
829



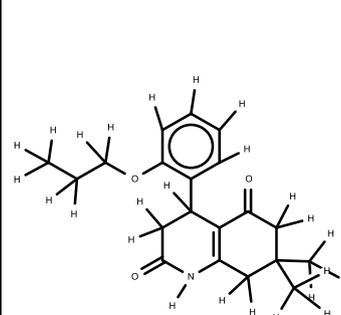
830



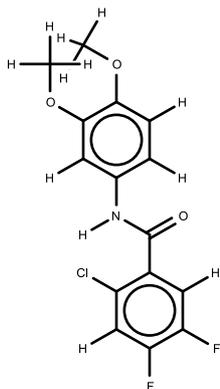
831



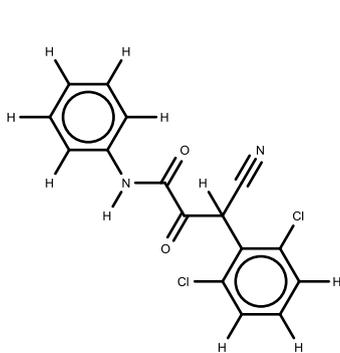
832



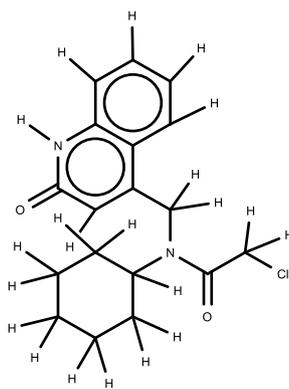
833



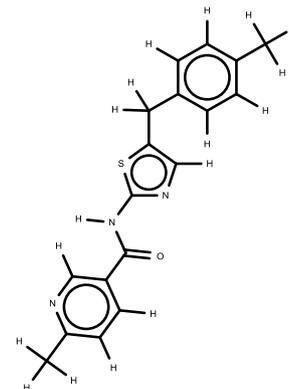
834



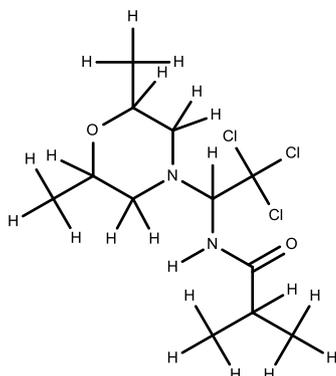
835



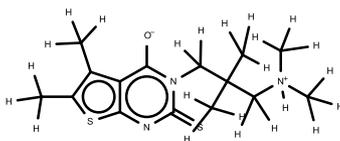
836



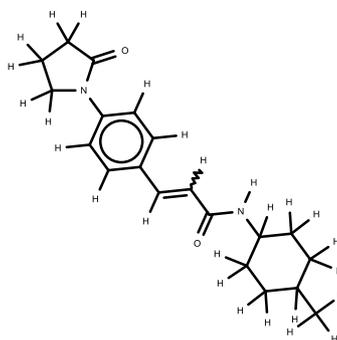
837



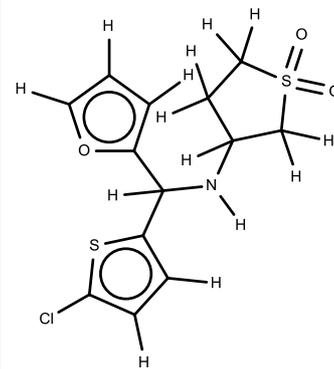
838



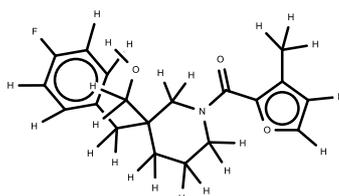
839



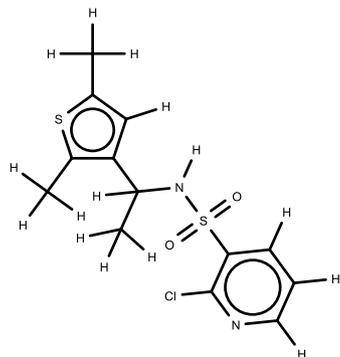
840



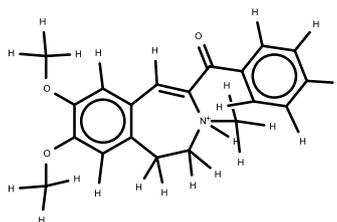
841



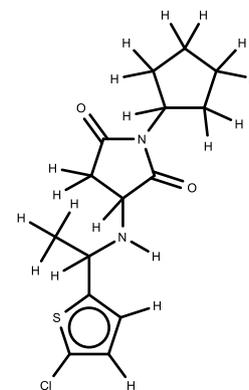
842



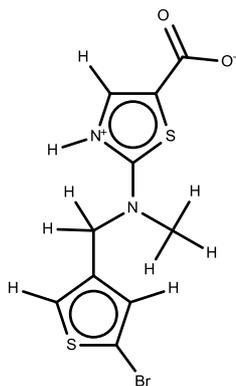
843



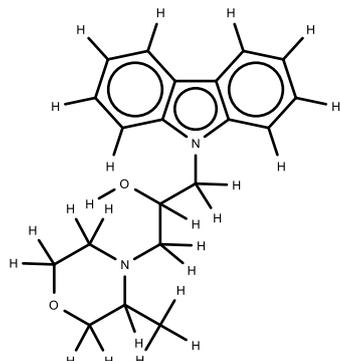
844



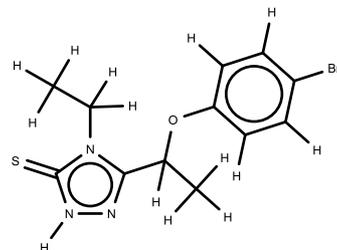
845



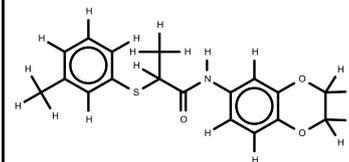
846



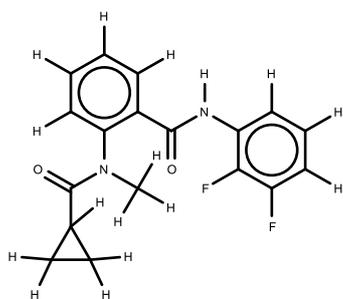
847



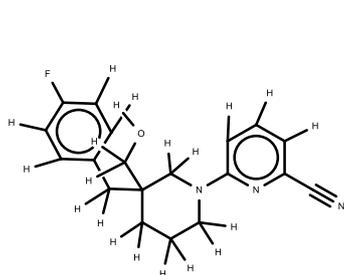
848



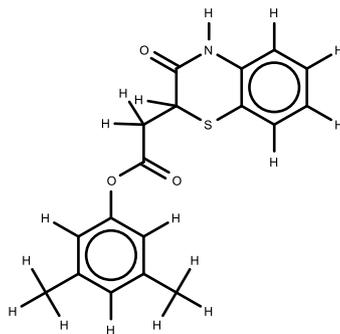
849



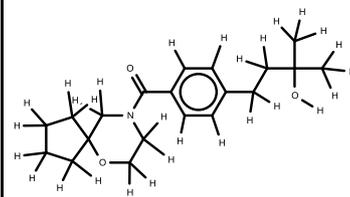
850



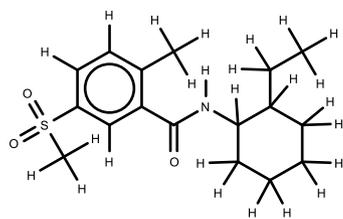
851



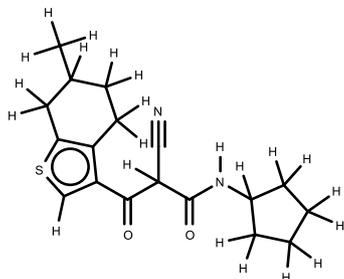
852



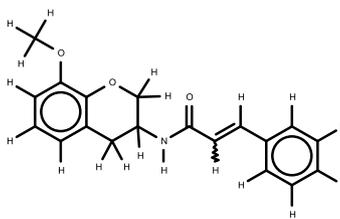
853



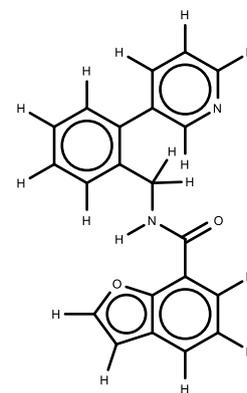
854



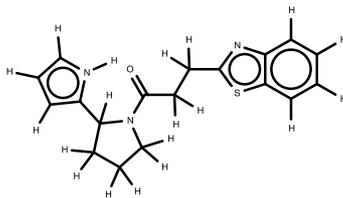
855



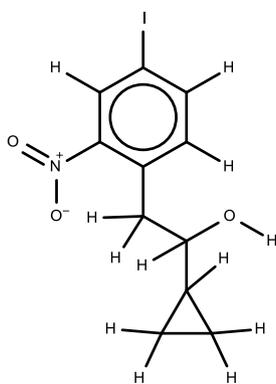
856



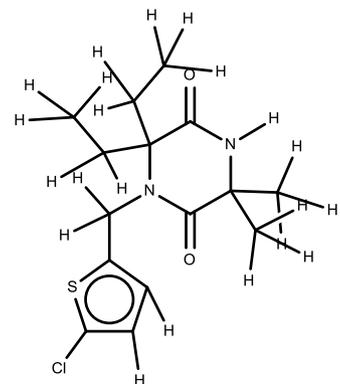
857



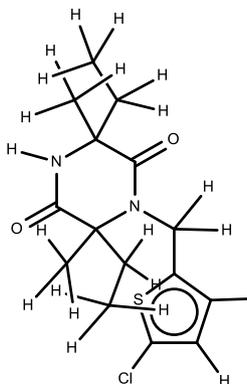
858



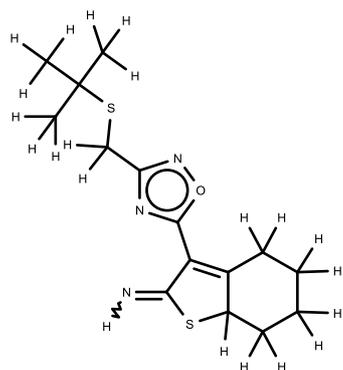
859



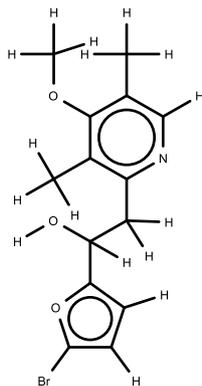
860



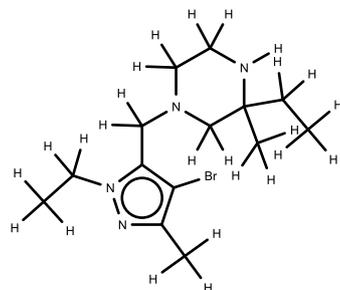
861



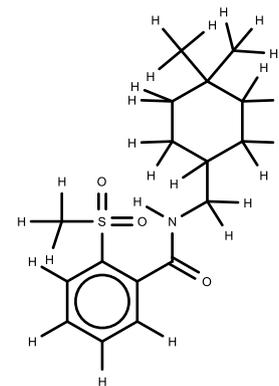
862



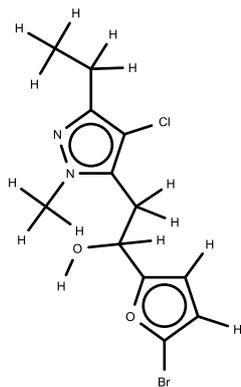
863



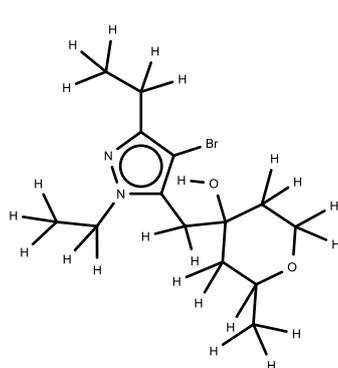
864



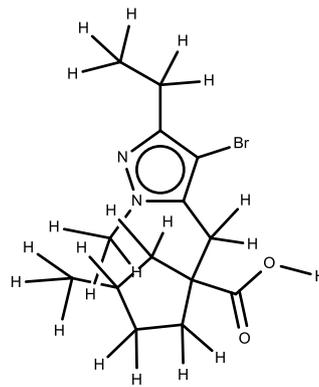
865



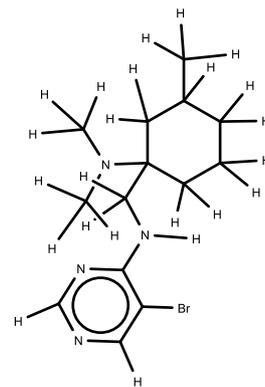
866



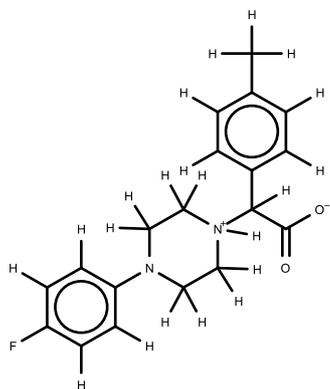
867



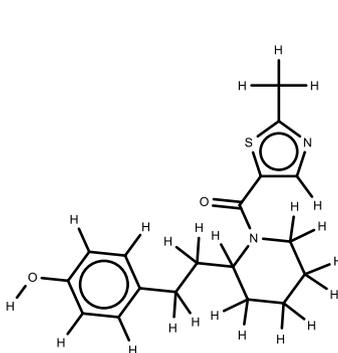
868



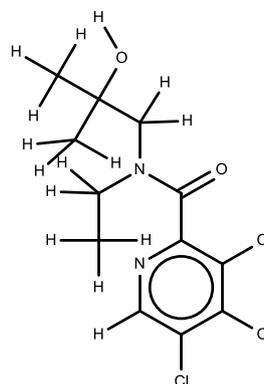
869



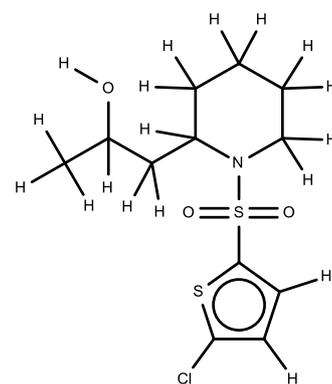
870



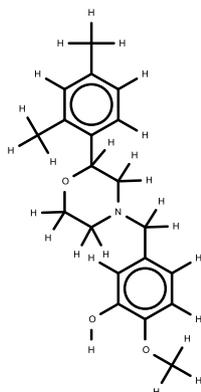
871



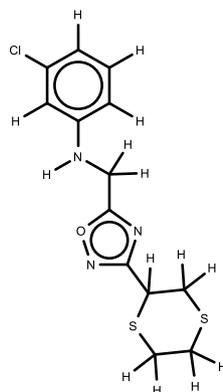
872



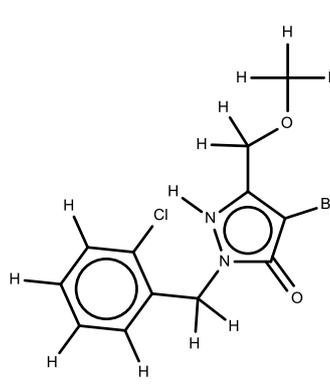
873



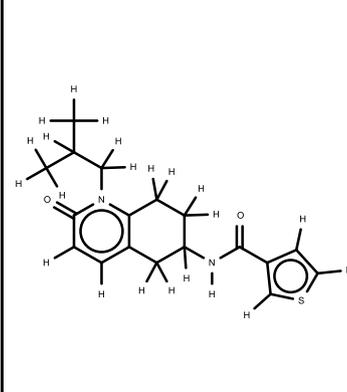
874



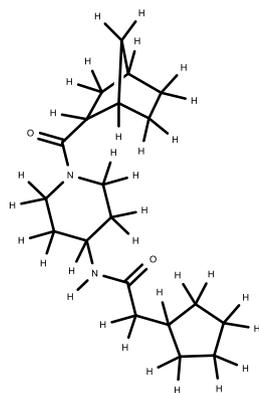
875



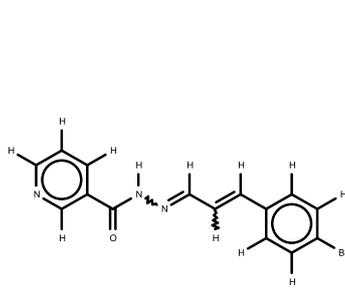
876



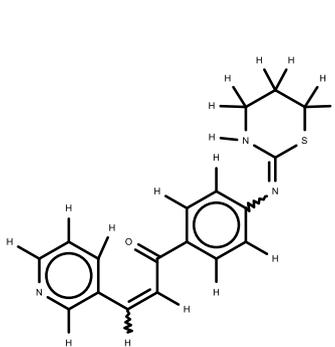
877



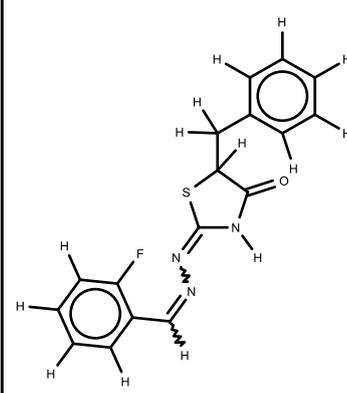
878



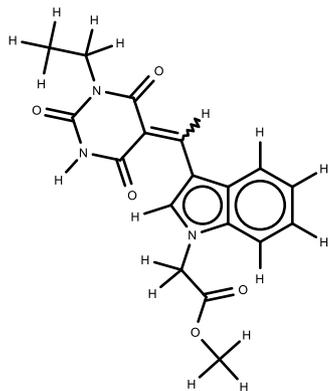
879



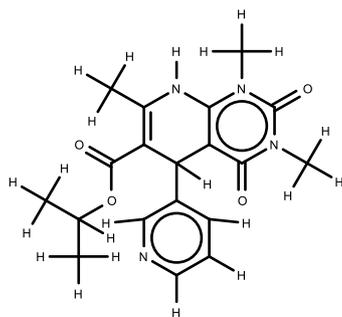
880



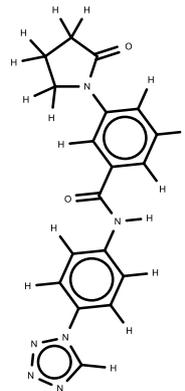
881



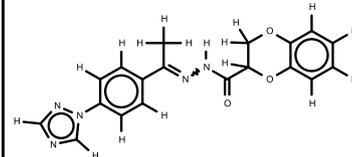
882



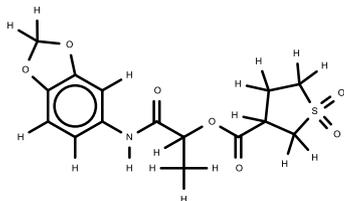
883



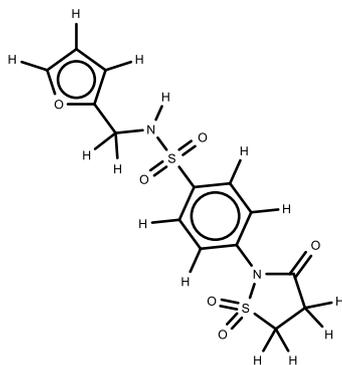
884



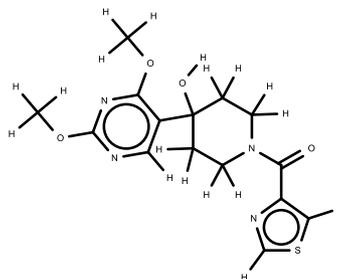
885



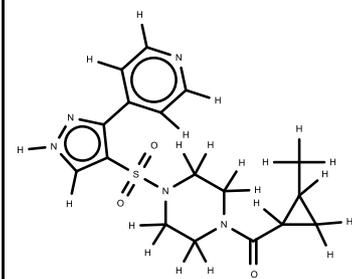
886



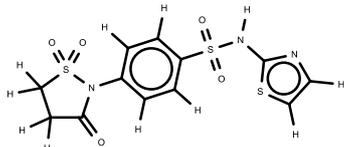
887



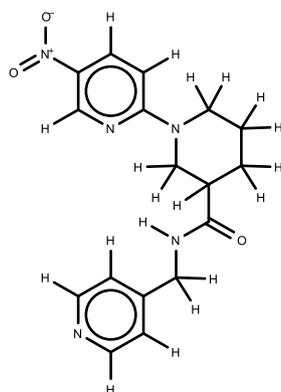
888



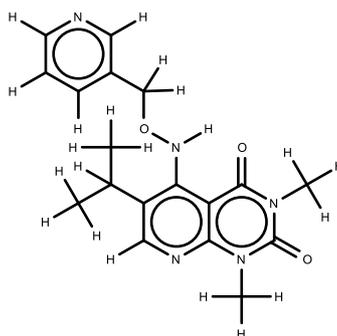
889



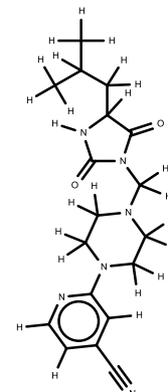
890



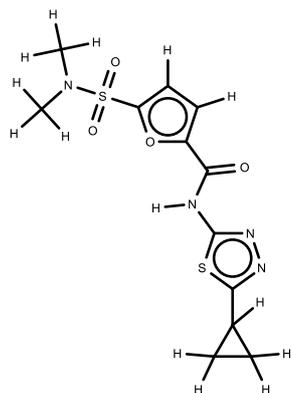
891



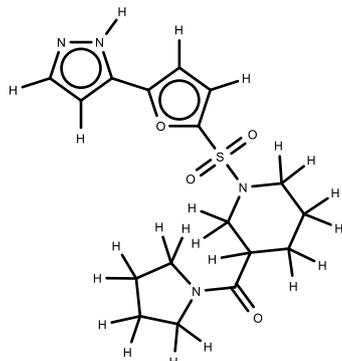
892



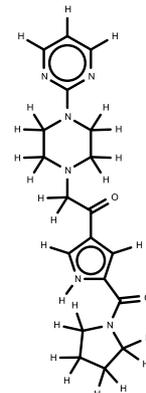
893



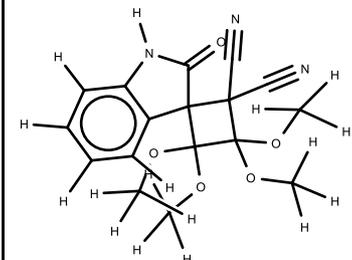
894



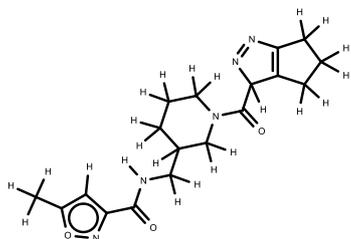
895



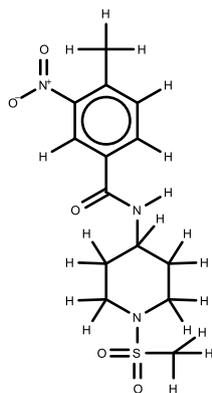
896



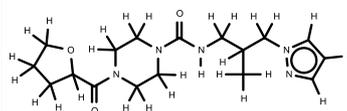
897



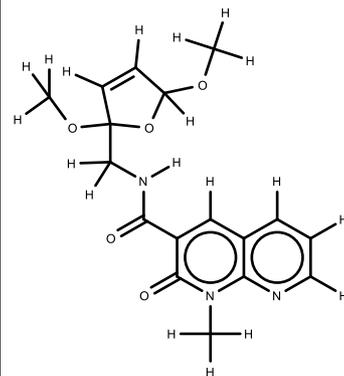
898



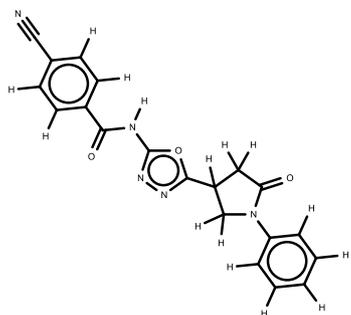
899



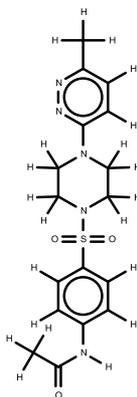
900



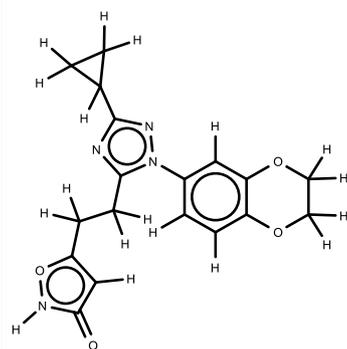
901



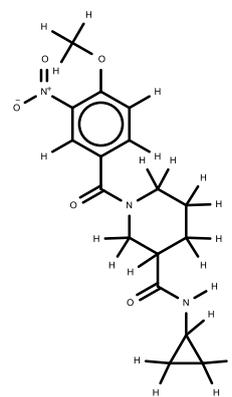
902



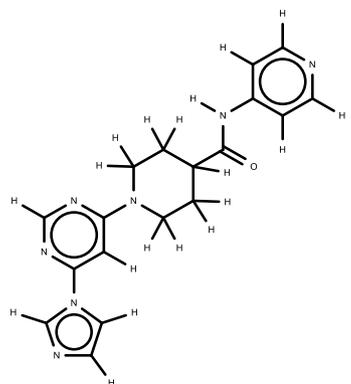
903



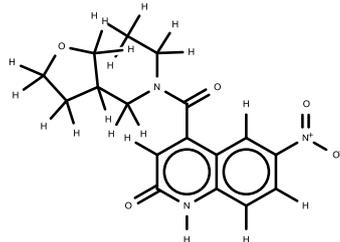
904



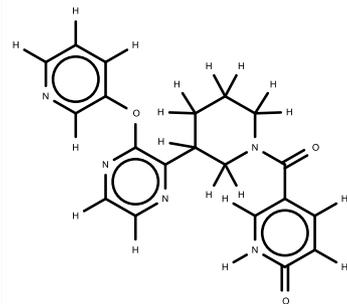
905



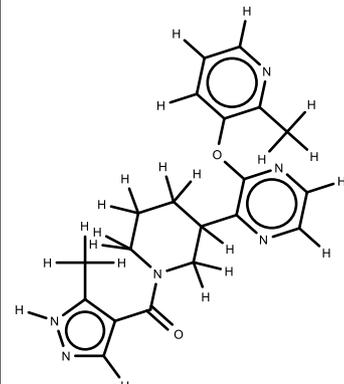
906



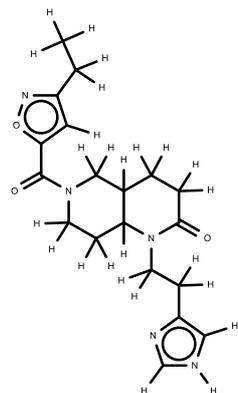
907



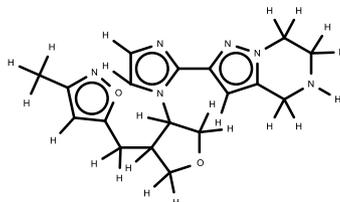
908



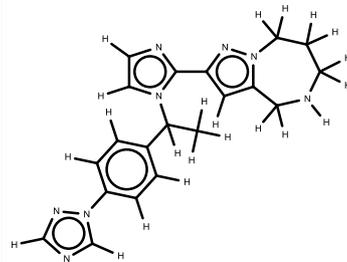
909



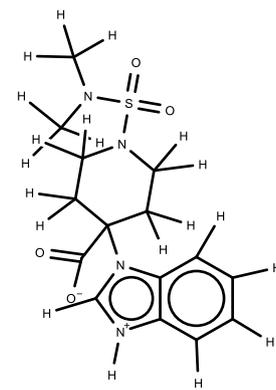
910



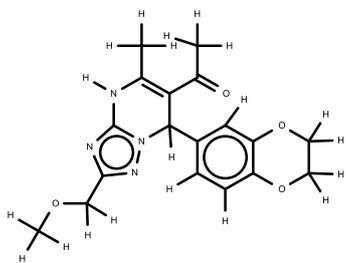
911



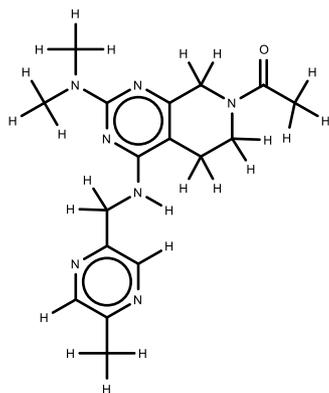
912



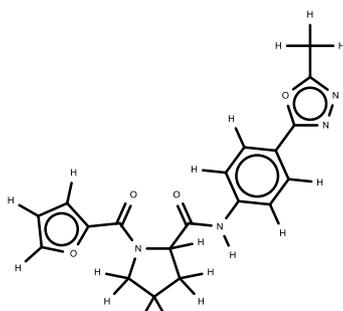
913



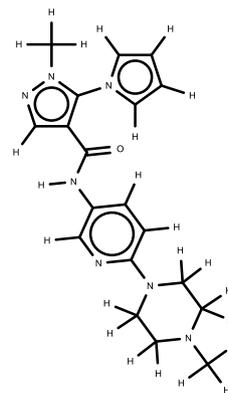
914



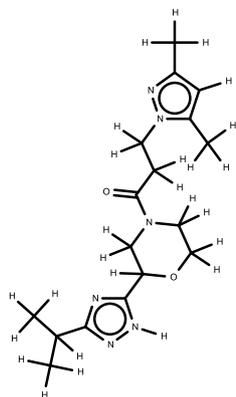
915



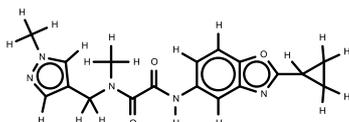
916



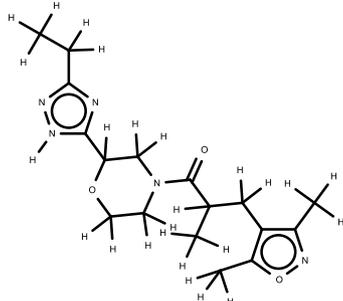
917



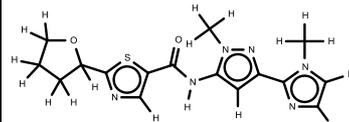
918



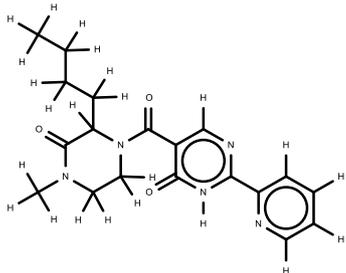
919



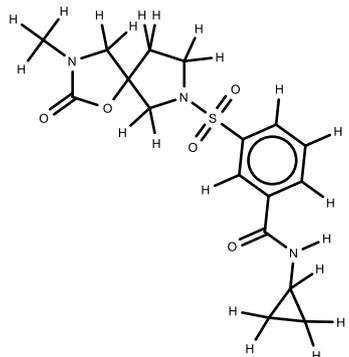
920



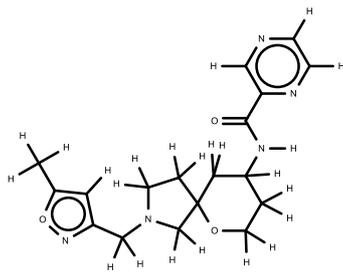
921



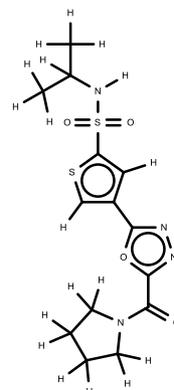
922



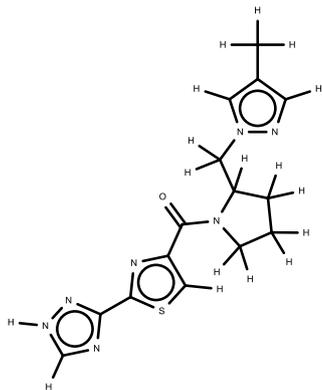
923



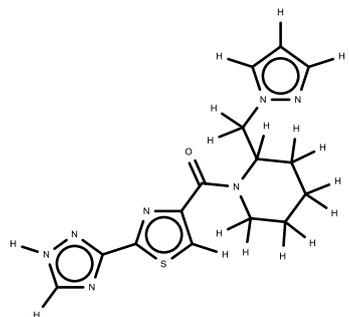
924



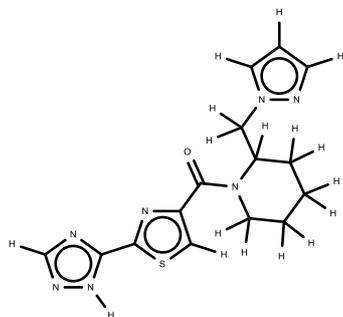
925



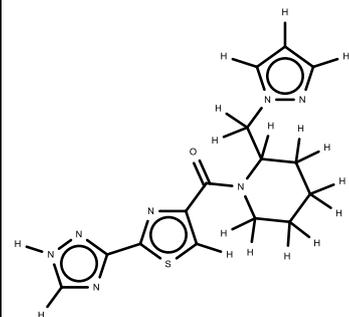
926



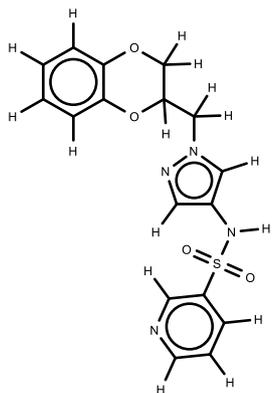
927



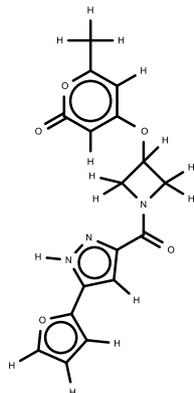
928



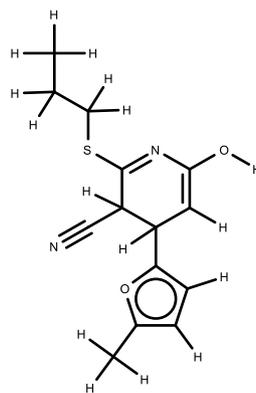
929



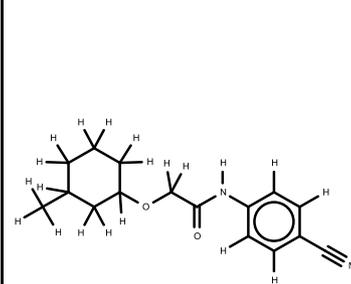
930



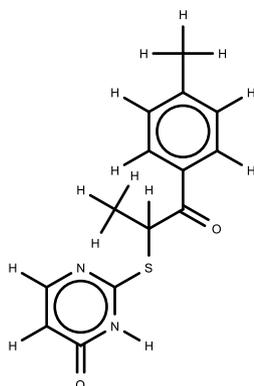
931



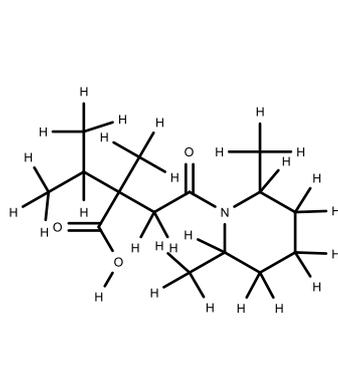
932



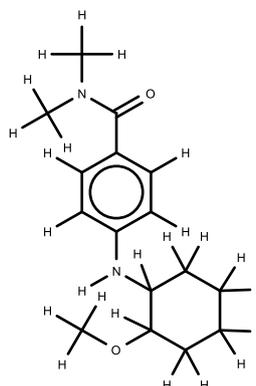
933



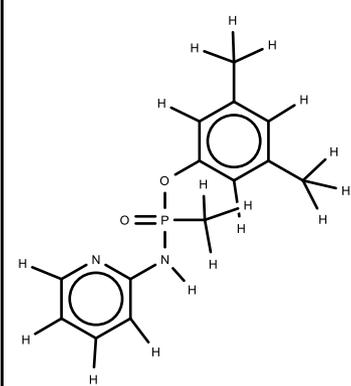
934



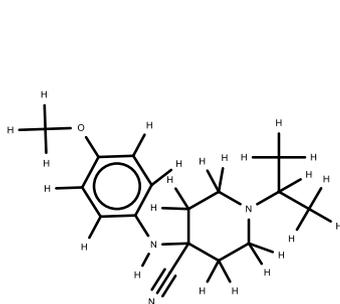
935



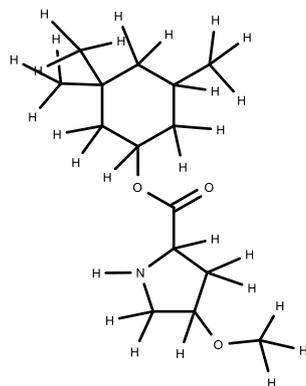
936



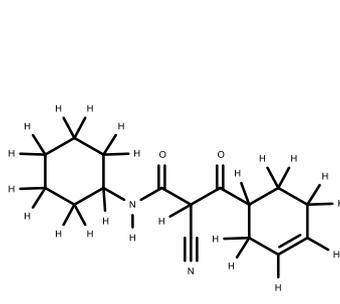
937



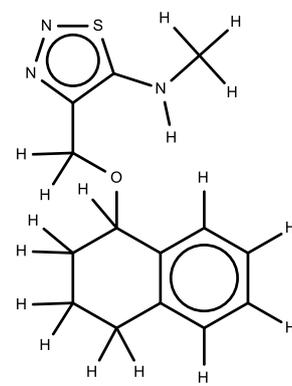
938



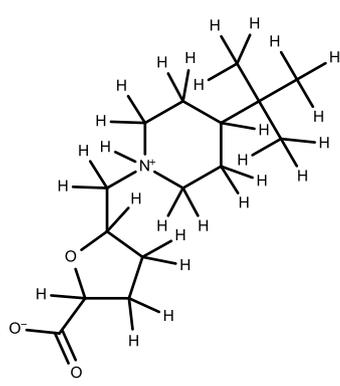
939



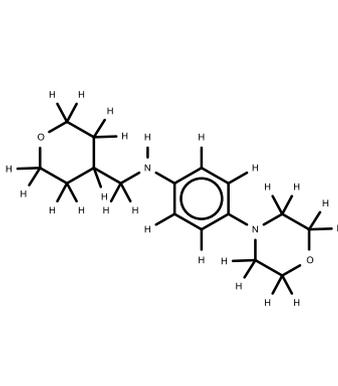
940



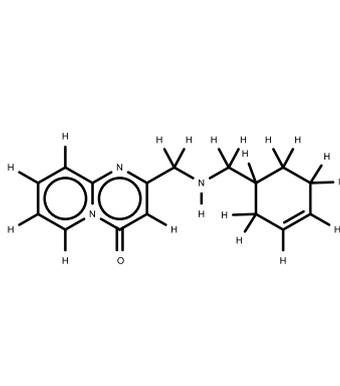
941



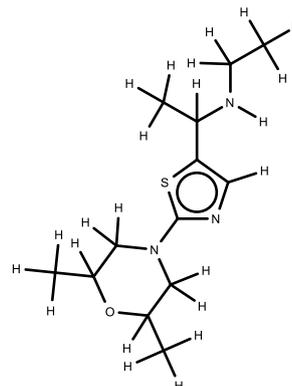
942



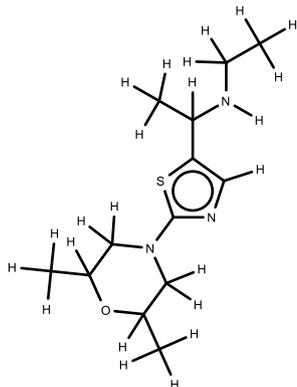
943



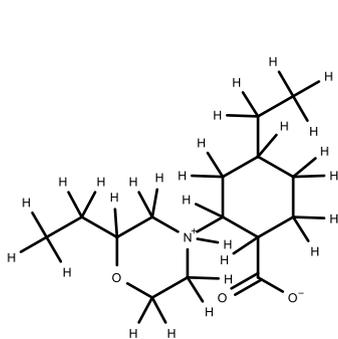
944



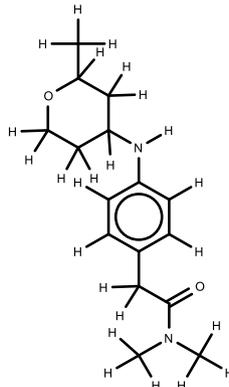
945



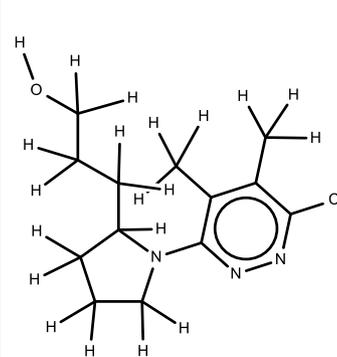
946



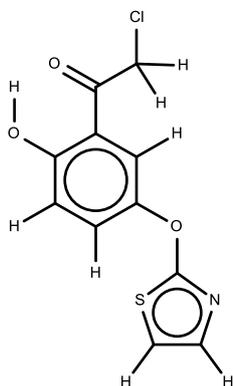
947



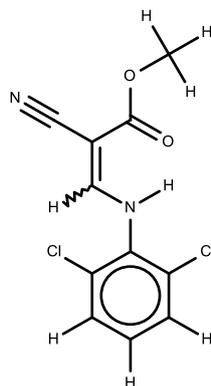
948



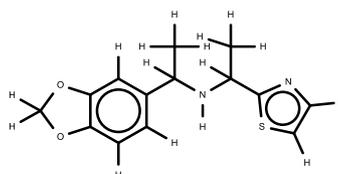
949



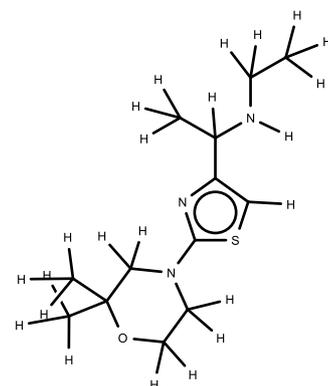
950



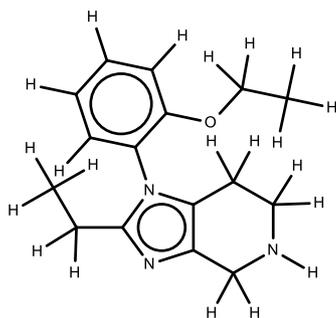
951



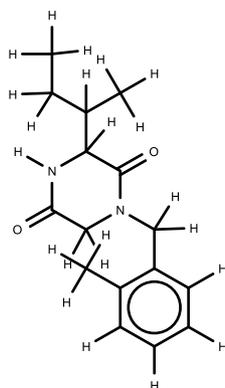
952



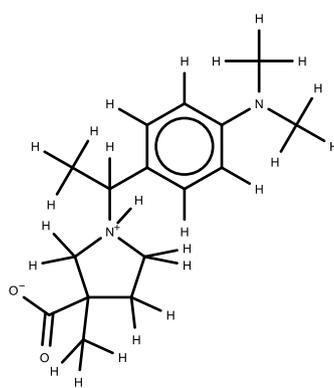
953



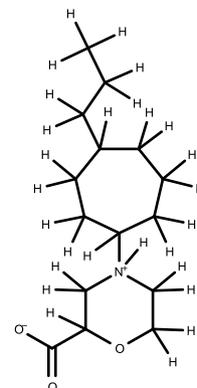
954



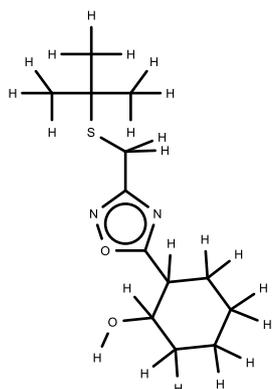
955



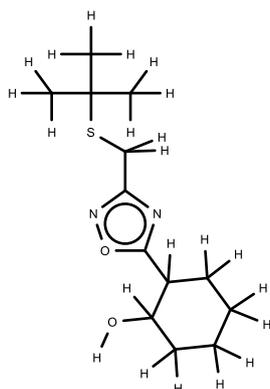
956



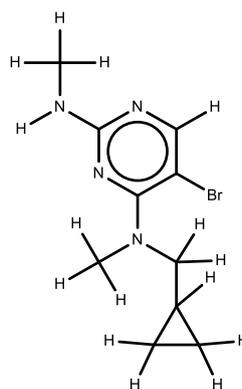
957



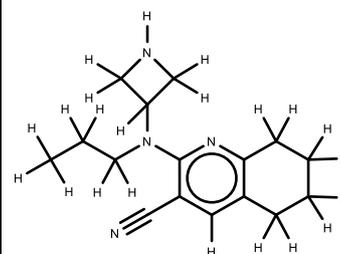
958



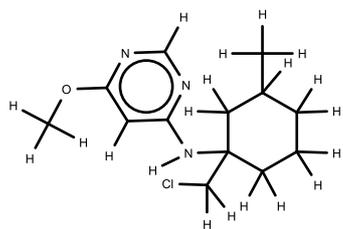
959



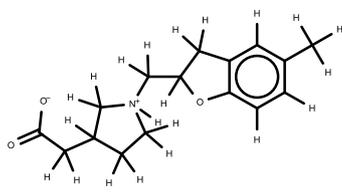
960



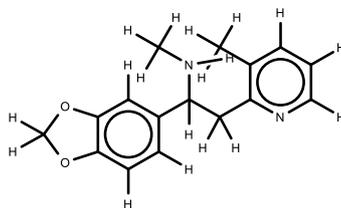
961



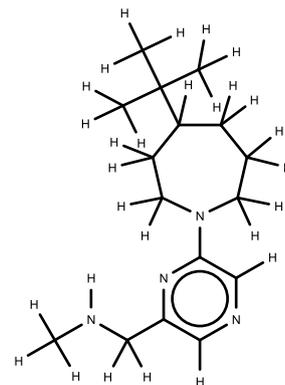
962



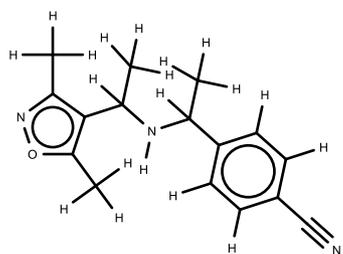
963



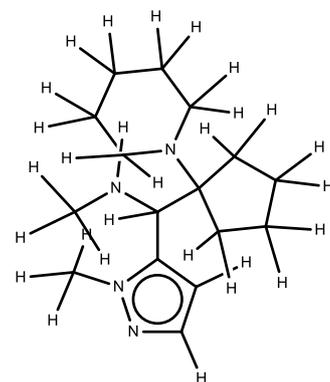
964



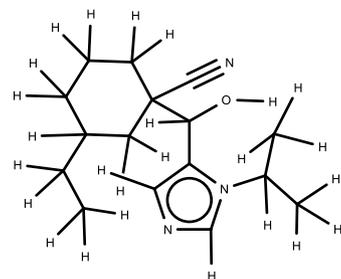
965



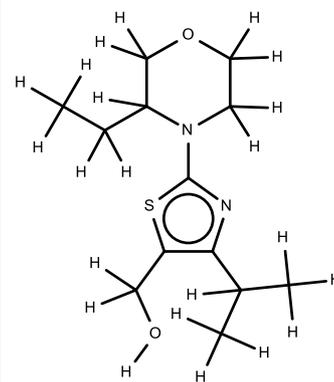
966



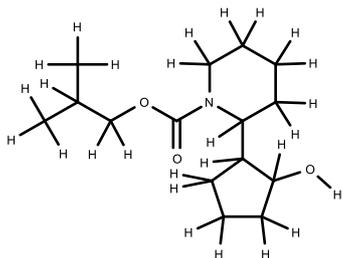
967



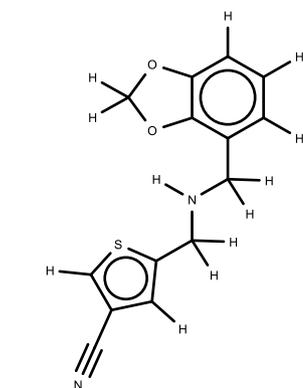
968



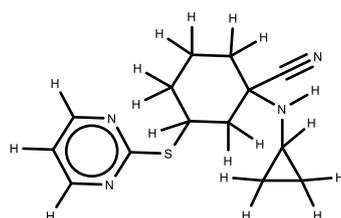
969



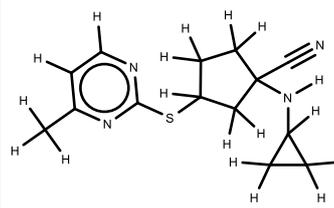
970



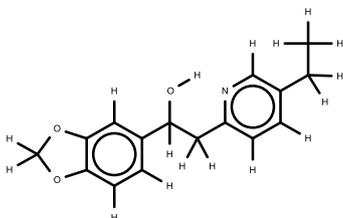
971



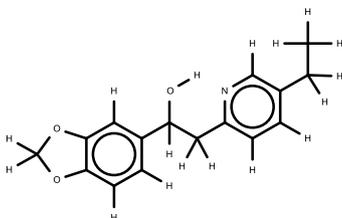
972



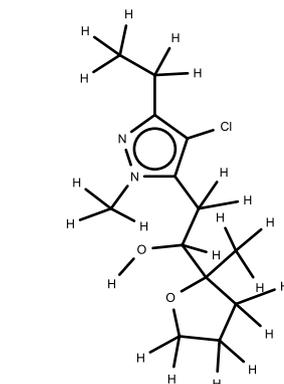
973



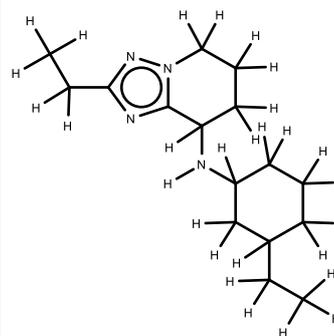
974



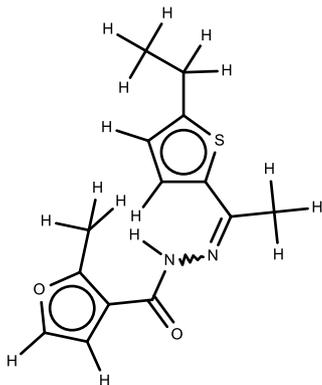
975



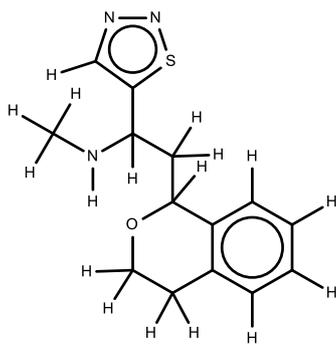
976



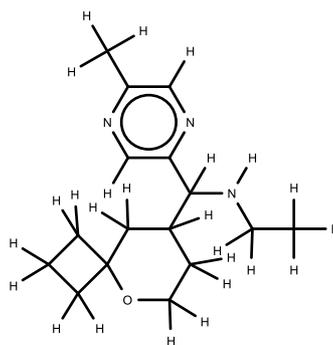
977



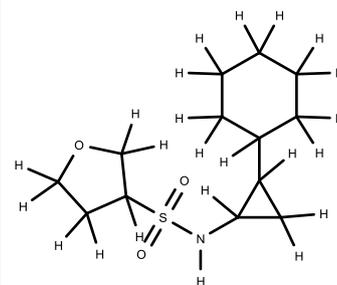
978



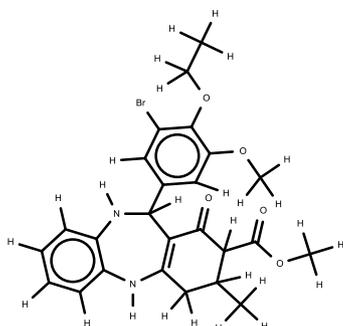
979



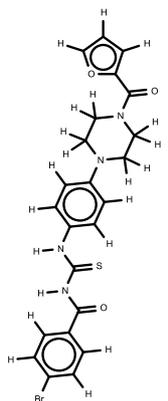
980



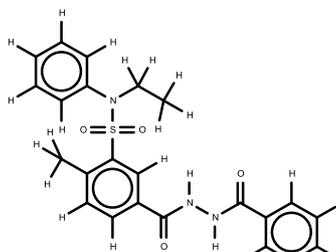
981



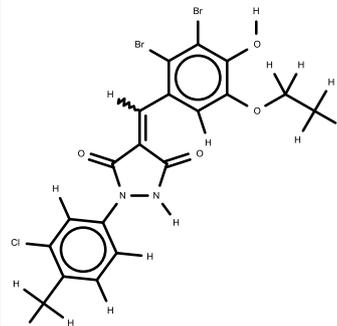
982



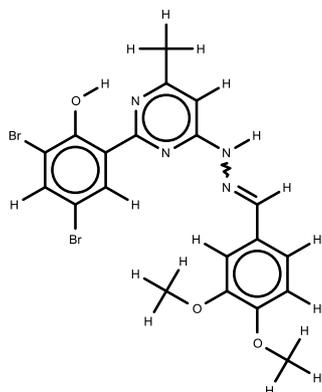
983



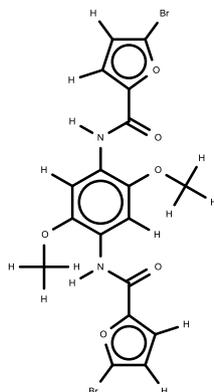
984



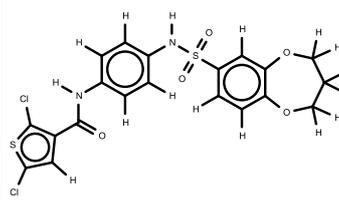
985



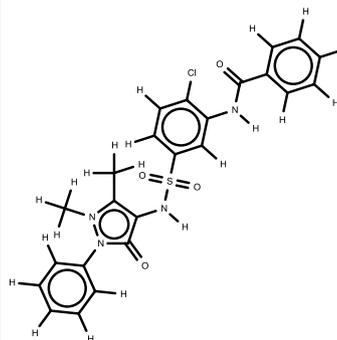
986



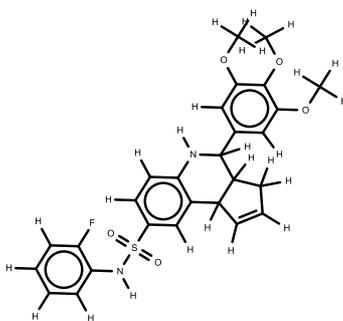
987



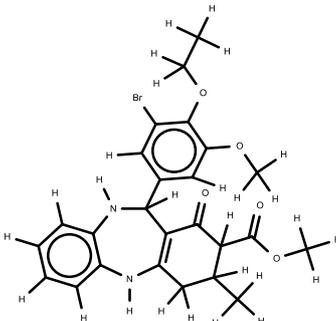
988



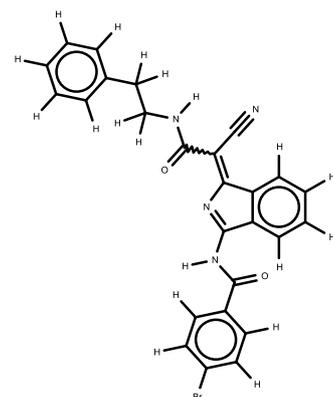
989



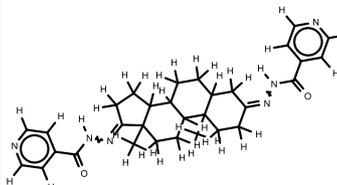
990



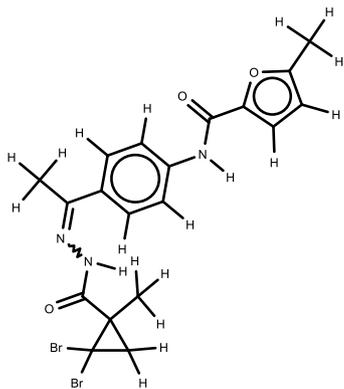
991



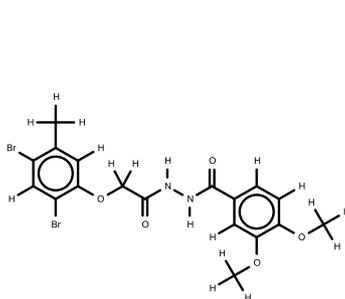
992



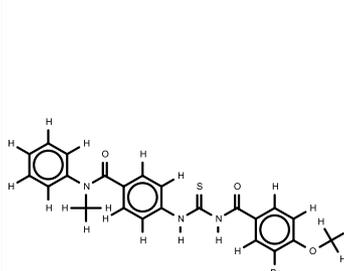
993



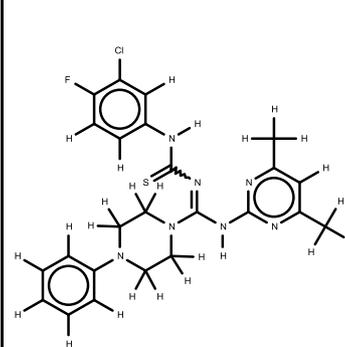
994



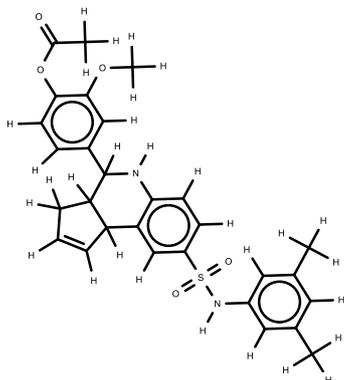
995



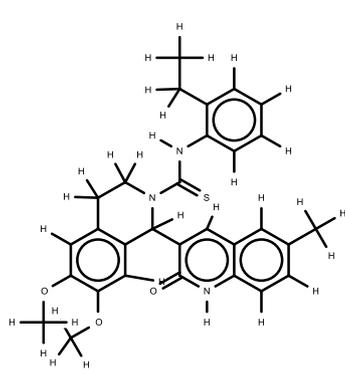
996



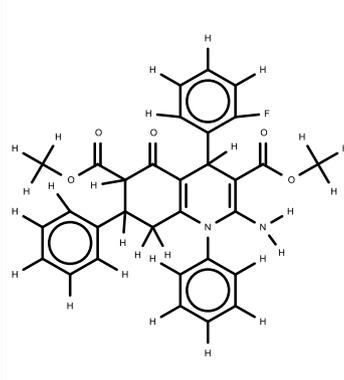
997



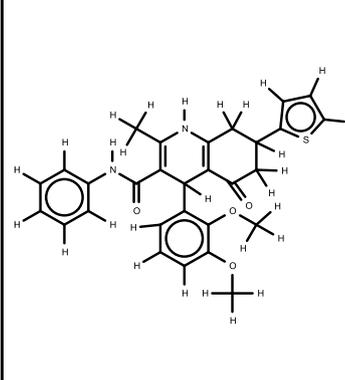
998



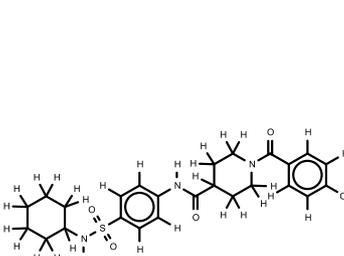
999



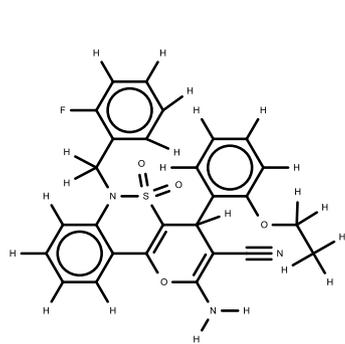
1000



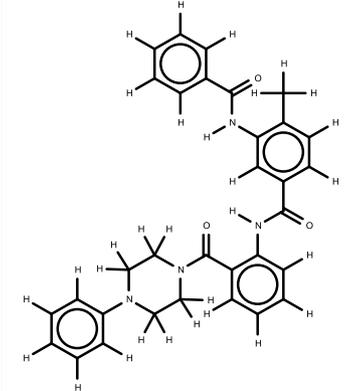
1001



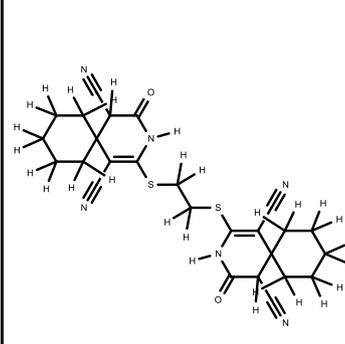
1002



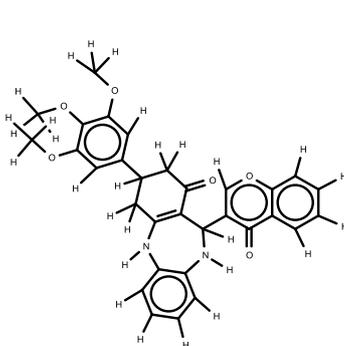
1003



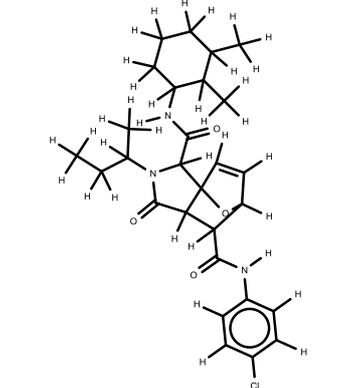
1004



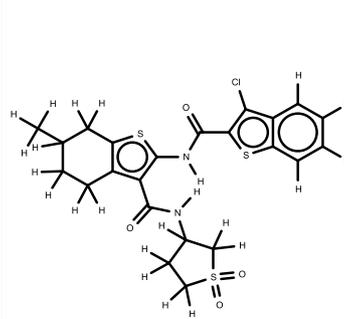
1005



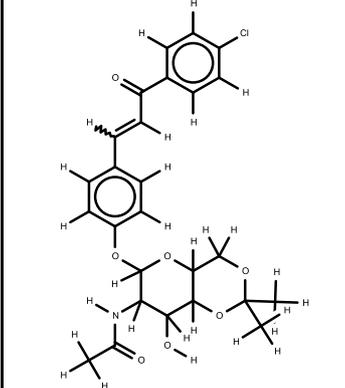
1006



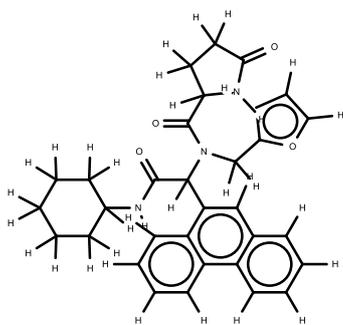
1007



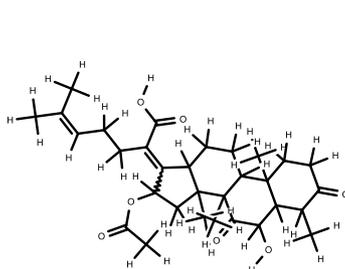
1008



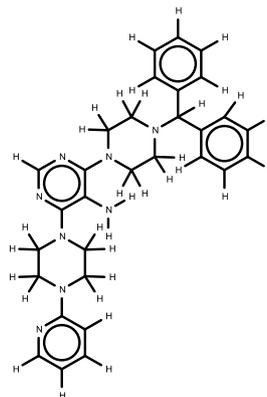
1009



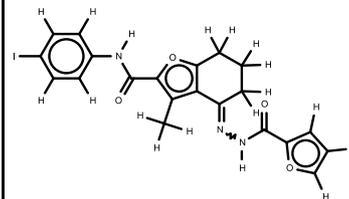
1010



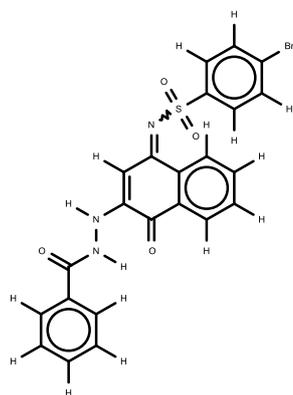
1011



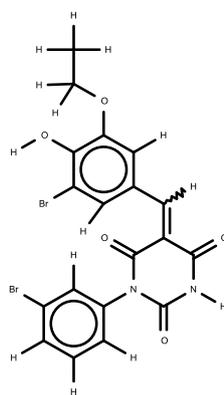
1012



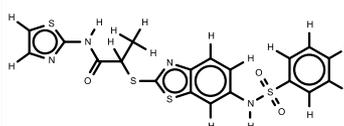
1013



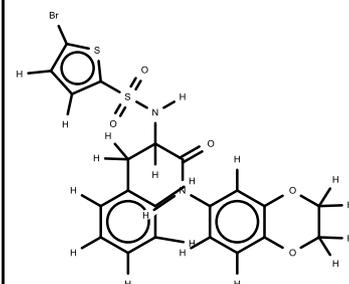
1014



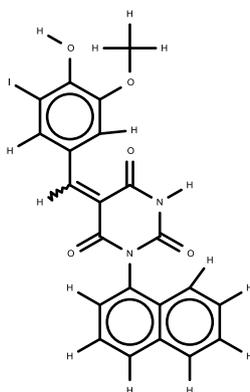
1015



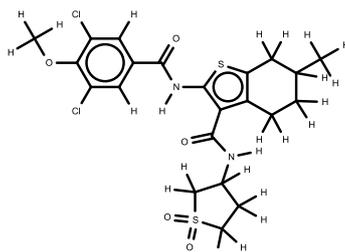
1016



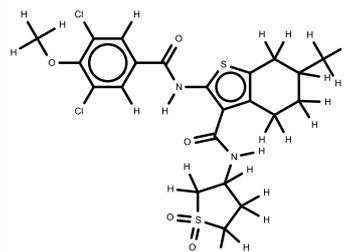
1017



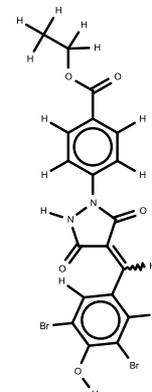
1018



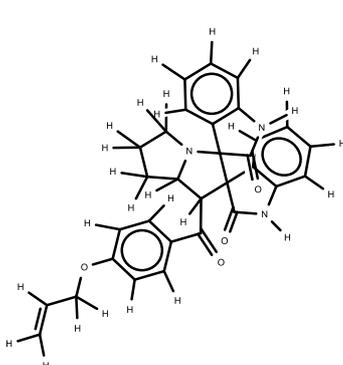
1019



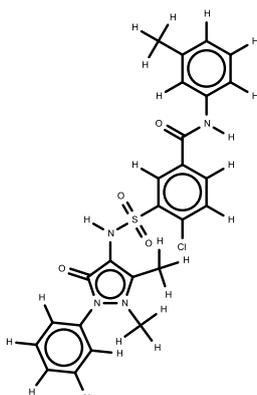
1020



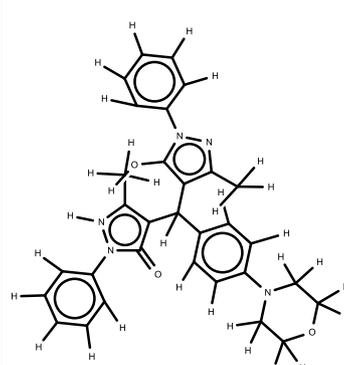
1021



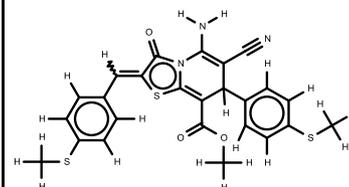
1022



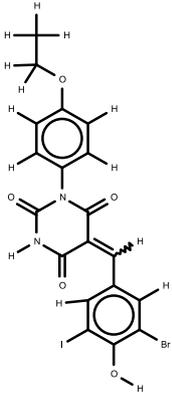
1023



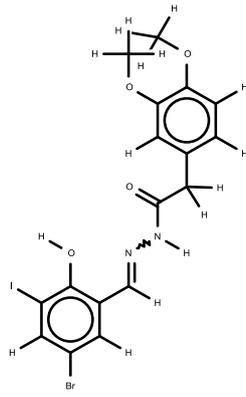
1024



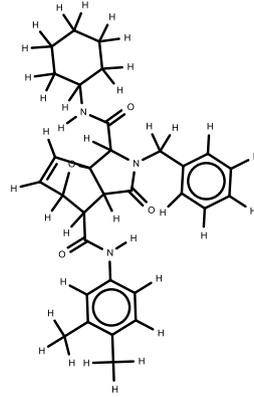
1025



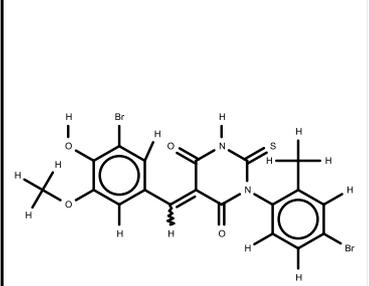
1026



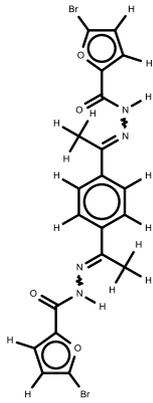
1027



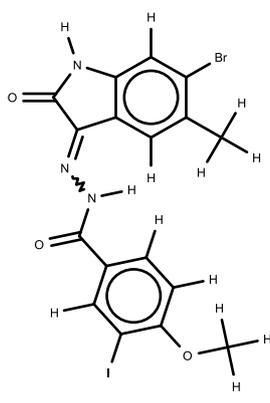
1028



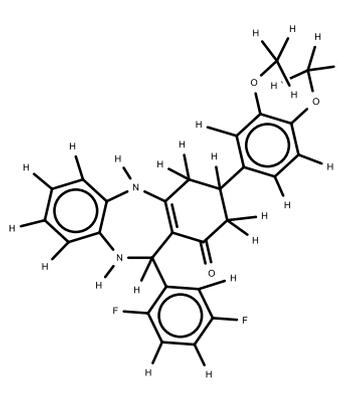
1029



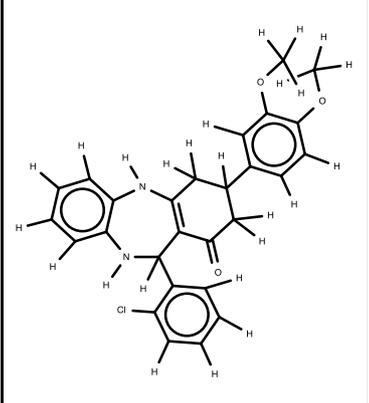
1030



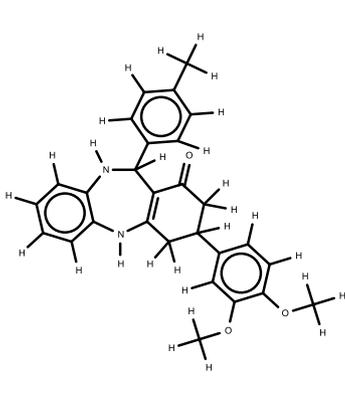
1031



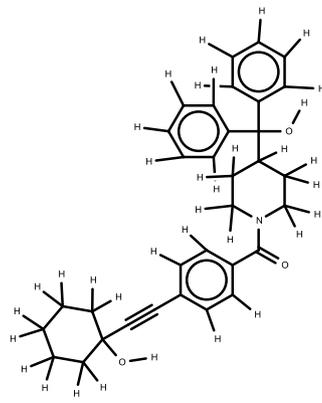
1032



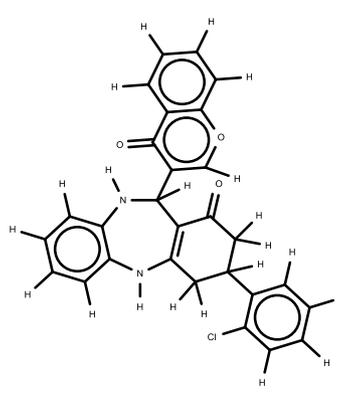
1033



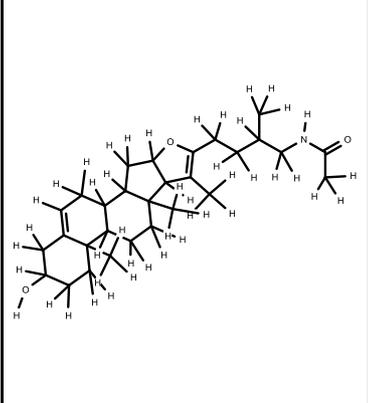
1034



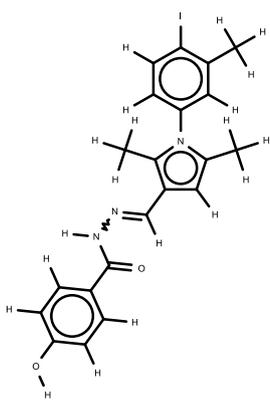
1035



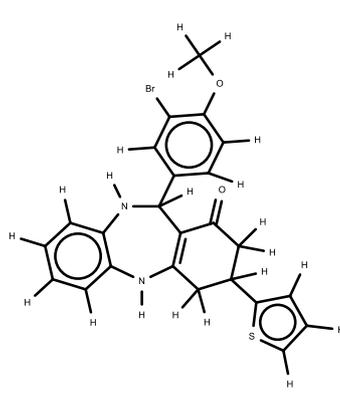
1036



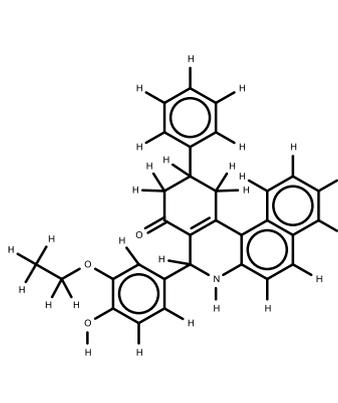
1037



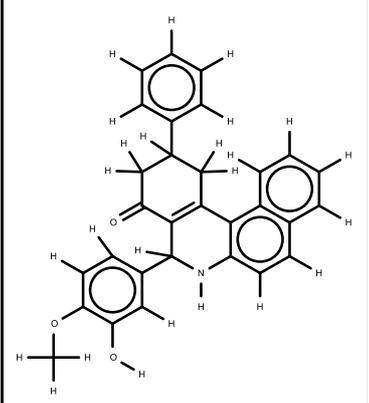
1038



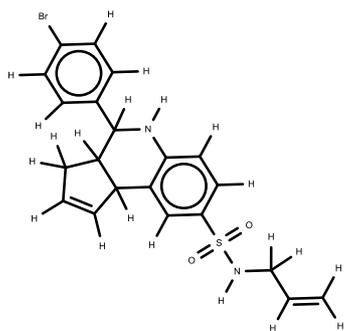
1039



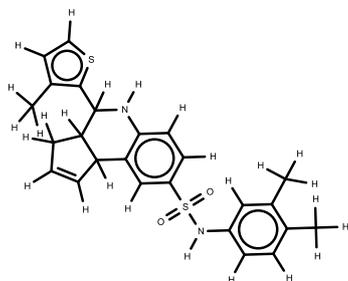
1040



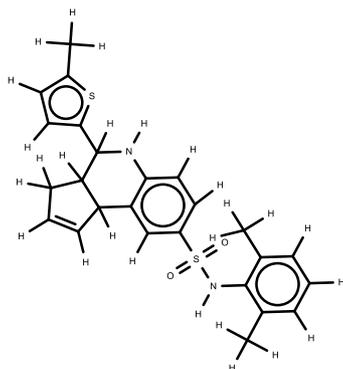
1041



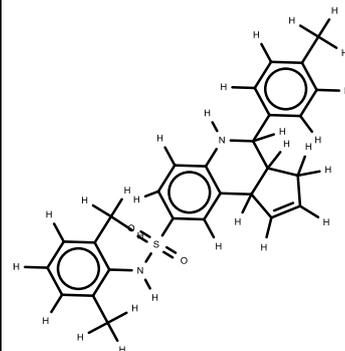
1042



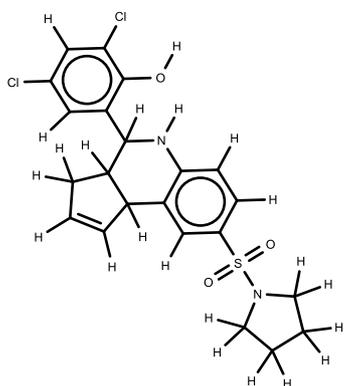
1043



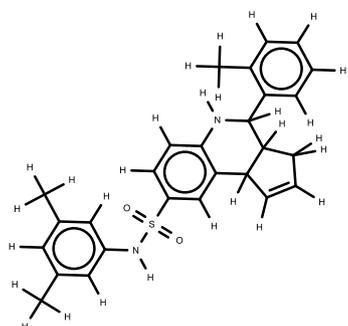
1044



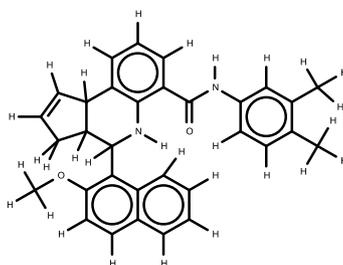
1045



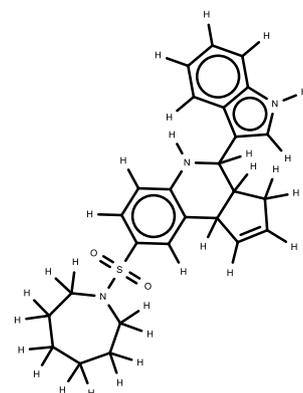
1046



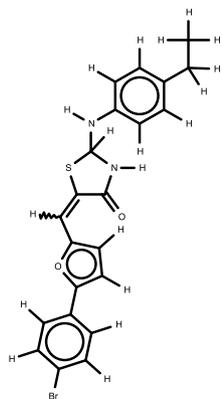
1047



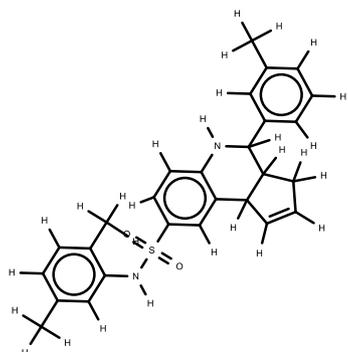
1048



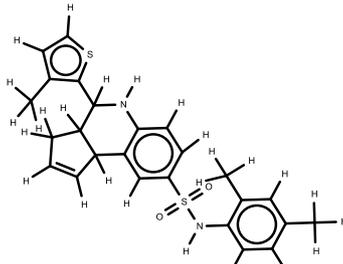
1049



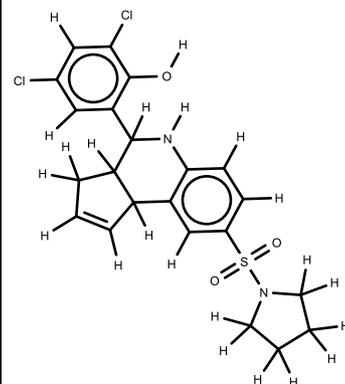
1050



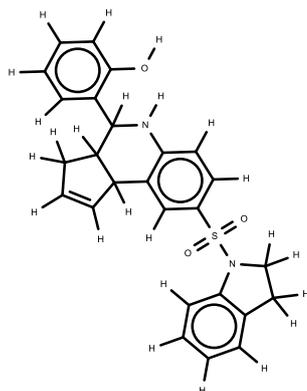
1051



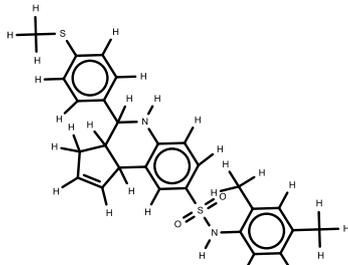
1052



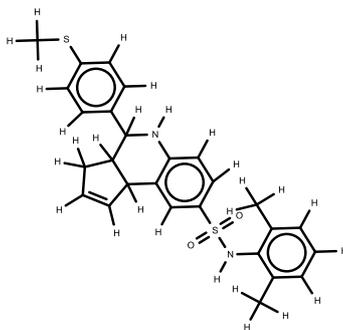
1053



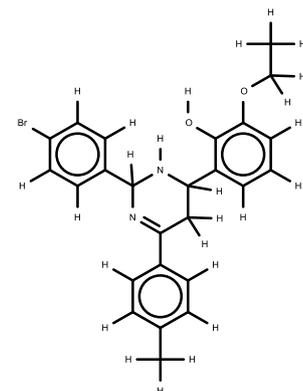
1054



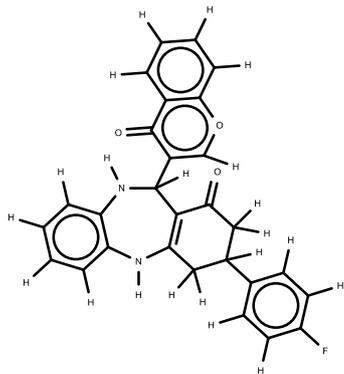
1055



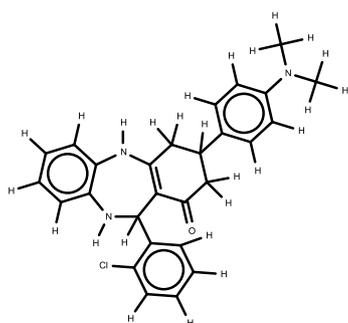
1056



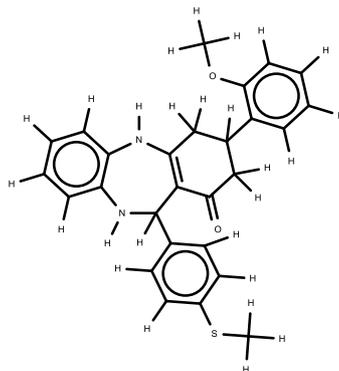
1057



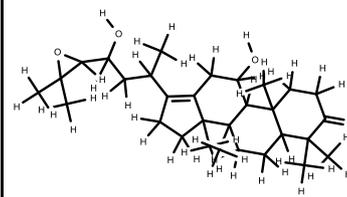
1058



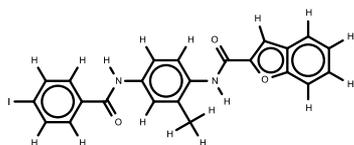
1059



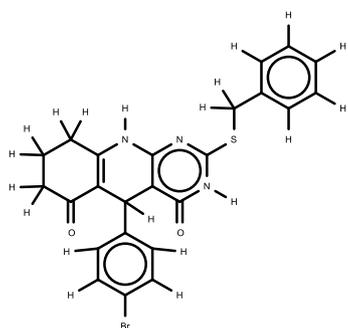
1060



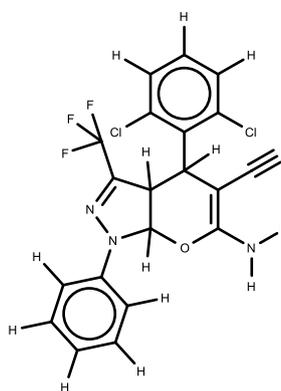
1061



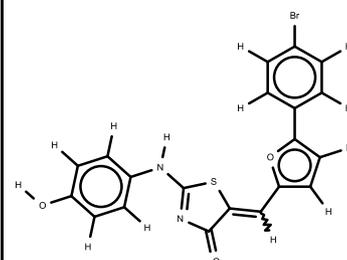
1062



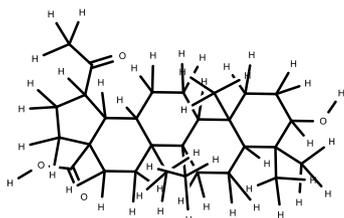
1063



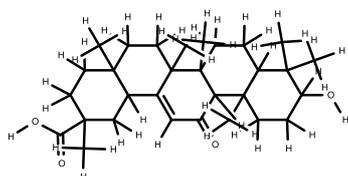
1064



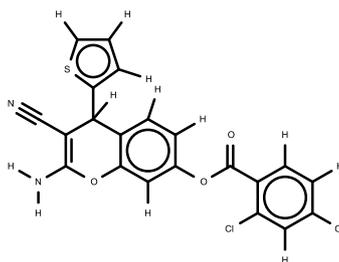
1065



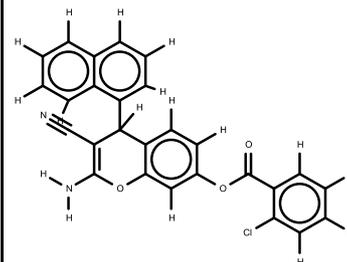
1066



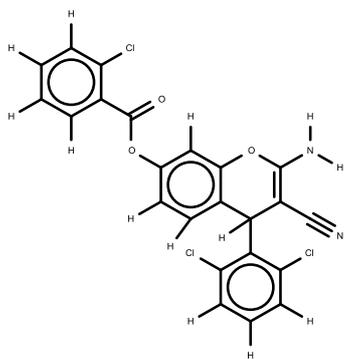
1067



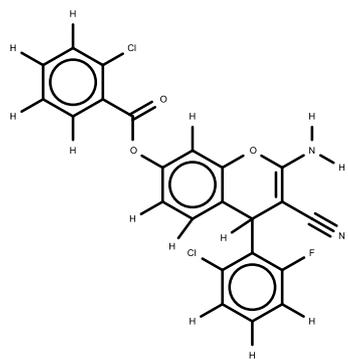
1068



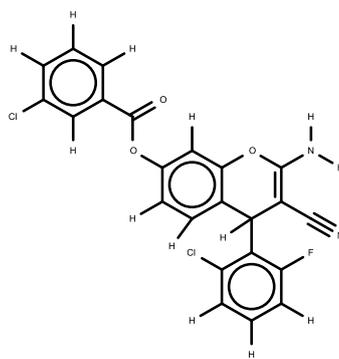
1069



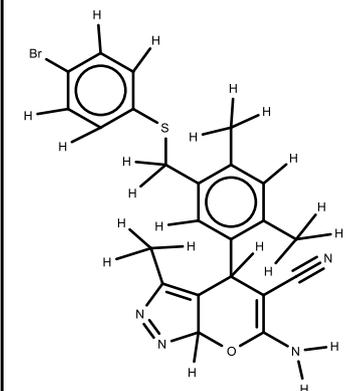
1070



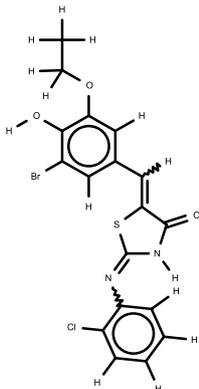
1071



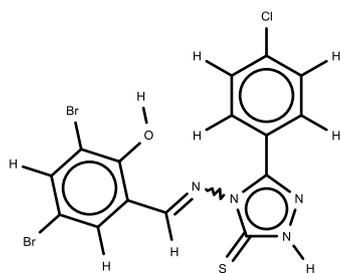
1072



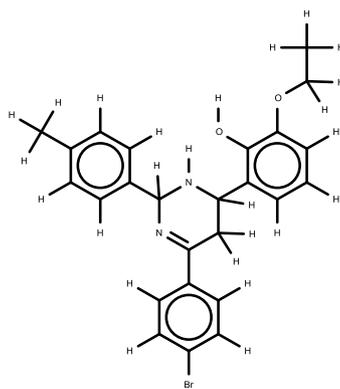
1073



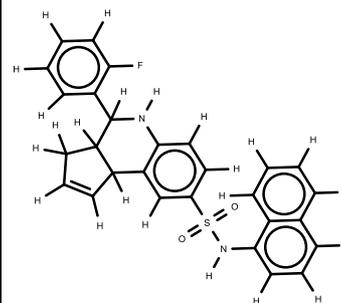
1074



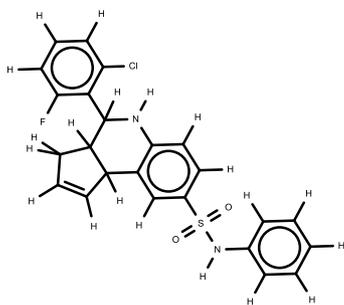
1075



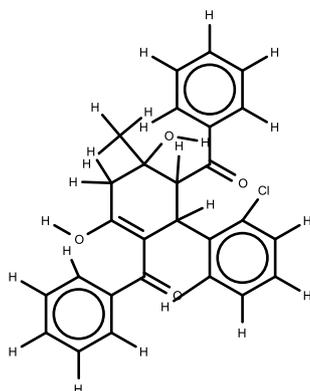
1076



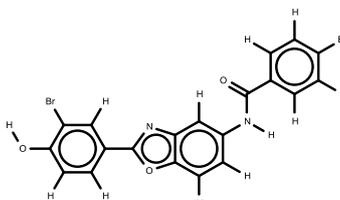
1077



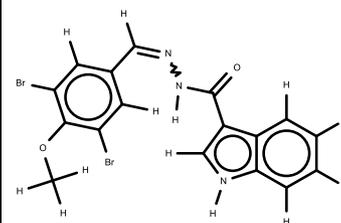
1078



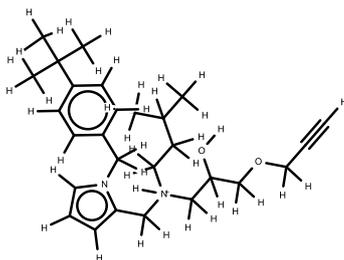
1079



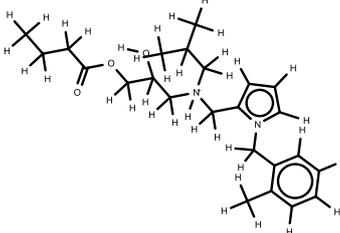
1080



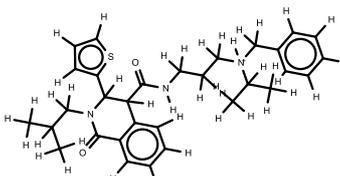
1081



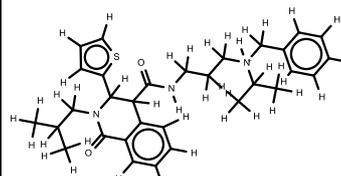
1082



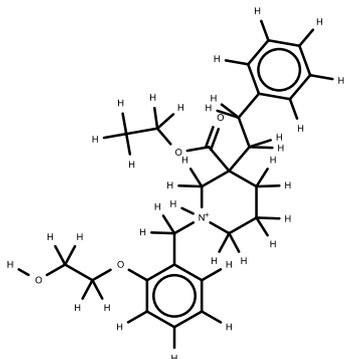
1083



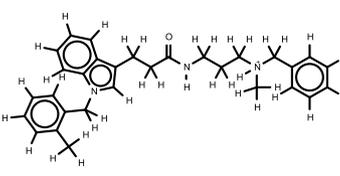
1084



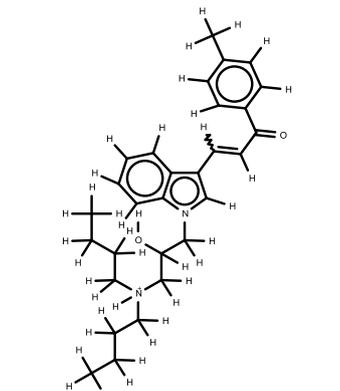
1085



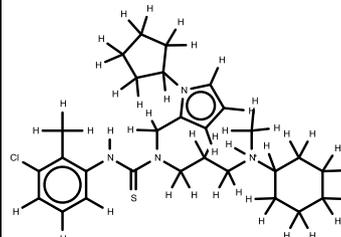
1086



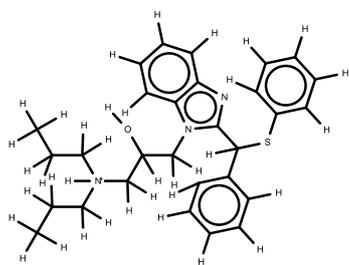
1087



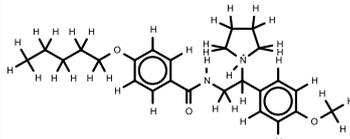
1088



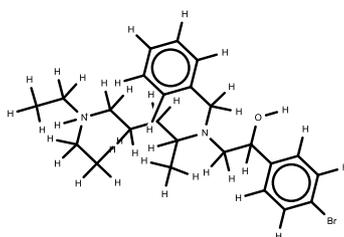
1089



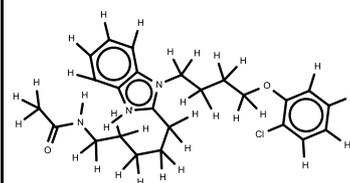
1090



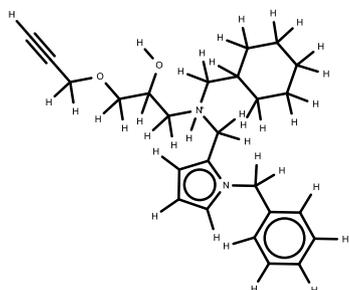
1091



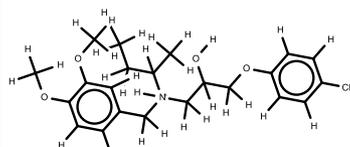
1092



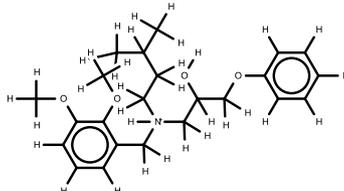
1093



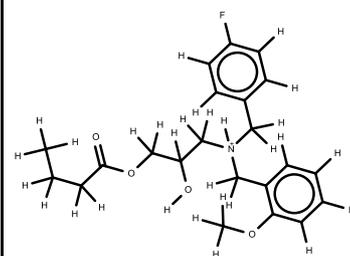
1094



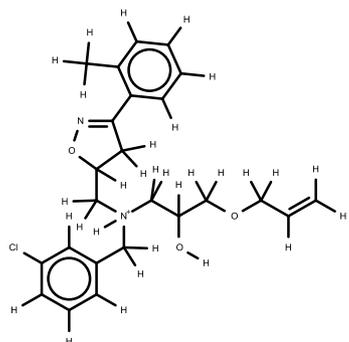
1095



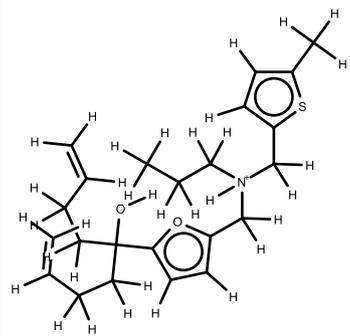
1096



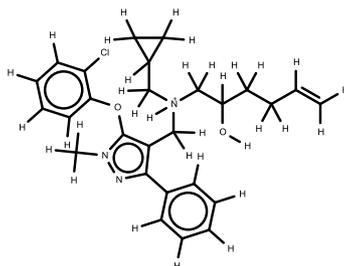
1097



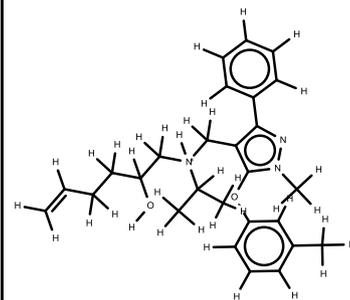
1098



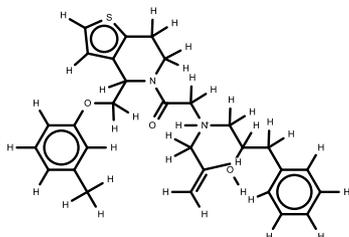
1099



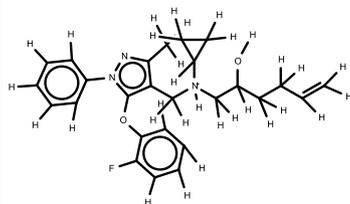
1100



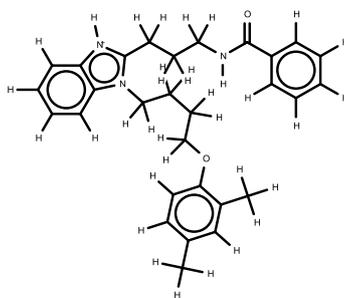
1101



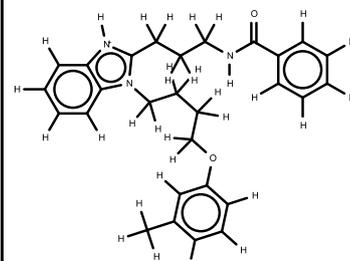
1102



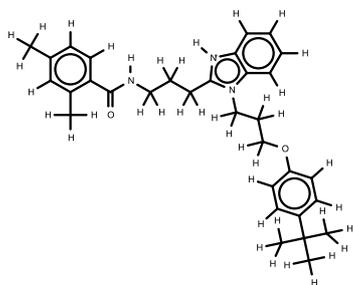
1103



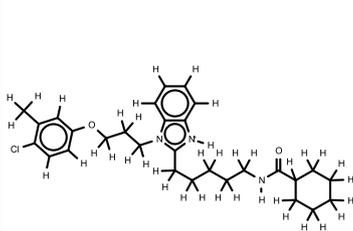
1104



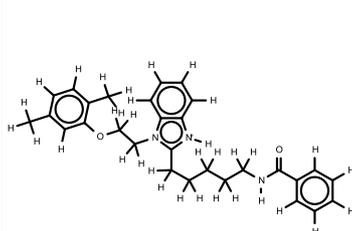
1105



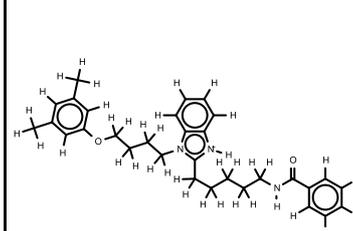
1106



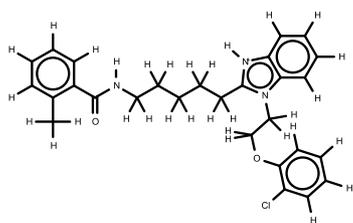
1107



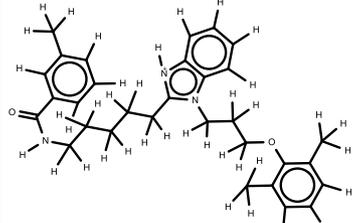
1108



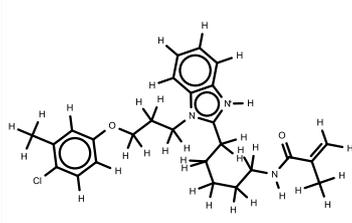
1109



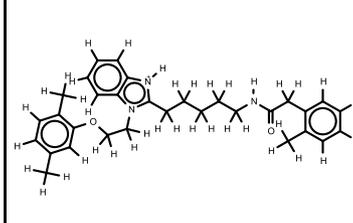
1110



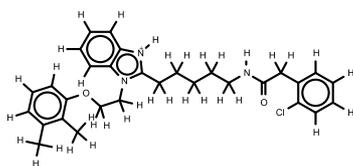
1111



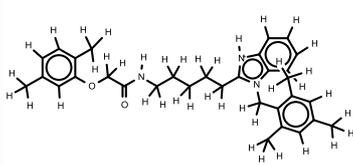
1112



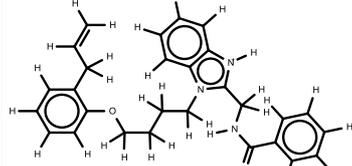
1113



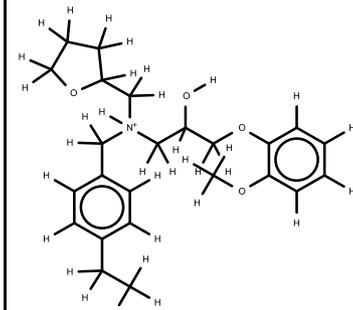
1114



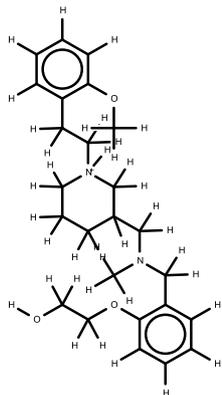
1115



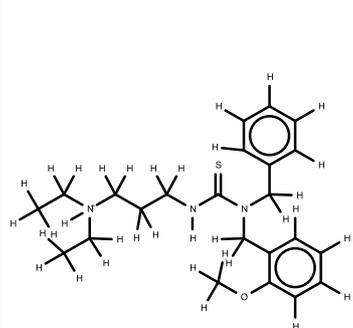
1116



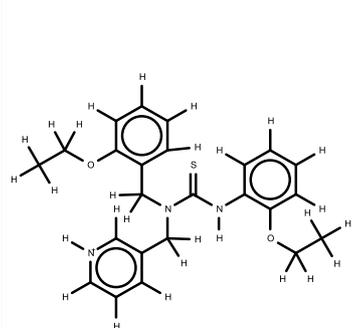
1117



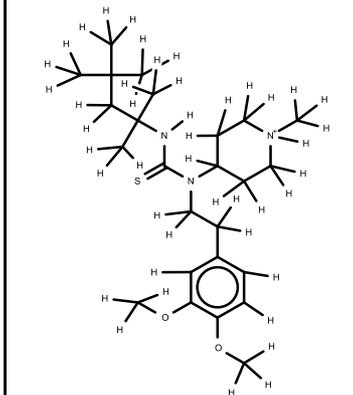
1118



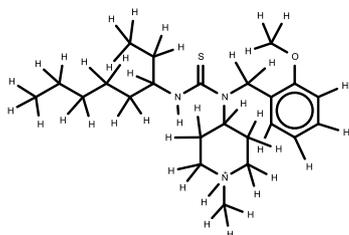
1119



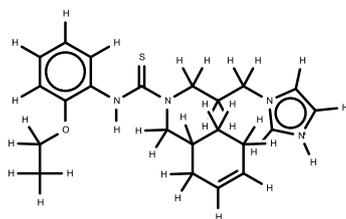
1120



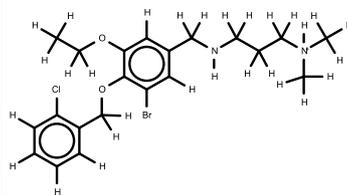
1121



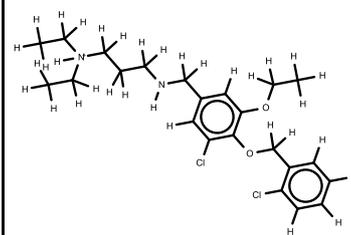
1122



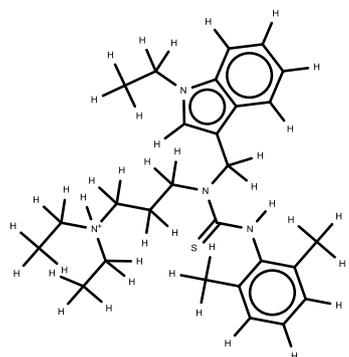
1123



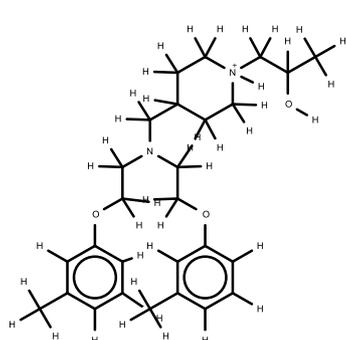
1124



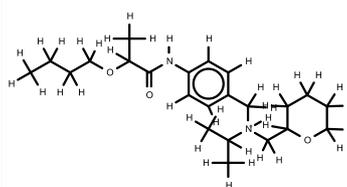
1125



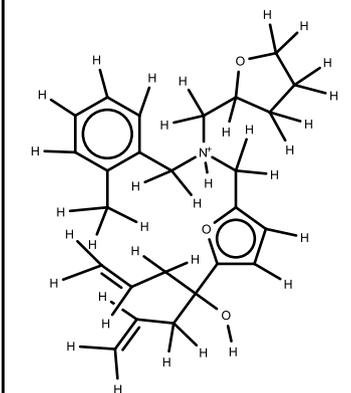
1126



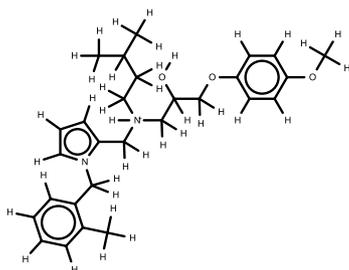
1127



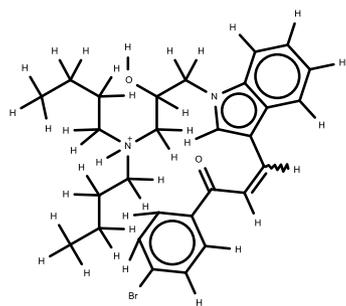
1128

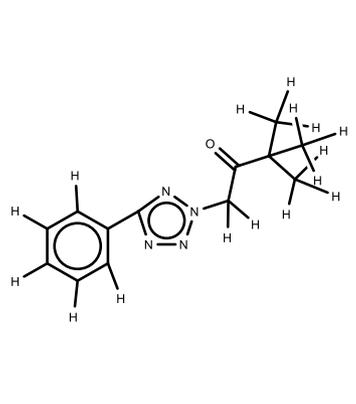
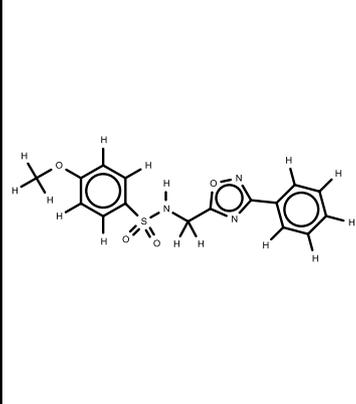
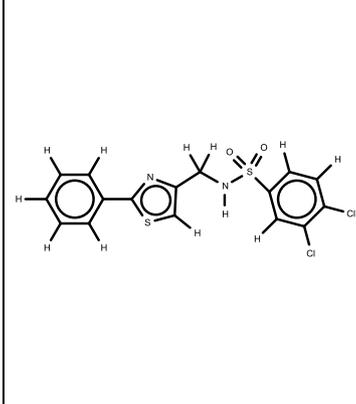
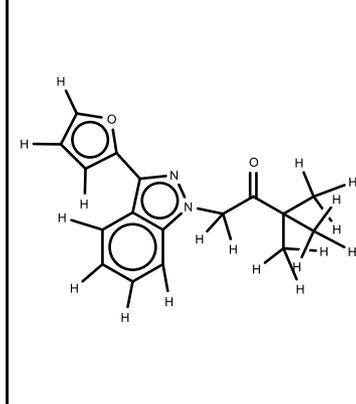
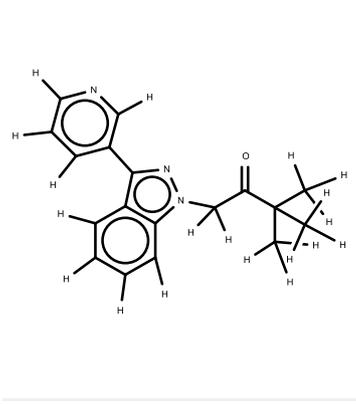
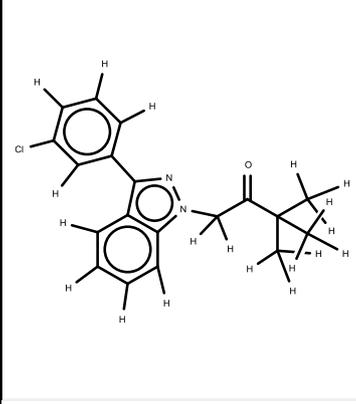
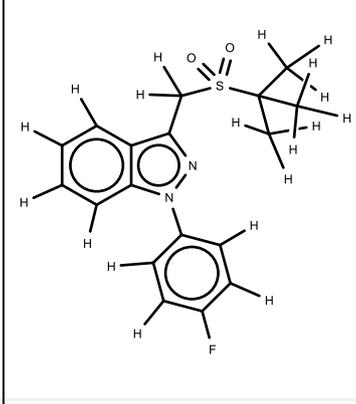
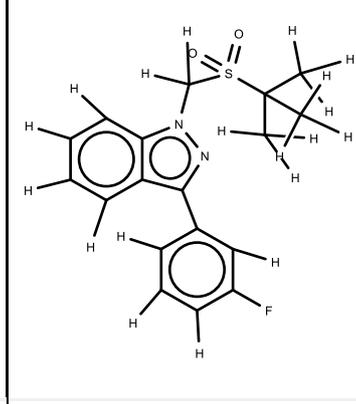
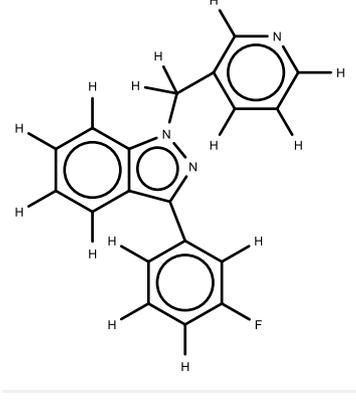
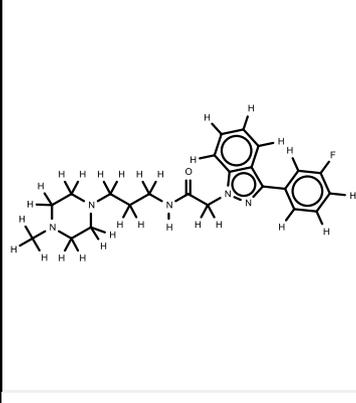
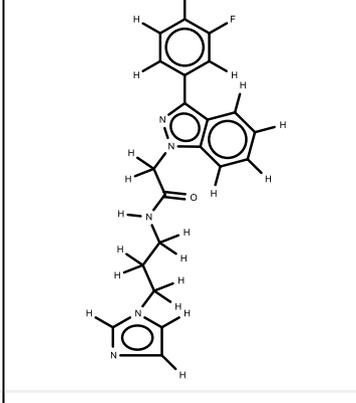
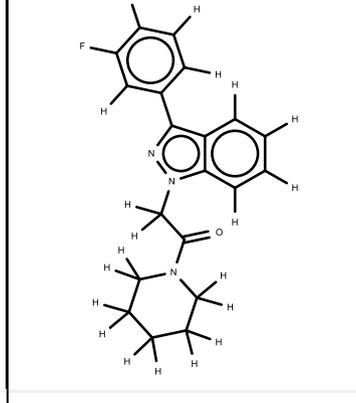
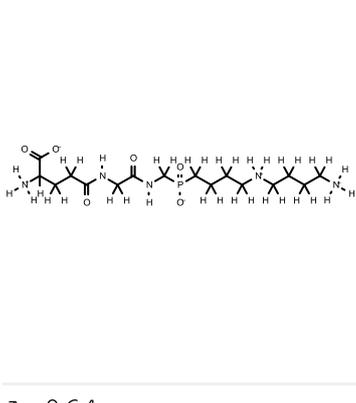
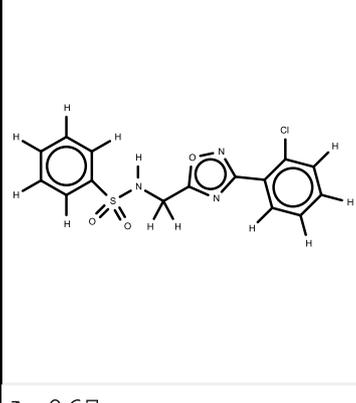
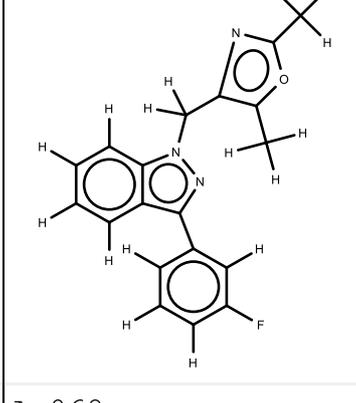
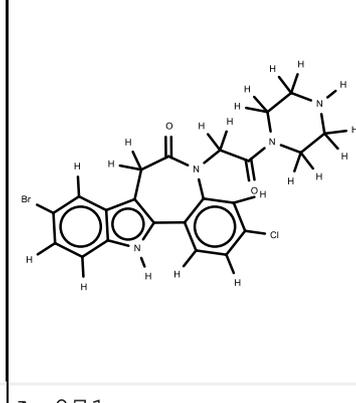


1129

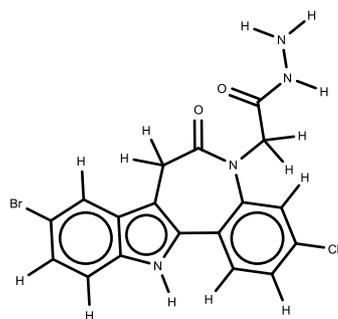


1130



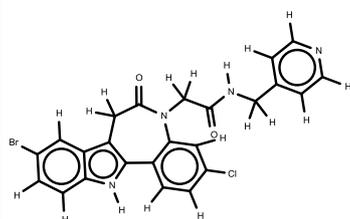
| 1  | 2   | 3  | 4   |
|--|---|--|---|
|    |    |    |    |
| A 011  | A 032   | A 039  | A 044   |
| 5  | 6   | 7  | 8   |
|   |   |   |   |
| A 045  | A 047   | A 050  | A 051   |
| 9  | 10  | 11   | 12  |
|  |  |  |  |
| A 053  | A 054   | A 055  | A 062   |
| 13   | 14  | 15   | 16  |
|  |  |  |  |
| A 064  | A 067   | A 068  | A 071   |

17



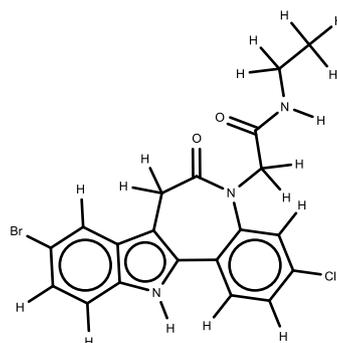
A 075

18



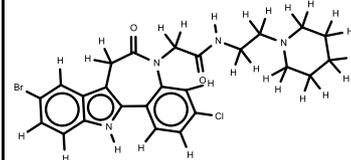
A 078

19



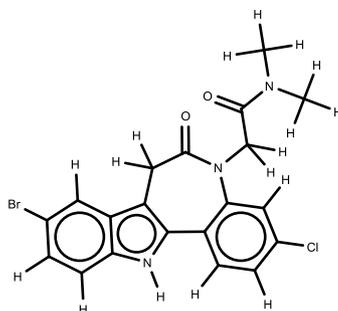
A 087

20



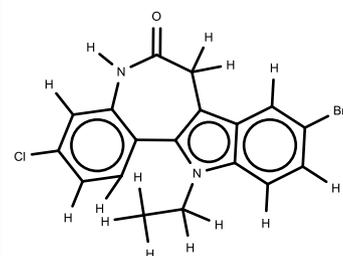
A 092

21



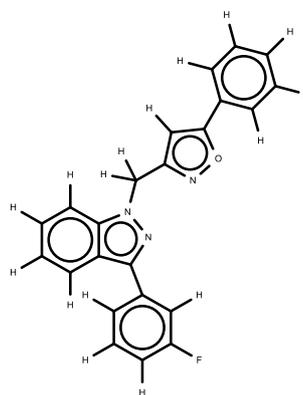
A 094

22



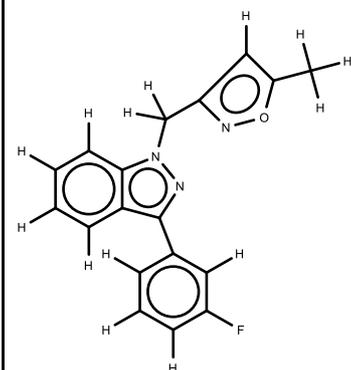
I023

23



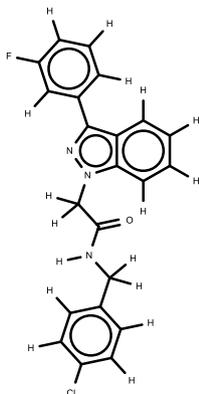
I024

24



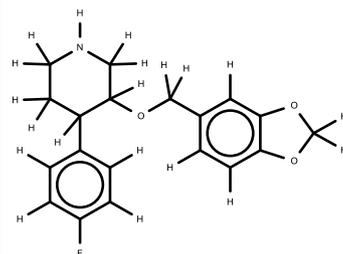
I025

25



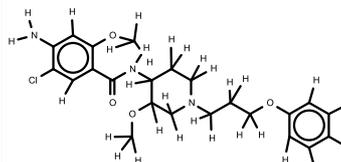
I026

26



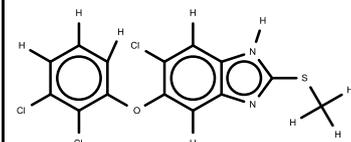
I028

27



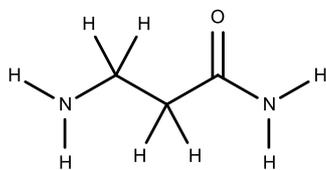
I032

28



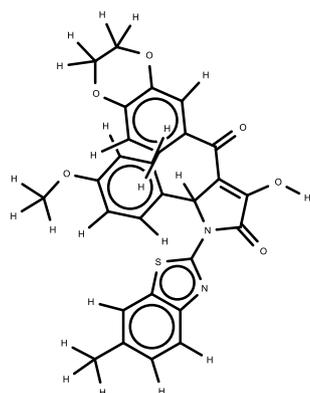
I033

29



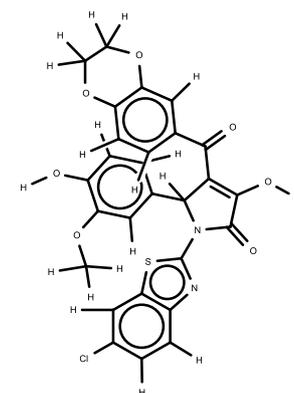
I034

30



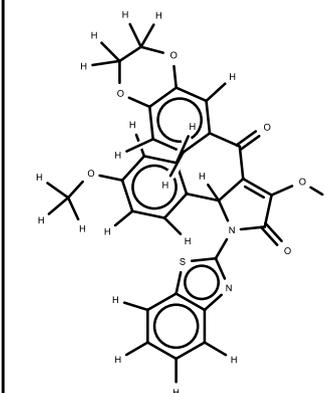
I036

31

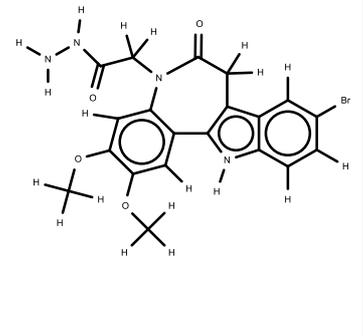
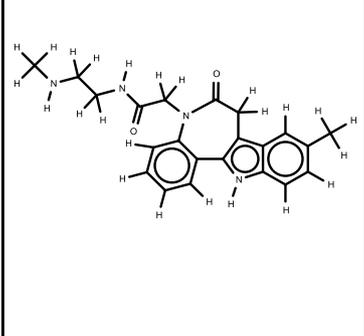
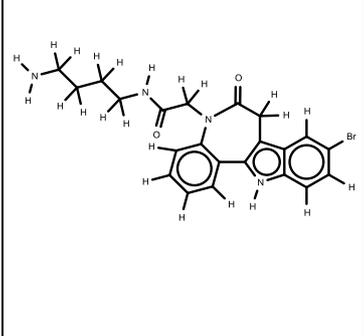
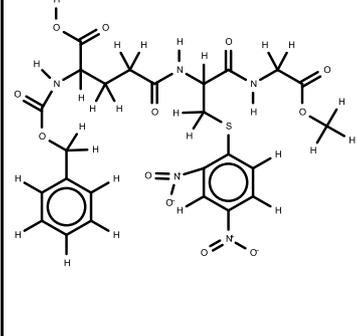
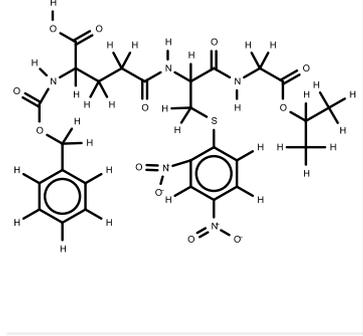
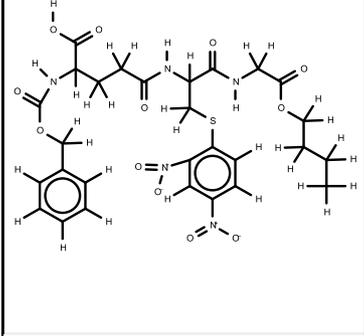
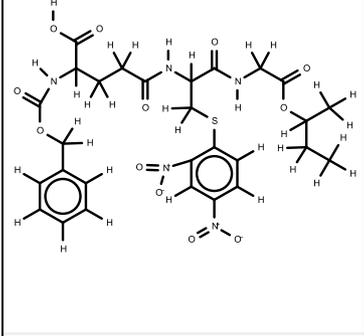
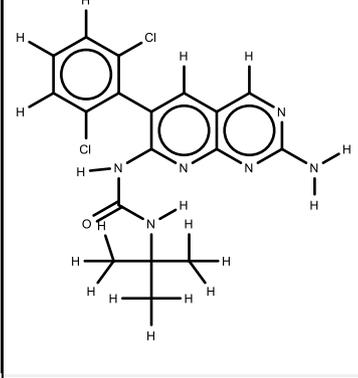
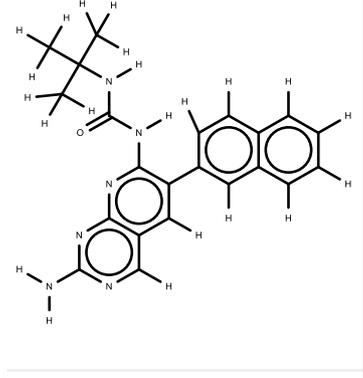
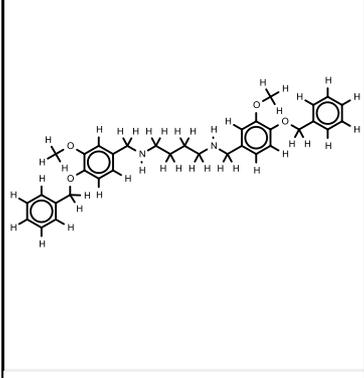
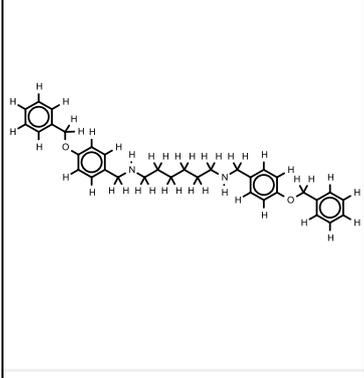
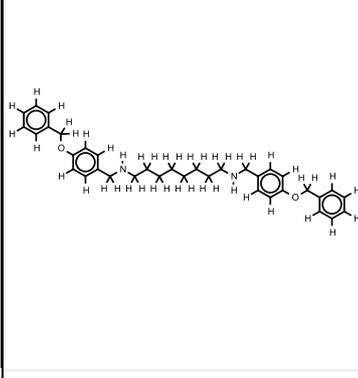
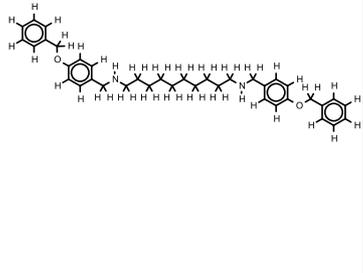
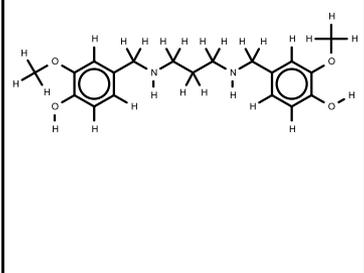
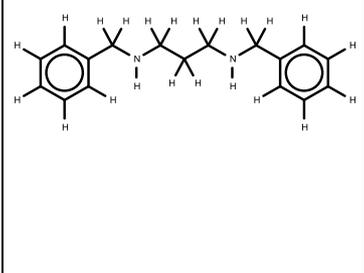
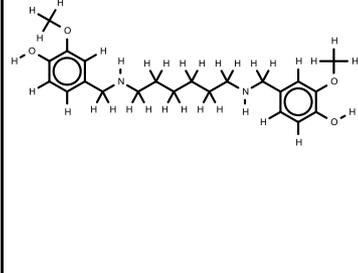


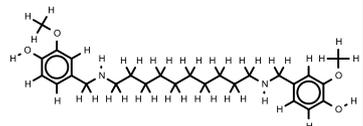
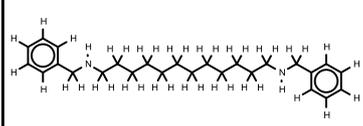
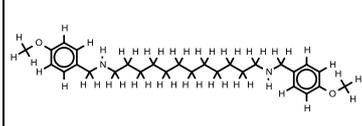
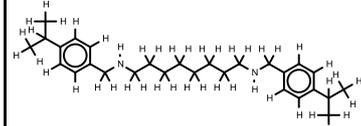
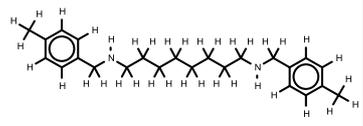
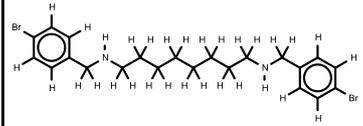
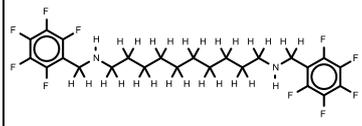
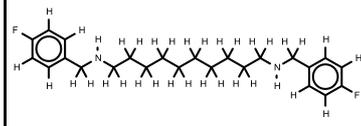
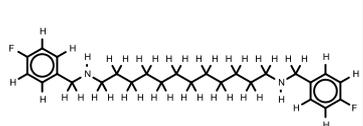
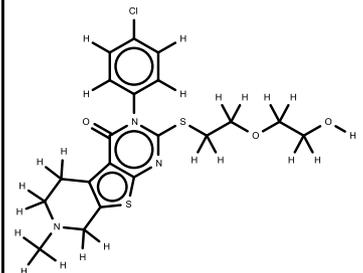
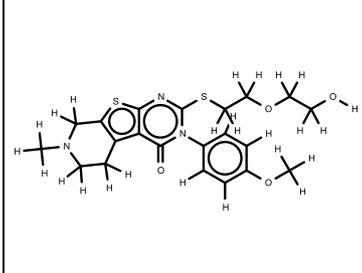
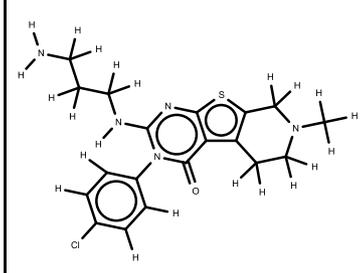
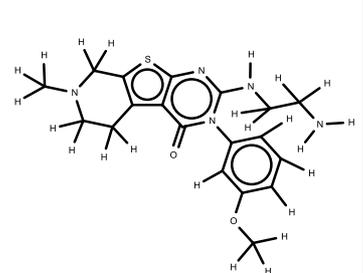
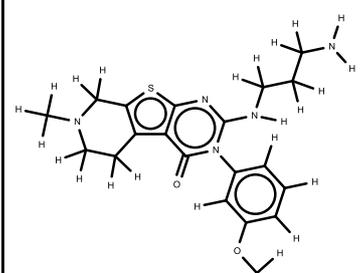
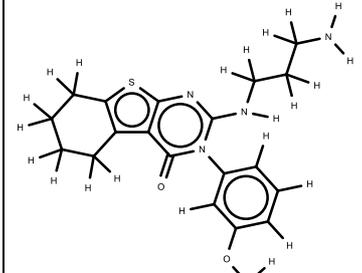
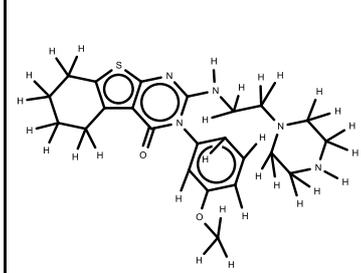
I038

32

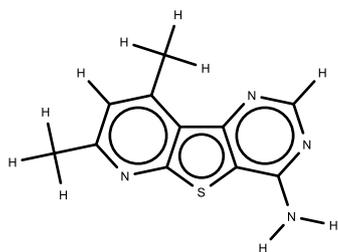


I042

|   |  |   |  |
|---|--|---|--|
| <p style="text-align: center;"><b>33</b></p>    | <p style="text-align: center;"><b>34</b></p>    | <p style="text-align: center;"><b>35</b></p>    | <p style="text-align: center;"><b>36</b></p>    |
| I046  | I048   | I049  | I054   |
| <p style="text-align: center;"><b>37</b></p>   | <p style="text-align: center;"><b>38</b></p>   | <p style="text-align: center;"><b>39</b></p>   | <p style="text-align: center;"><b>40</b></p>   |
| I057  | I058   | I059  | I086   |
| <p style="text-align: center;"><b>41</b></p>  | <p style="text-align: center;"><b>42</b></p>  | <p style="text-align: center;"><b>43</b></p>  | <p style="text-align: center;"><b>44</b></p>  |
| I087  | I090   | I092  | I094   |
| <p style="text-align: center;"><b>45</b></p>  | <p style="text-align: center;"><b>46</b></p>  | <p style="text-align: center;"><b>47</b></p>  | <p style="text-align: center;"><b>48</b></p>  |
| I095  | I099   | I100  | I104   |

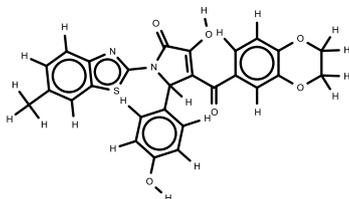
|   |  |   |  |
|---|--|---|--|
| <p style="text-align: center;"><b>49</b></p>  <p>I105</p>   | <p style="text-align: center;"><b>50</b></p>  <p>I108</p>   | <p style="text-align: center;"><b>51</b></p>  <p>I109</p>   | <p style="text-align: center;"><b>52</b></p>  <p>I110</p>   |
| <p style="text-align: center;"><b>53</b></p>  <p>I113</p>   | <p style="text-align: center;"><b>54</b></p>  <p>I114</p>   | <p style="text-align: center;"><b>55</b></p>  <p>I116</p>   | <p style="text-align: center;"><b>56</b></p>  <p>I117</p>   |
| <p style="text-align: center;"><b>57</b></p>  <p>I122</p> | <p style="text-align: center;"><b>58</b></p>  <p>I123</p> | <p style="text-align: center;"><b>59</b></p>  <p>I125</p> | <p style="text-align: center;"><b>60</b></p>  <p>I126</p> |
| <p style="text-align: center;"><b>61</b></p>  <p>I131</p> | <p style="text-align: center;"><b>62</b></p>  <p>I133</p> | <p style="text-align: center;"><b>63</b></p>  <p>I136</p> | <p style="text-align: center;"><b>64</b></p>  <p>I137</p> |

65



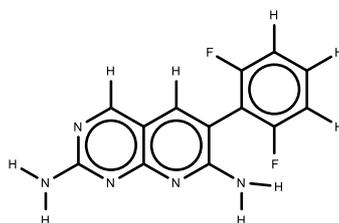
I138

66



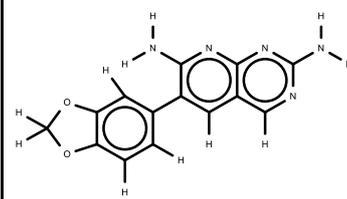
I139

67



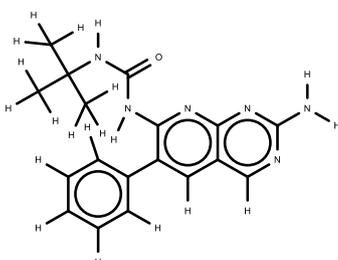
I143

68



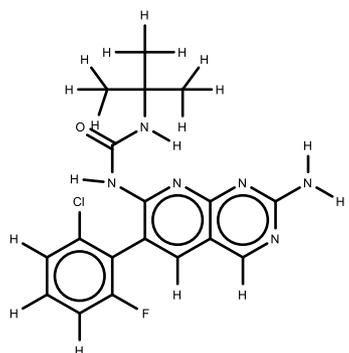
I144

69



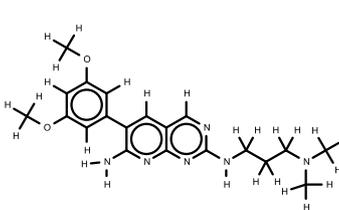
I149

70



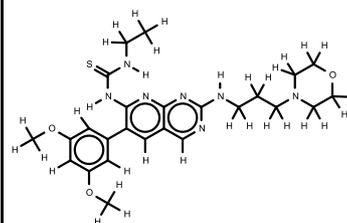
I152

71



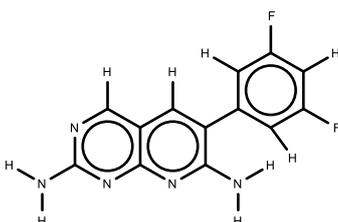
I154

72



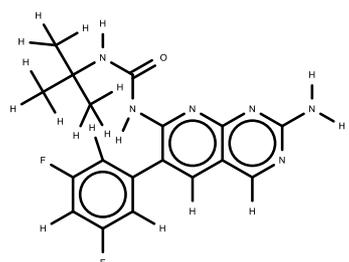
I157

73



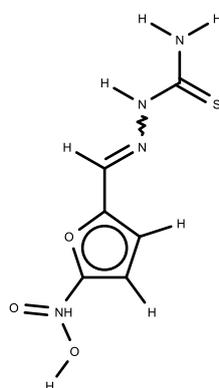
I158

74



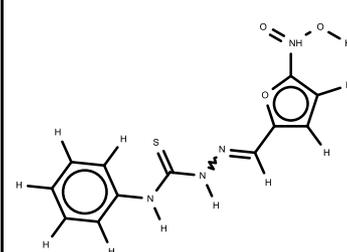
I159

75



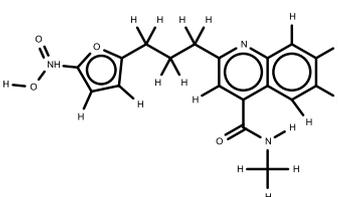
I163

76



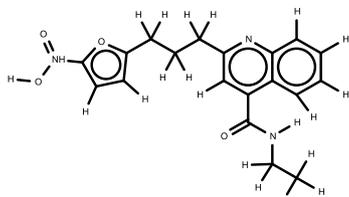
I164

77



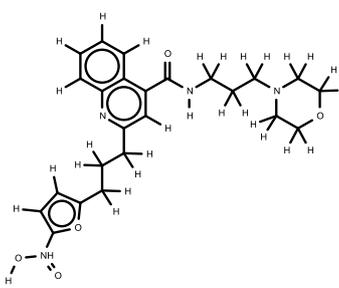
I167

78



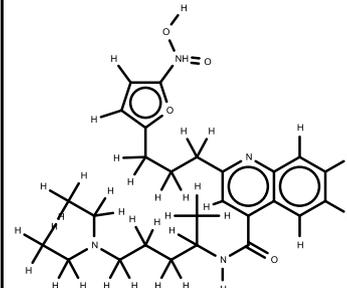
I168

79



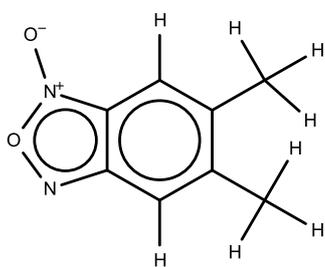
I170

80



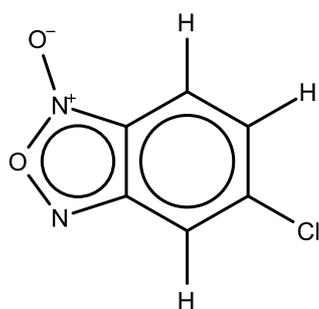
I171

81



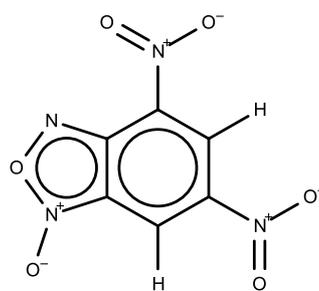
I174

82



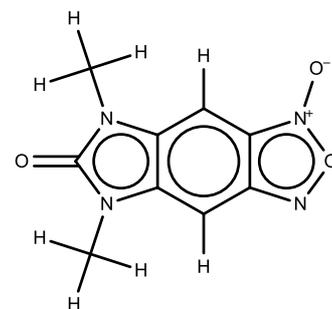
I175

83



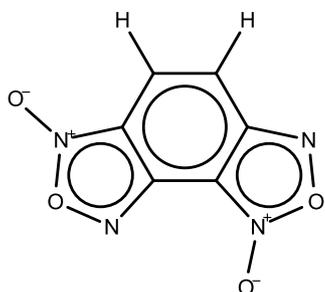
I180

84



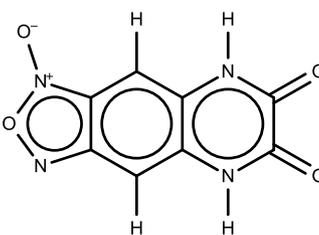
I181

85



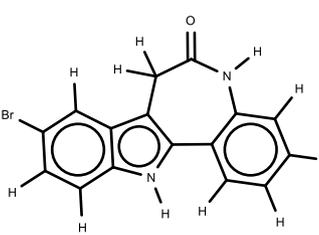
I182

86



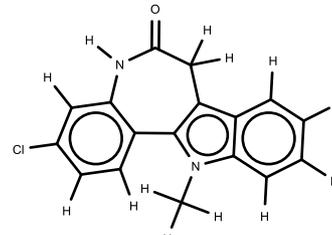
I185

87



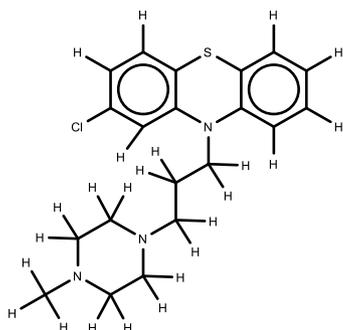
I187

88



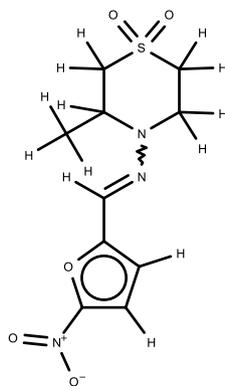
I189

89



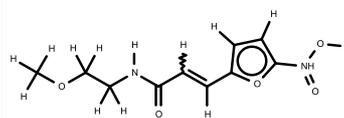
I194

90



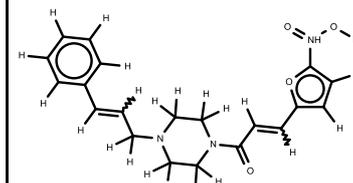
I196

91



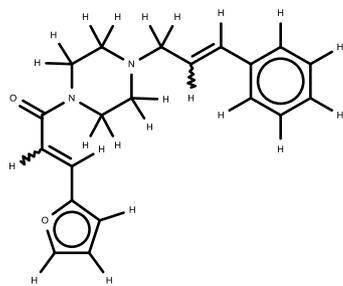
I197

92



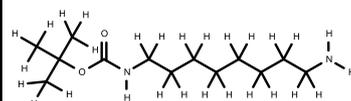
I198

93



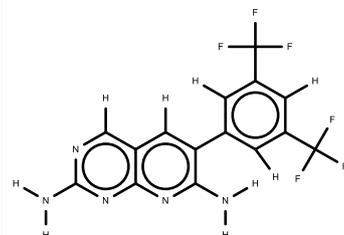
I199

94



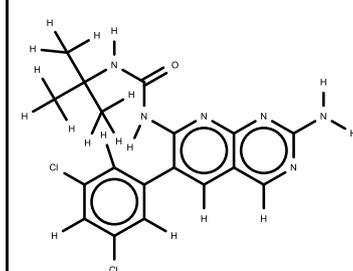
I202

95

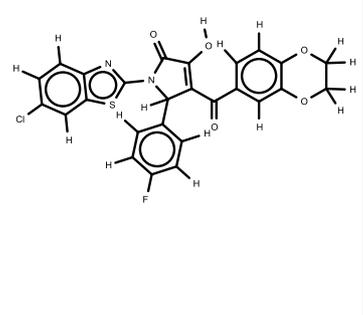
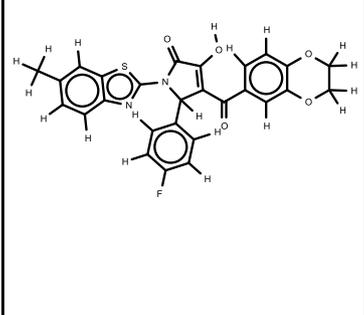
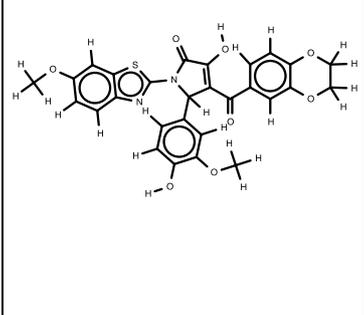
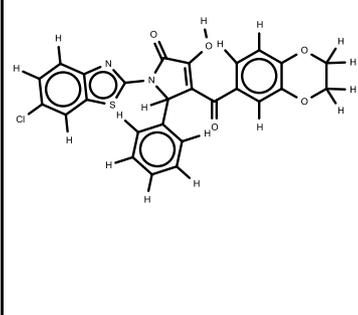
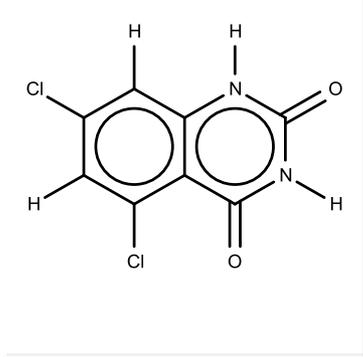
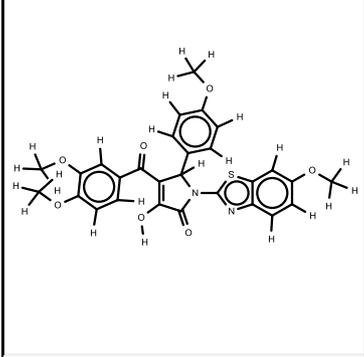
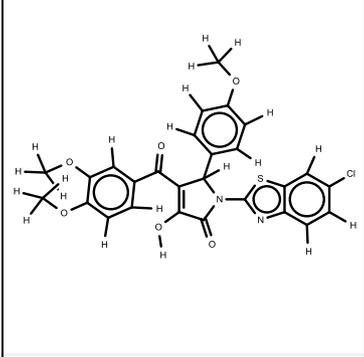
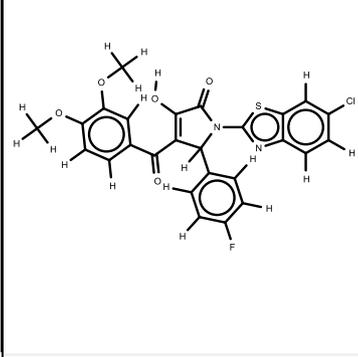
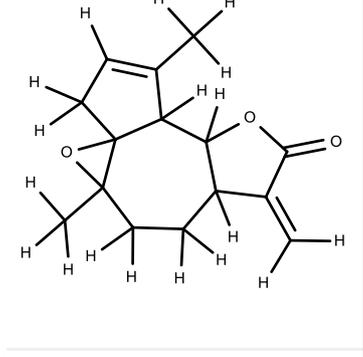
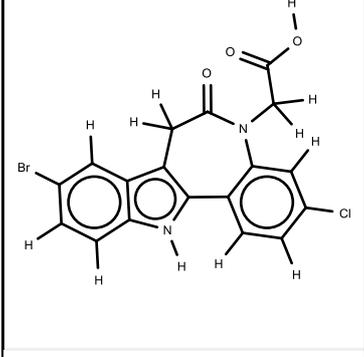
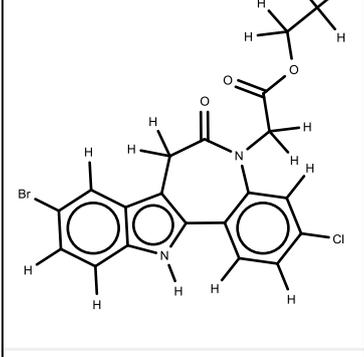
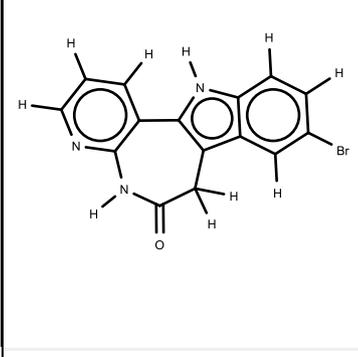
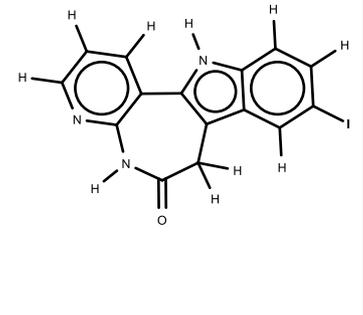
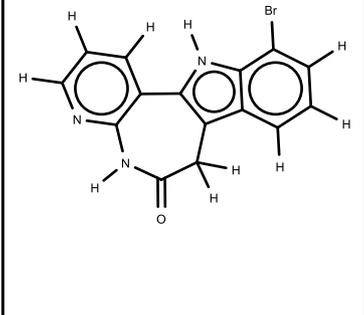
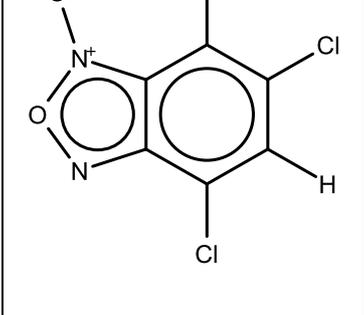
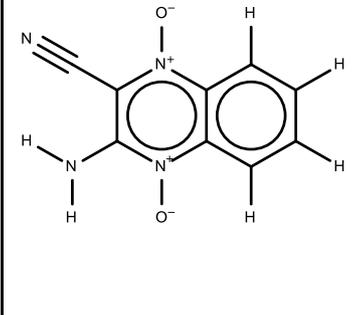


I203

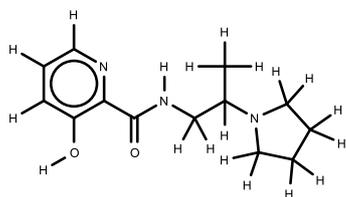
96



I205

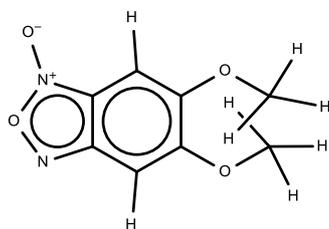
|  |   |  |   |
|--|---|--|---|
| <p style="text-align: center;"><b>97</b></p>     | <p style="text-align: center;"><b>98</b></p>     | <p style="text-align: center;"><b>99</b></p>     | <p style="text-align: center;"><b>100</b></p>    |
| <p>I206</p>  | <p>I210</p>   | <p>I212</p>  | <p>I215</p>   |
| <p style="text-align: center;"><b>101</b></p>   | <p style="text-align: center;"><b>102</b></p>   | <p style="text-align: center;"><b>103</b></p>   | <p style="text-align: center;"><b>104</b></p>   |
| <p>I216</p>  | <p>I218</p>   | <p>I219</p>  | <p>I220</p>   |
| <p style="text-align: center;"><b>105</b></p>  | <p style="text-align: center;"><b>106</b></p>  | <p style="text-align: center;"><b>107</b></p>  | <p style="text-align: center;"><b>108</b></p>  |
| <p>I225</p>  | <p>I226</p>   | <p>I227</p>  | <p>I232</p>   |
| <p style="text-align: center;"><b>109</b></p>  | <p style="text-align: center;"><b>110</b></p>  | <p style="text-align: center;"><b>111</b></p>  | <p style="text-align: center;"><b>112</b></p>  |
| <p>I233</p>  | <p>I236</p>   | <p>I238</p>  | <p>I241</p>   |

113



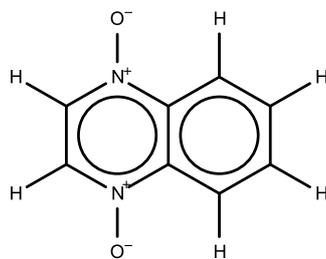
I243

114



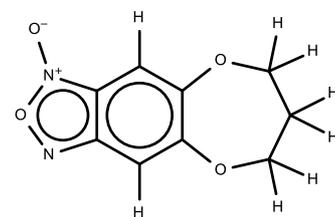
I244

115



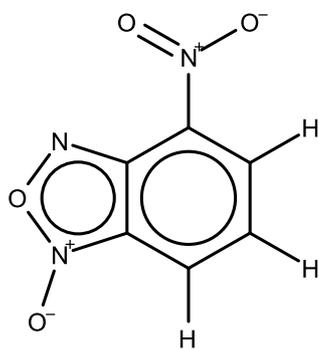
I246

116



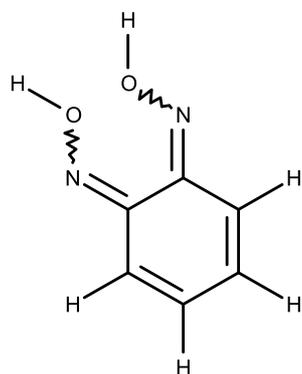
I248

117



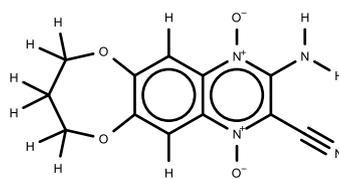
I249

118



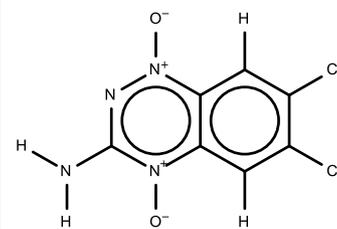
I250

119



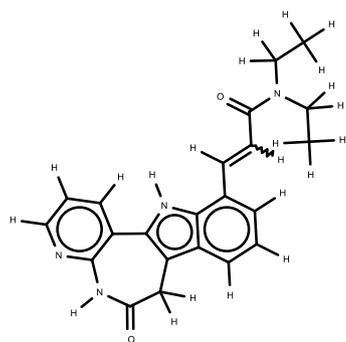
I252

120



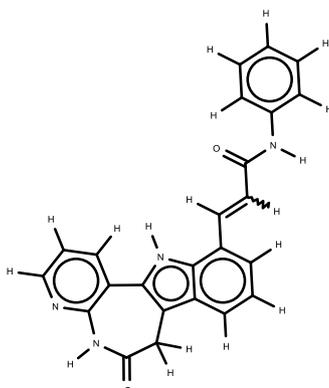
I253

121



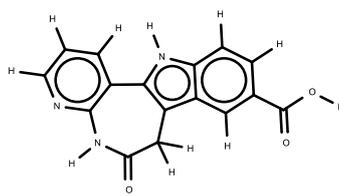
I257

122



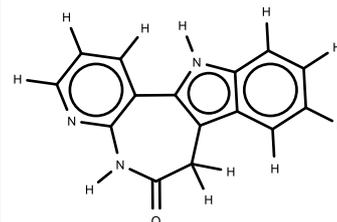
I260

123



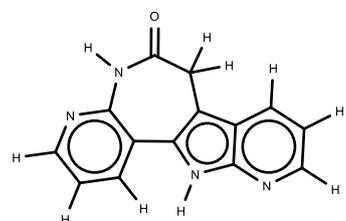
I264

124



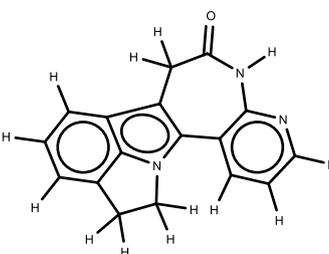
I265

125



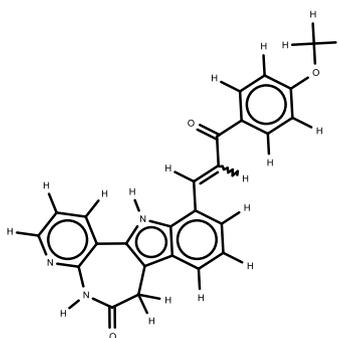
I267

126



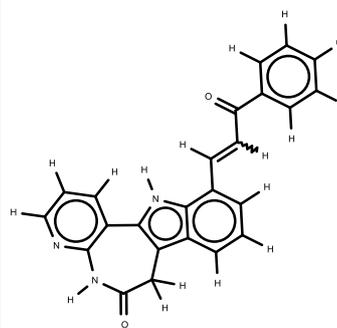
I268

127



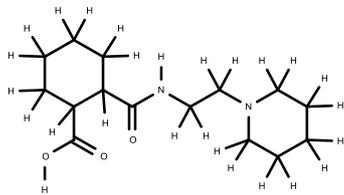
I272

128

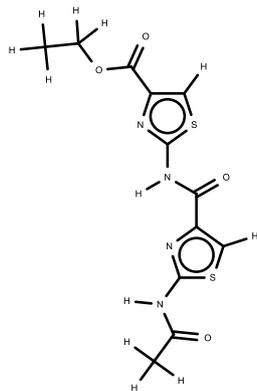


I273

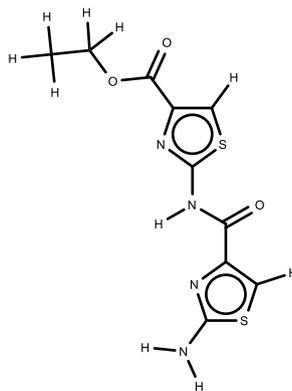
129



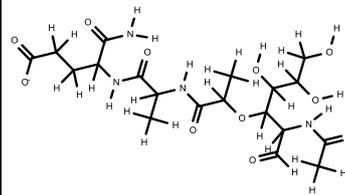
130



131



132

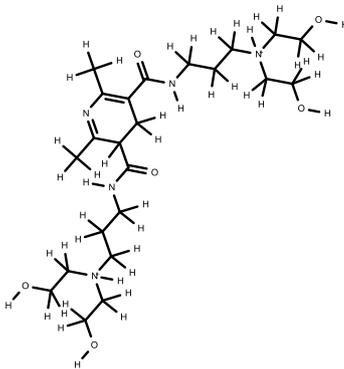


I274

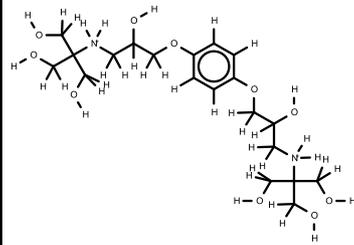
I285

I286

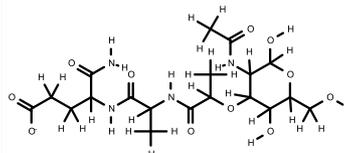
133



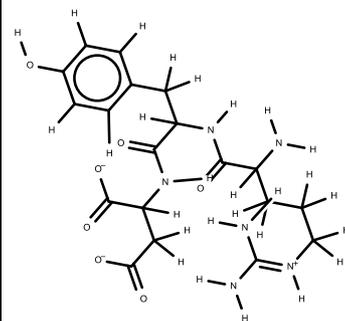
134



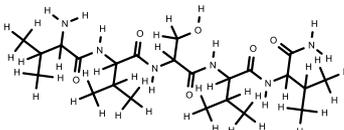
135



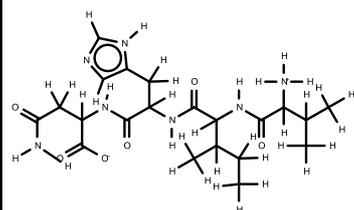
136



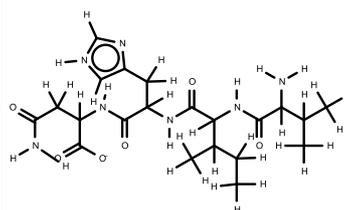
137



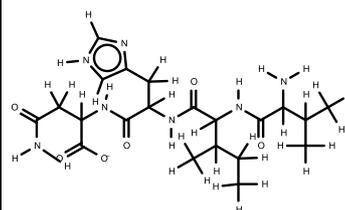
138



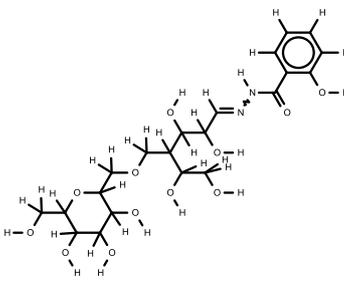
139



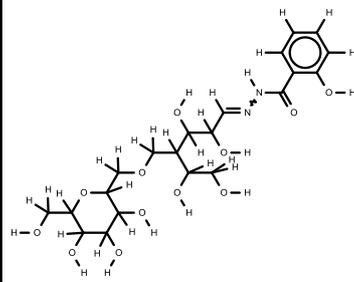
140



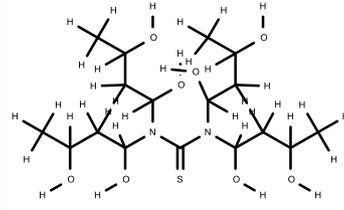
141



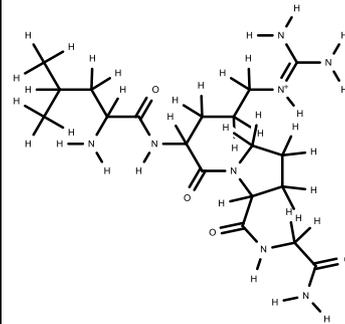
142



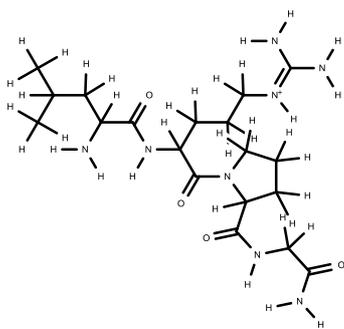
143



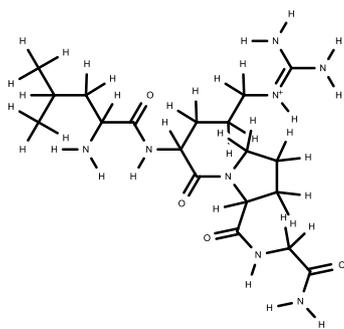
144



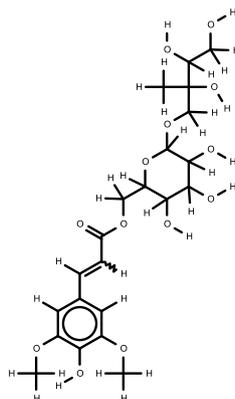
145



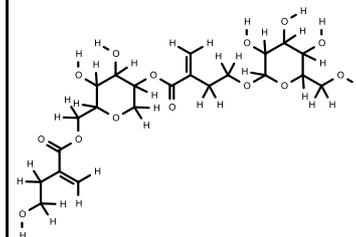
146



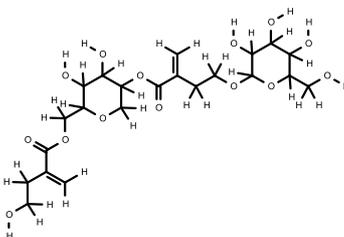
147



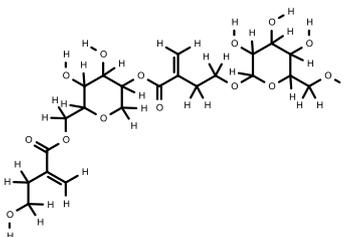
148



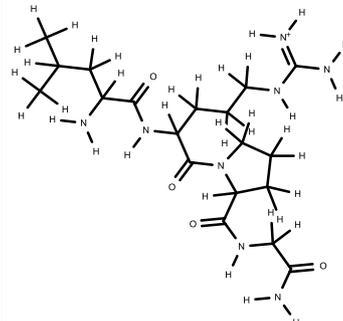
149



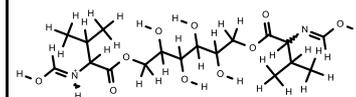
150



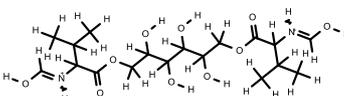
151



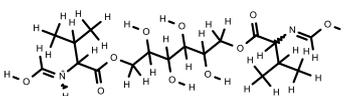
152



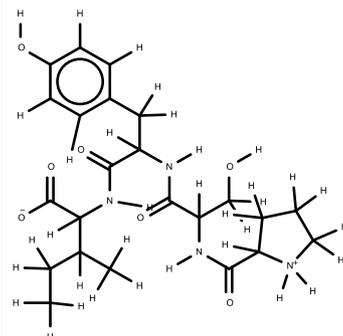
153



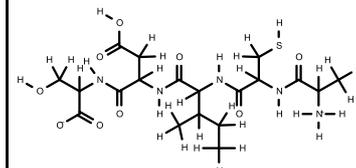
154



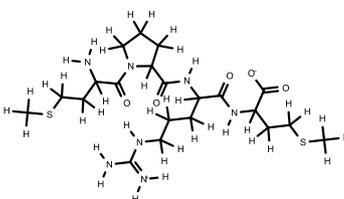
155



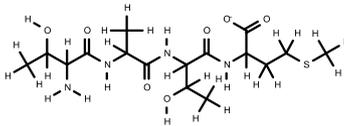
156



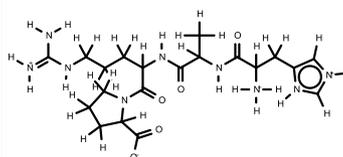
157



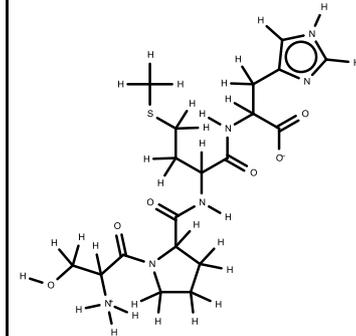
158



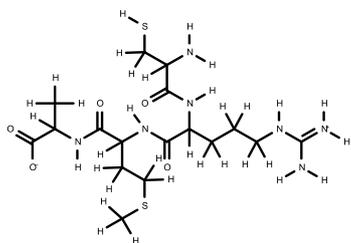
159



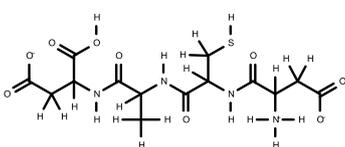
160



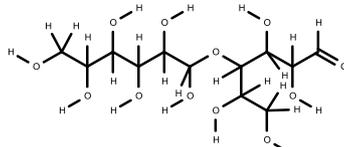
161



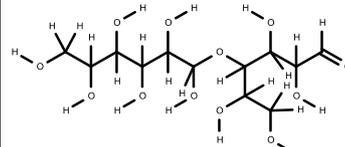
162



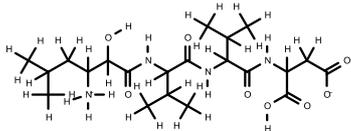
163



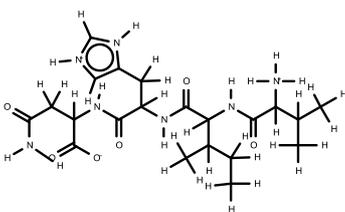
164



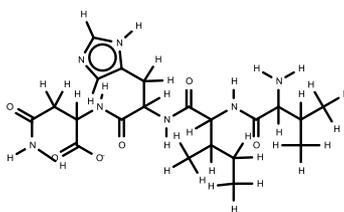
165



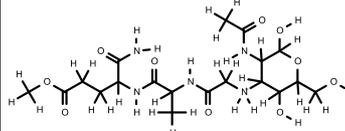
166



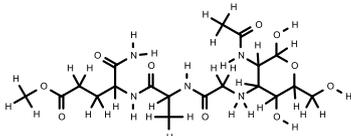
167



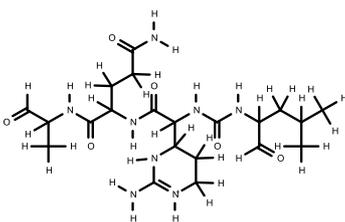
168



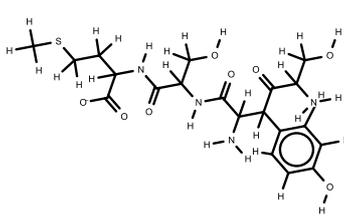
169



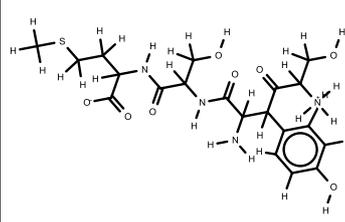
170



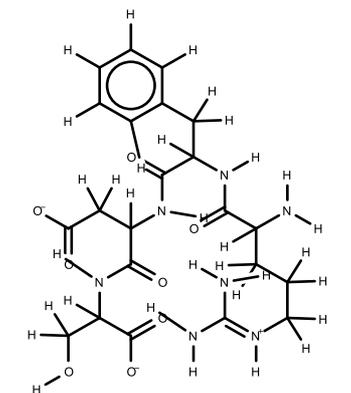
171



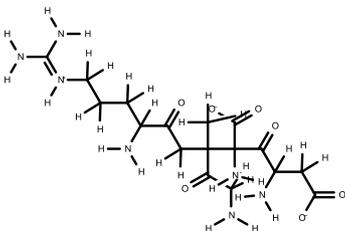
172



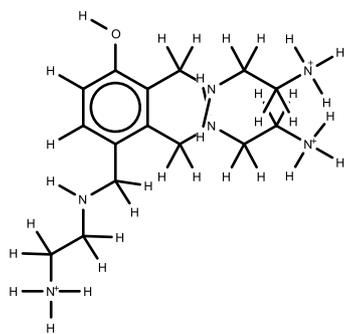
173



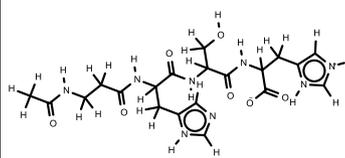
174



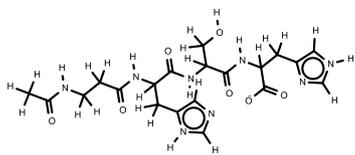
175



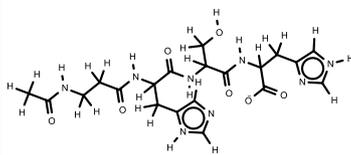
176



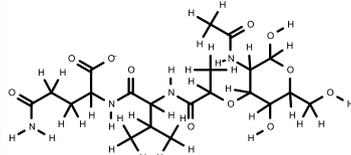
177



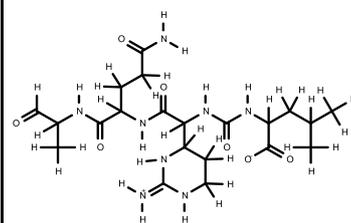
178



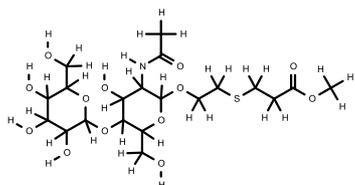
179



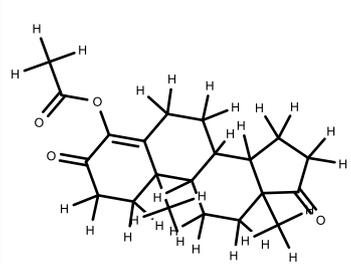
180



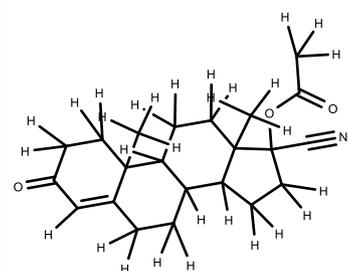
181



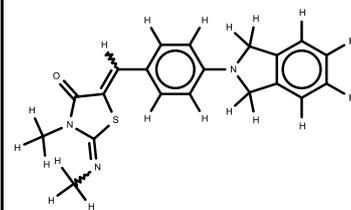
182



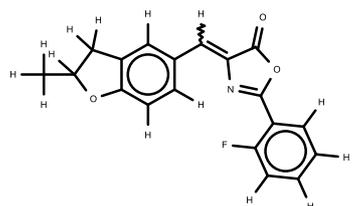
183



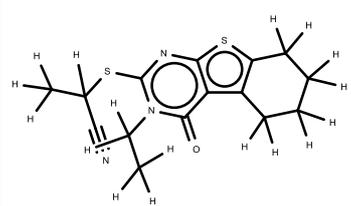
184



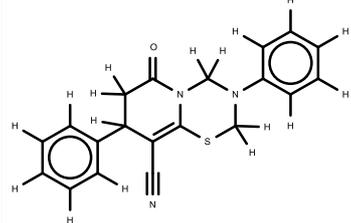
185



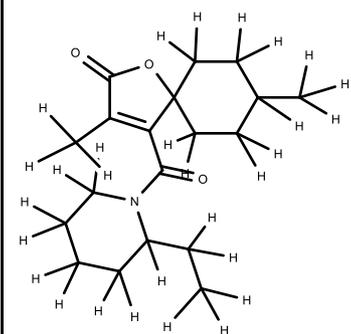
186



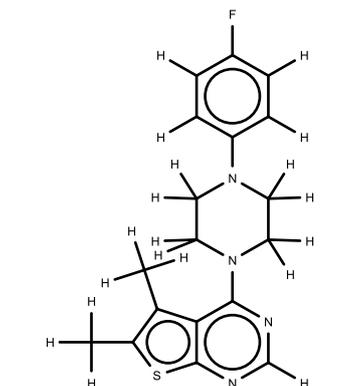
187



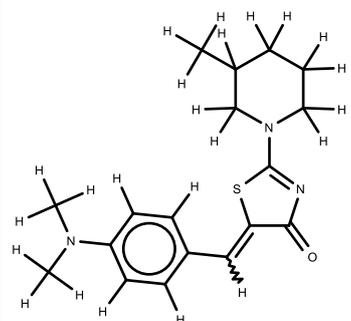
188



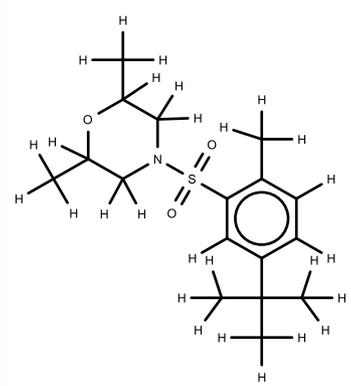
189



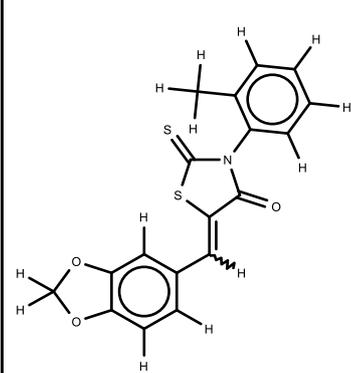
190



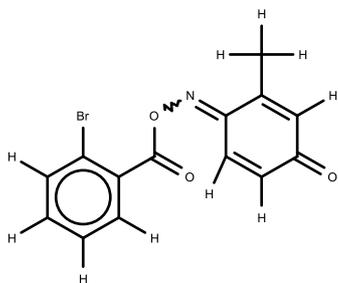
191



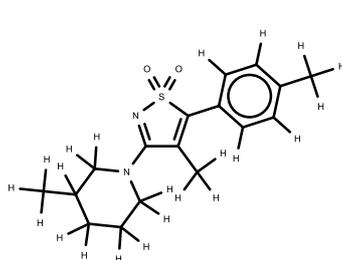
192



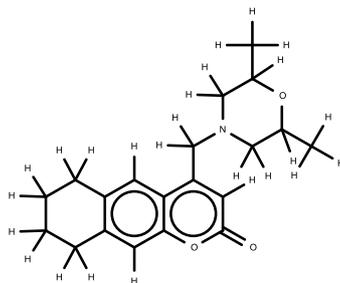
193



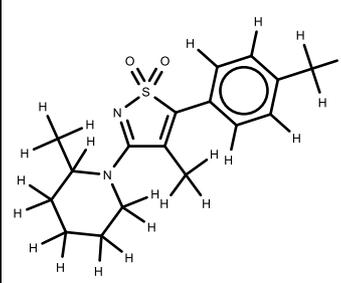
194



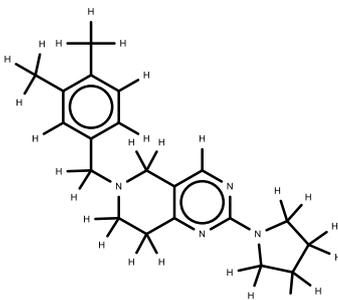
195



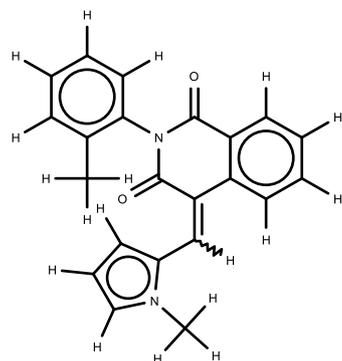
196



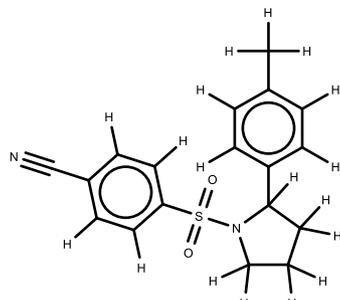
197



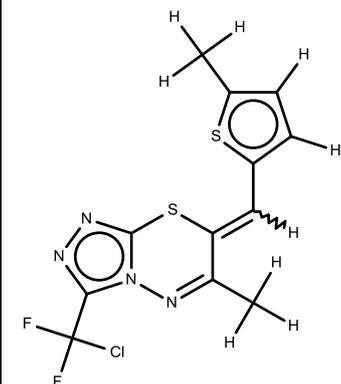
198



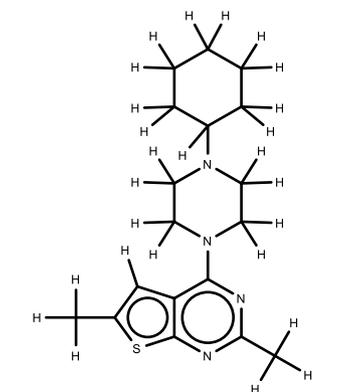
199



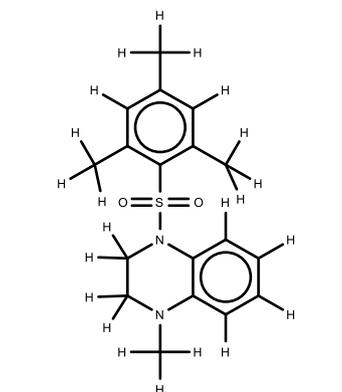
200



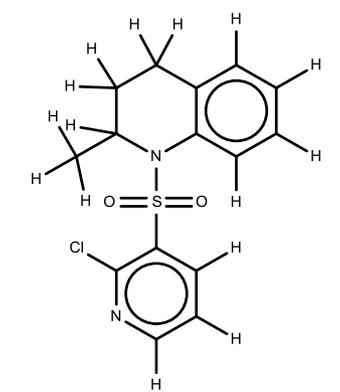
201



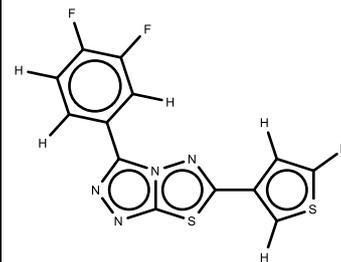
202



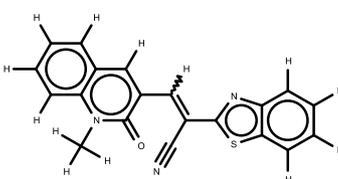
203



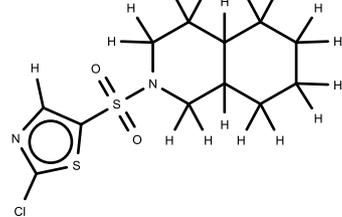
204



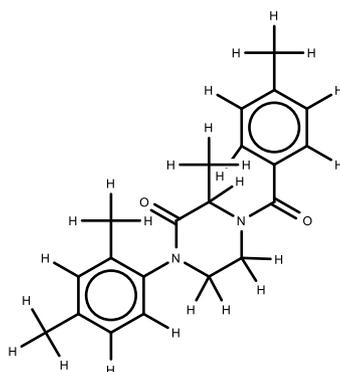
205



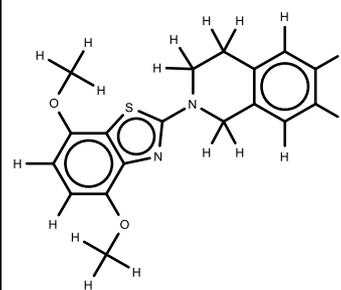
206



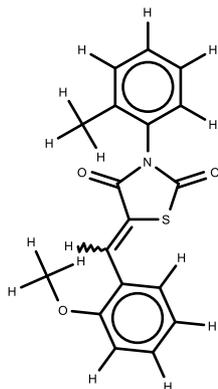
207



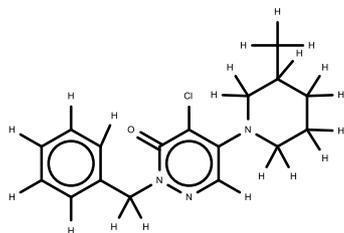
208



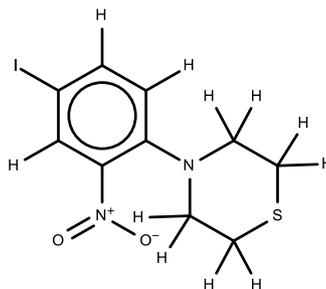
209



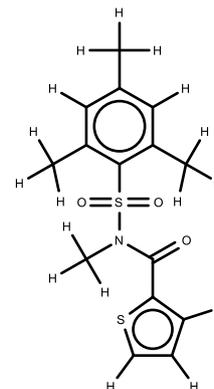
210



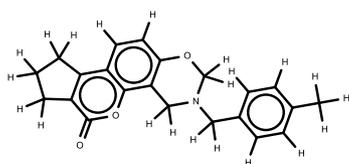
211



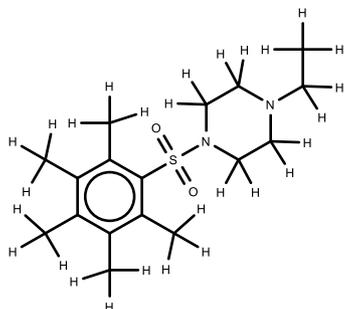
212



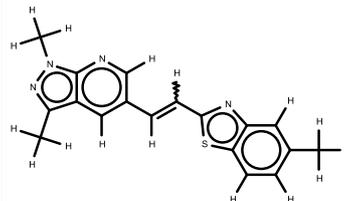
213



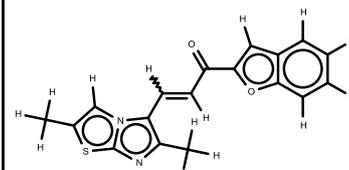
214



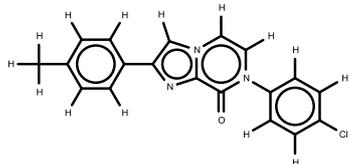
215



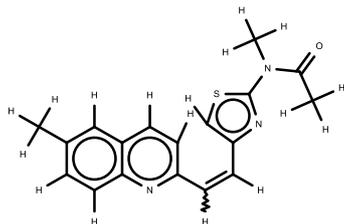
216



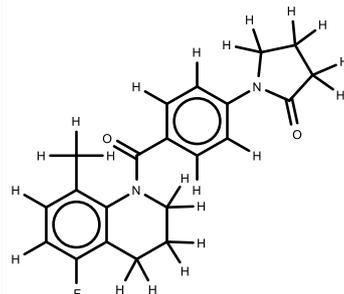
217



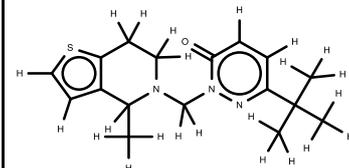
218



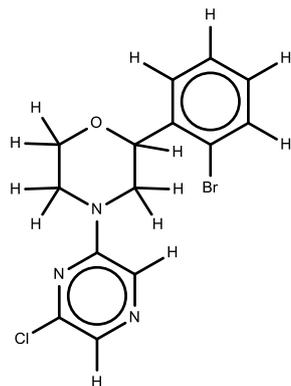
219



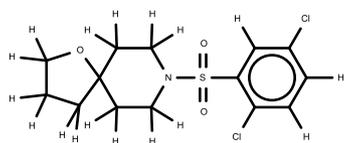
220



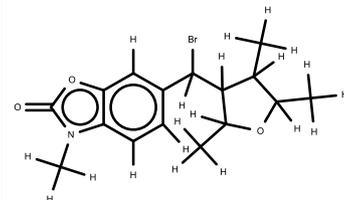
221



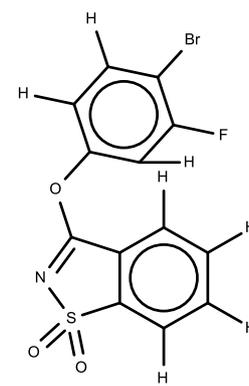
222



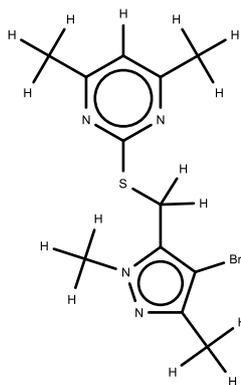
223



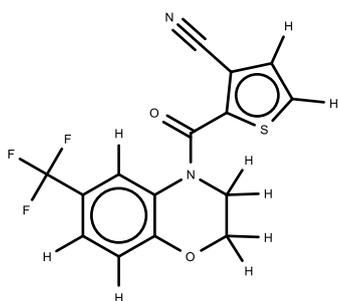
224



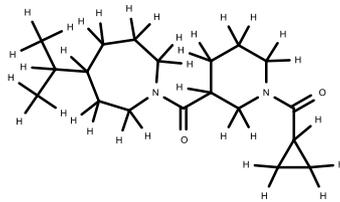
225



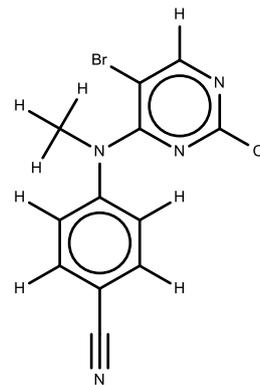
226



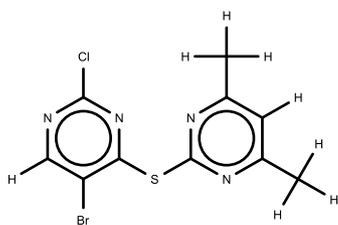
227



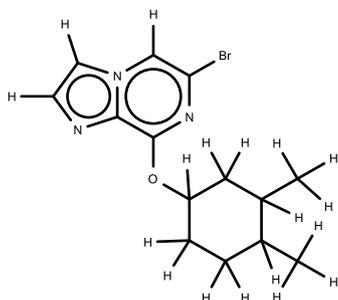
228



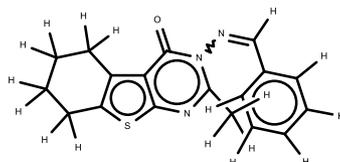
229



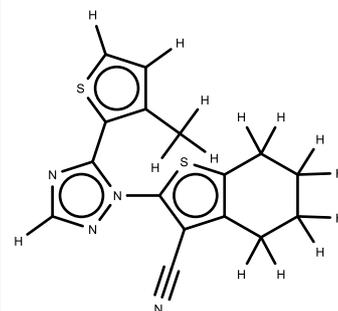
230



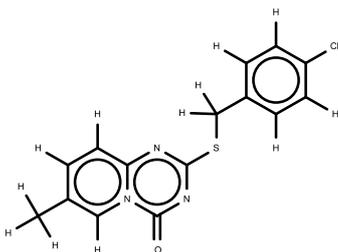
231



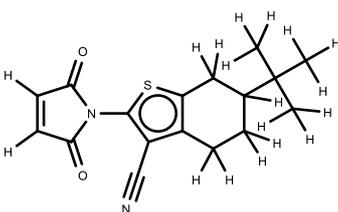
232



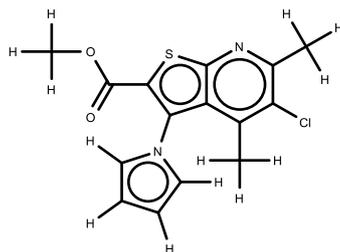
233



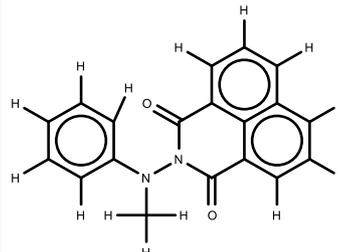
234



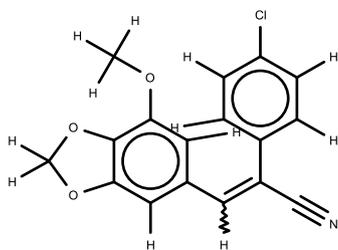
235



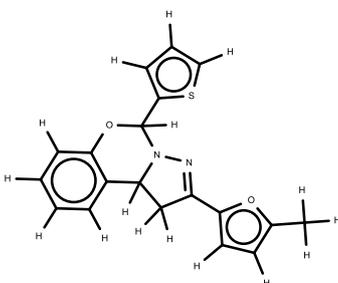
236



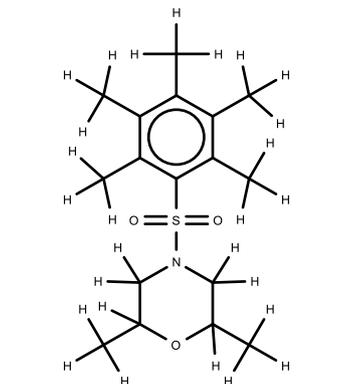
237



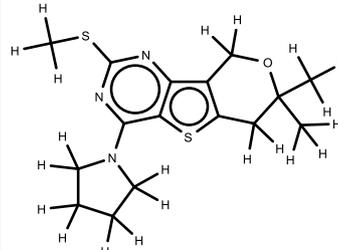
238



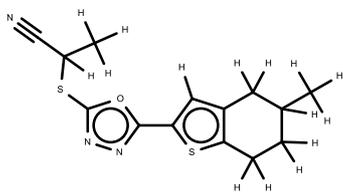
239



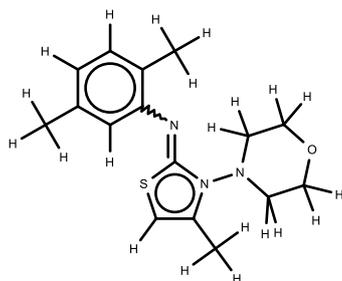
240



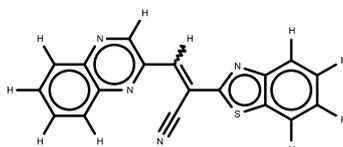
241



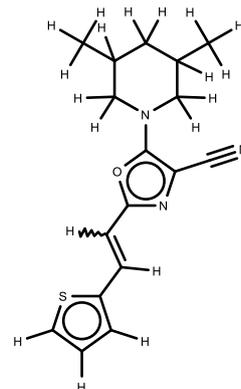
242



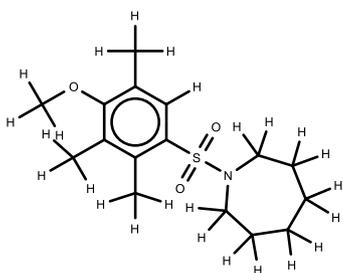
243



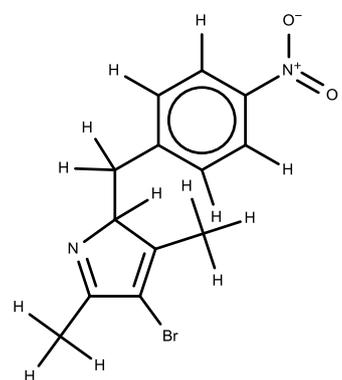
244



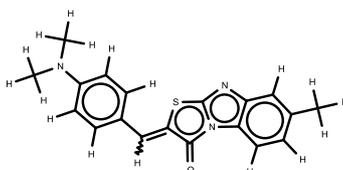
245



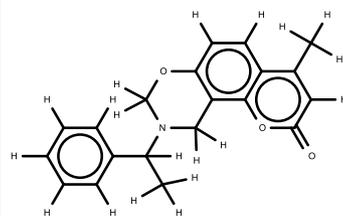
246



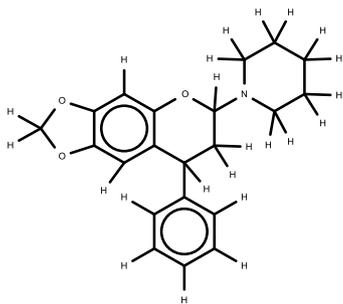
247



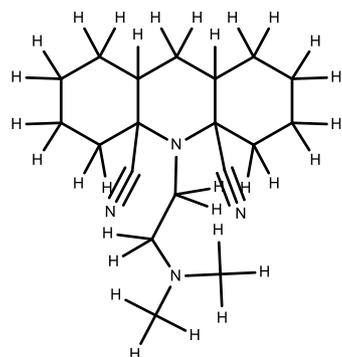
248



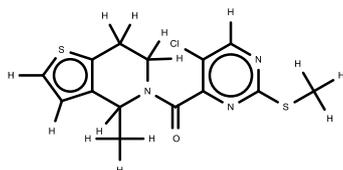
249



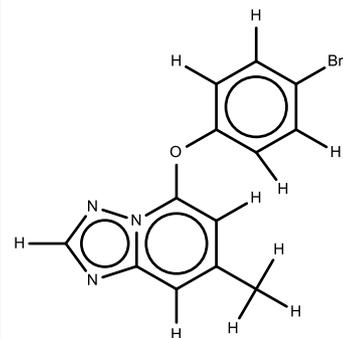
250



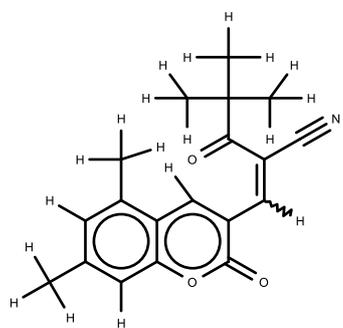
251



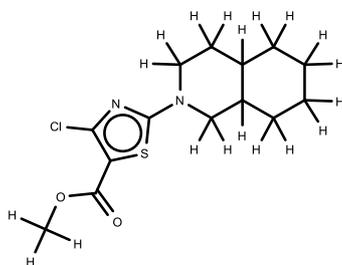
252



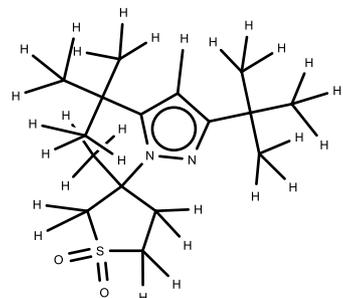
253



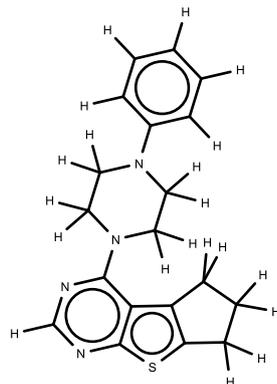
254



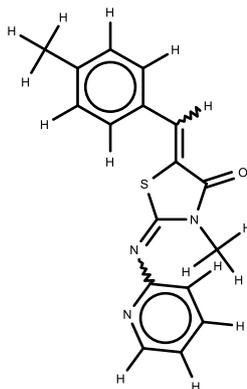
255



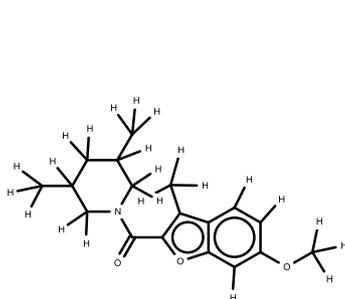
256



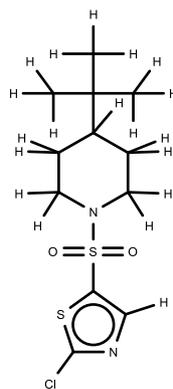
257



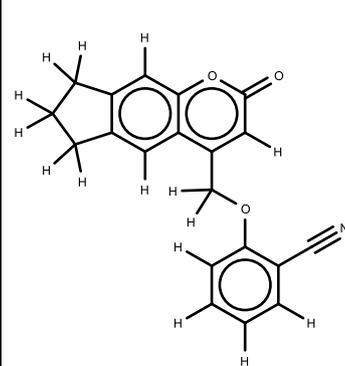
258



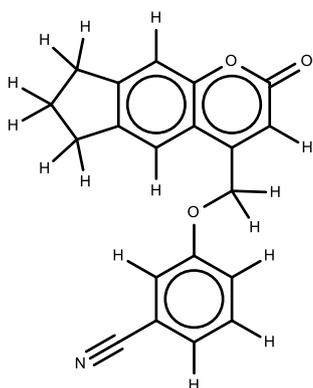
259



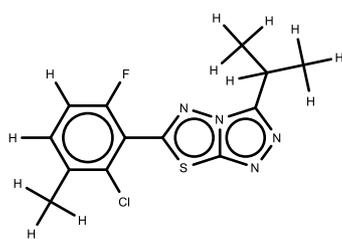
260



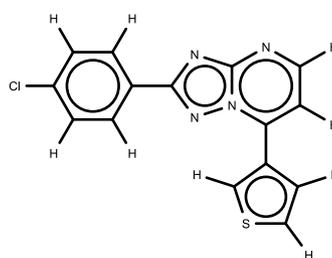
261



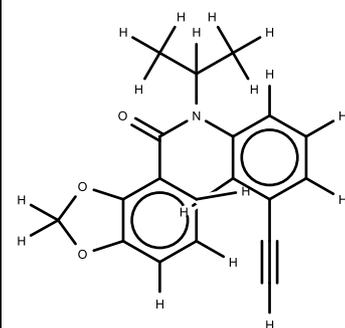
262



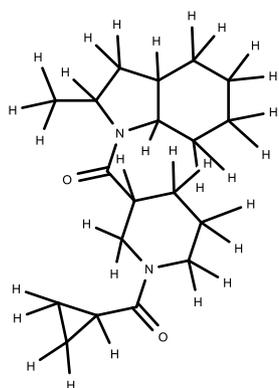
263



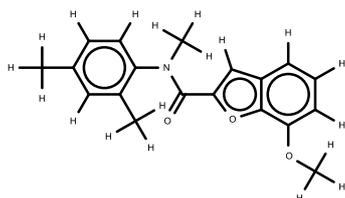
264



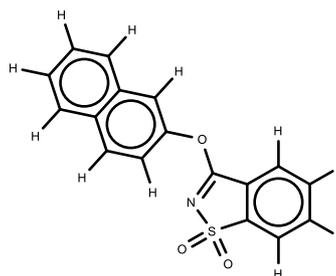
265



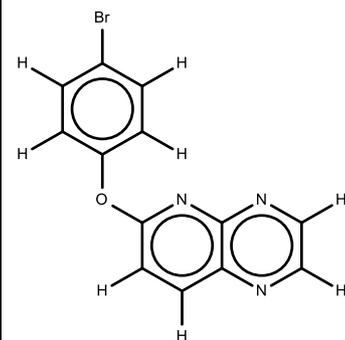
266



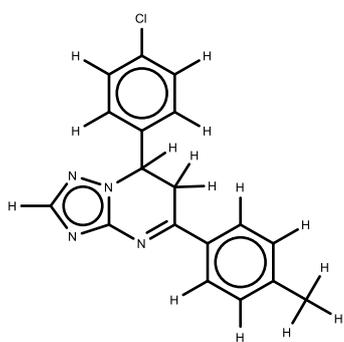
267



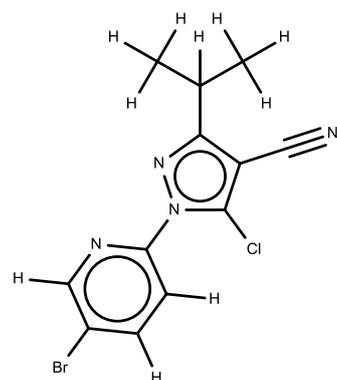
268



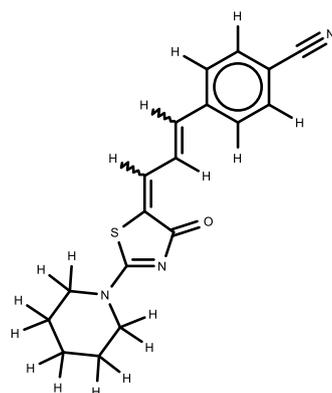
269



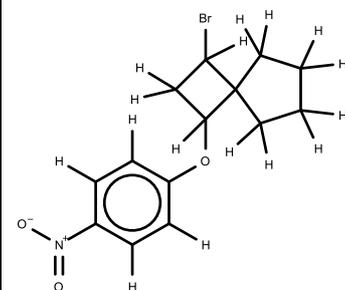
270



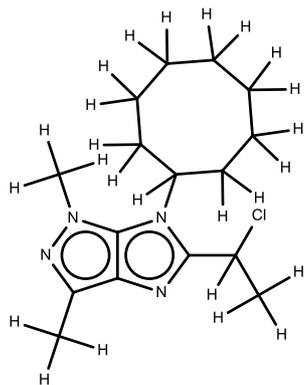
271



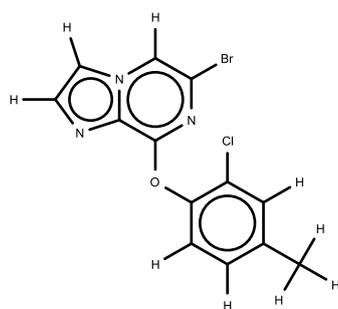
272



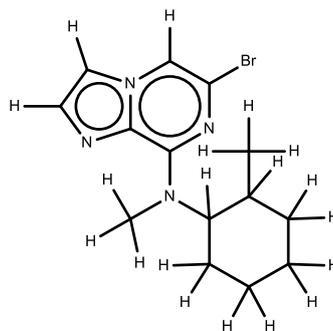
273



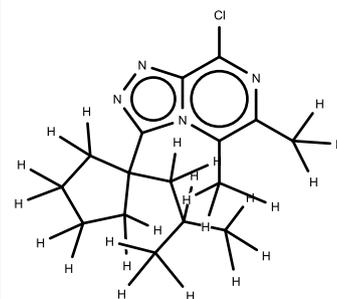
274



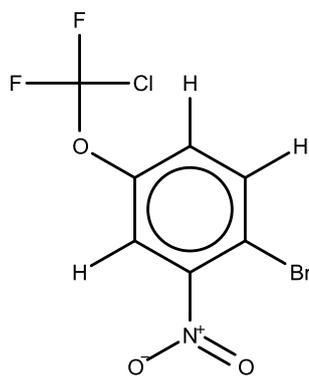
275



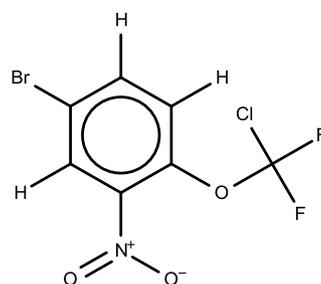
276



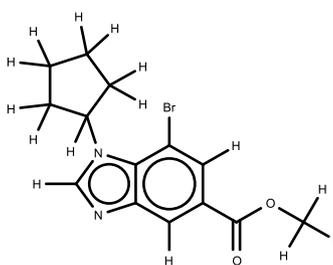
277



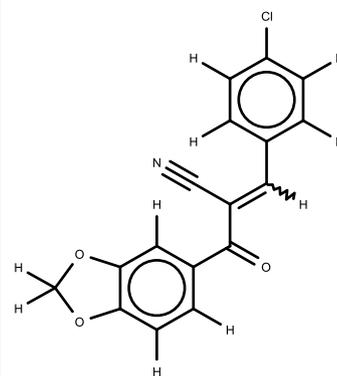
278



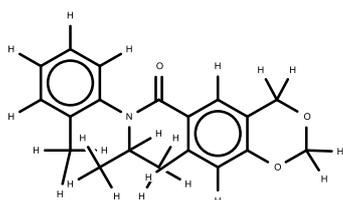
279



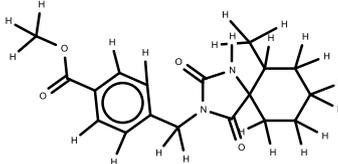
280



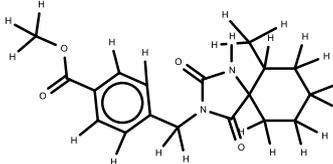
281



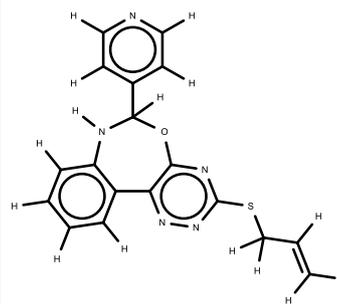
282



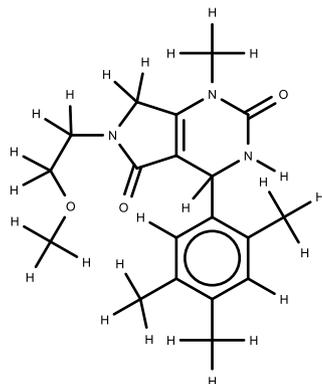
283



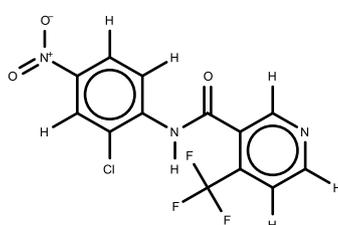
284



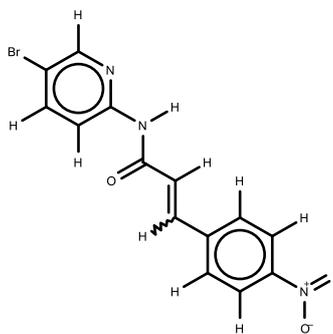
285



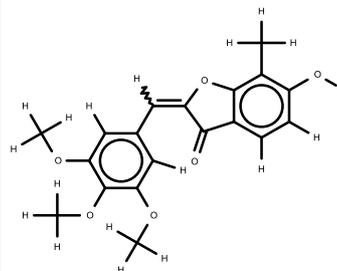
286



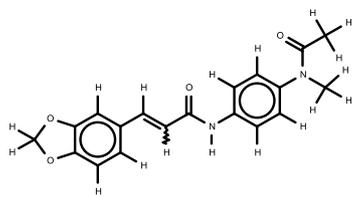
287



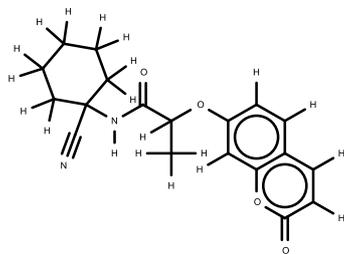
288



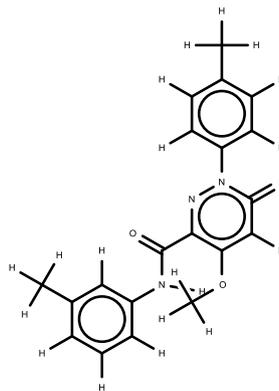
289



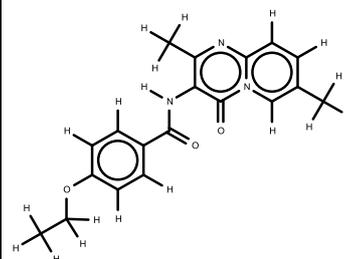
290



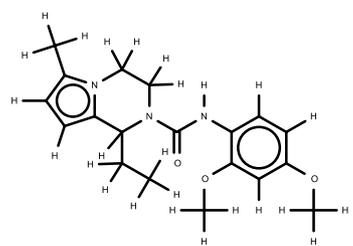
291



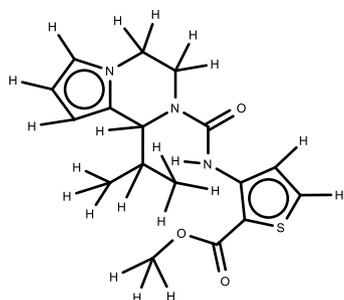
292



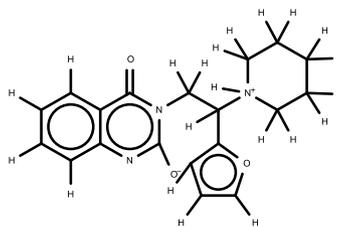
293



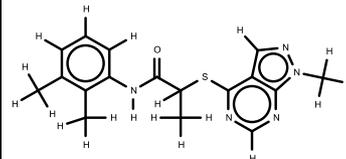
294



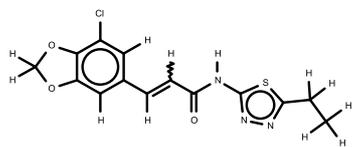
295



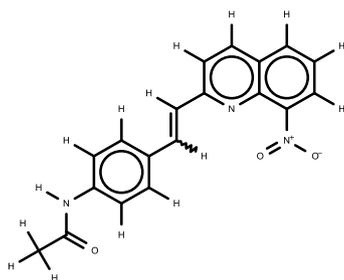
296



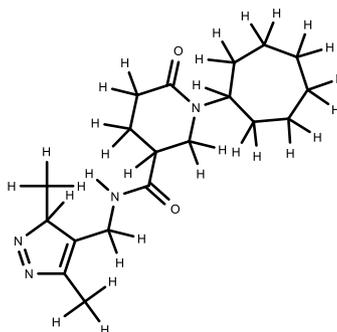
297



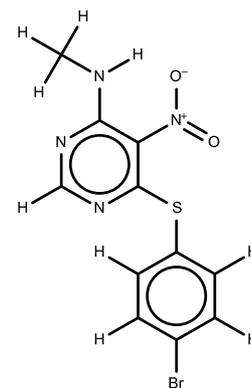
298



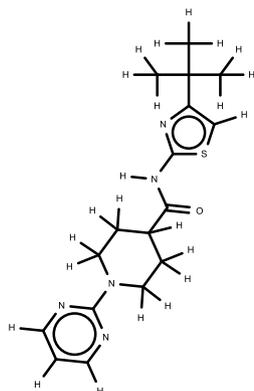
299



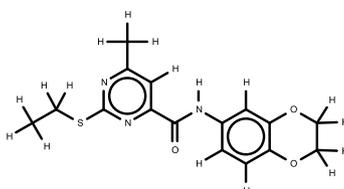
300



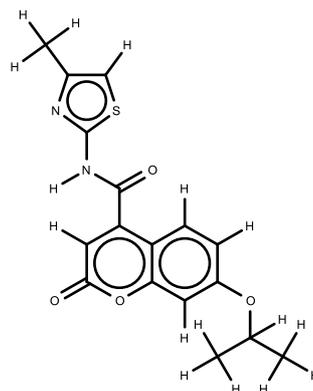
301



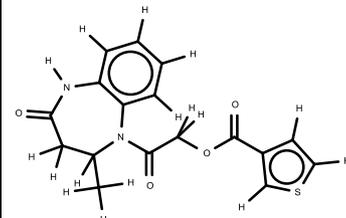
302



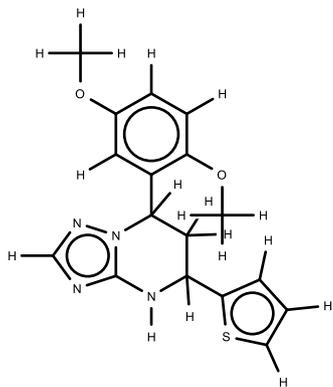
303



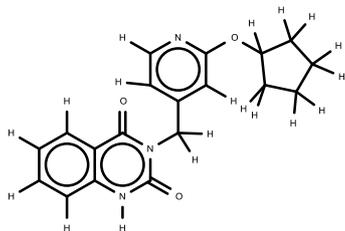
304



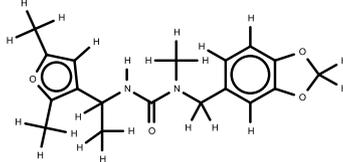
305



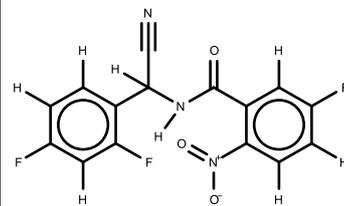
306



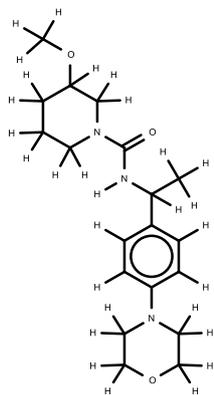
307



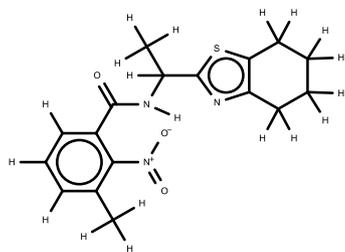
308



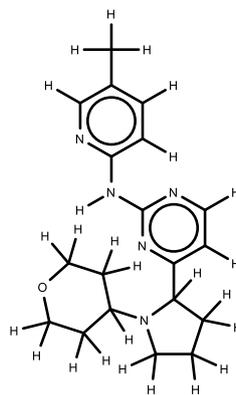
309



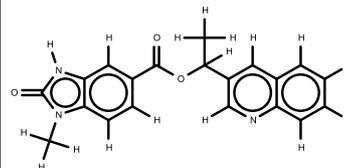
310



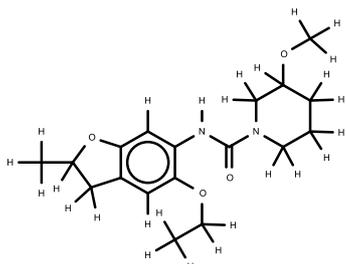
311



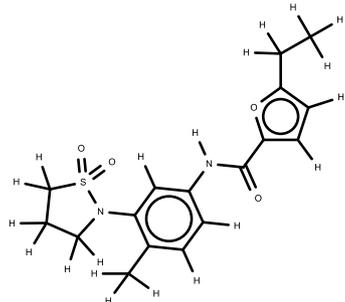
312



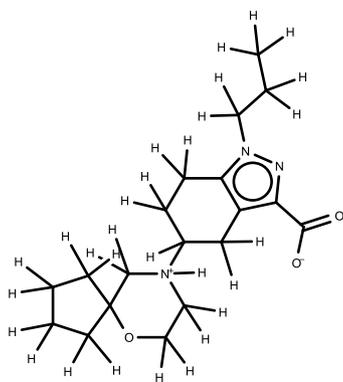
313



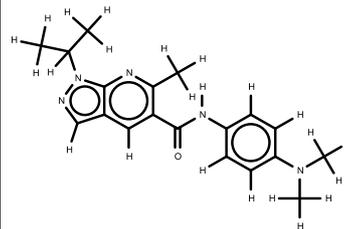
314



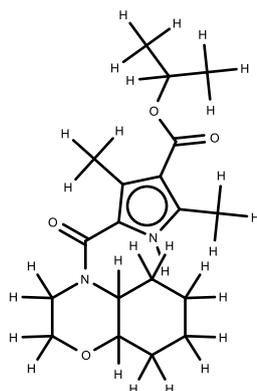
315



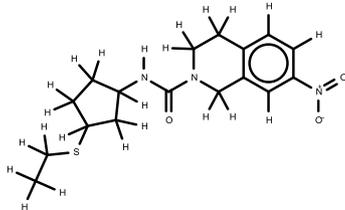
316



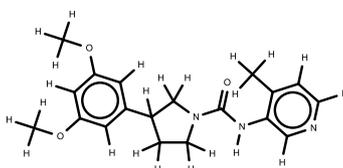
317



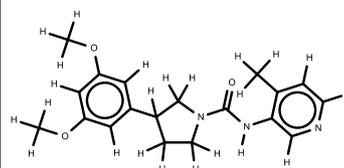
318



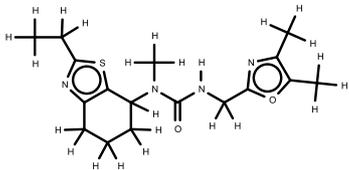
319



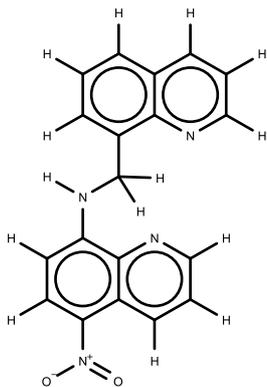
320



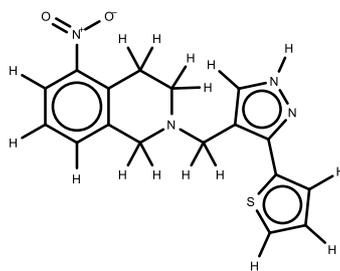
321



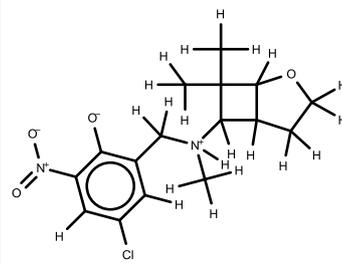
322



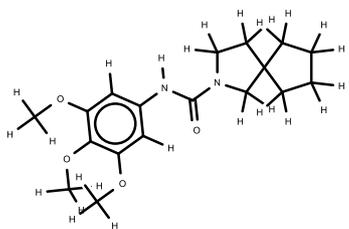
323



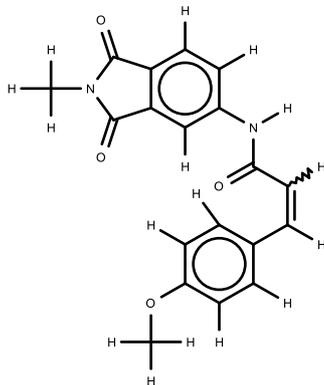
324



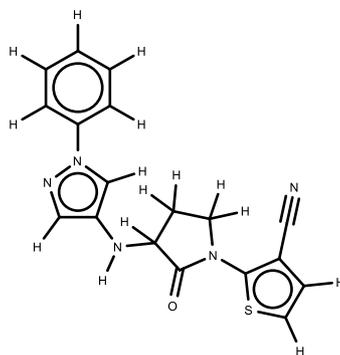
325



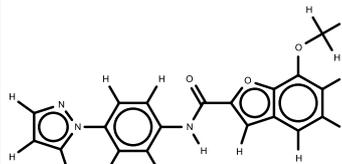
326



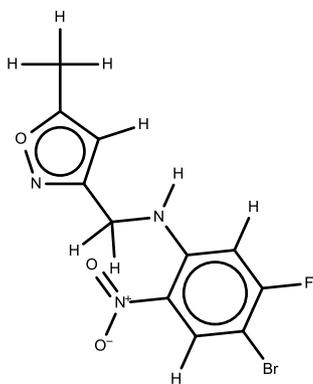
327



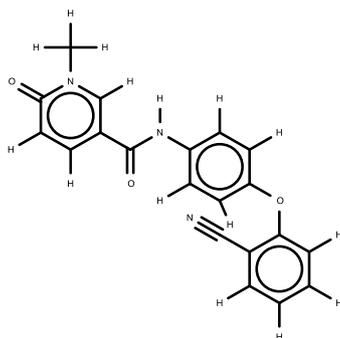
328



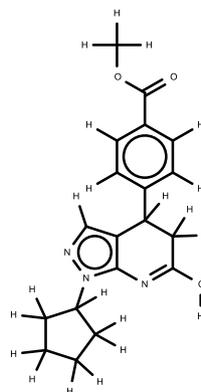
329



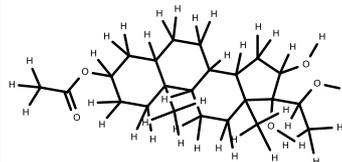
330



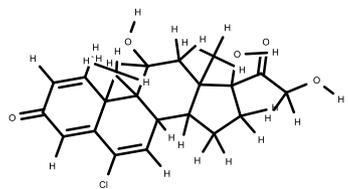
331



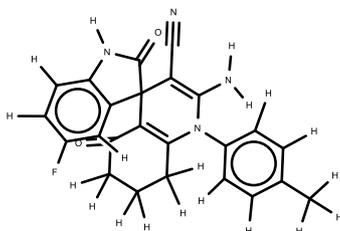
332



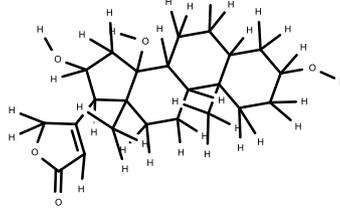
333



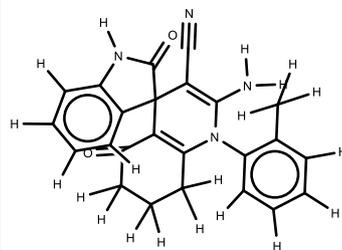
334



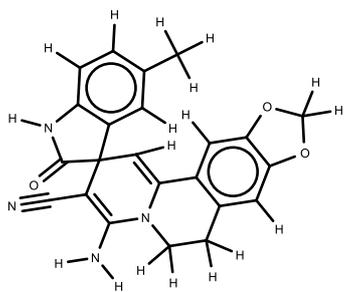
335



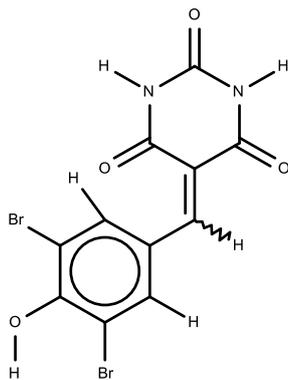
336



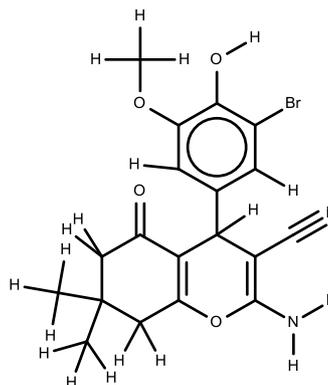
337



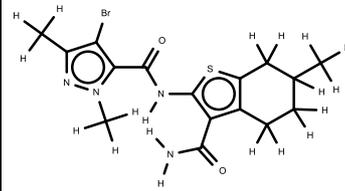
338



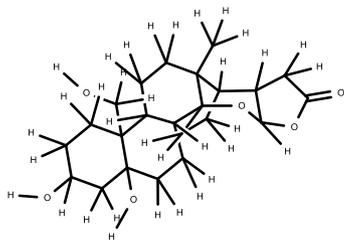
339



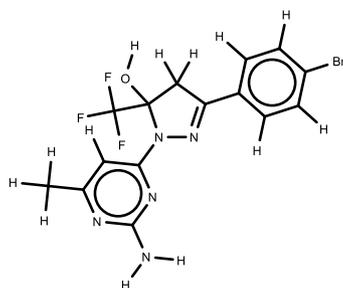
340



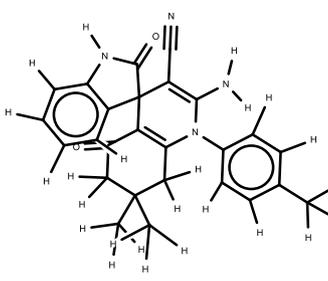
341



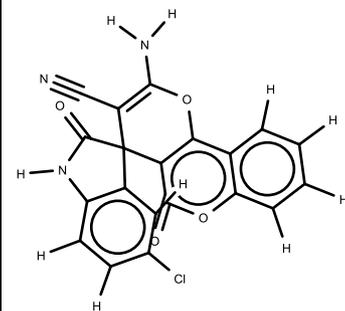
342



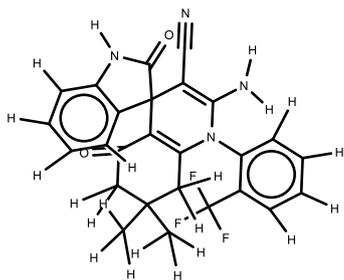
343



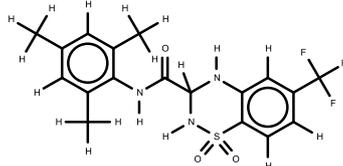
344



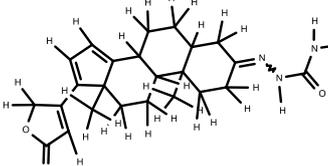
345



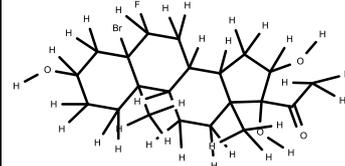
346



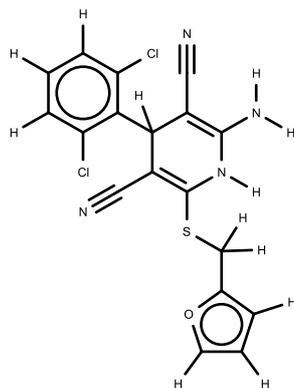
347



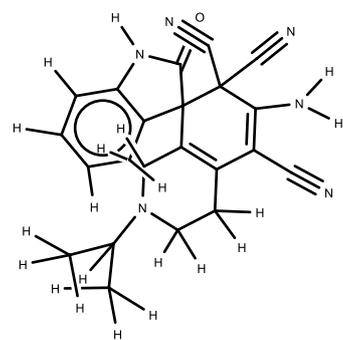
348



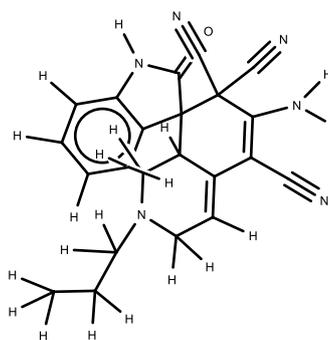
349



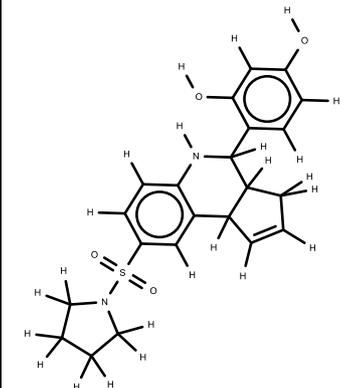
350



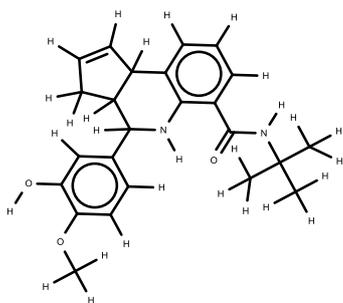
351



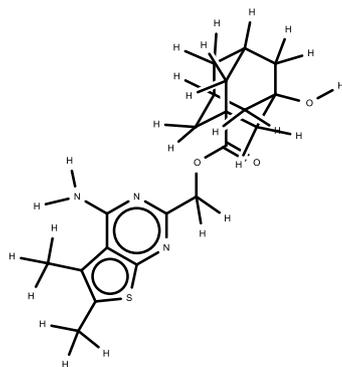
352



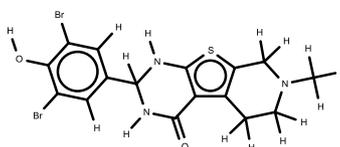
353



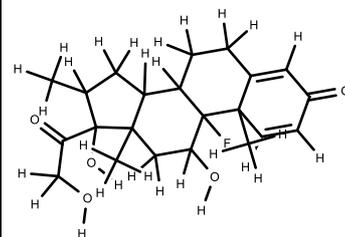
354



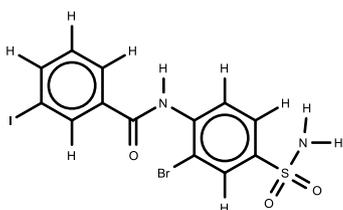
355



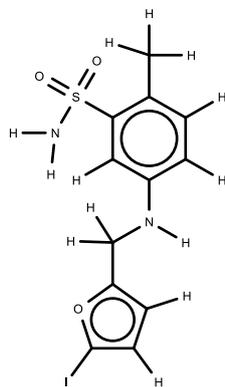
356



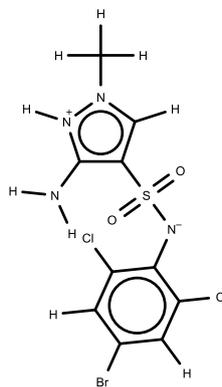
357



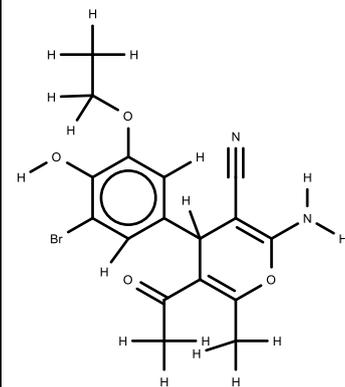
358



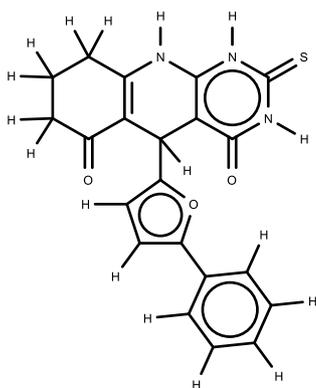
359



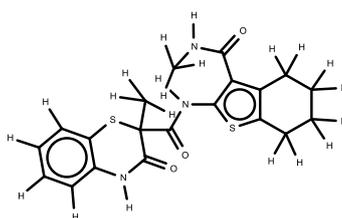
360



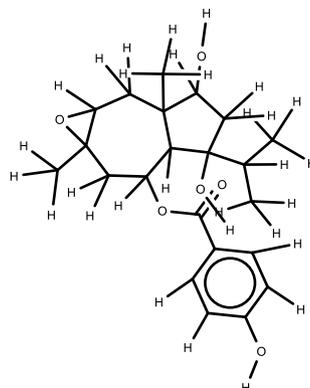
361



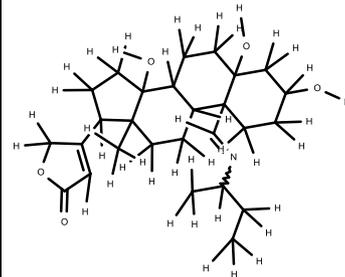
362



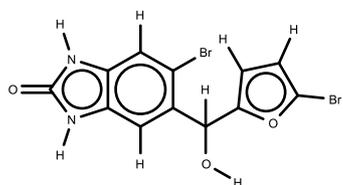
363



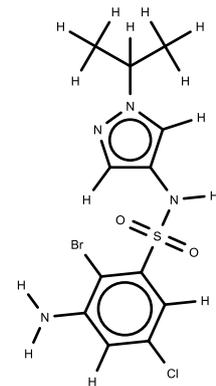
364



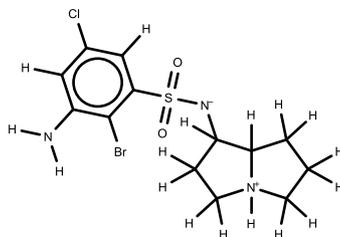
365



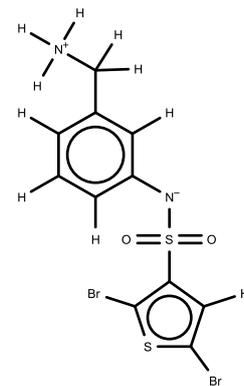
366



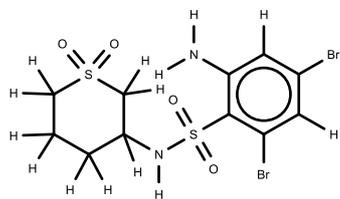
367



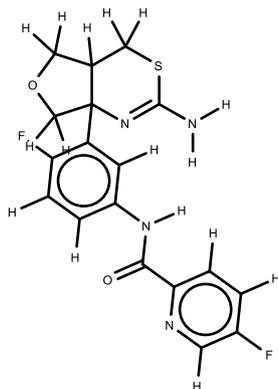
368



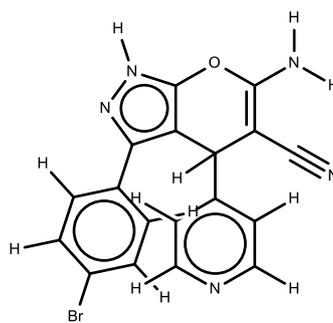
369



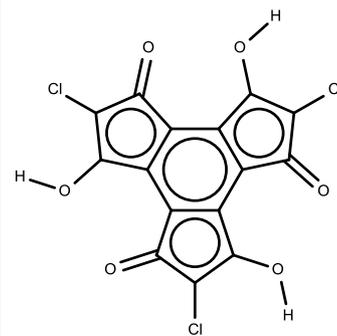
370



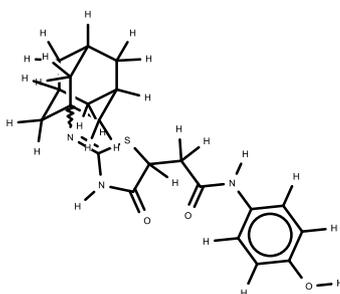
371



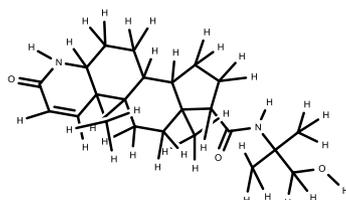
372



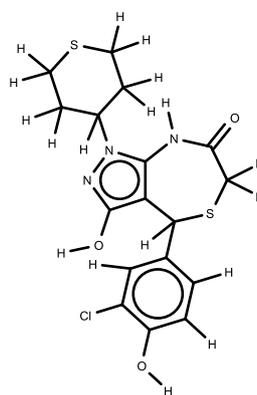
373



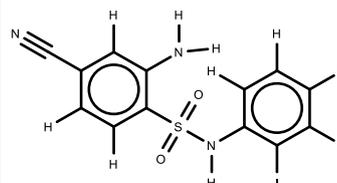
374



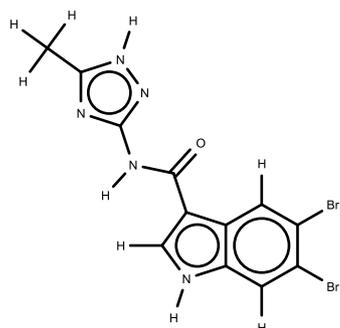
375



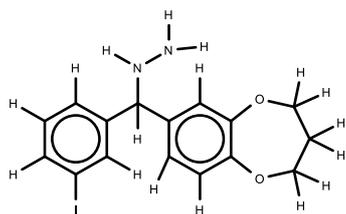
376



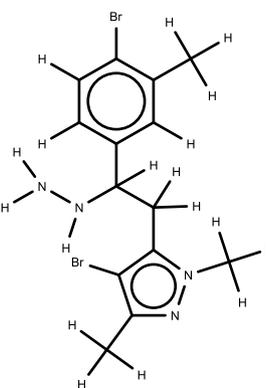
377



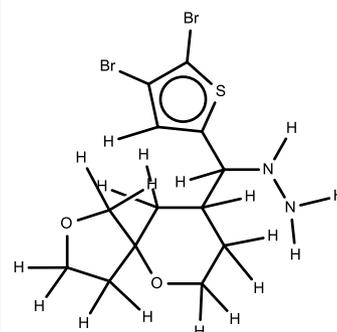
378



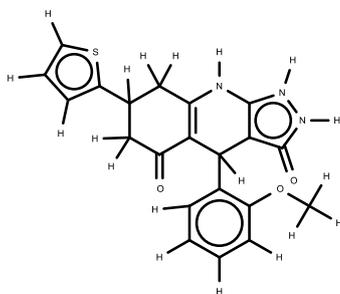
379



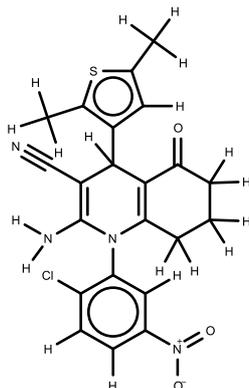
380



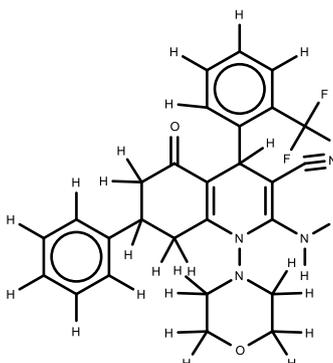
381



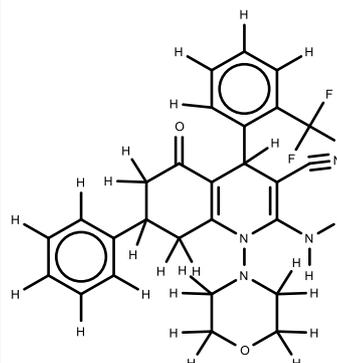
382



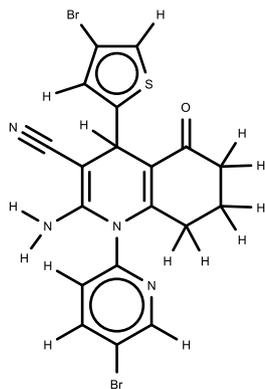
383



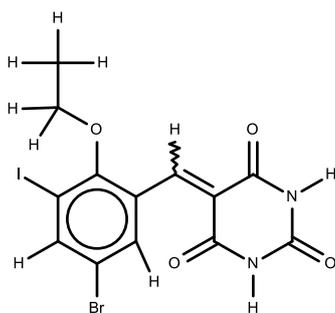
384



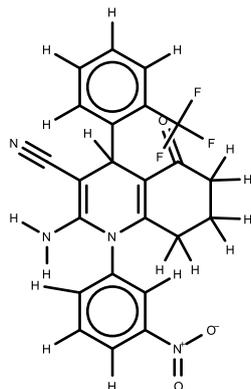
385



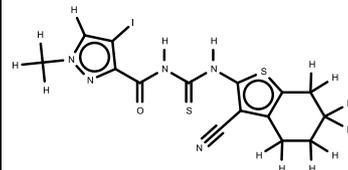
386



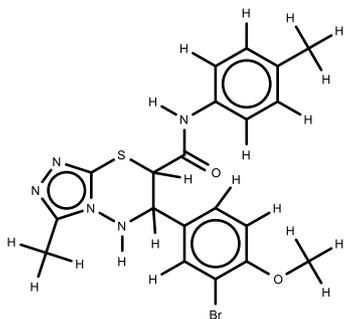
387



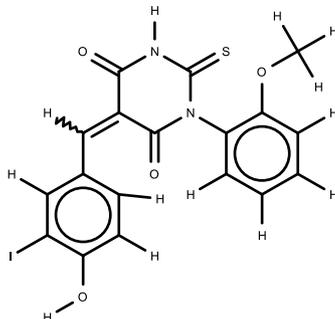
388



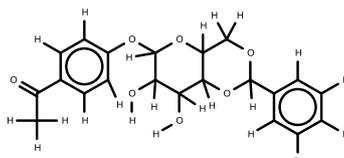
389



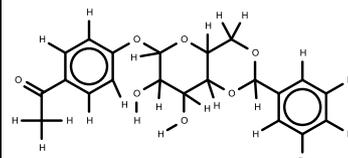
390



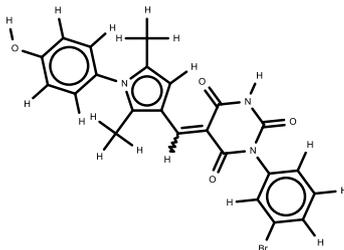
391



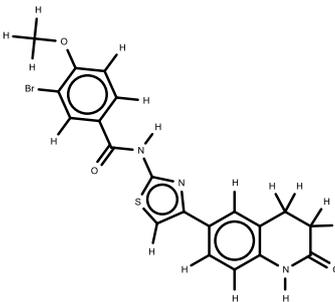
392



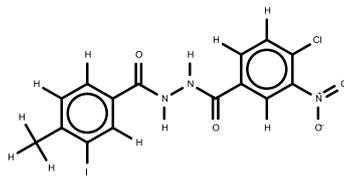
393



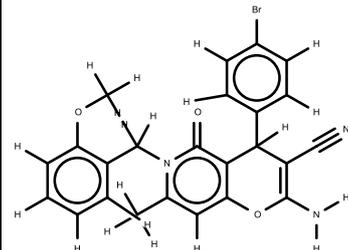
394



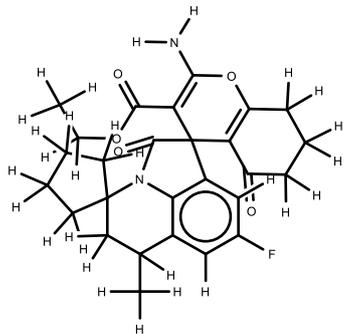
395



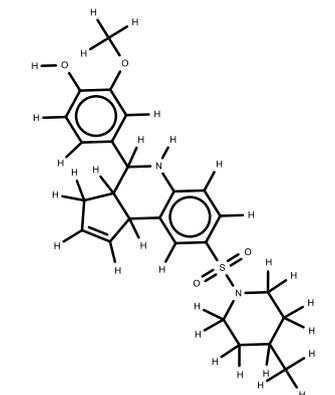
396



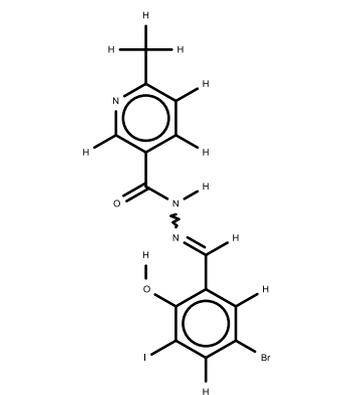
397



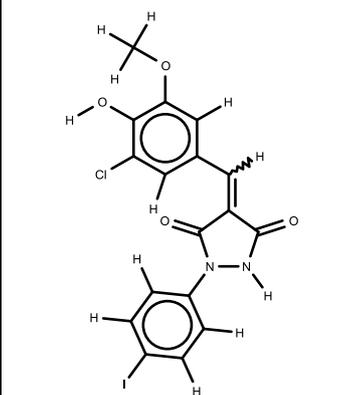
398



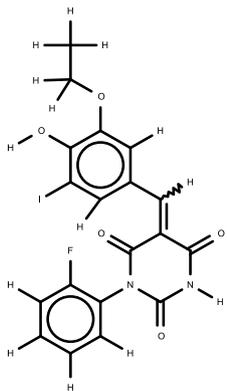
399



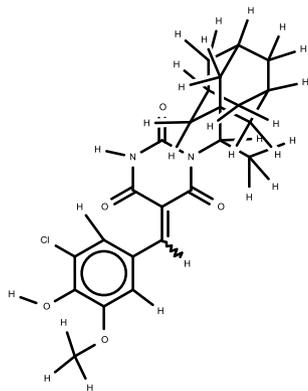
400



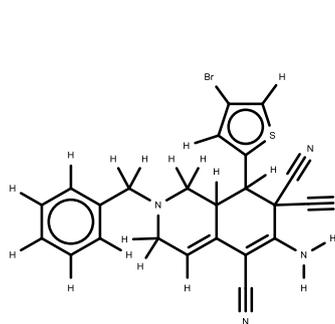
401



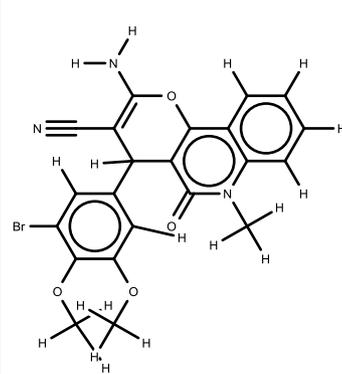
402



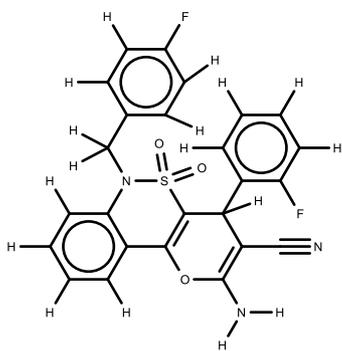
403



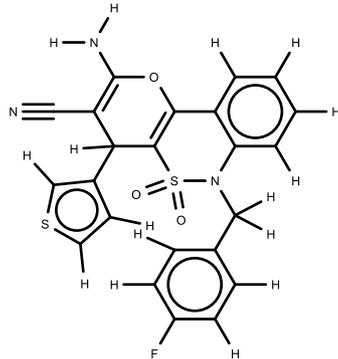
404



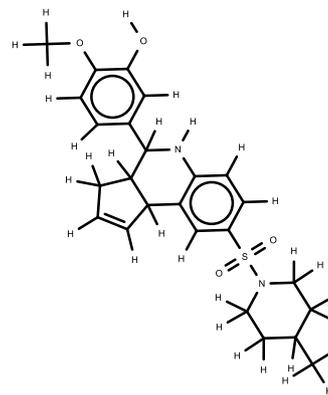
405



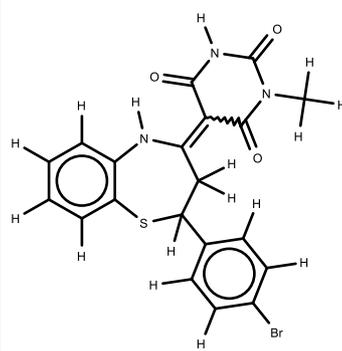
406



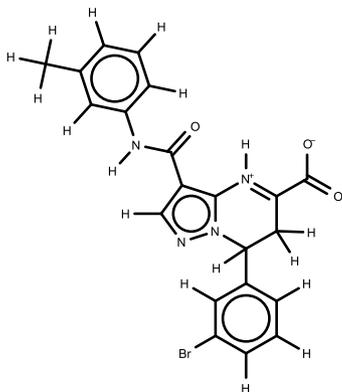
407



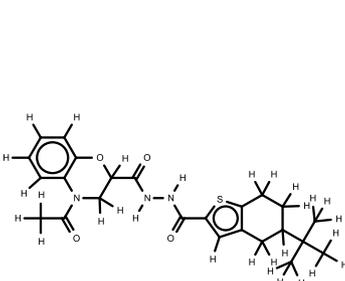
408



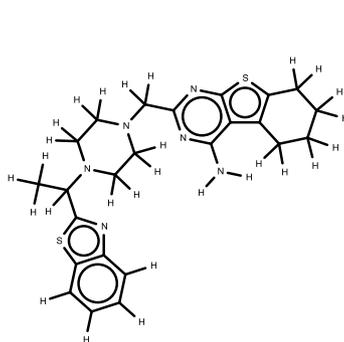
409



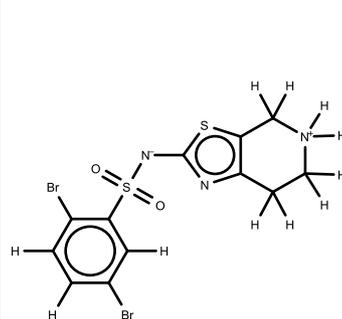
410



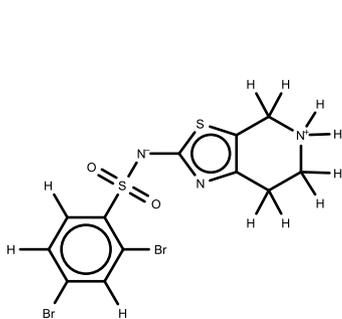
411



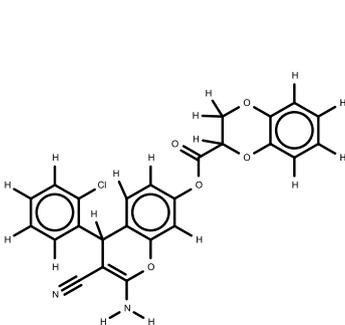
412



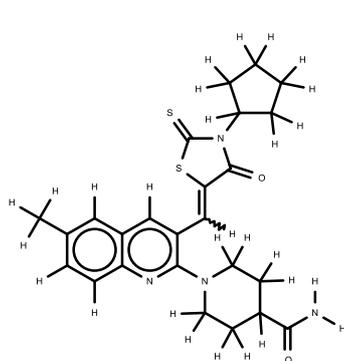
413



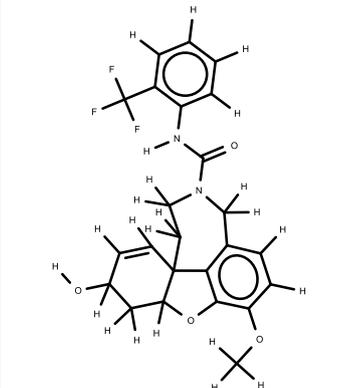
414



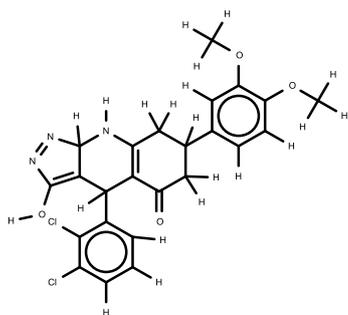
415



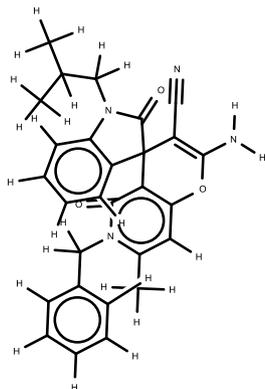
416



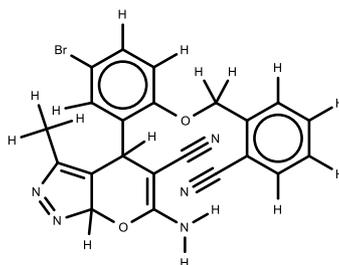
417



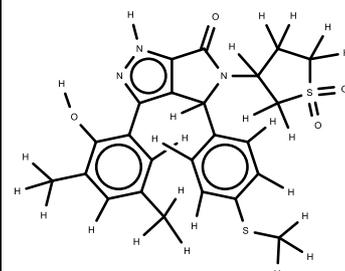
418



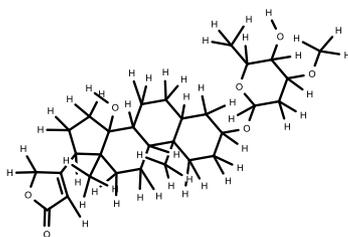
419



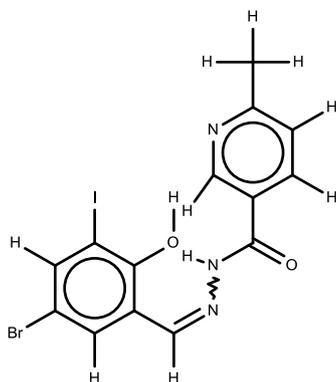
420



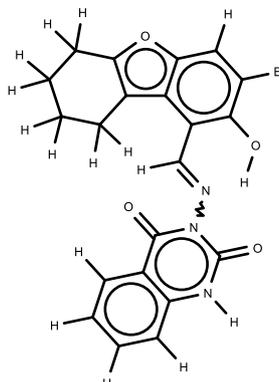
421



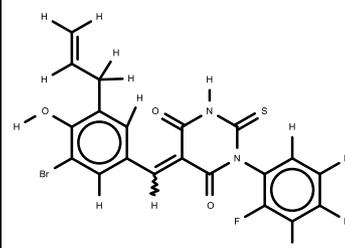
422



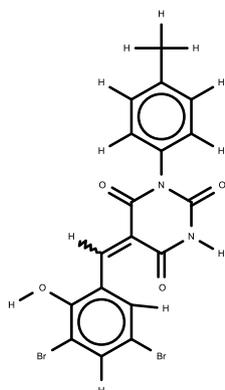
423



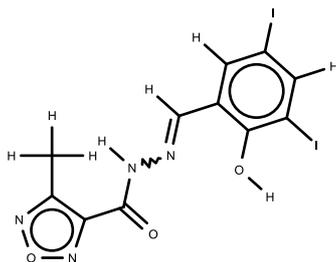
424



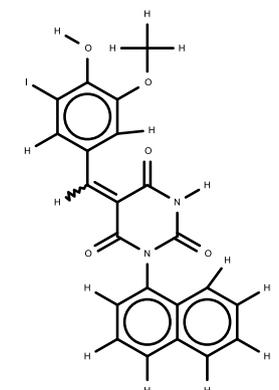
425



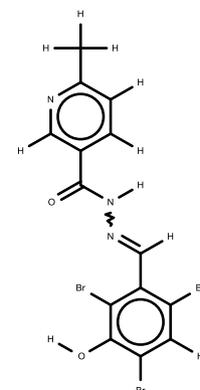
426



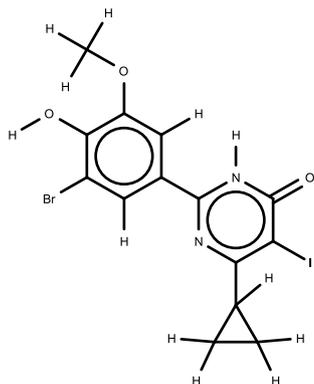
427



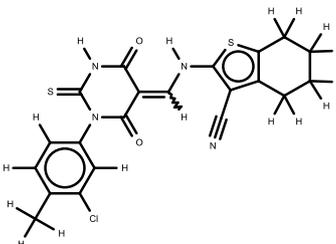
428



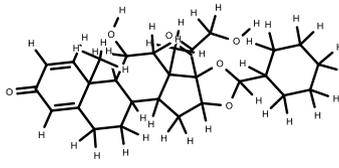
429



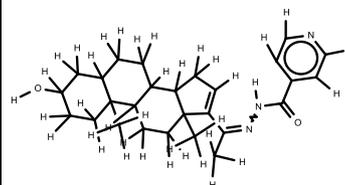
430



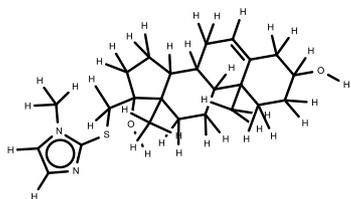
431



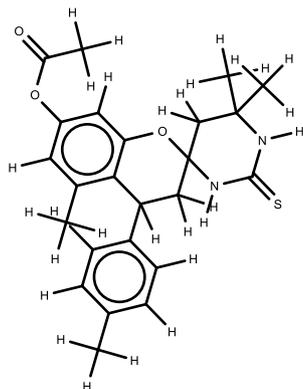
432



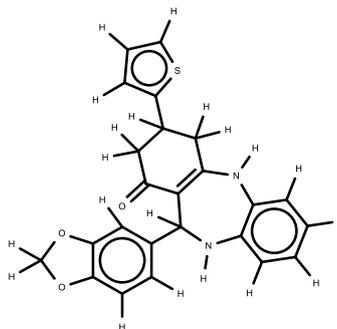
433



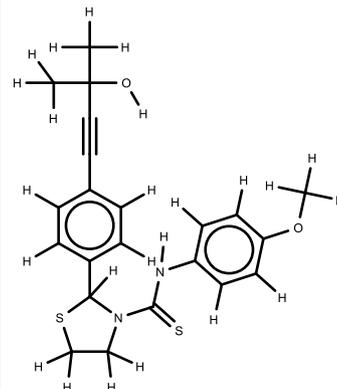
434



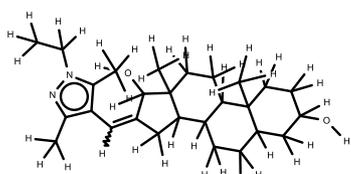
435



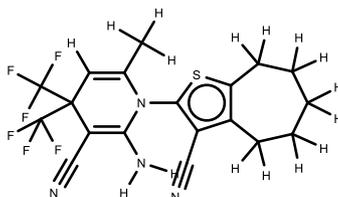
436



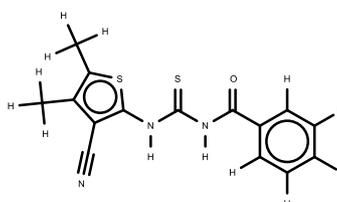
437



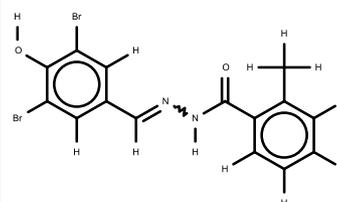
438



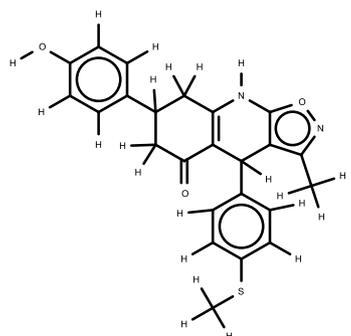
439



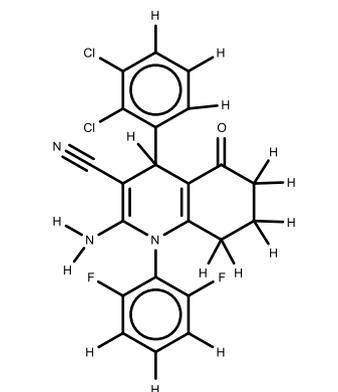
440



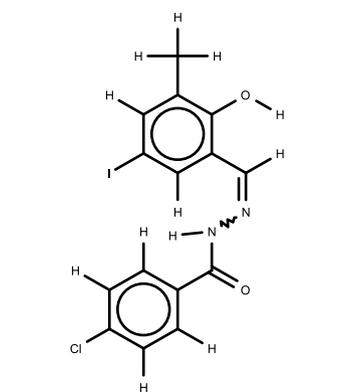
441



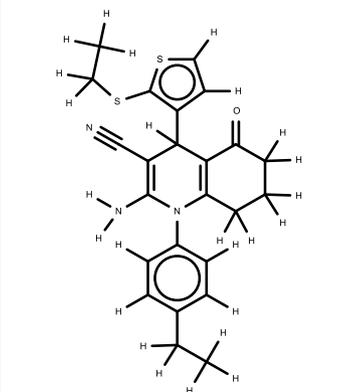
442



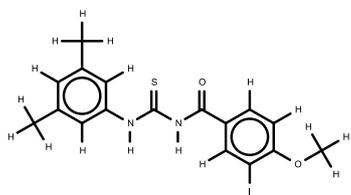
443



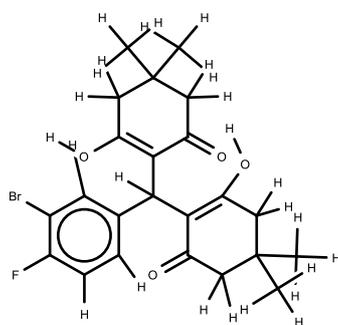
444



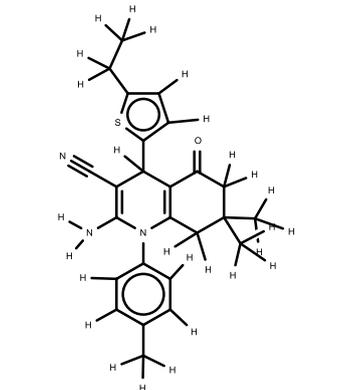
445



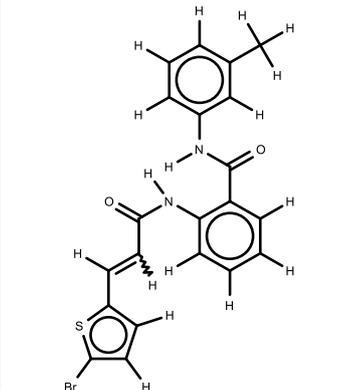
446



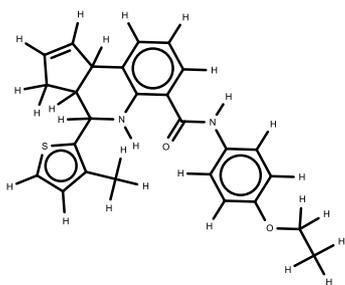
447



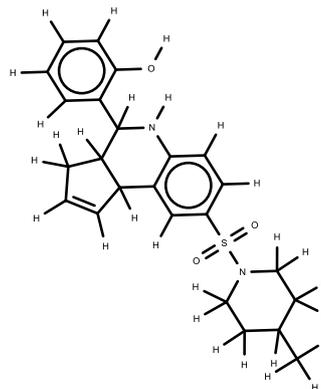
448



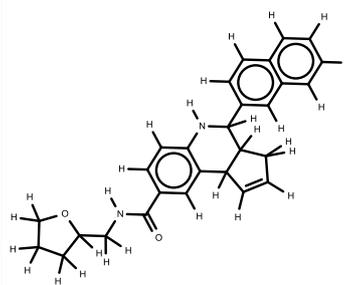
449



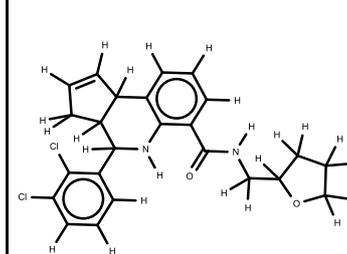
450



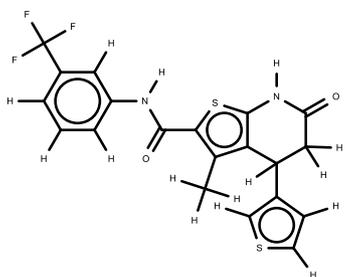
451



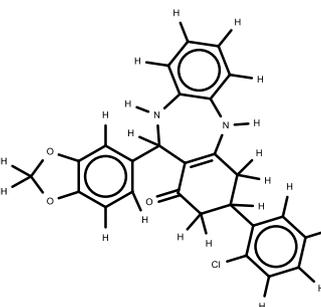
452



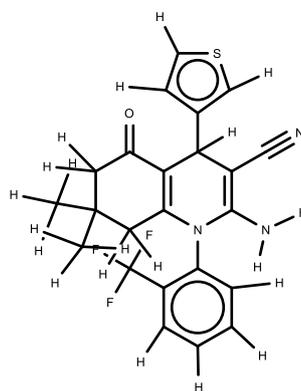
453



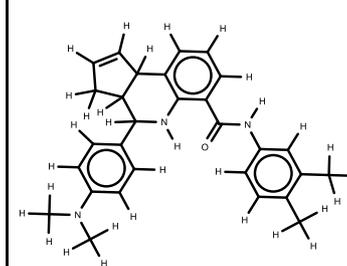
454



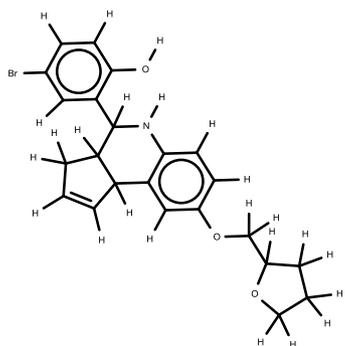
455



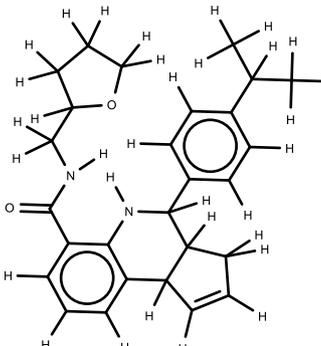
456



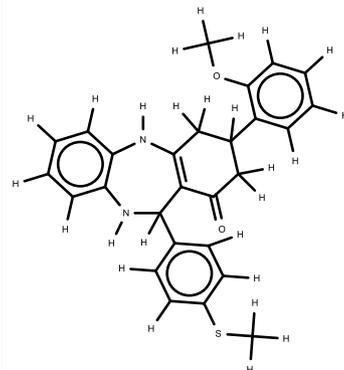
457



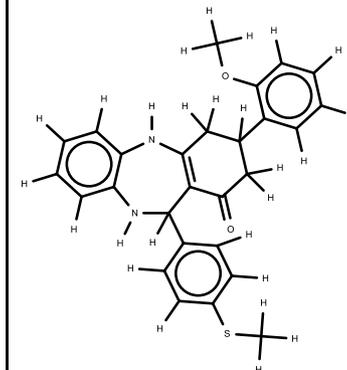
458



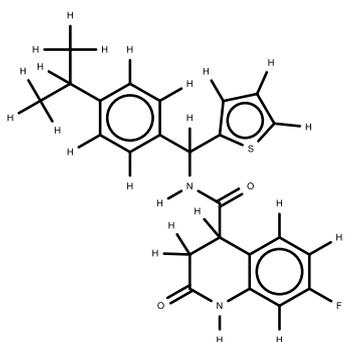
459



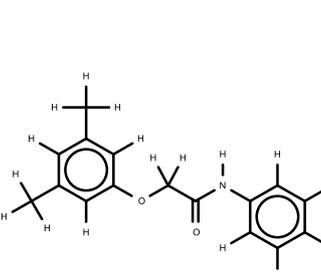
460



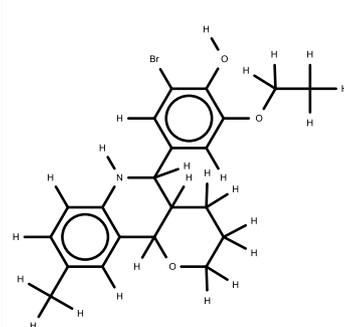
461



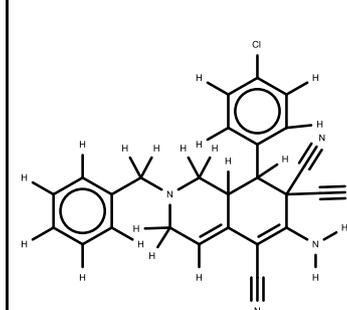
462



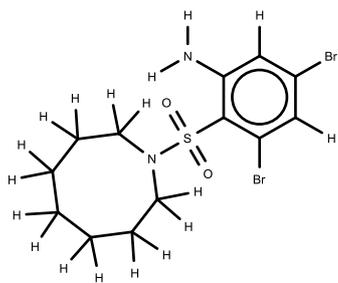
463



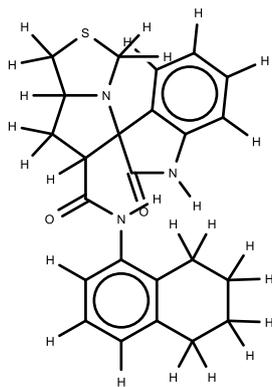
464



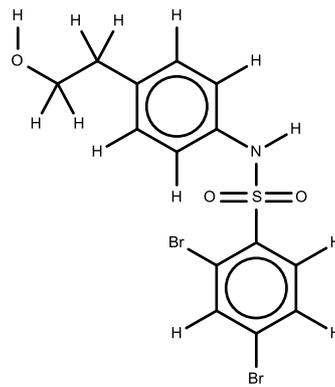
465



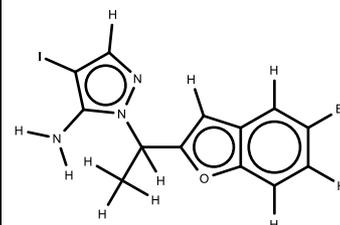
466



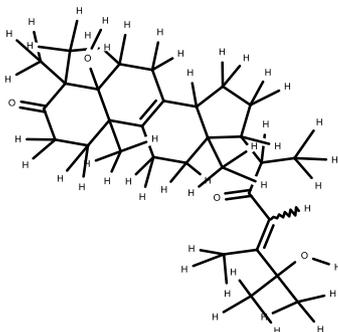
467



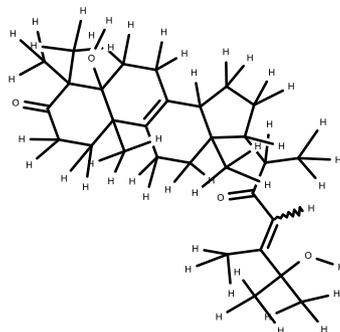
468



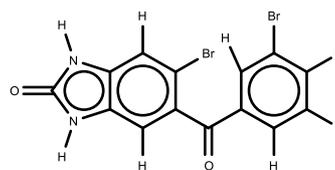
469



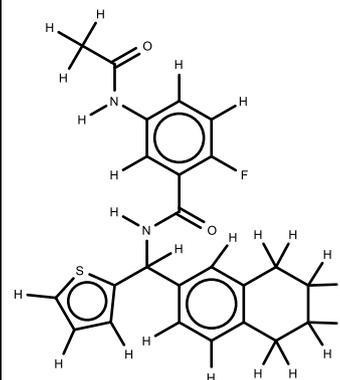
470



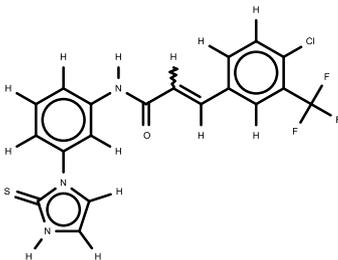
471



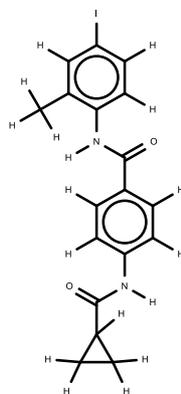
472



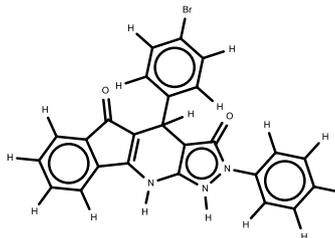
473



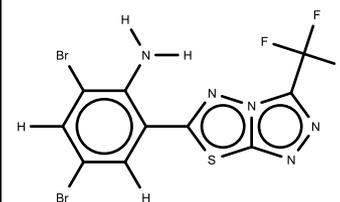
474



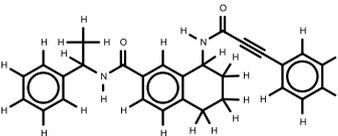
475



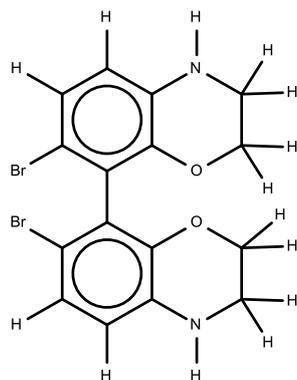
476



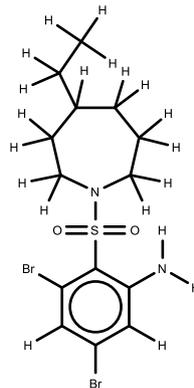
477



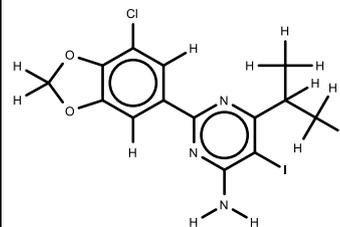
478



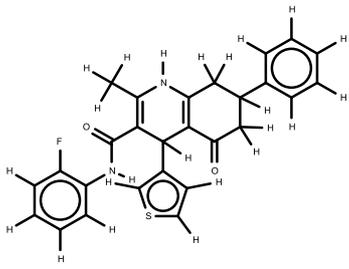
479



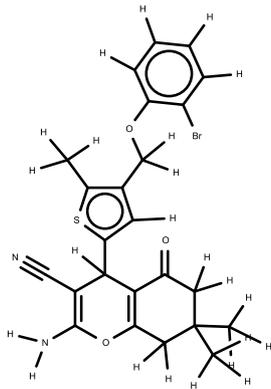
480



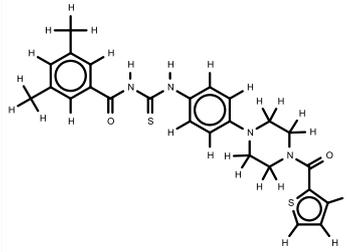
481



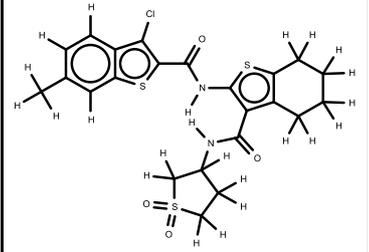
482



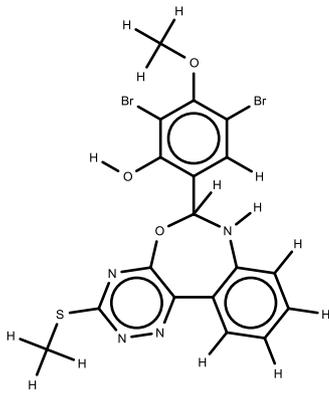
483



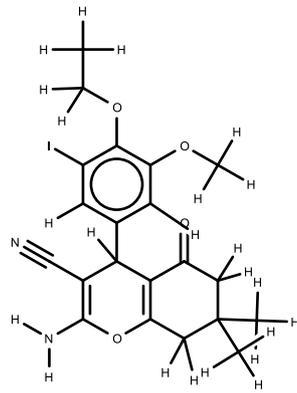
484



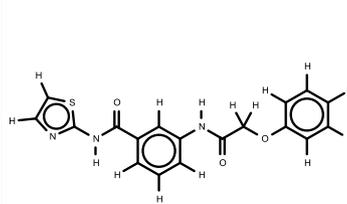
485



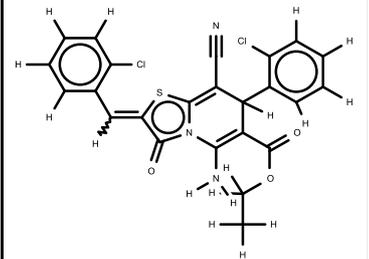
486



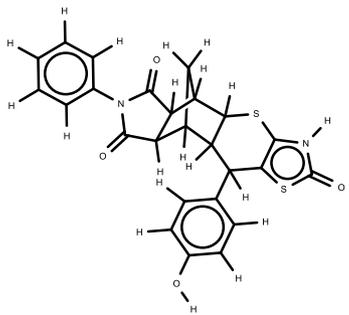
487



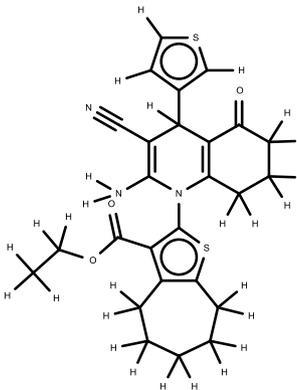
488



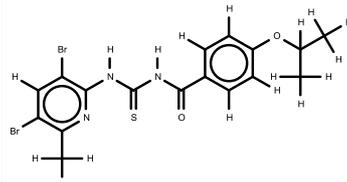
489



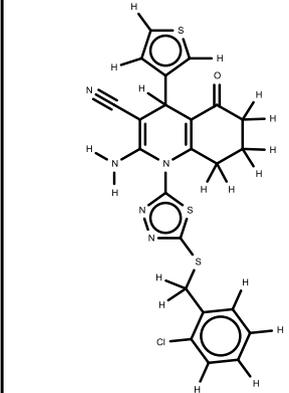
490



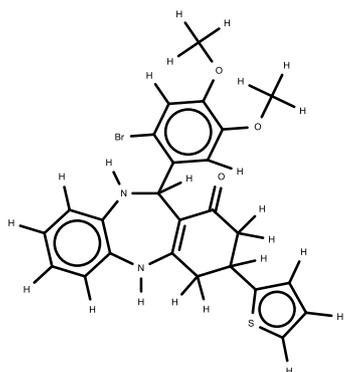
491



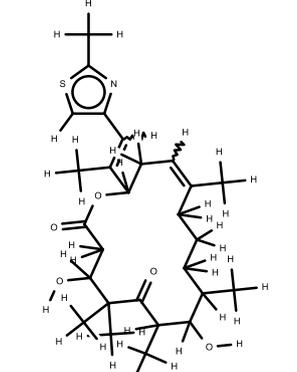
492



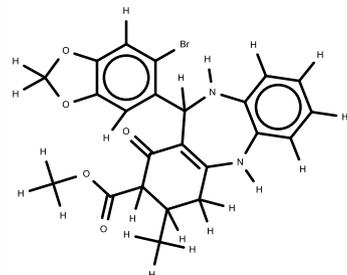
493



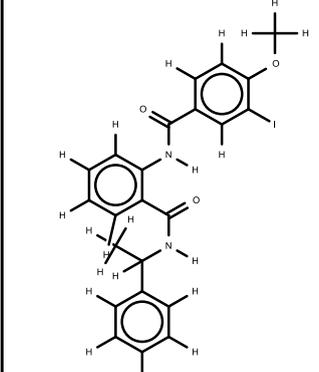
494



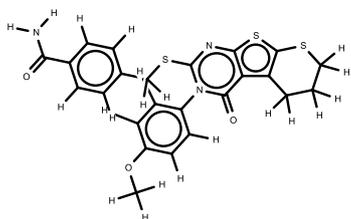
495



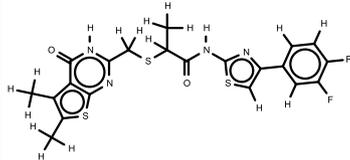
496



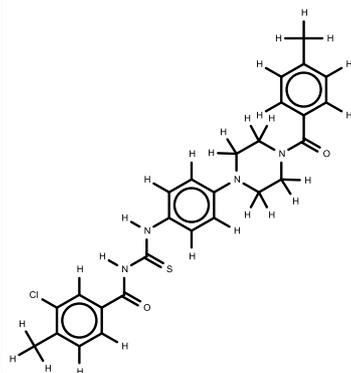
497



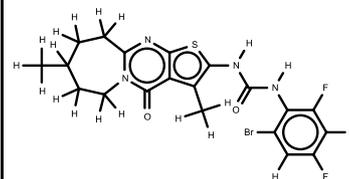
498



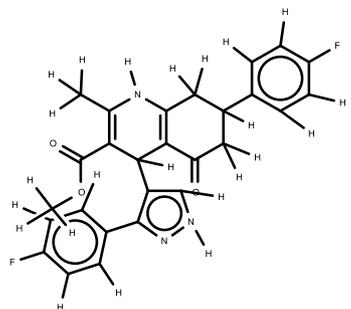
499



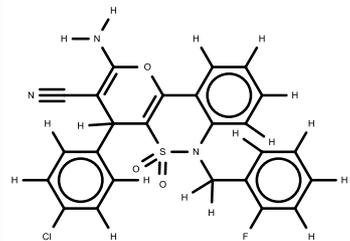
500



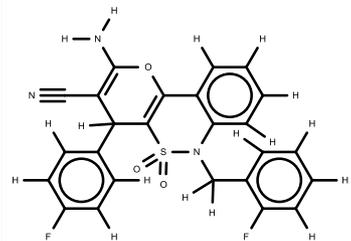
501



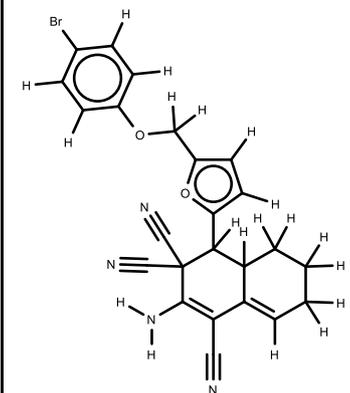
502



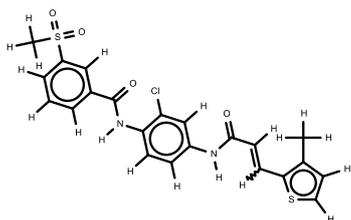
503



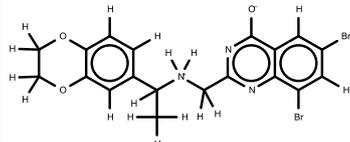
504



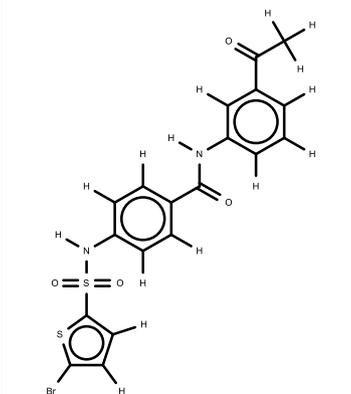
505



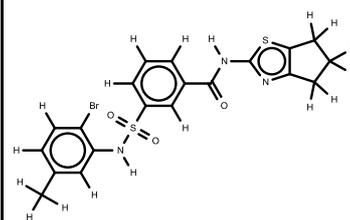
506



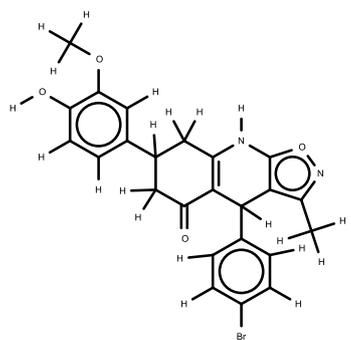
507



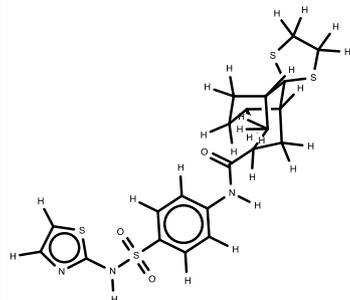
508



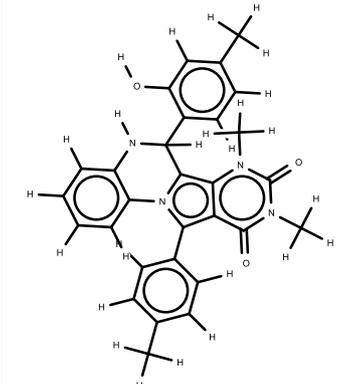
509



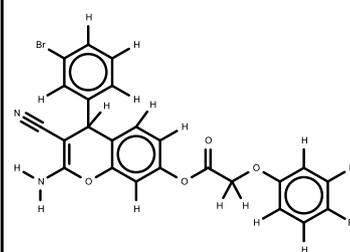
510



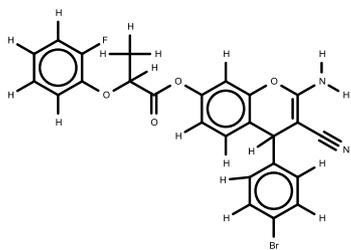
511



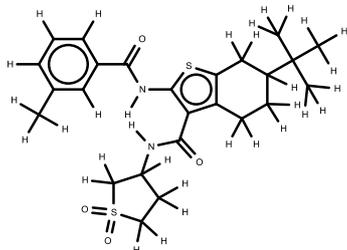
512



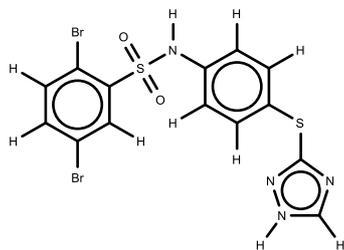
513



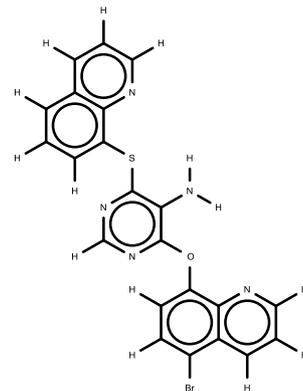
514



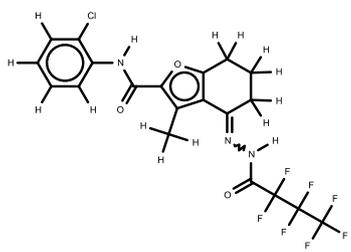
515



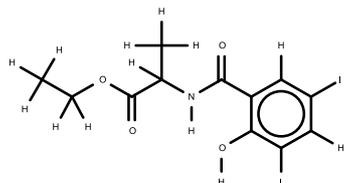
516



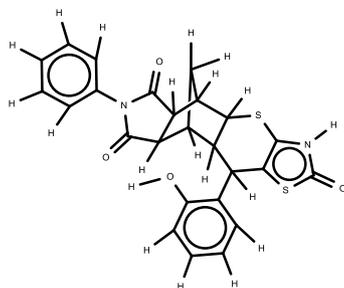
517



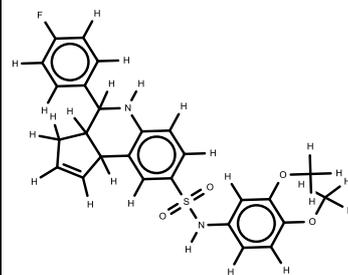
518



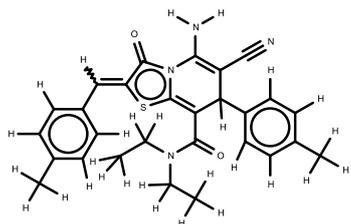
519



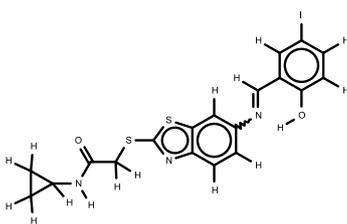
520



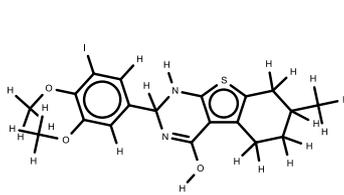
521



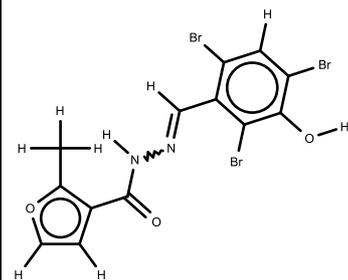
522



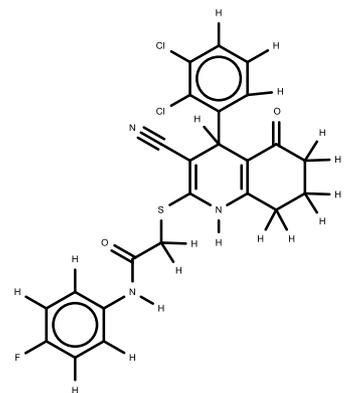
523



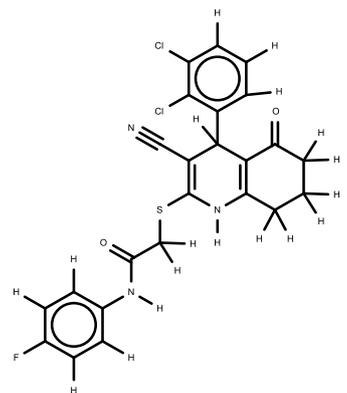
524



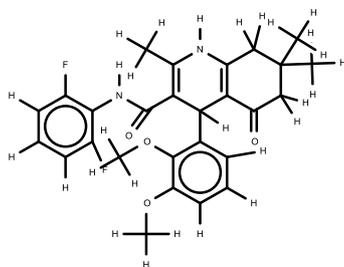
525



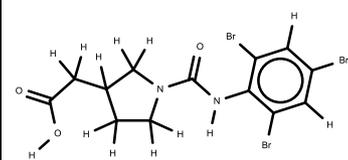
526



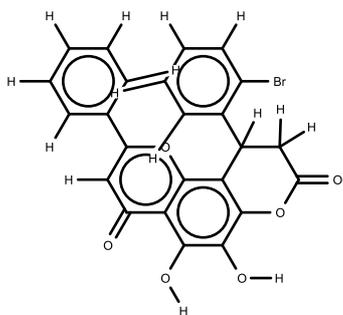
527



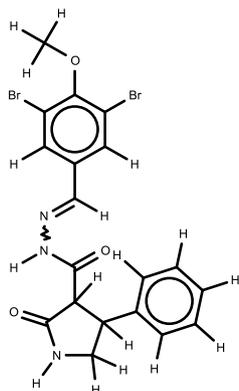
528



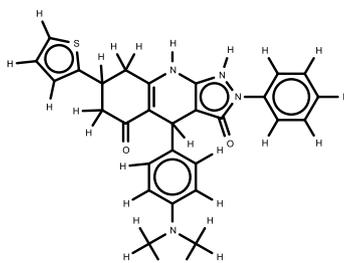
529



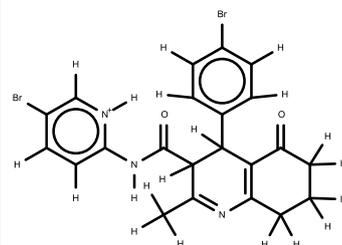
530



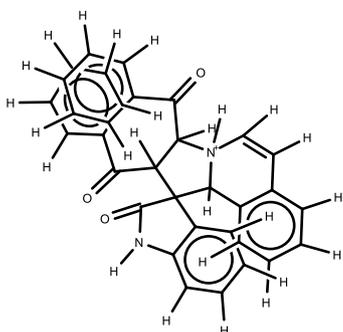
531



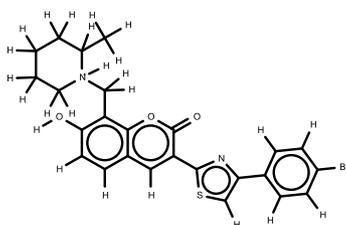
532



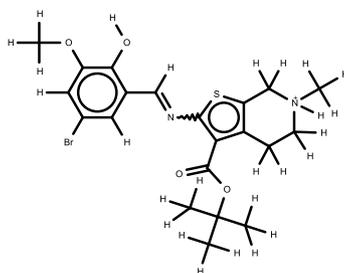
533



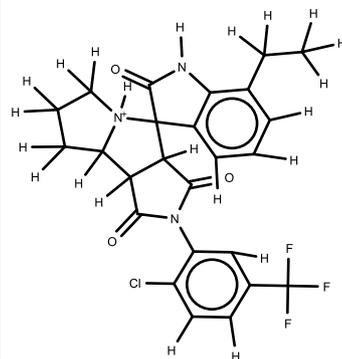
534



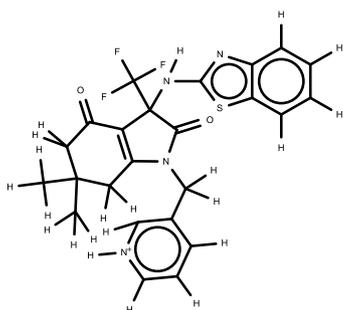
535



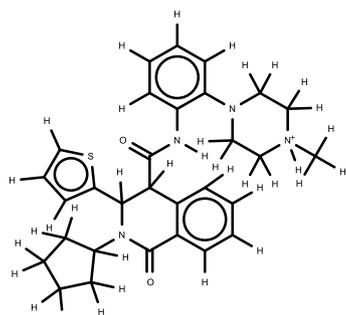
536



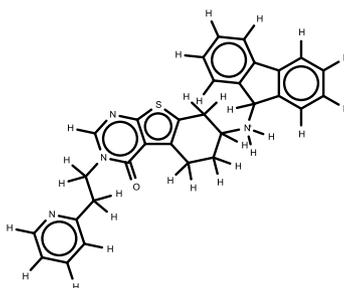
537



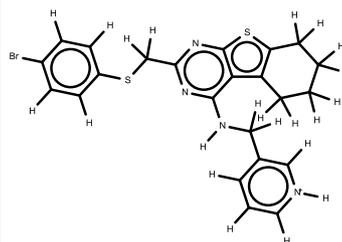
538



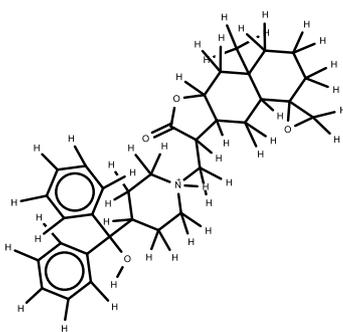
539



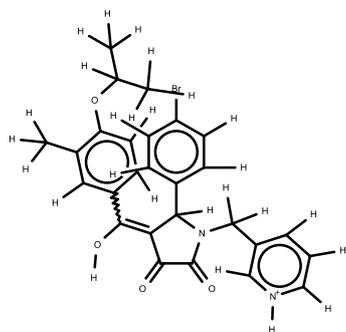
540



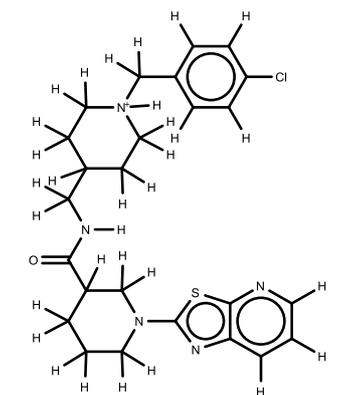
541



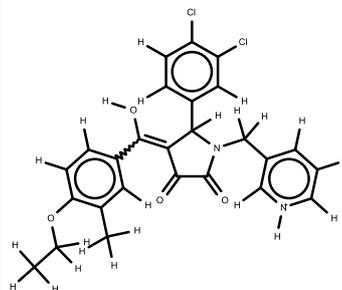
542



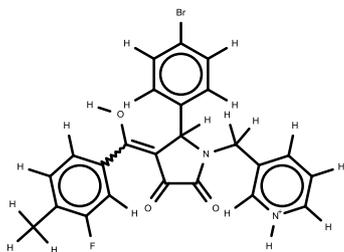
543



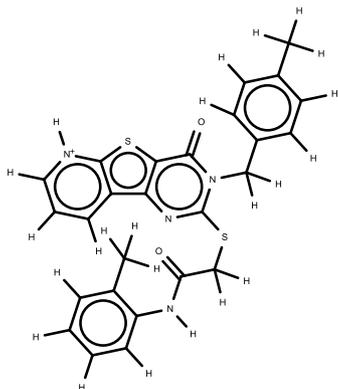
544



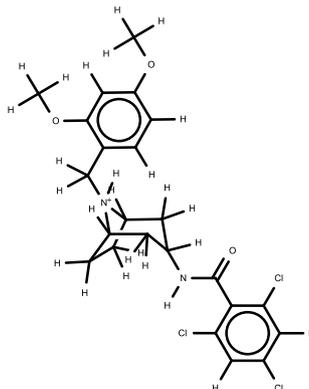
545



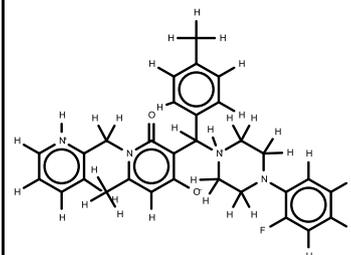
546



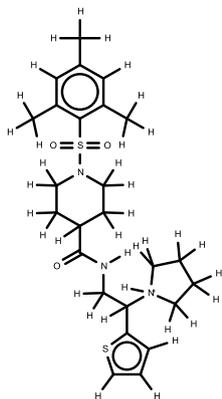
547



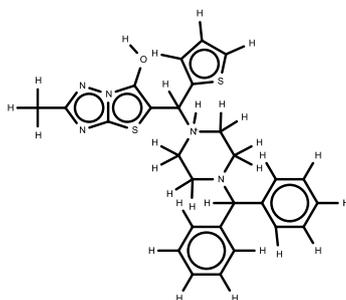
548



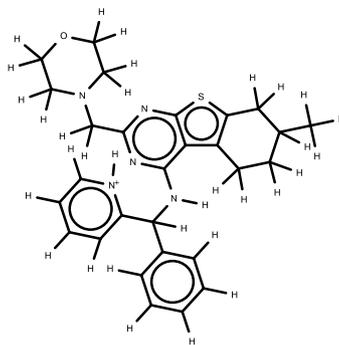
549



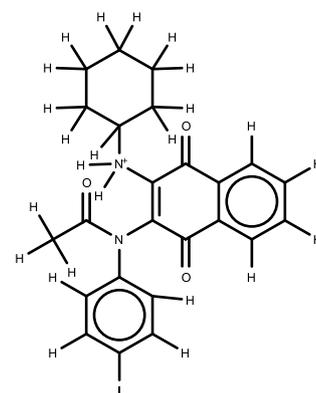
550



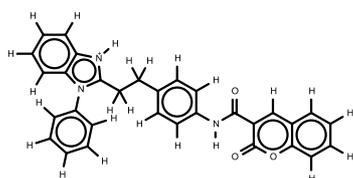
551



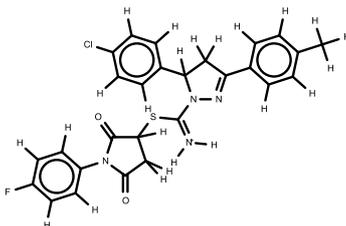
552



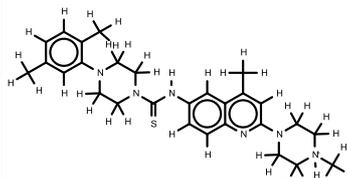
553



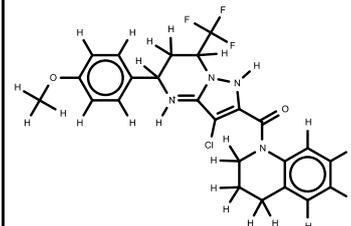
554



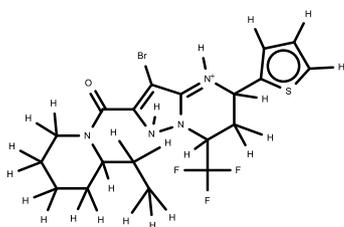
555



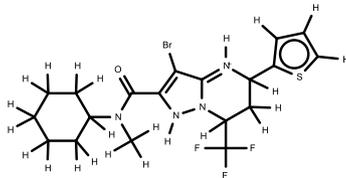
556



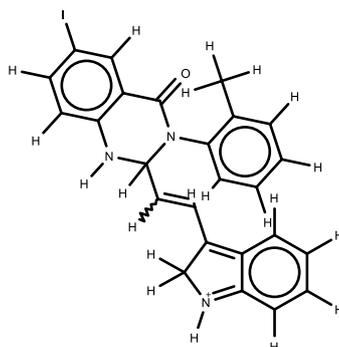
557



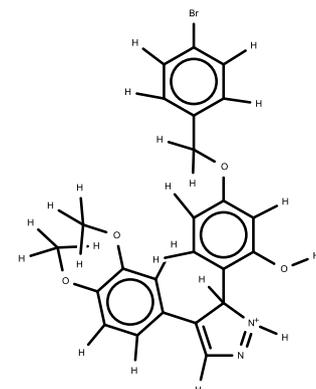
558



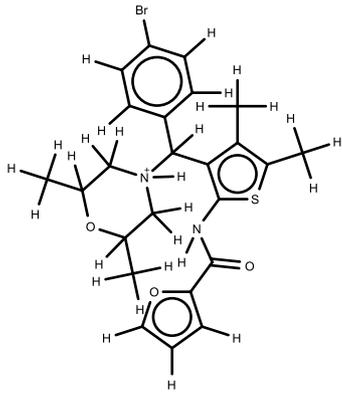
559



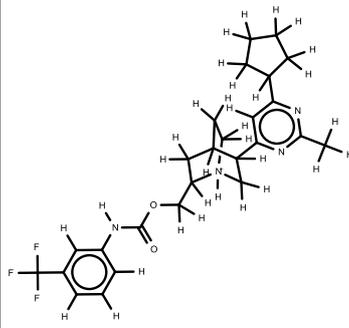
560



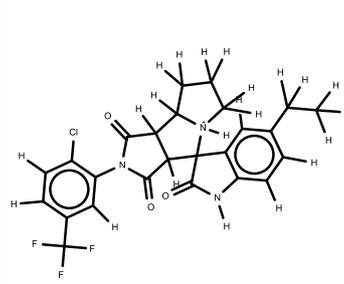
561



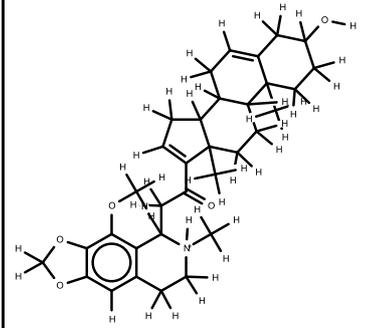
562



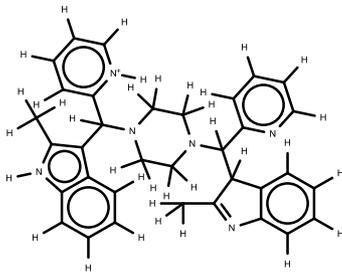
563



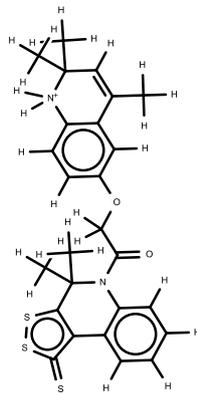
564



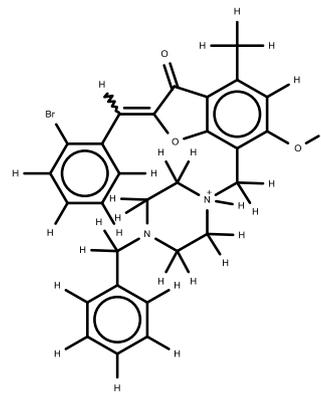
565



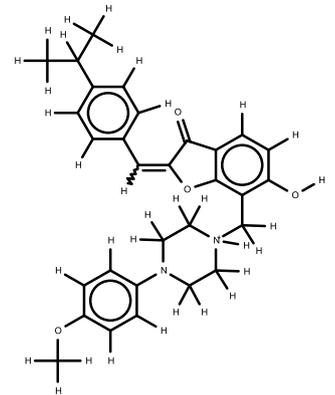
566



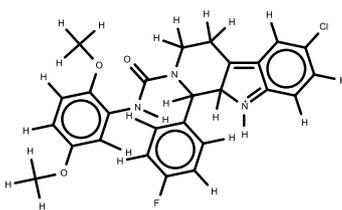
567



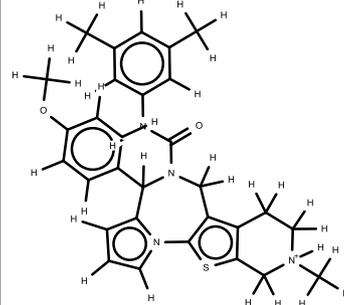
568



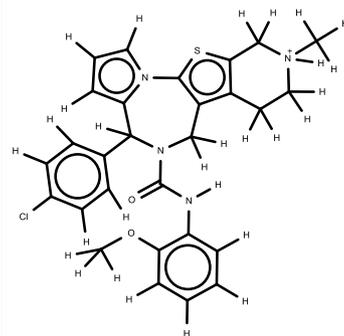
569



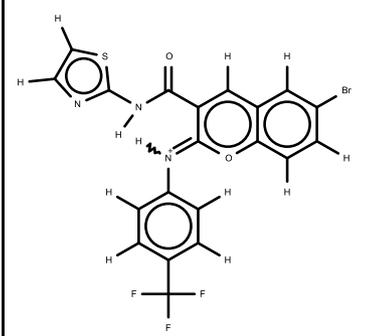
570



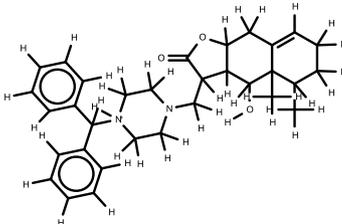
571



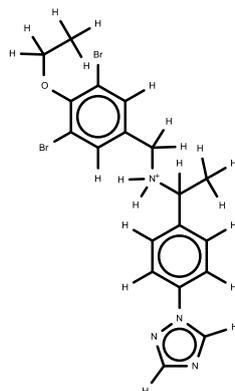
572



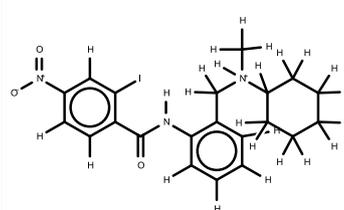
573



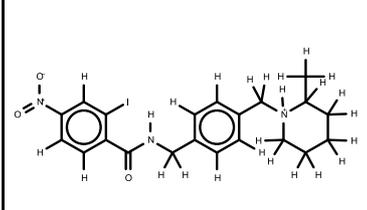
574



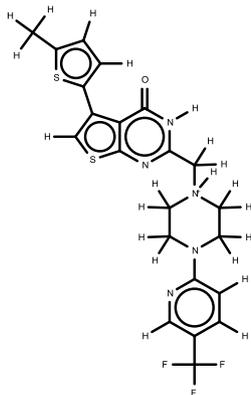
575



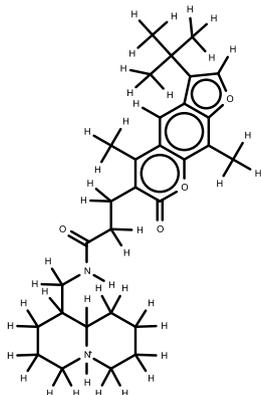
576



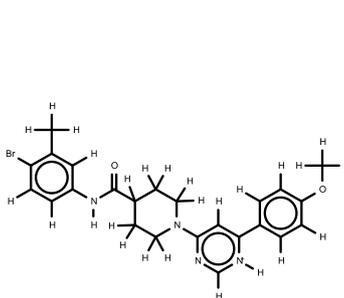
577



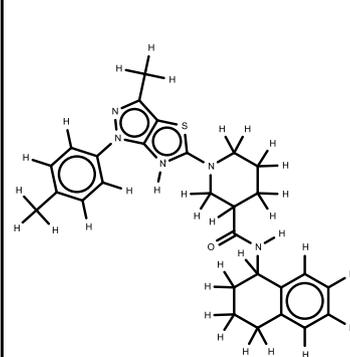
578



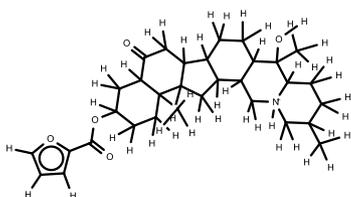
579



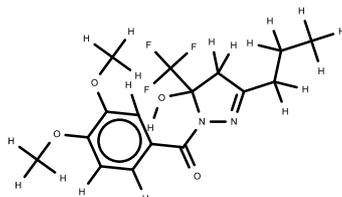
580



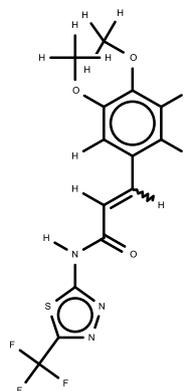
581



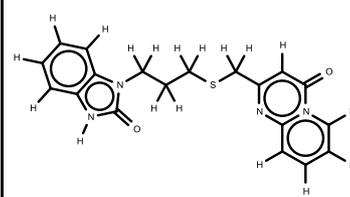
582



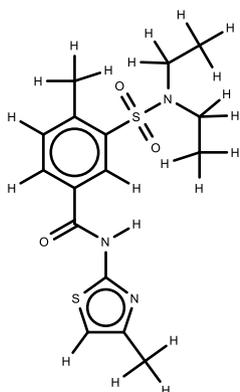
583



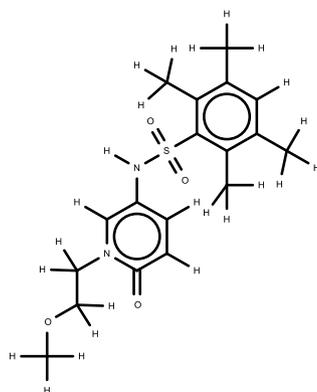
584



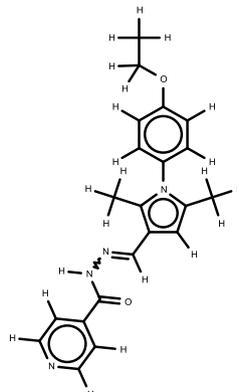
585



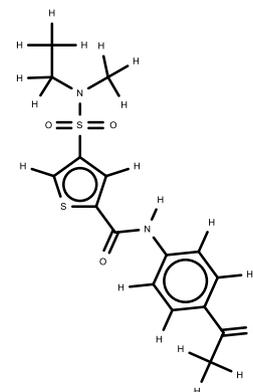
586



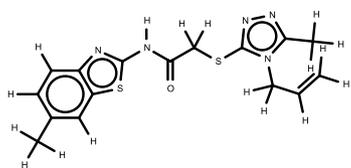
587



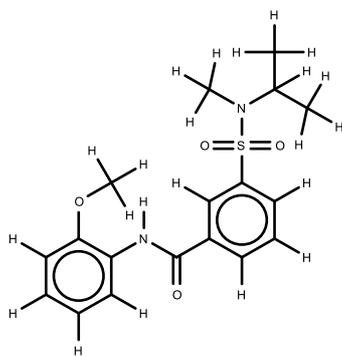
588



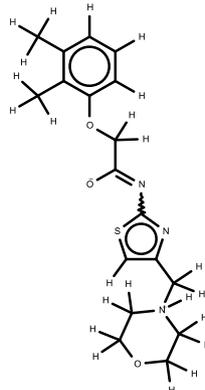
589



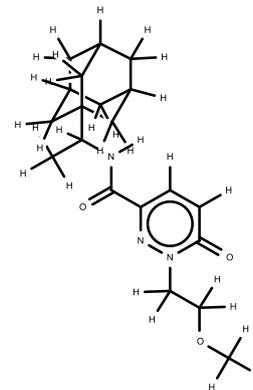
590



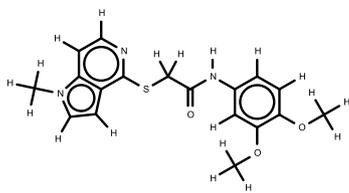
591



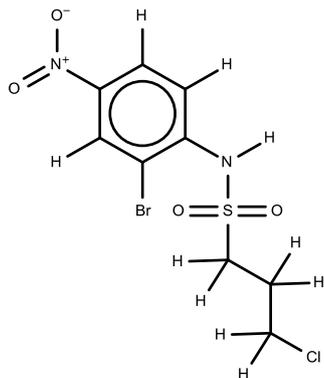
592



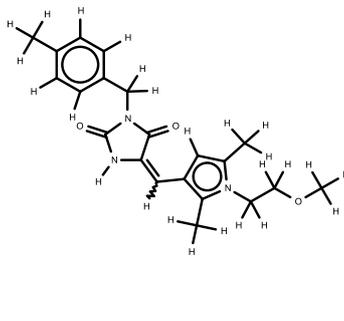
593



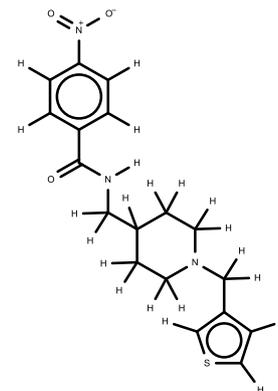
594



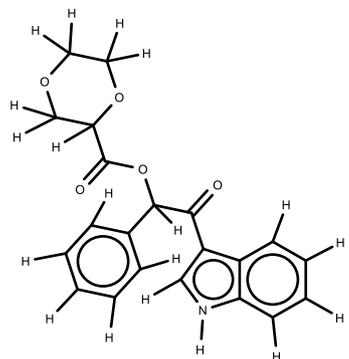
595



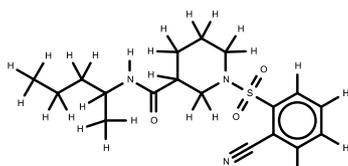
596



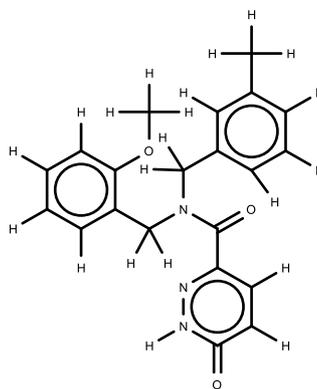
597



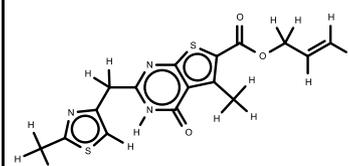
598



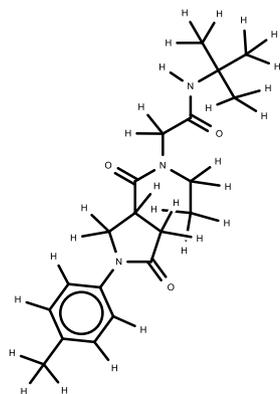
599



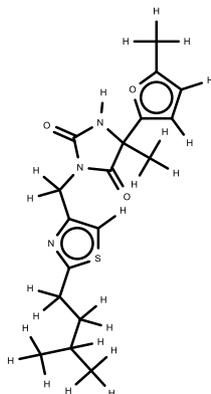
600



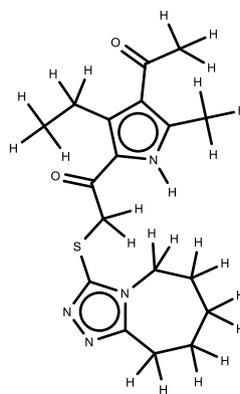
601



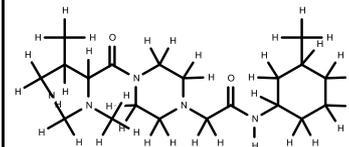
602



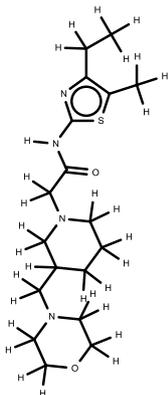
603



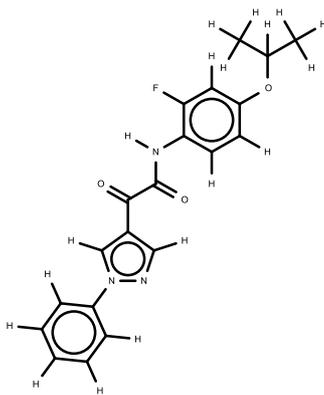
604



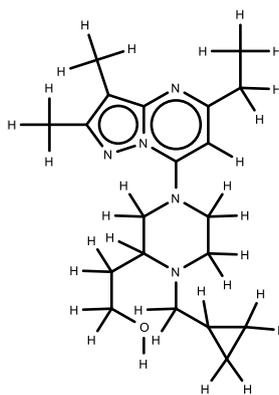
605



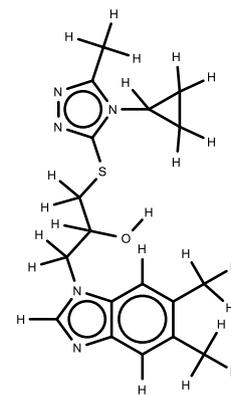
606



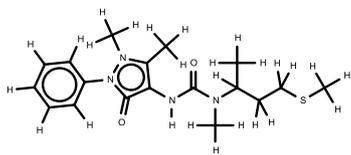
607



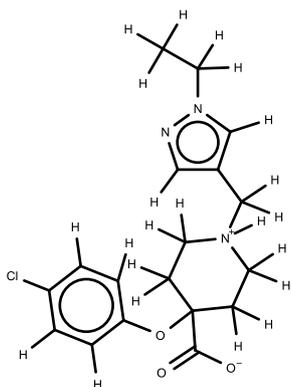
608



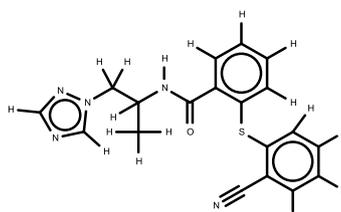
609



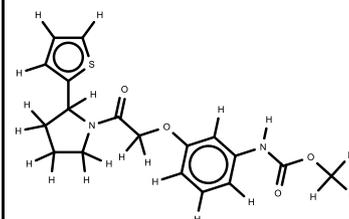
610



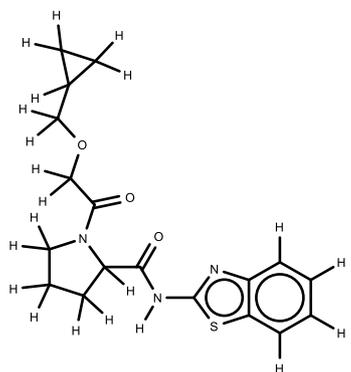
611



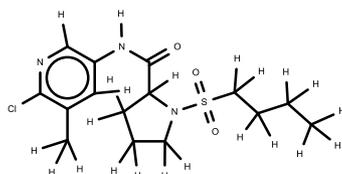
612



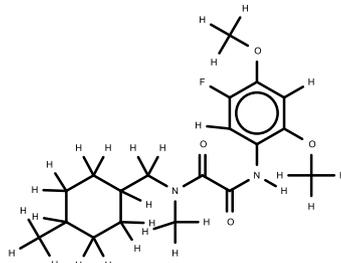
613



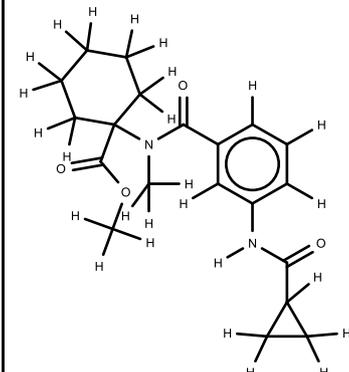
614



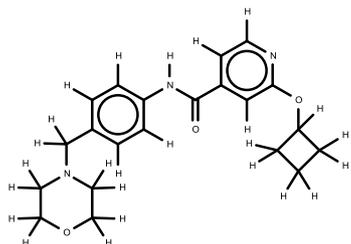
615



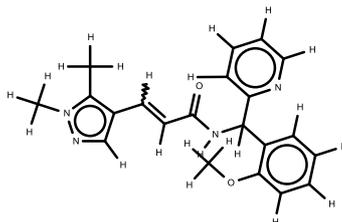
616



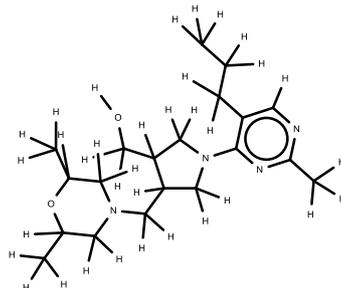
617



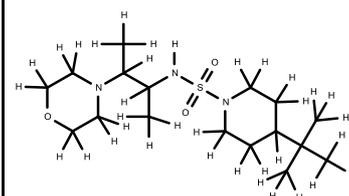
618



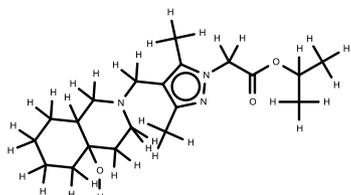
619



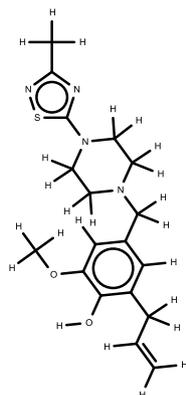
620



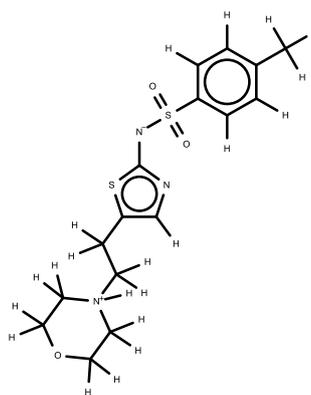
621



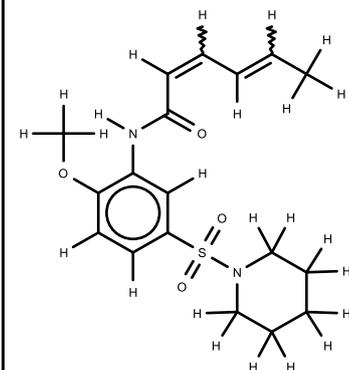
622



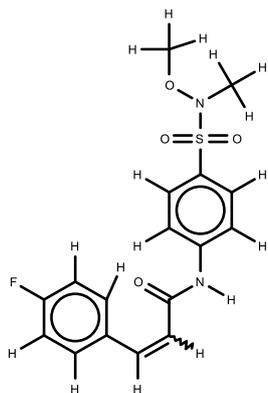
623



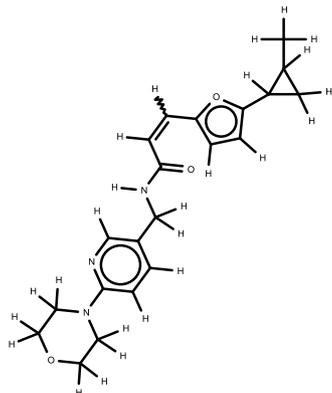
624



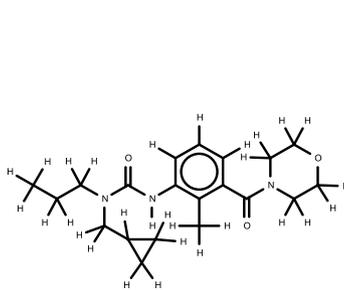
625



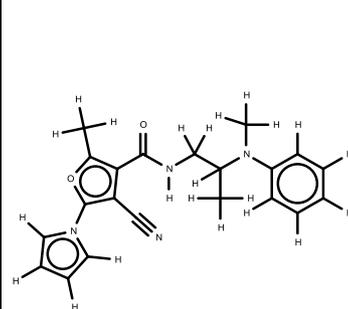
626



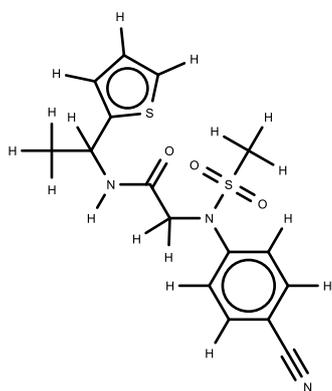
627



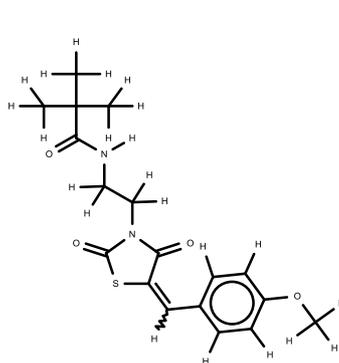
628



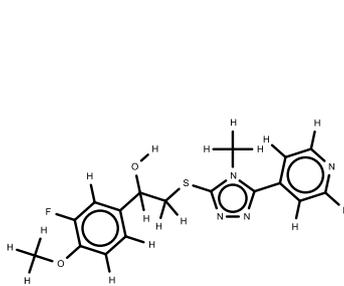
629



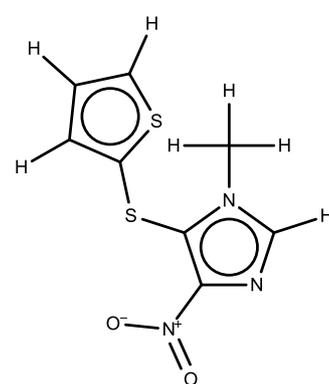
630



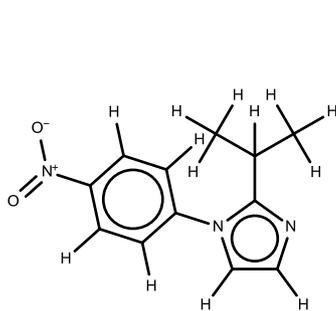
631



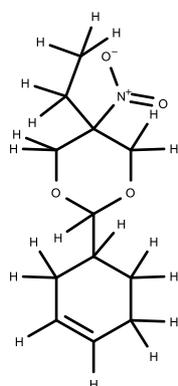
632



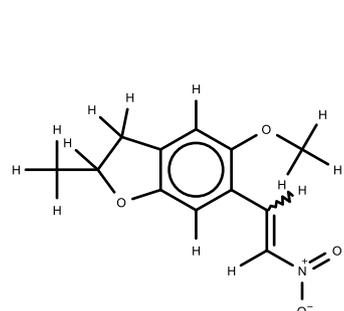
633



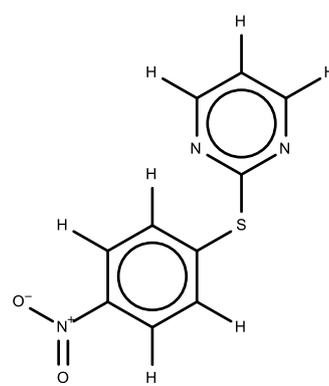
634



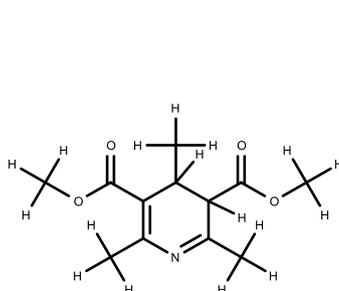
635



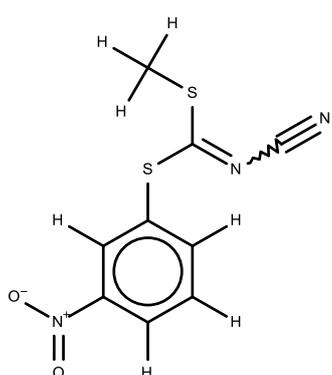
636



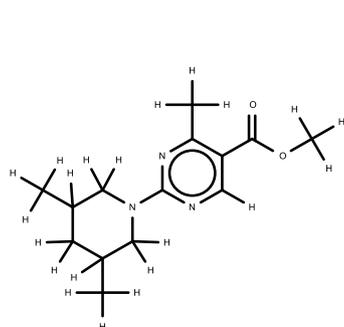
637



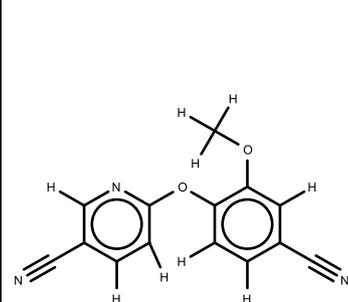
638



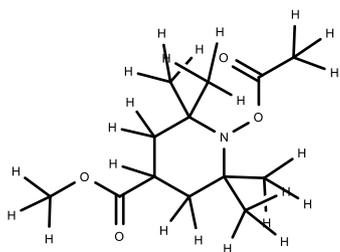
639



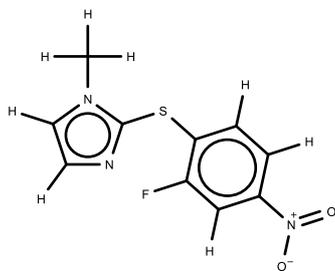
640



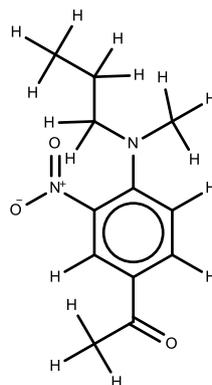
641



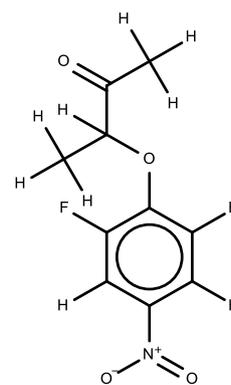
642



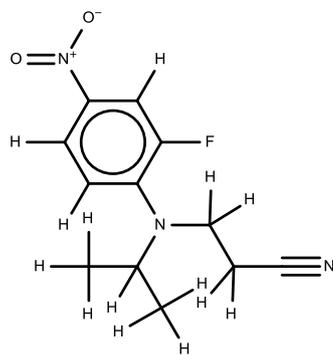
643



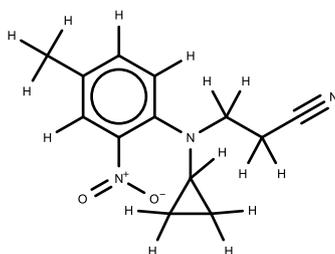
644



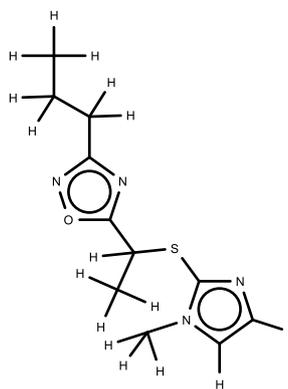
645



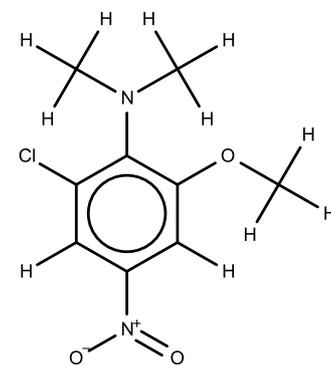
646



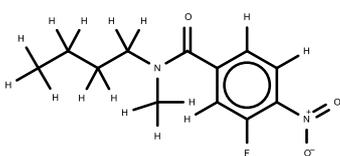
647



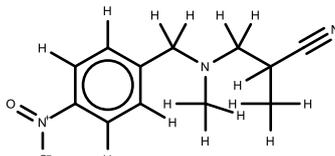
648



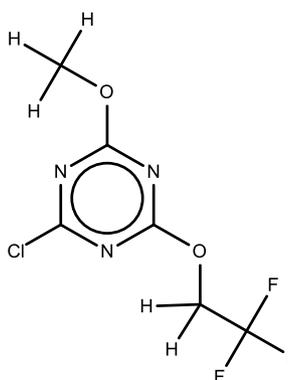
649



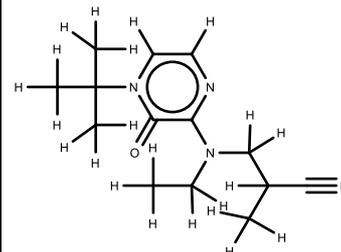
650



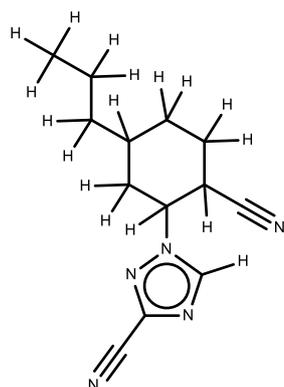
651



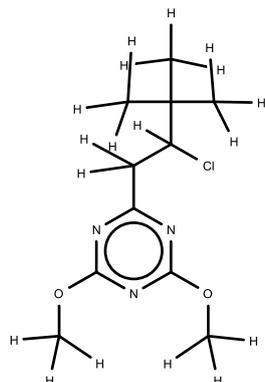
652



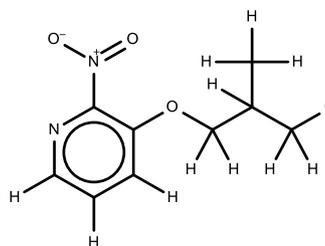
653



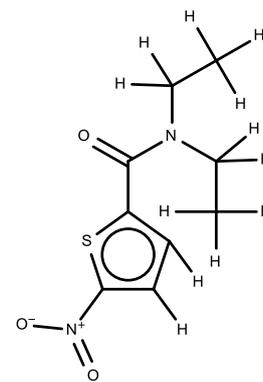
654



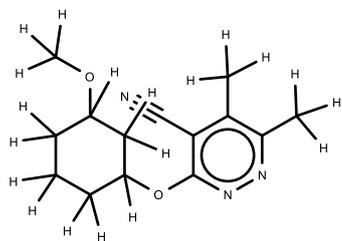
655



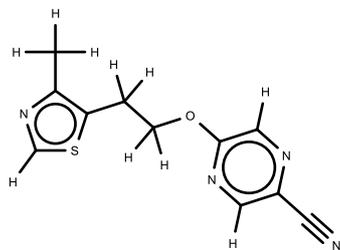
656



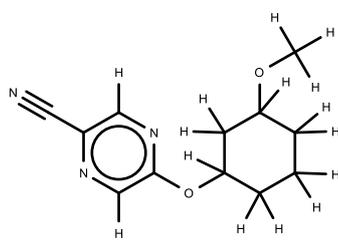
657



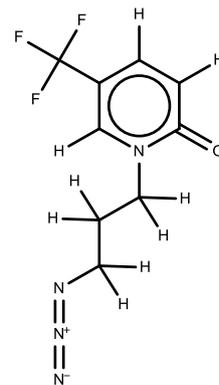
658



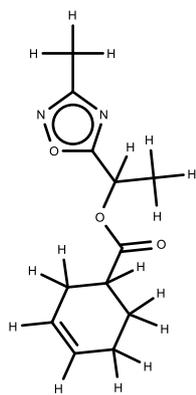
659



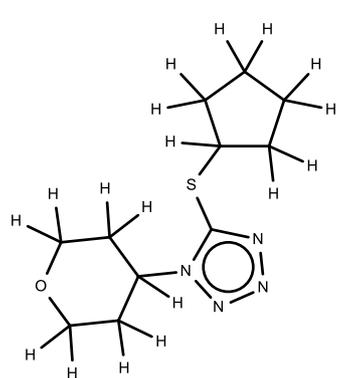
660



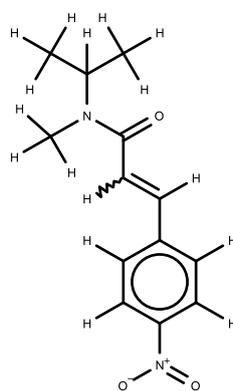
661



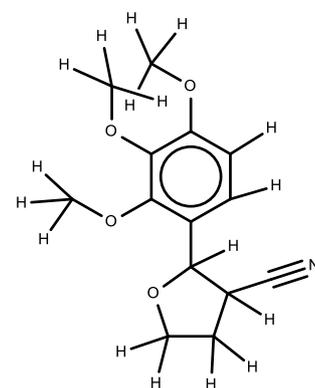
662



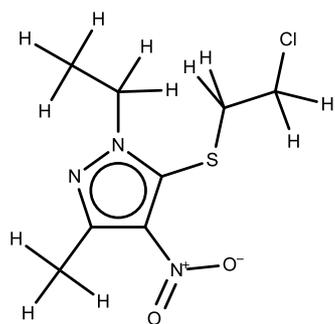
663



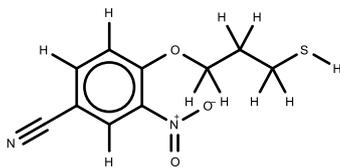
664



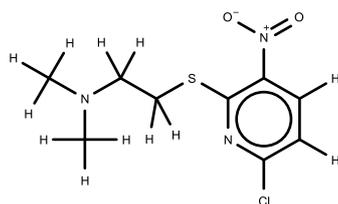
665



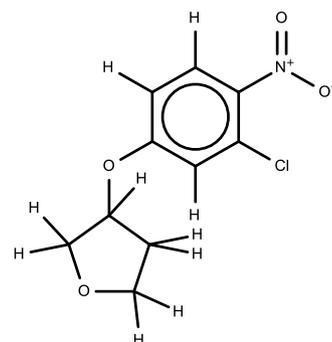
666



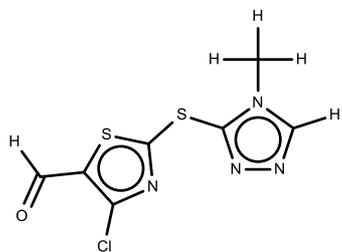
667



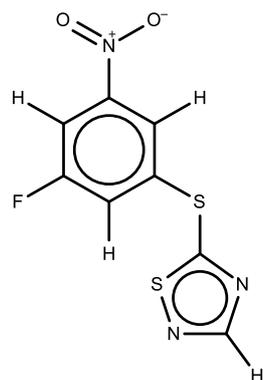
668



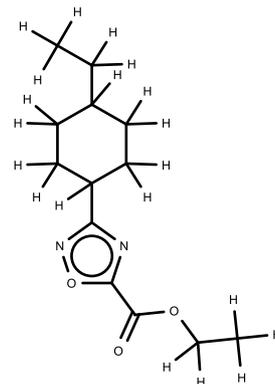
669



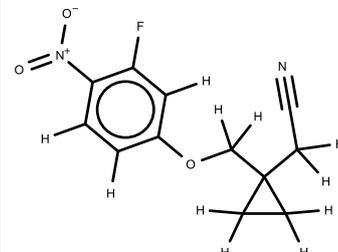
670



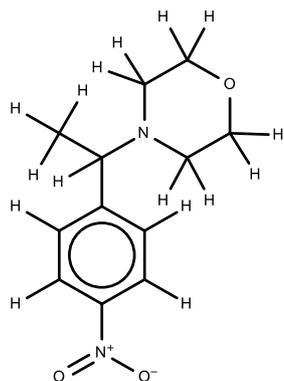
671



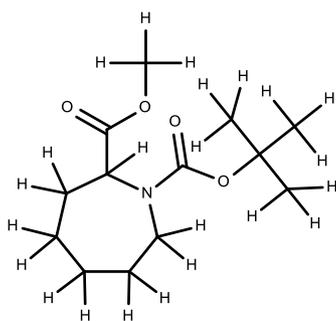
672



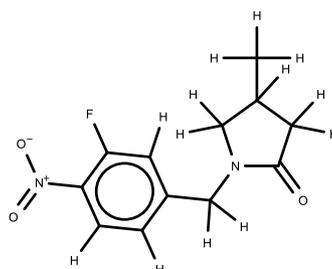
673



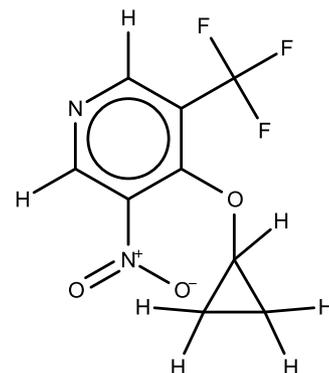
674



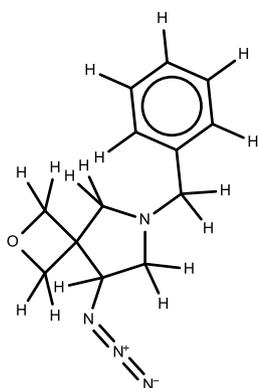
675



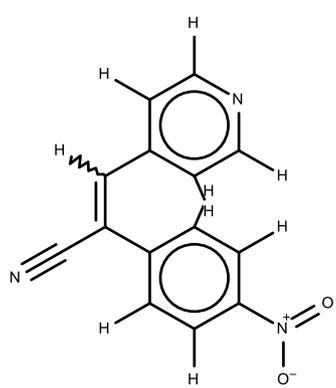
676



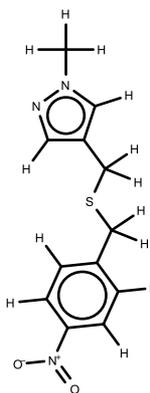
677



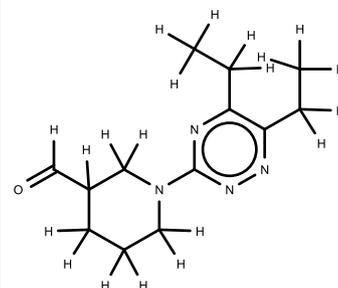
678



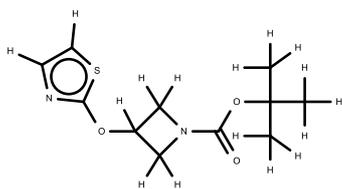
679



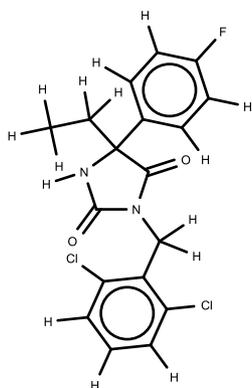
680



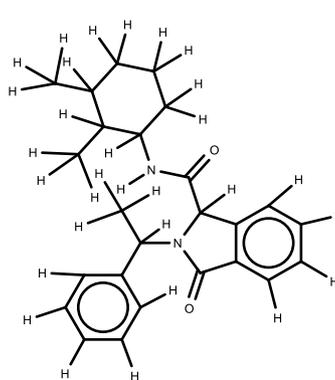
681



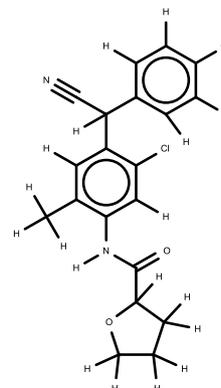
682



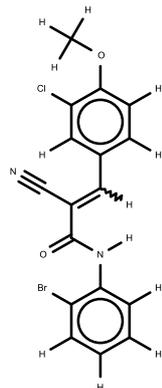
683



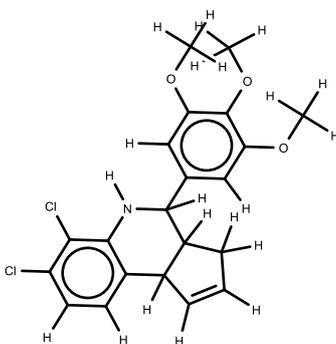
684



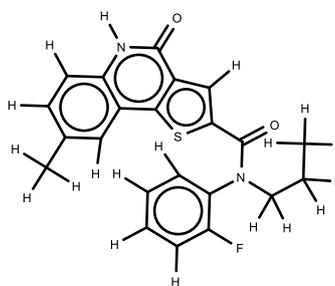
685



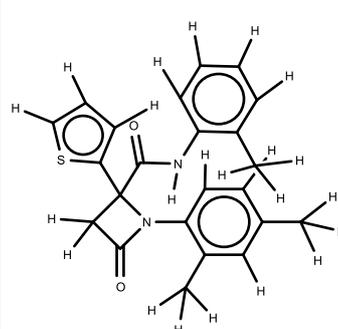
686



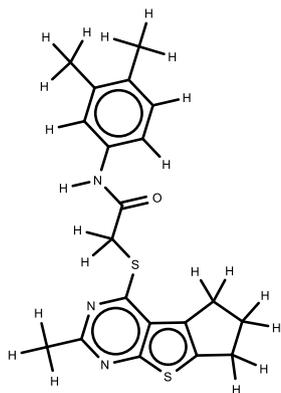
687



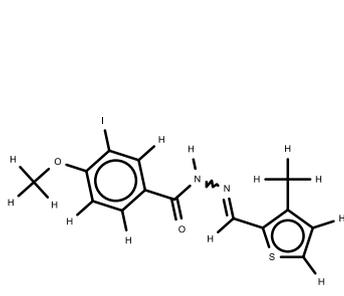
688



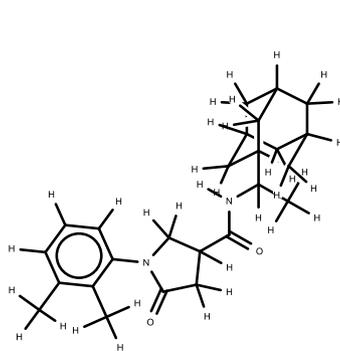
689



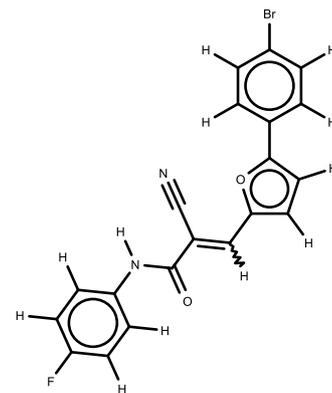
690



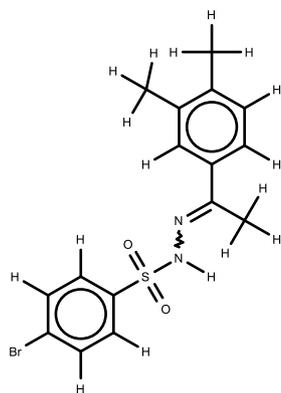
691



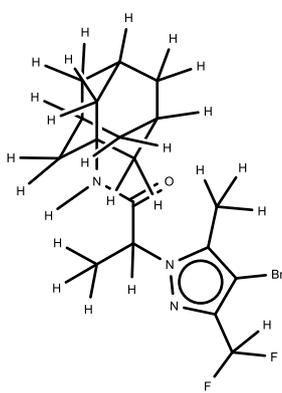
692



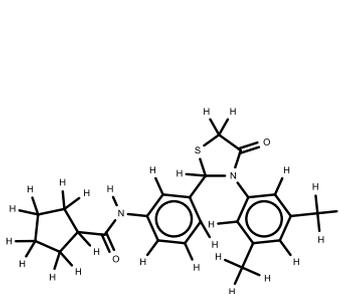
693



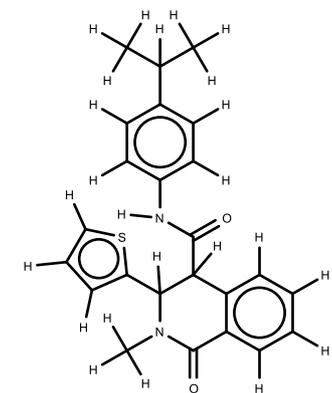
694



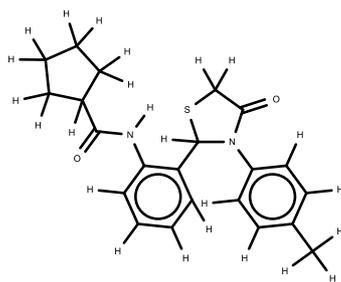
695



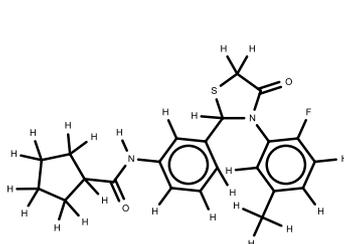
696



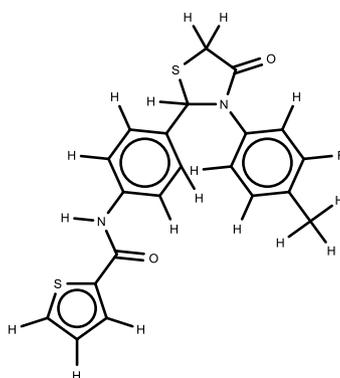
697



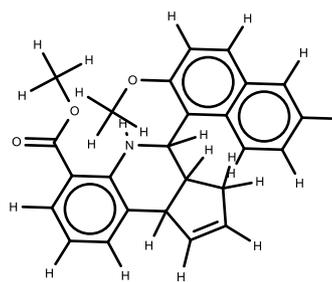
698



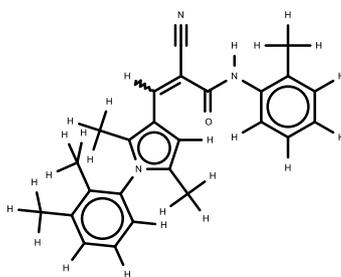
699



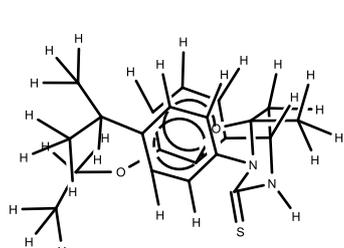
700



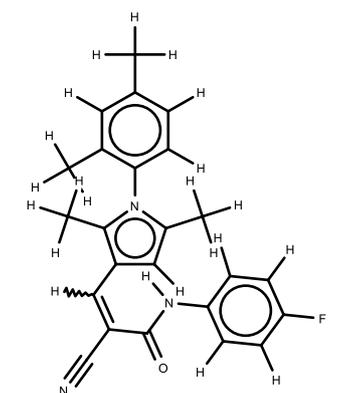
701



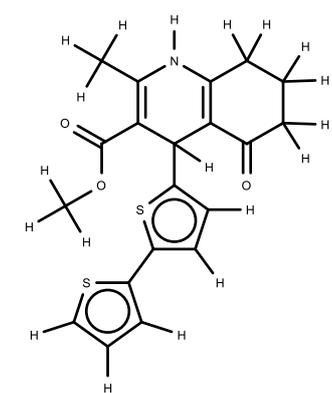
702



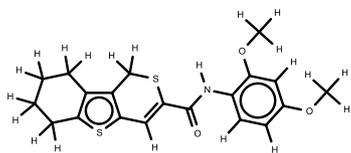
703



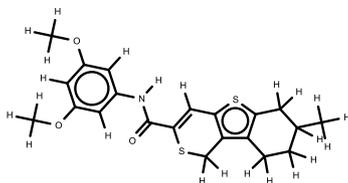
704



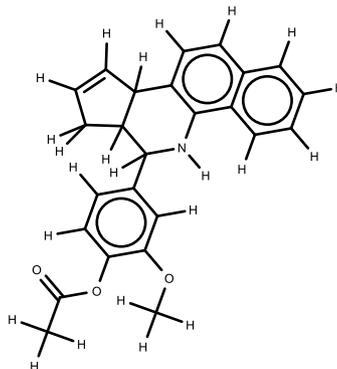
705



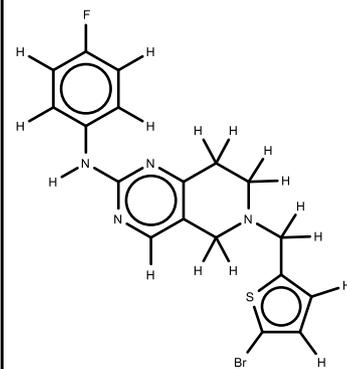
706



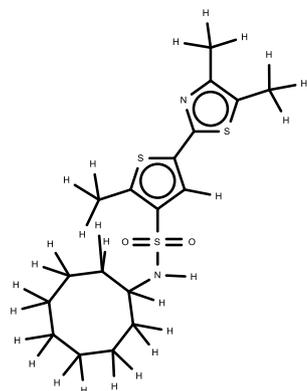
707



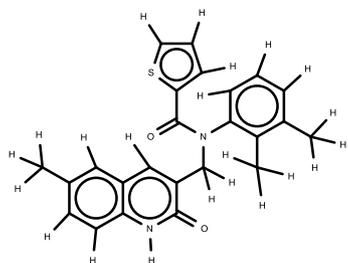
708



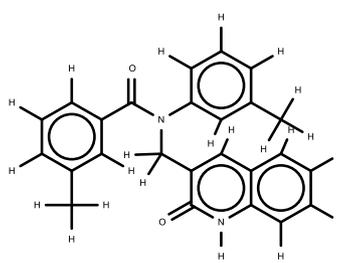
709



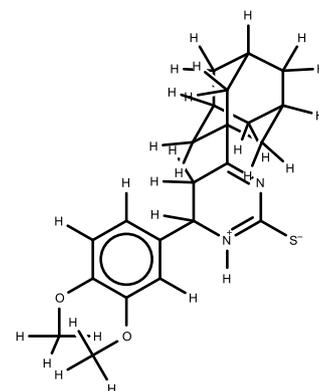
710



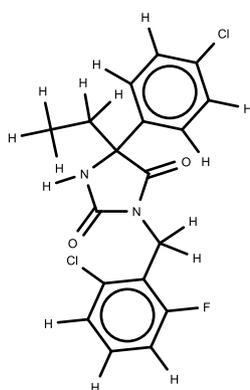
711



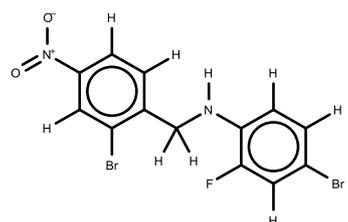
712



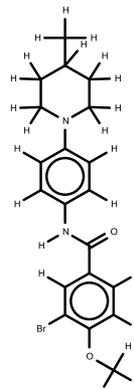
713



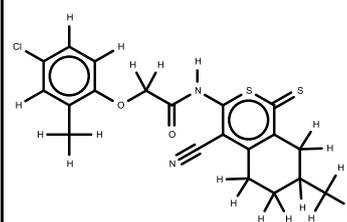
714



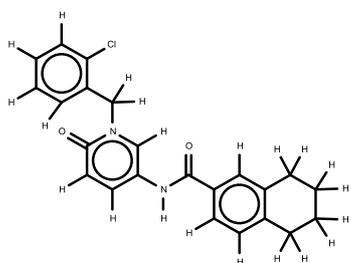
715



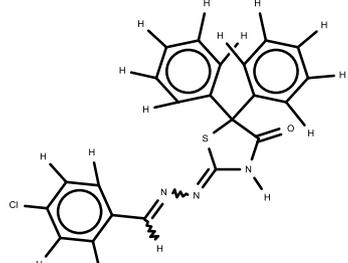
716



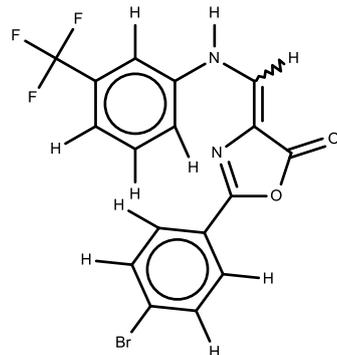
717



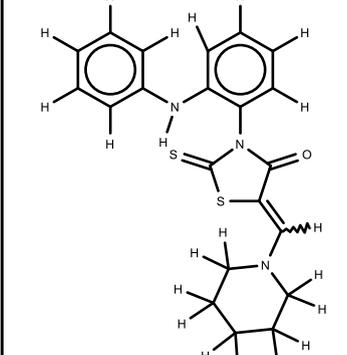
718



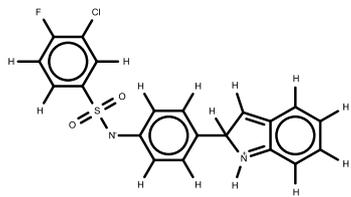
719



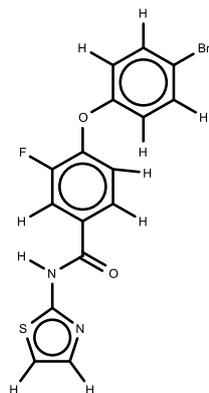
720



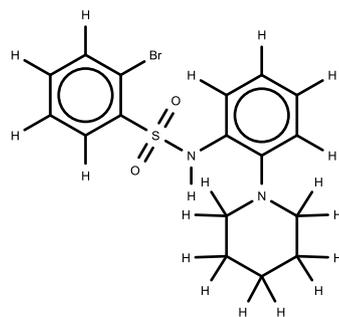
721



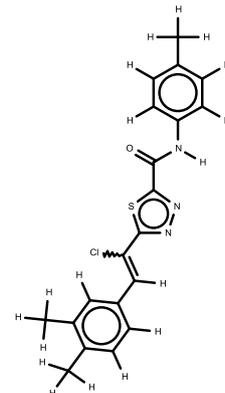
722



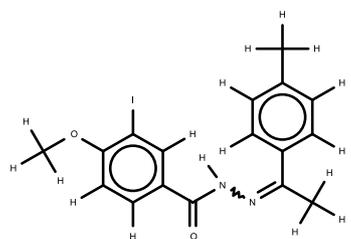
723



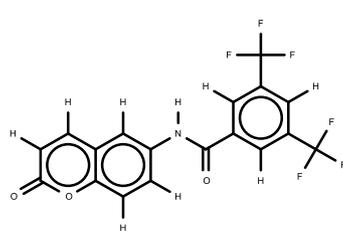
724



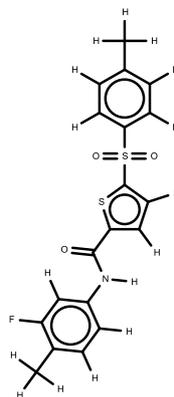
725



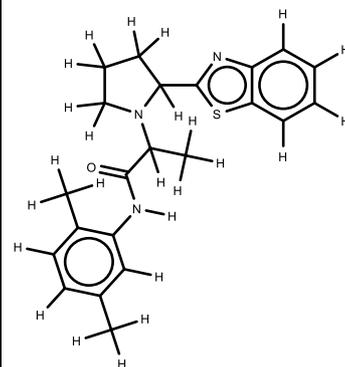
726



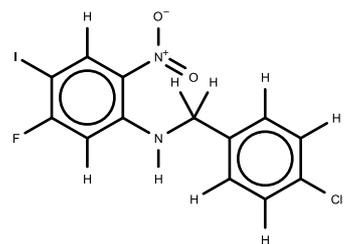
727



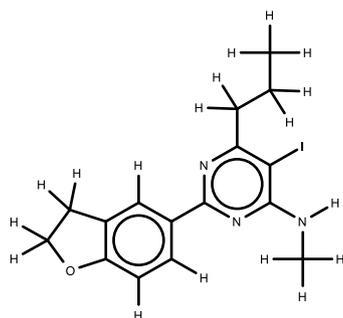
728



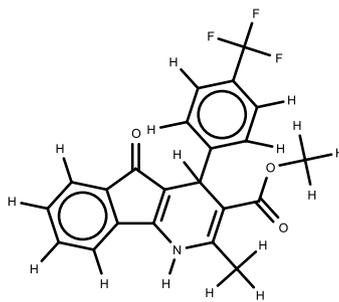
729



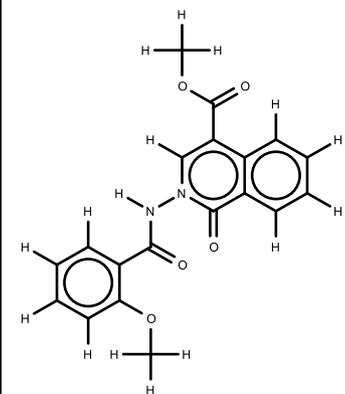
730



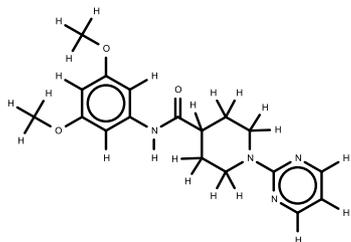
731



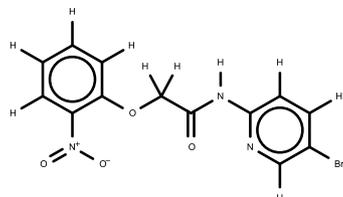
732



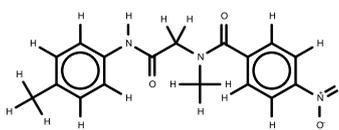
733



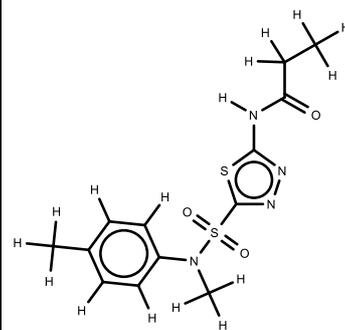
734



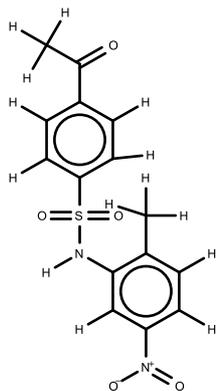
735



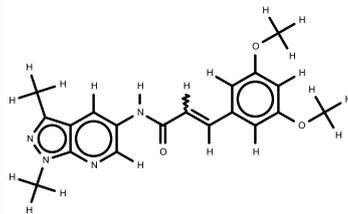
736



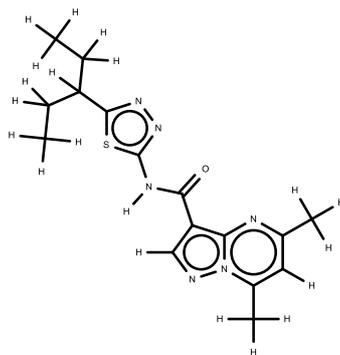
737



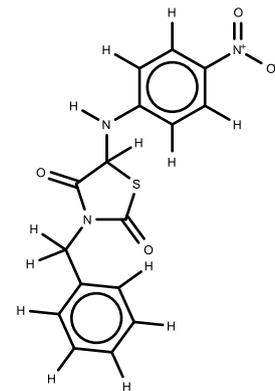
738



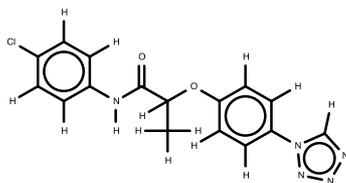
739



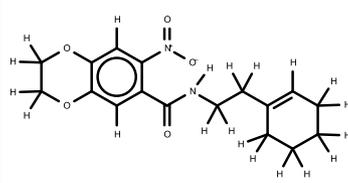
740



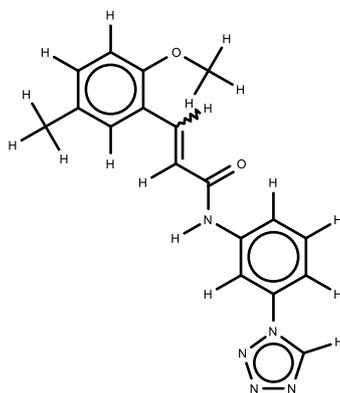
741



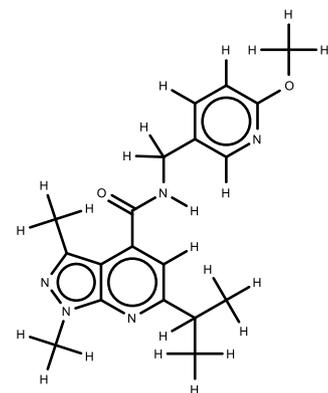
742



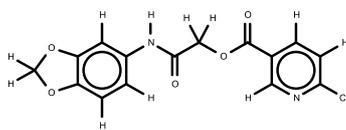
743



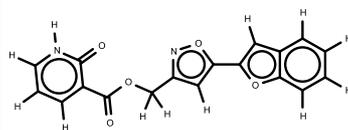
744



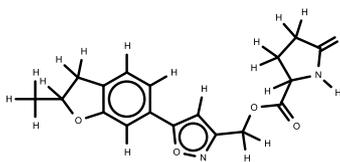
745



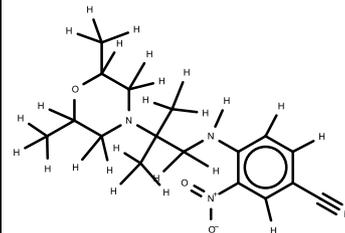
746



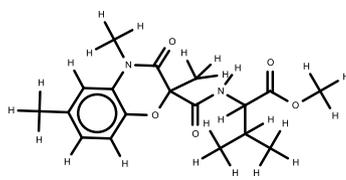
747



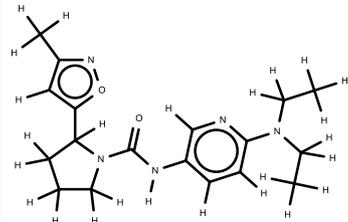
748



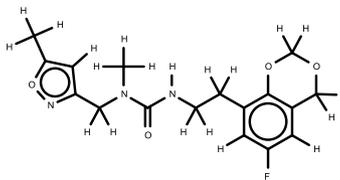
749



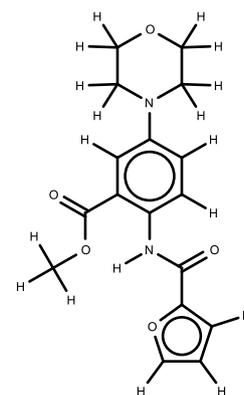
750



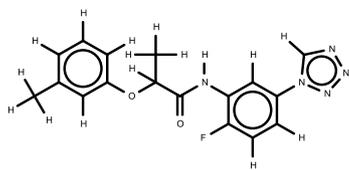
751



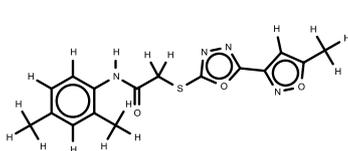
752



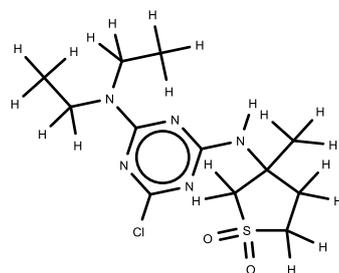
753



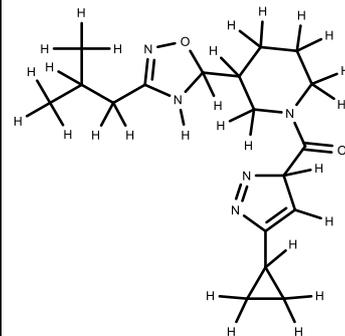
754



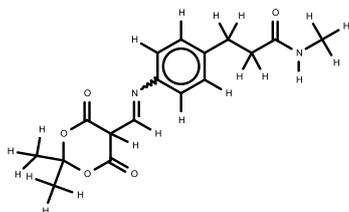
755



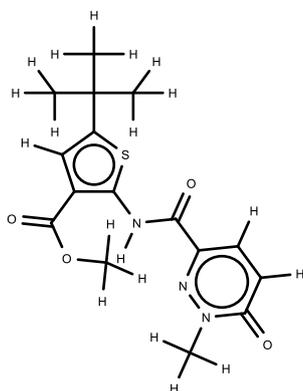
756



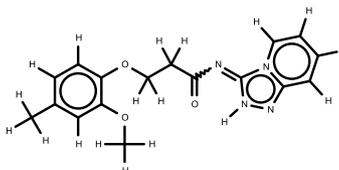
757



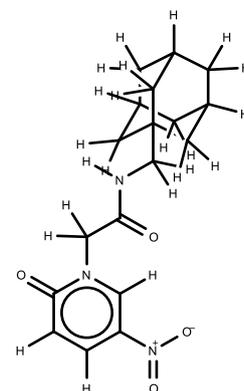
758



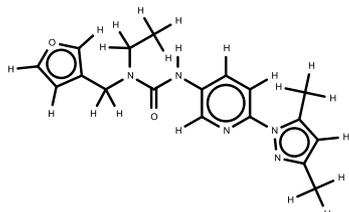
759



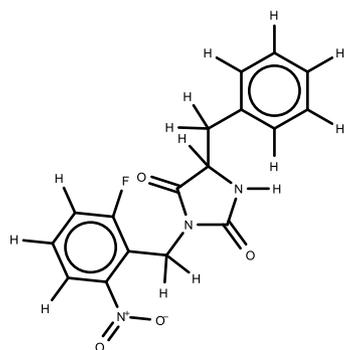
760



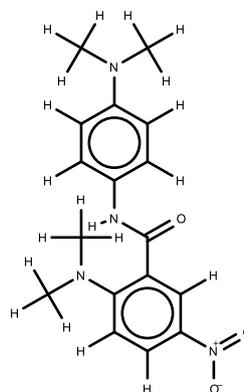
761



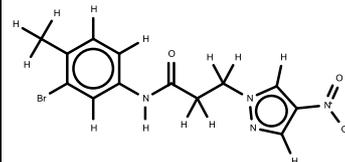
762



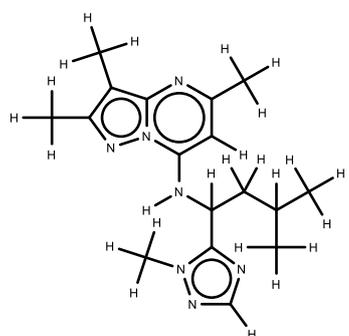
763



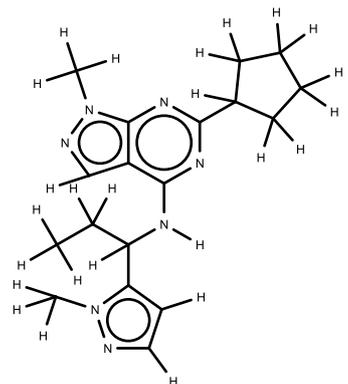
764



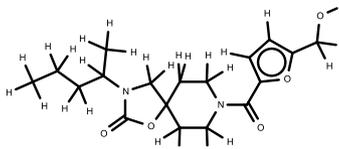
765



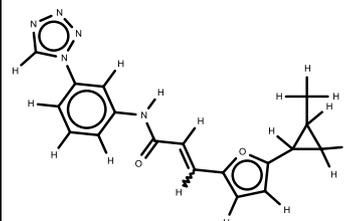
766



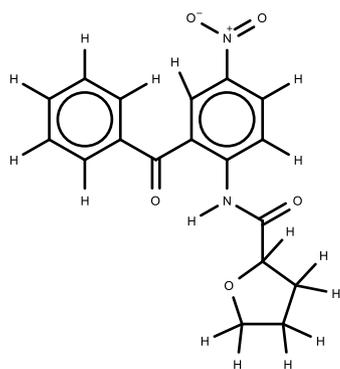
767



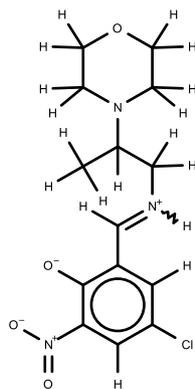
768



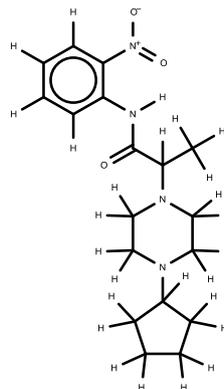
769



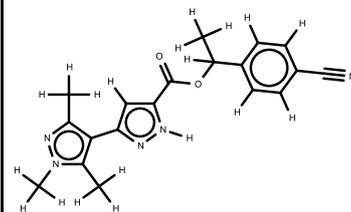
770



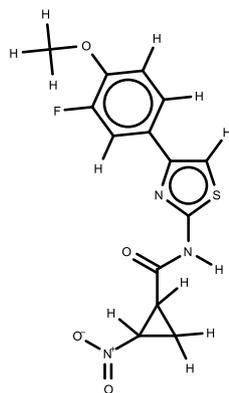
771



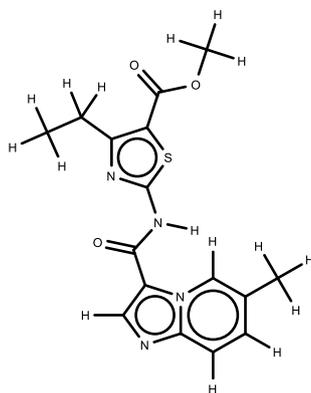
772



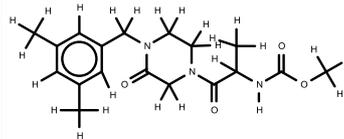
773



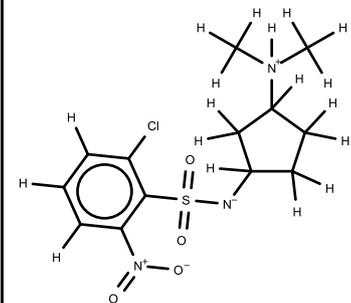
774



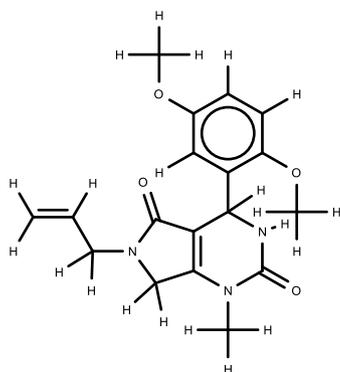
775



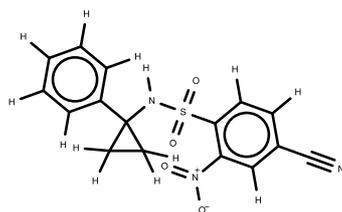
776



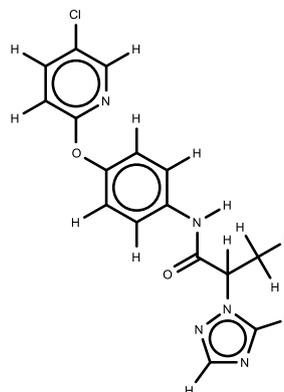
777



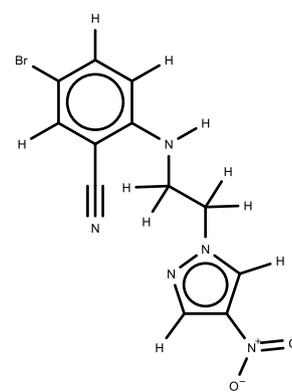
778



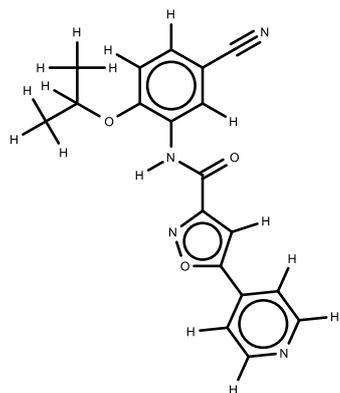
779



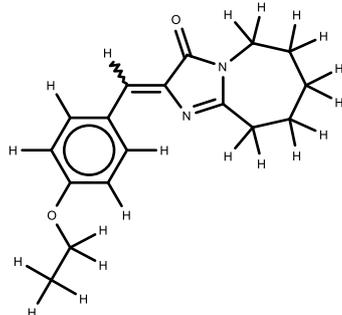
780



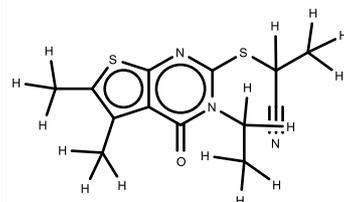
781



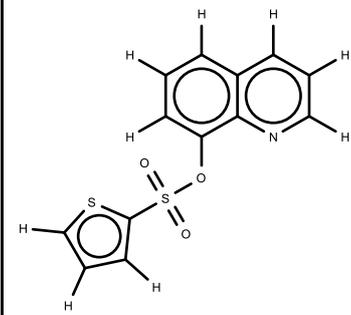
782



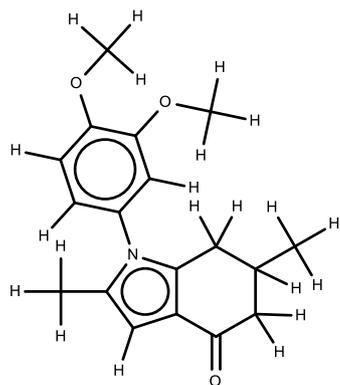
783



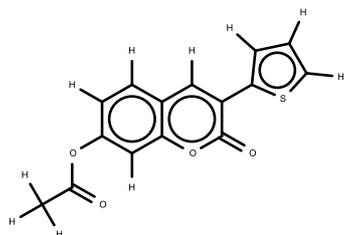
784



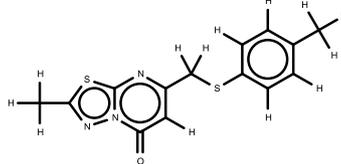
785



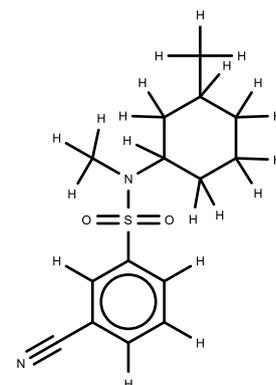
786



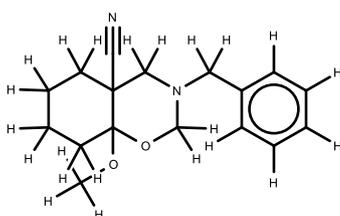
787



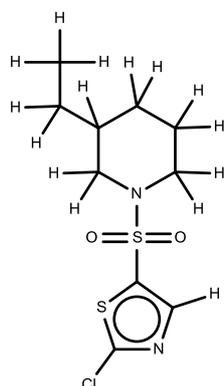
788



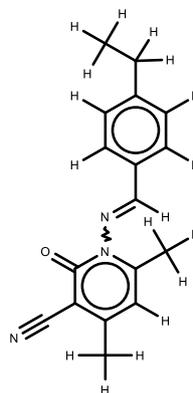
789



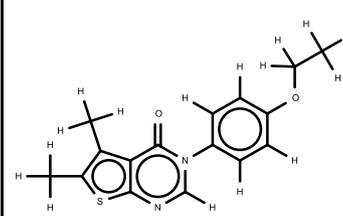
790



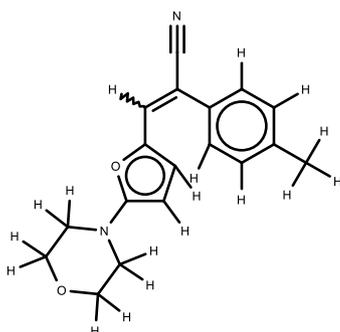
791



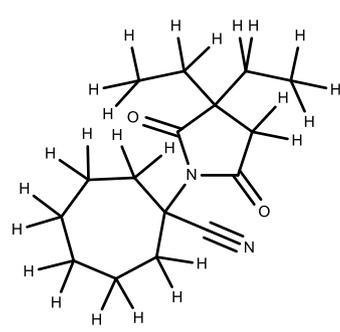
792



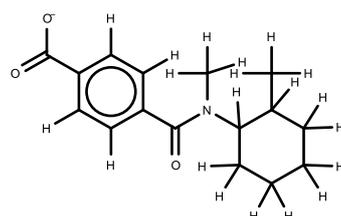
793



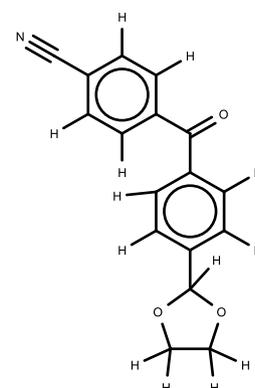
794



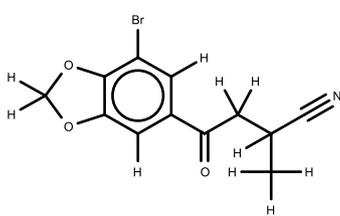
795



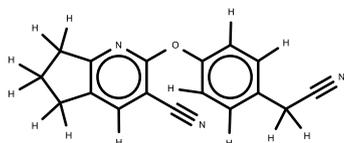
796



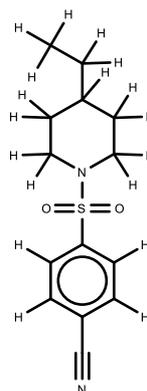
797



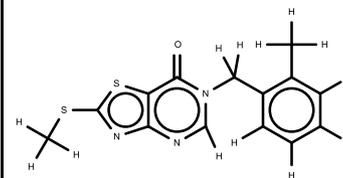
798



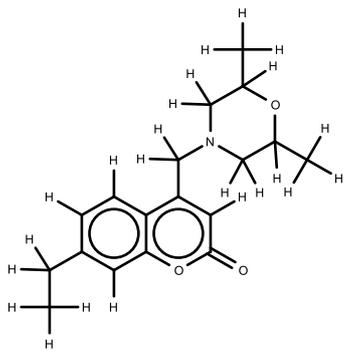
799



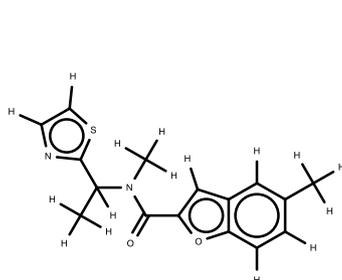
800



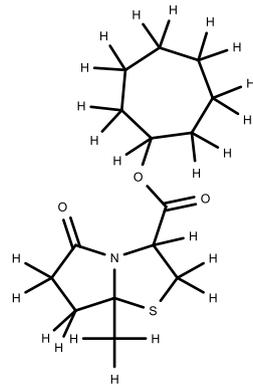
801



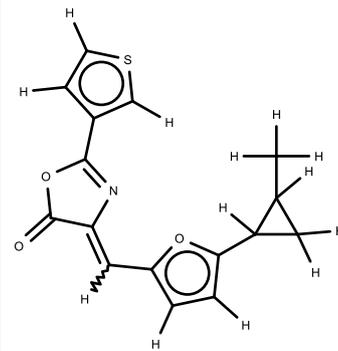
802



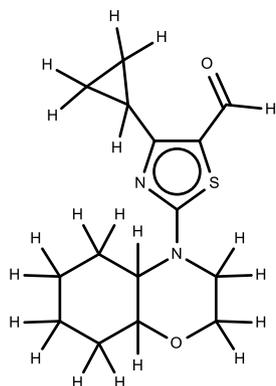
803



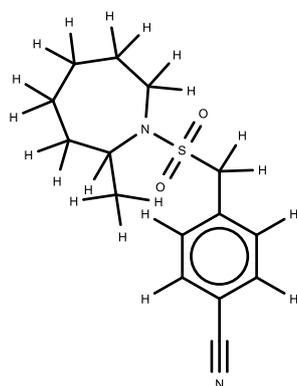
804



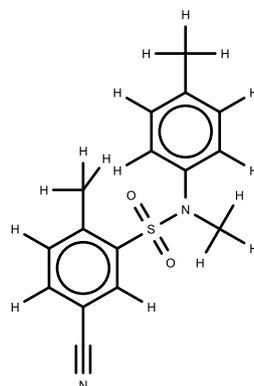
805



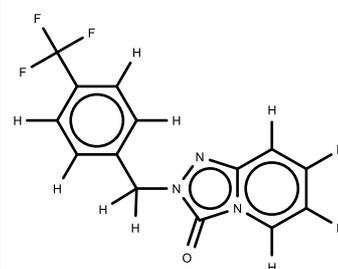
806



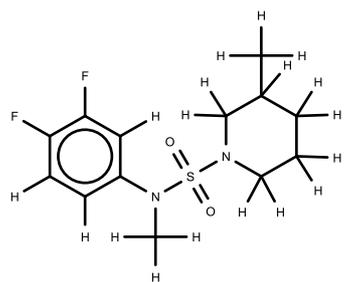
807



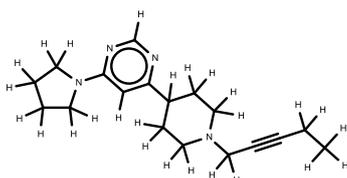
808



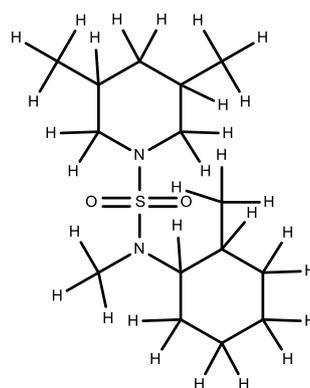
809



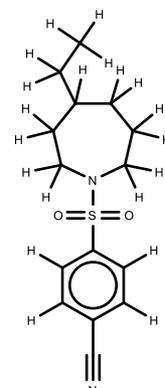
810



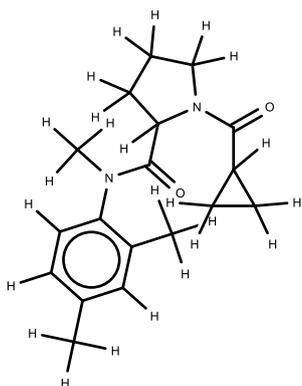
811



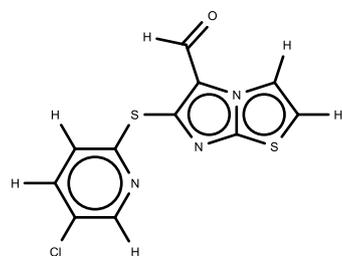
812



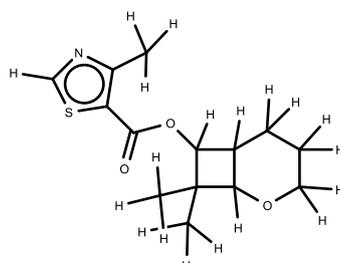
813



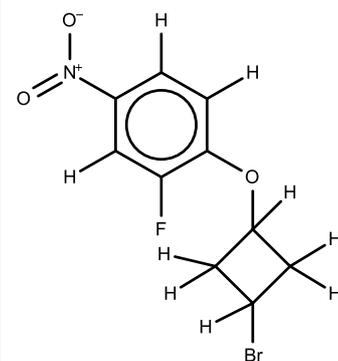
814



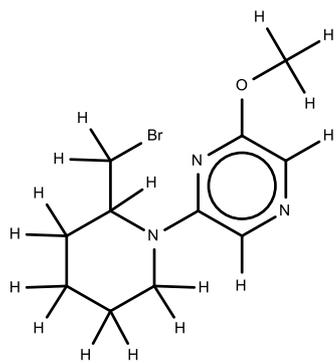
815



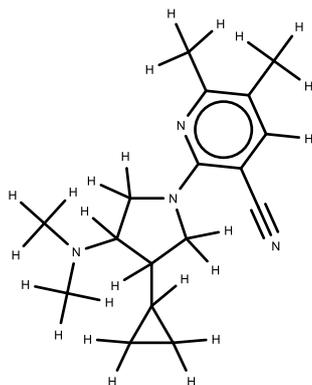
816



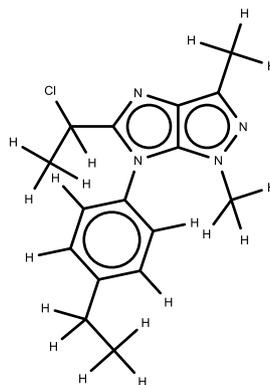
817



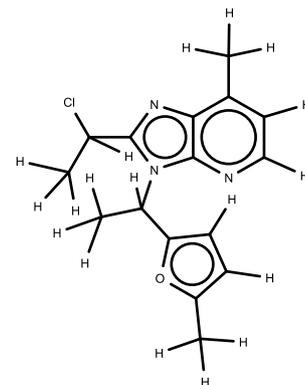
818



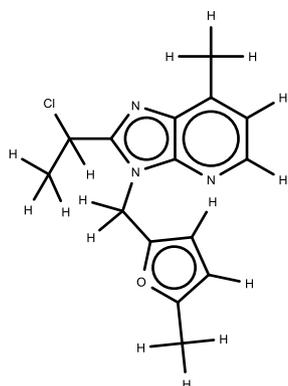
819



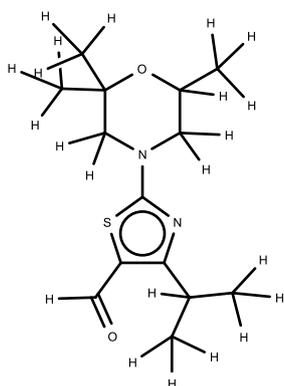
820



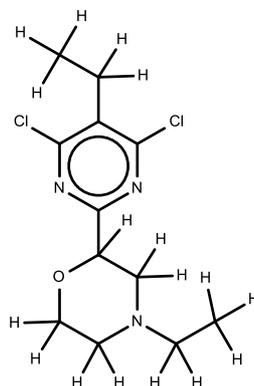
821



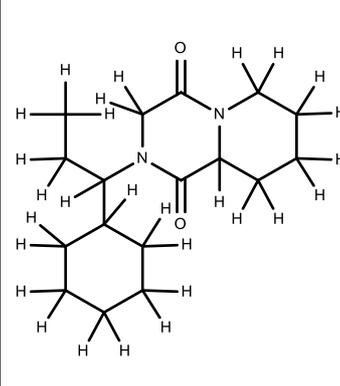
822



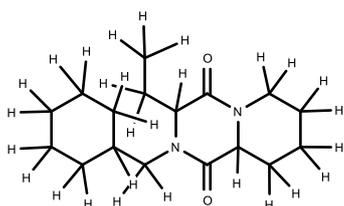
823



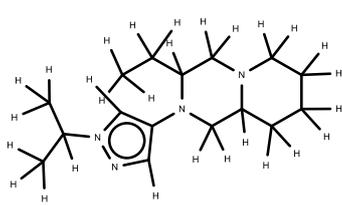
824



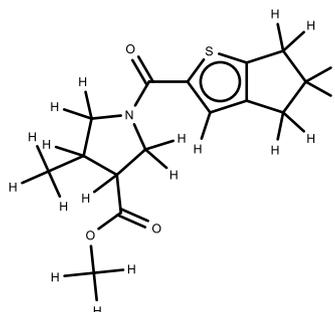
825



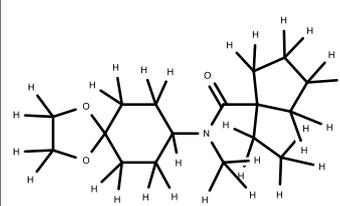
826



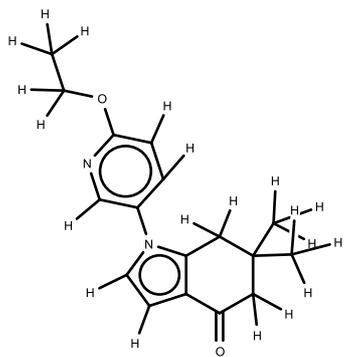
827



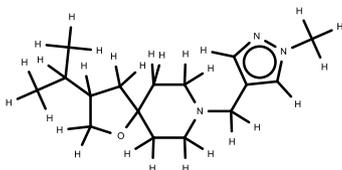
828



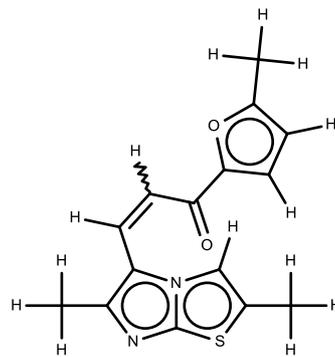
829



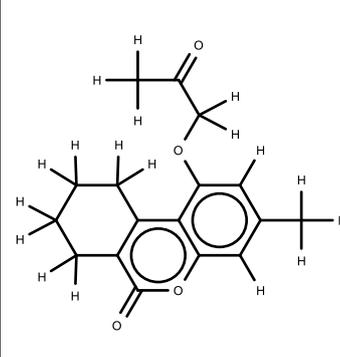
830



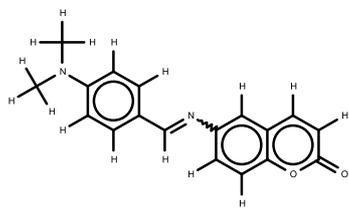
831



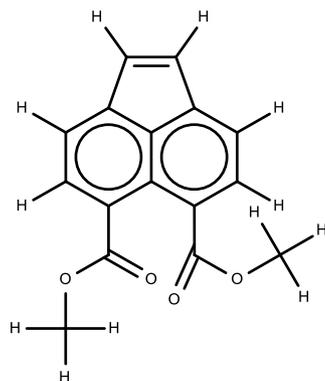
832



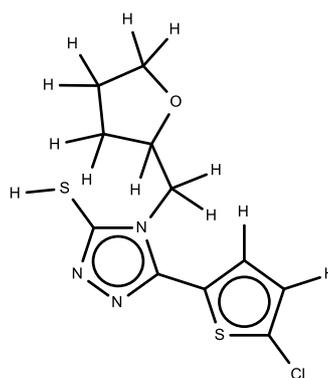
833



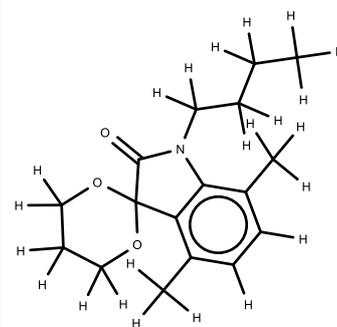
834



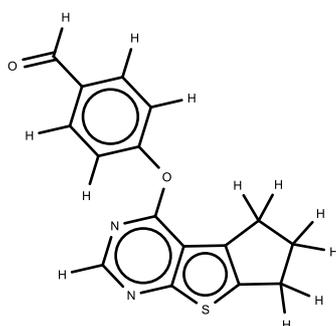
835



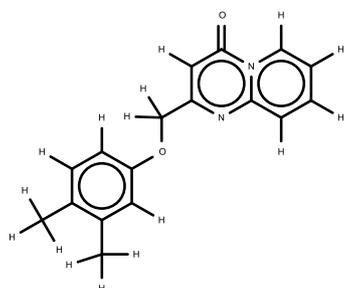
836



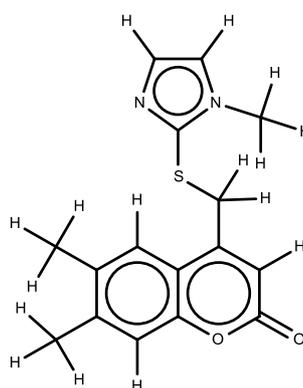
837



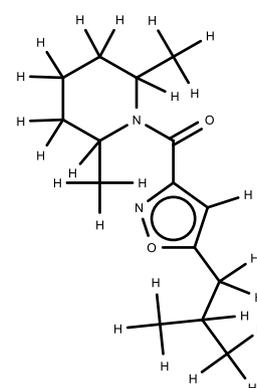
838



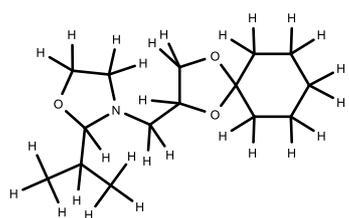
839



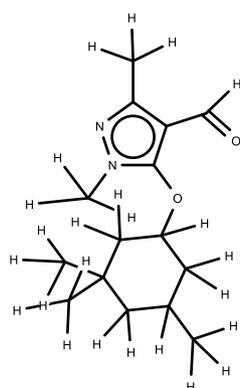
840



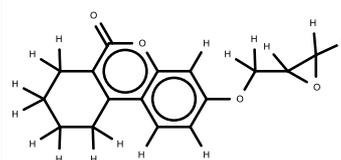
841



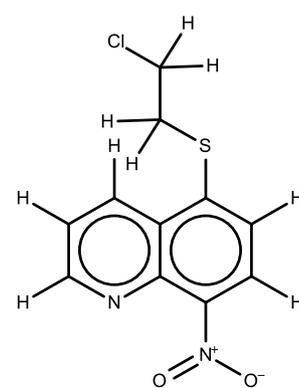
842



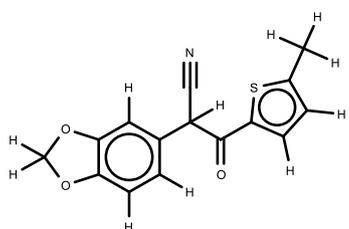
843



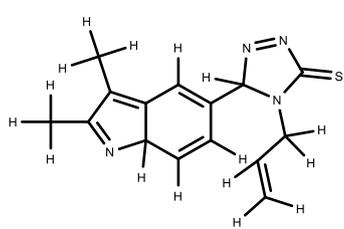
844



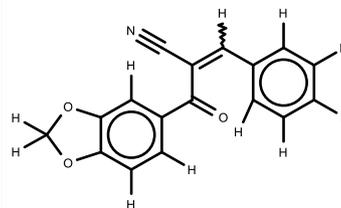
845



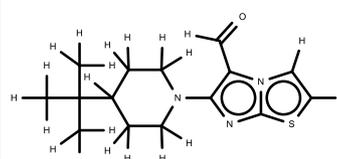
846



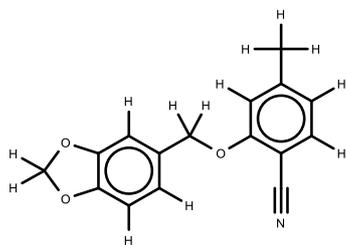
847



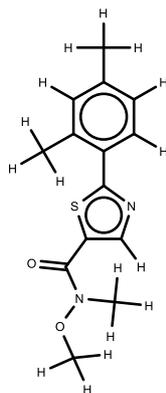
848



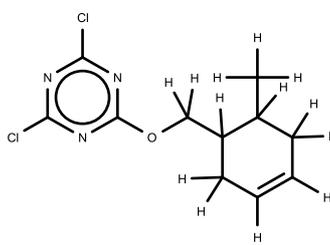
849



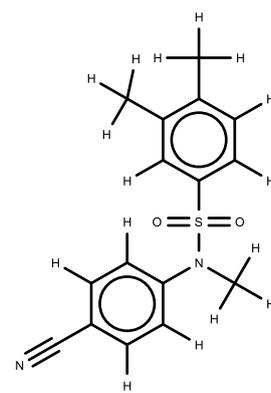
850



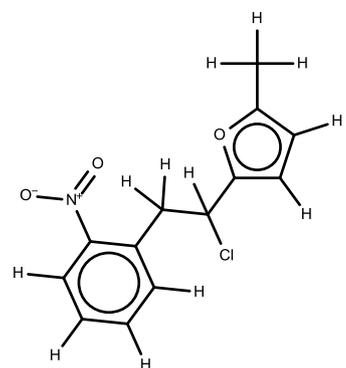
851



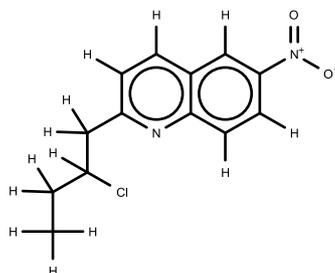
852



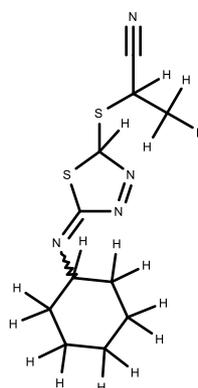
853



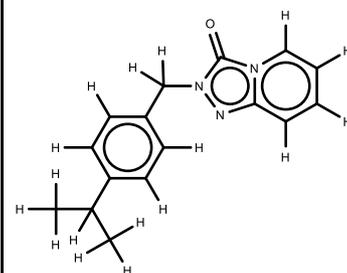
854



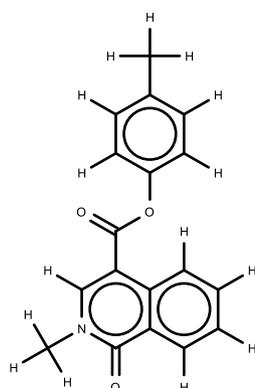
855



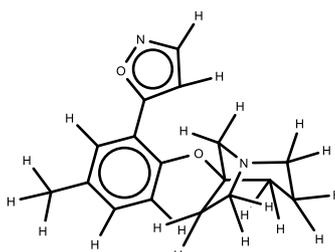
856



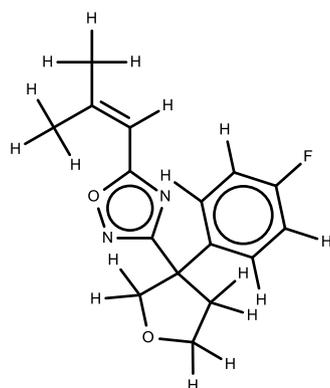
857



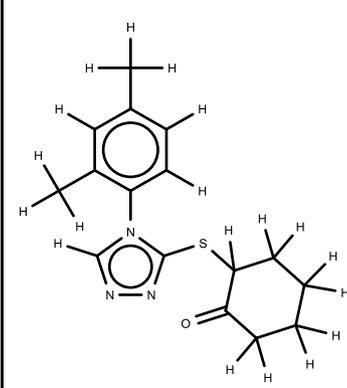
858



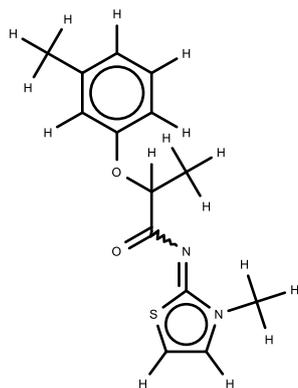
859



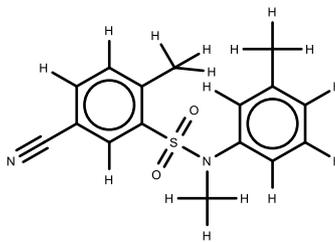
860



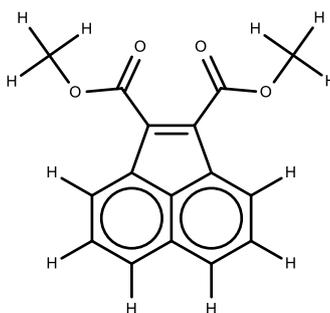
861



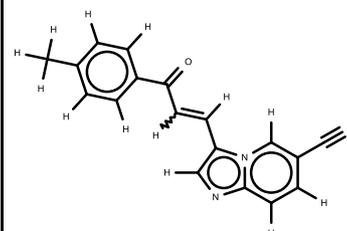
862



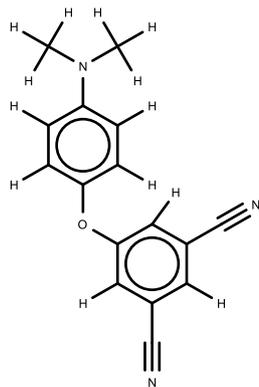
863



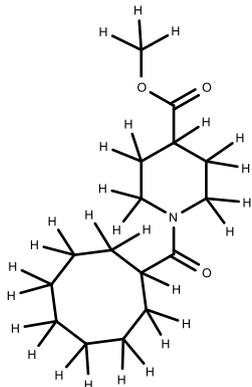
864



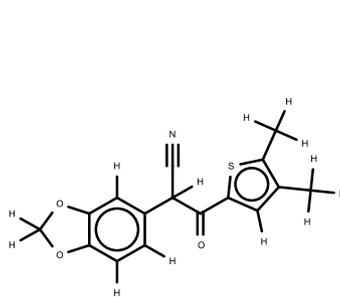
865



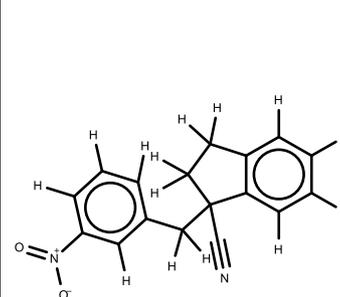
866



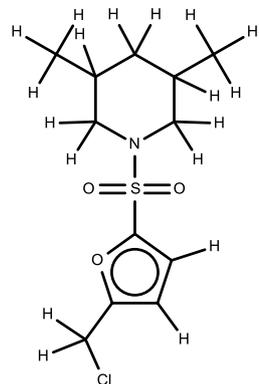
867



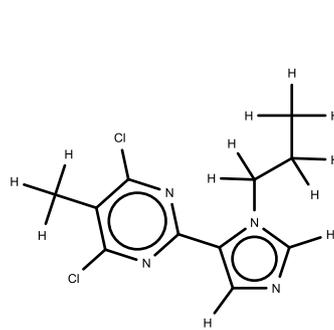
868



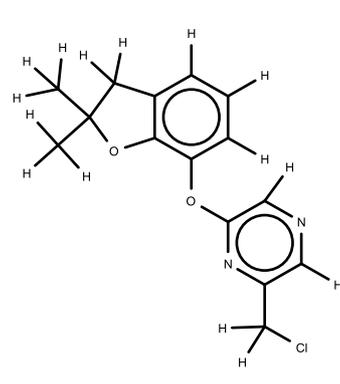
869



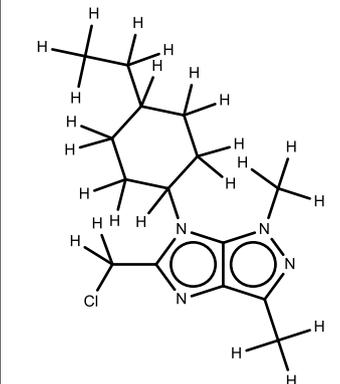
870



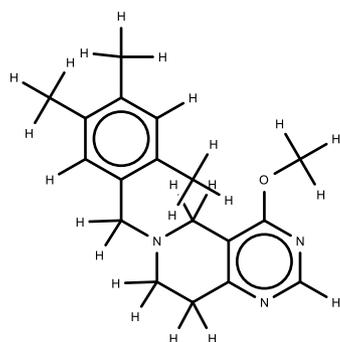
871



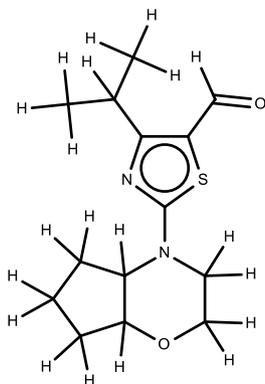
872



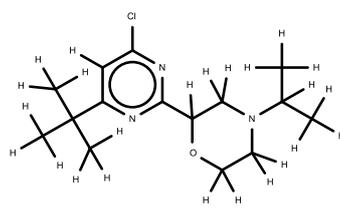
873



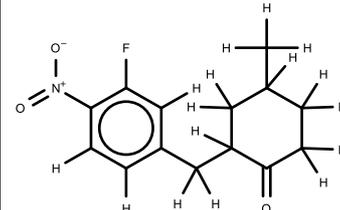
874



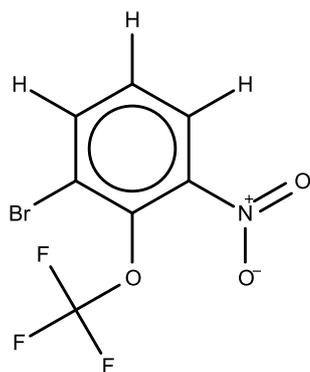
875



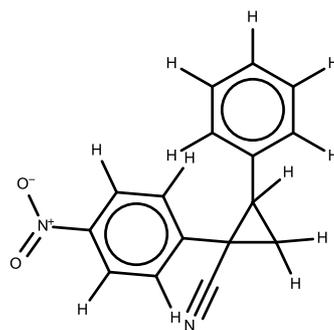
876



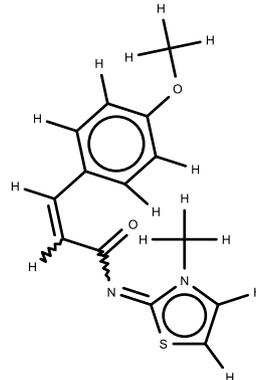
877



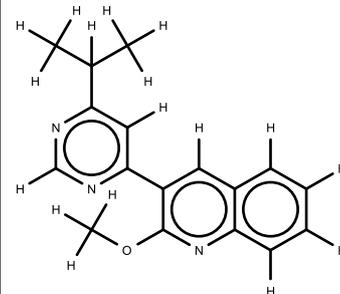
878



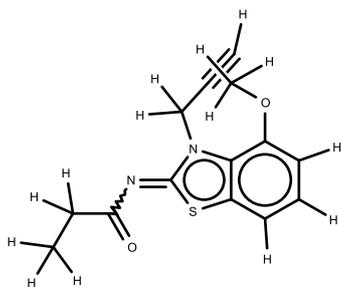
879



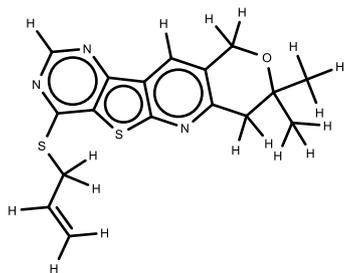
880



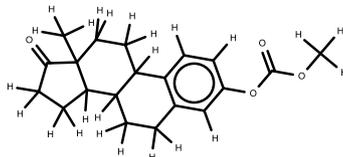
881



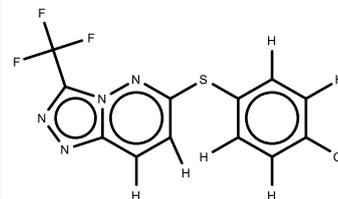
882



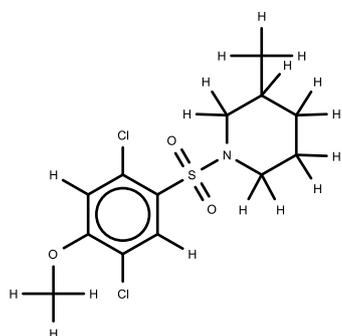
883



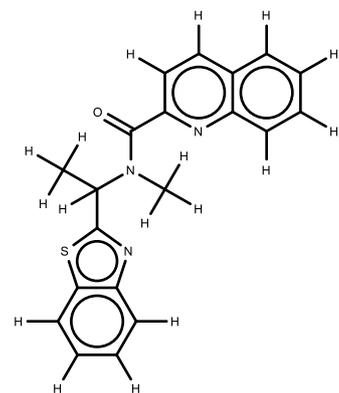
884



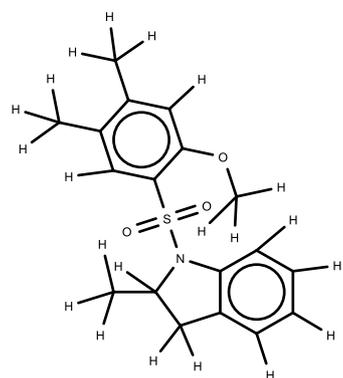
885



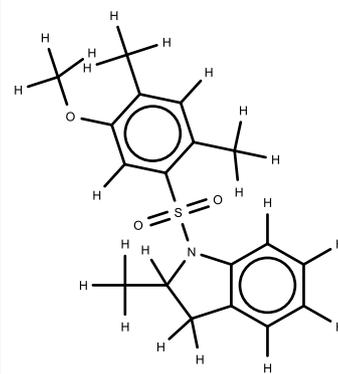
886



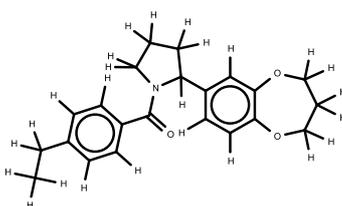
887



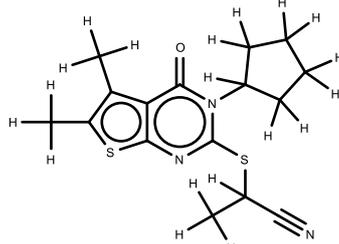
888



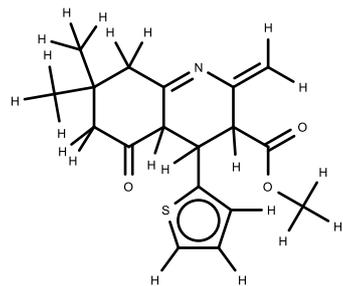
889



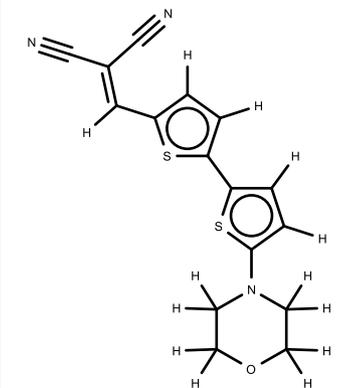
890



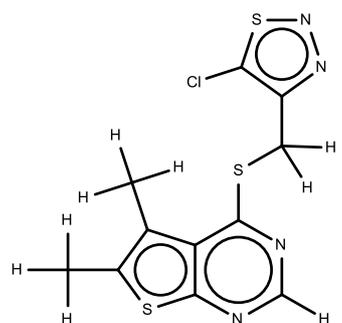
891



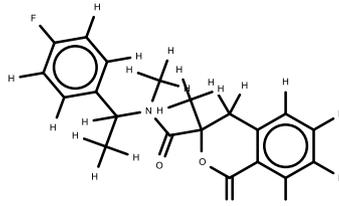
892



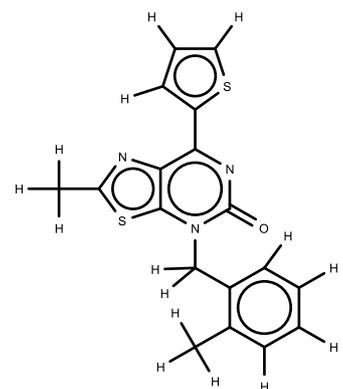
893



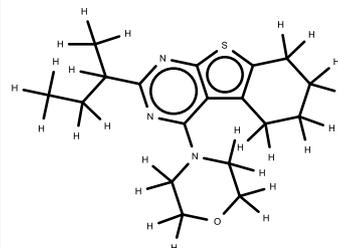
894



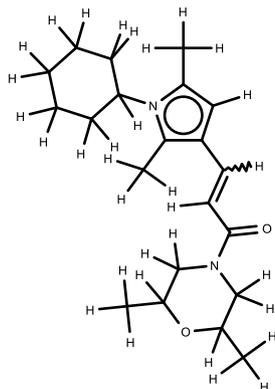
895



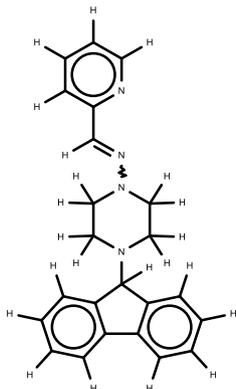
896



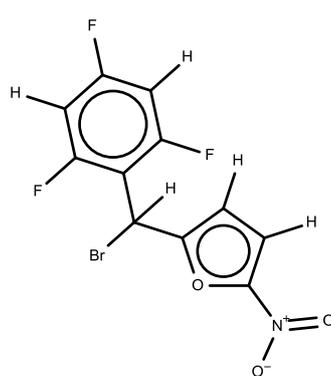
897



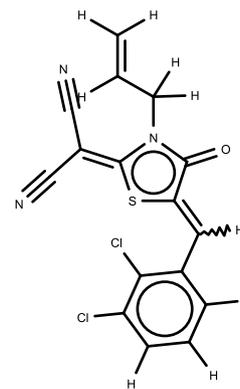
898



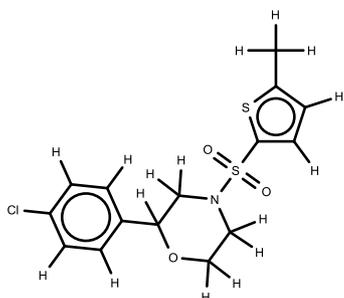
899



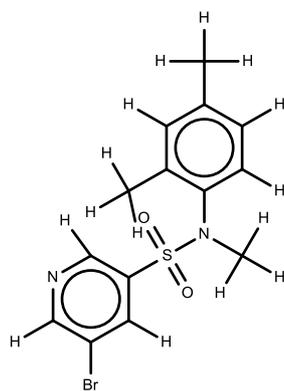
900



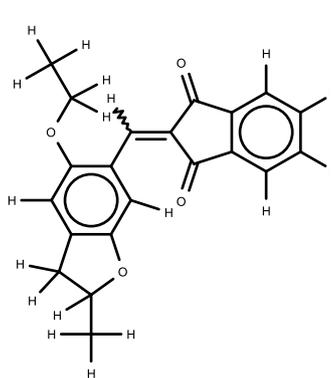
901



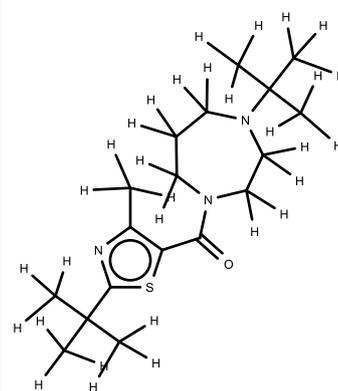
902



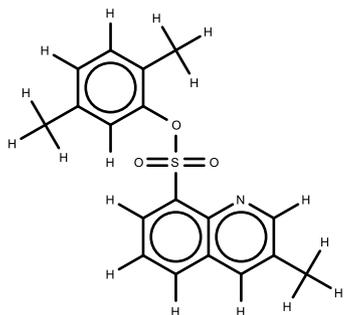
903



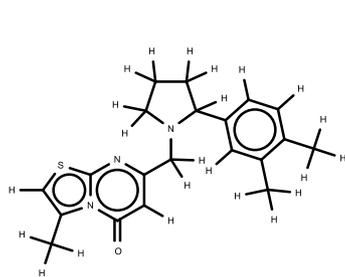
904



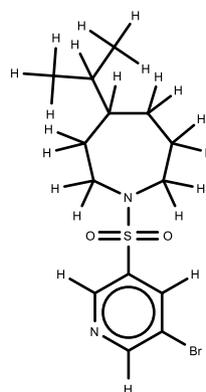
905



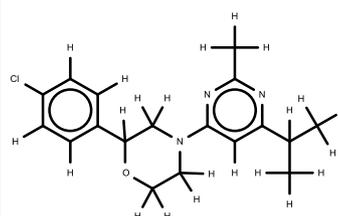
906



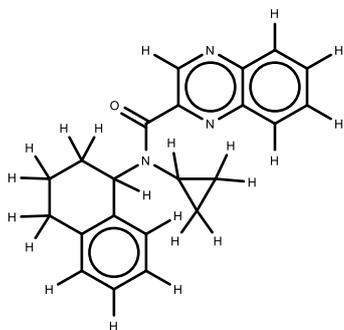
907



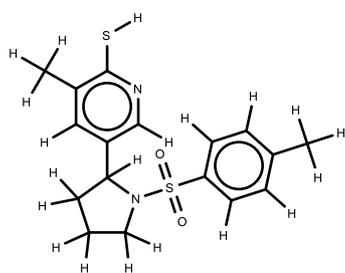
908



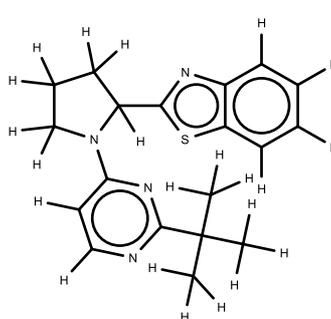
909



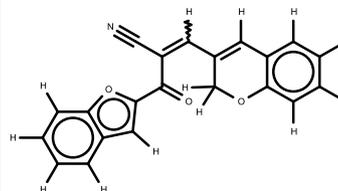
910



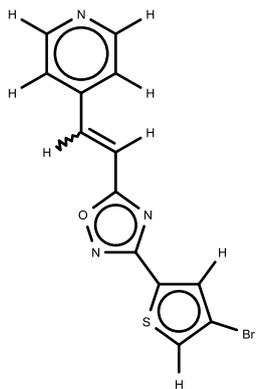
911



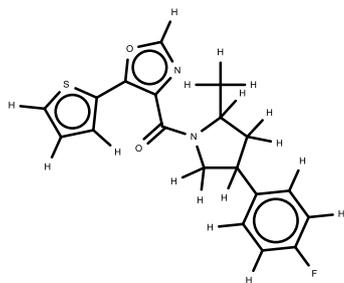
912



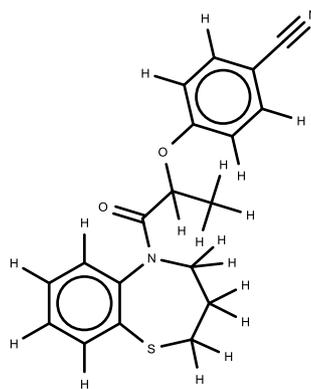
913



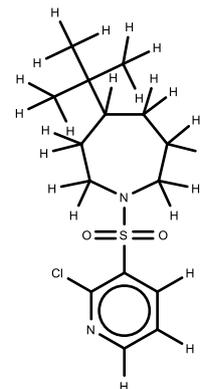
914



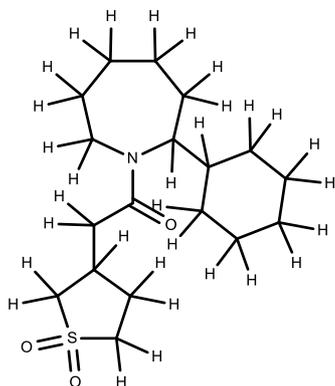
915



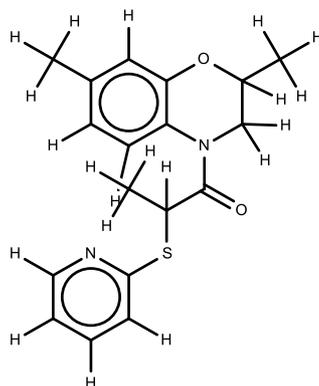
916



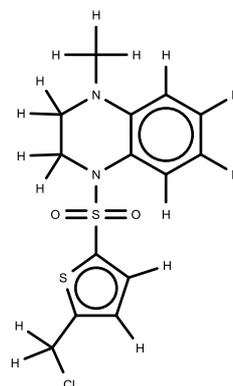
917



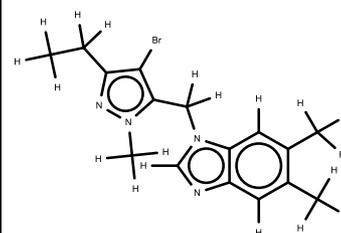
918



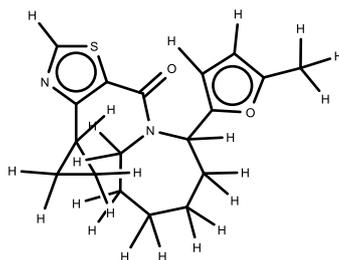
919



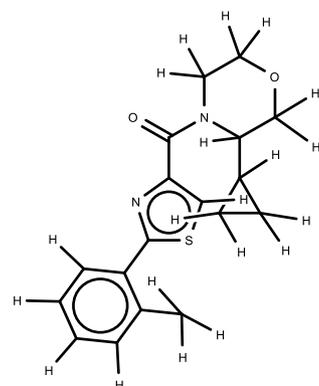
920



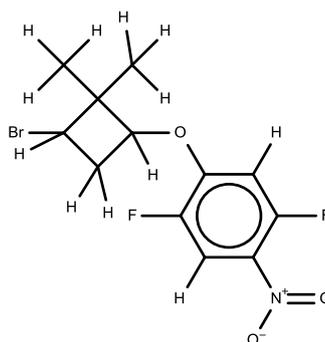
921



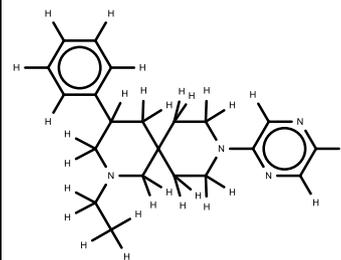
922



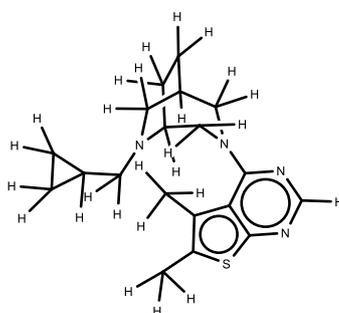
923



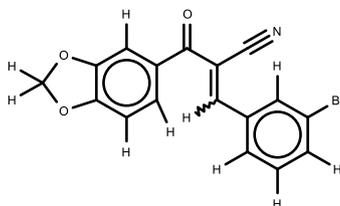
924



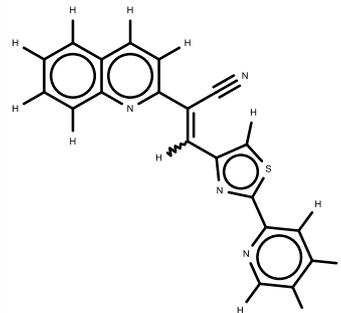
925



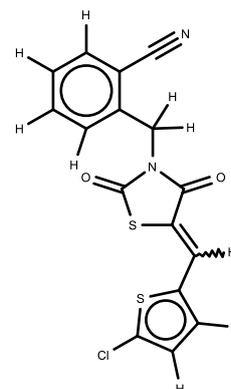
926

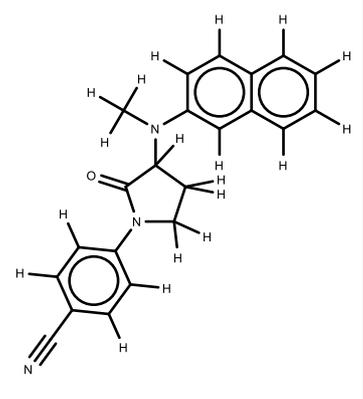
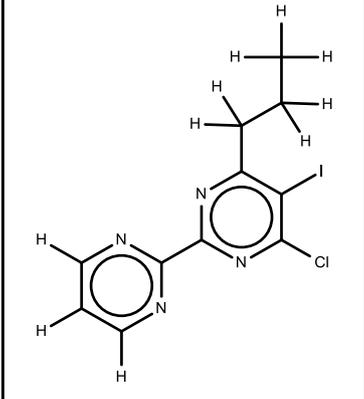
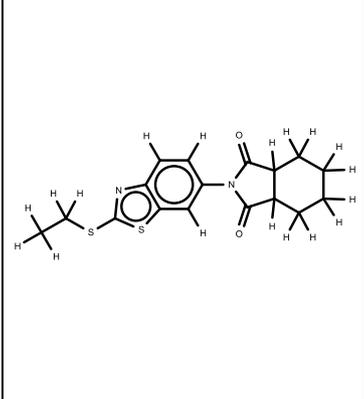
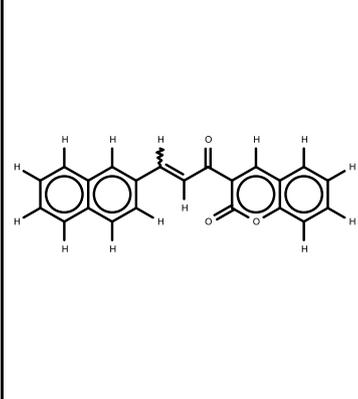
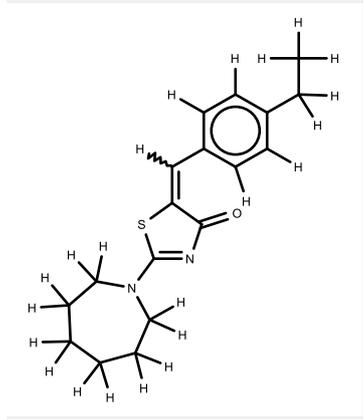
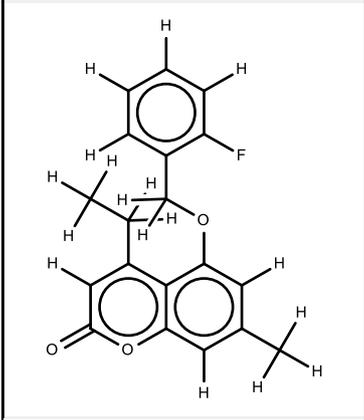
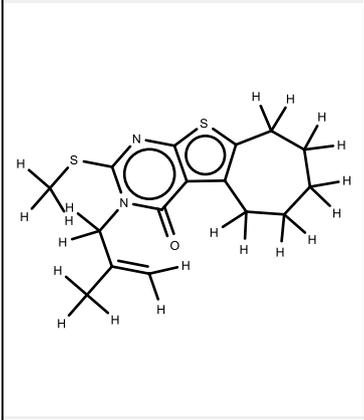
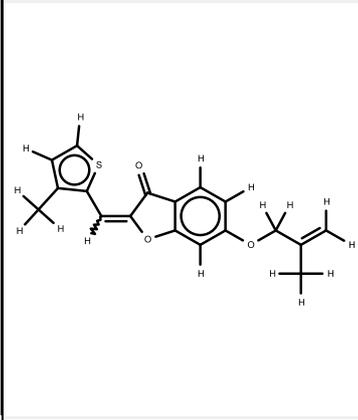
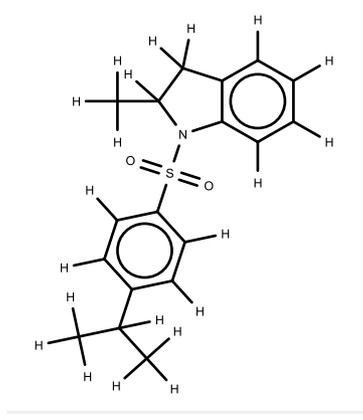
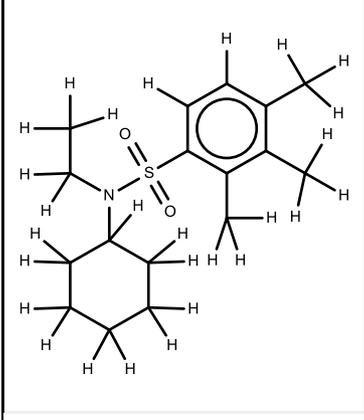
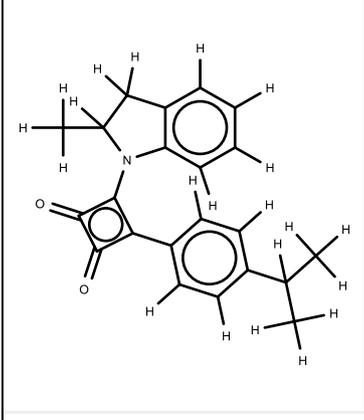
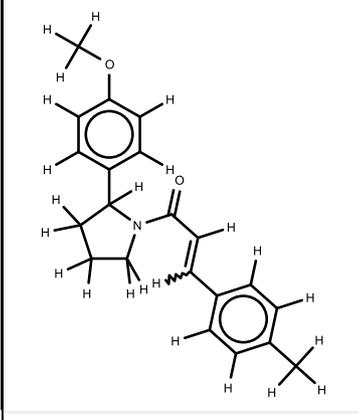
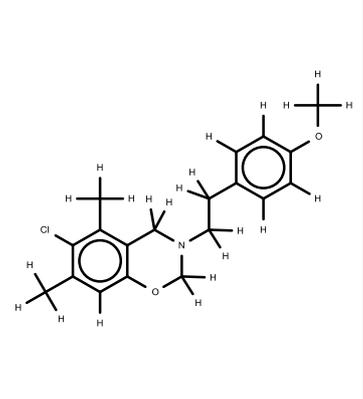
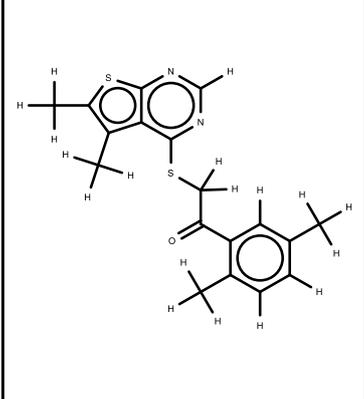
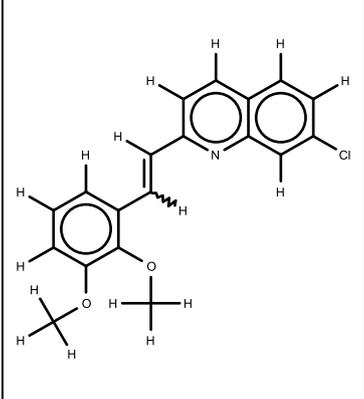
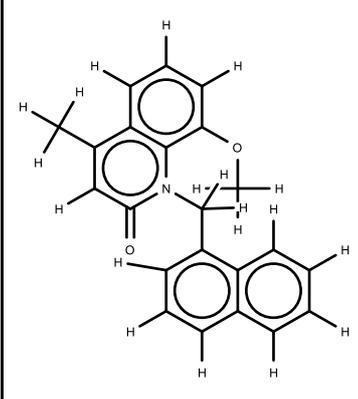


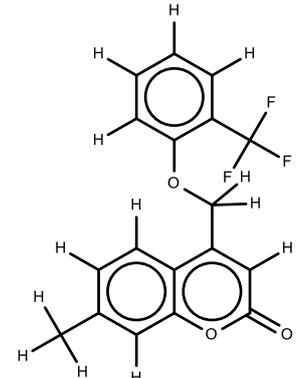
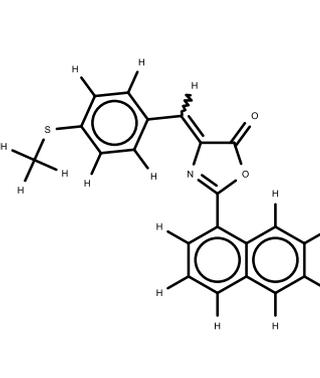
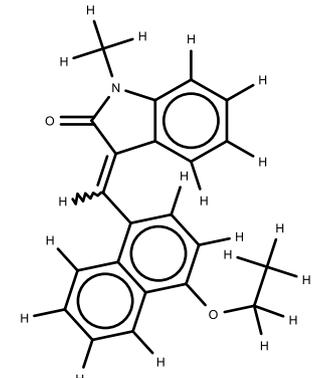
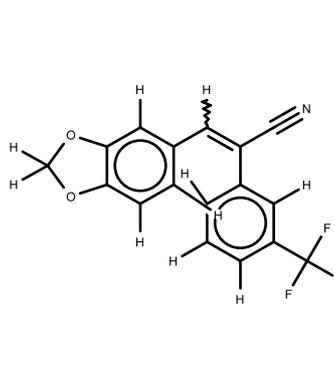
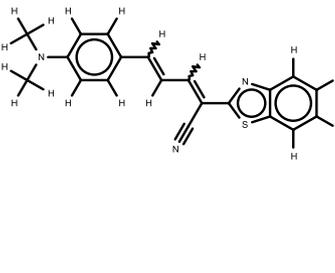
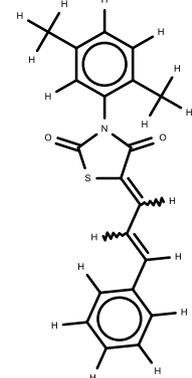
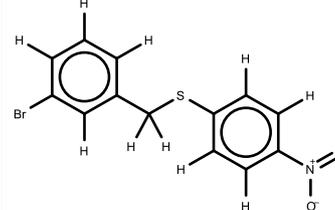
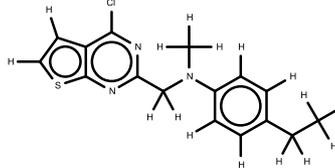
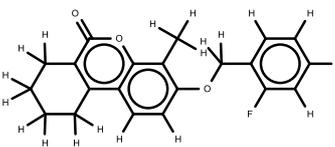
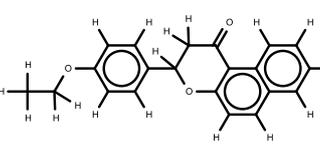
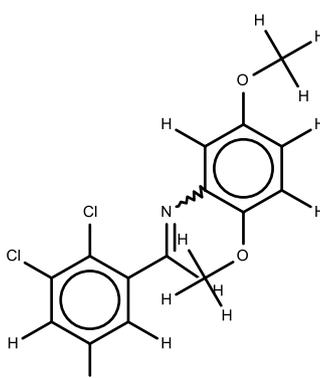
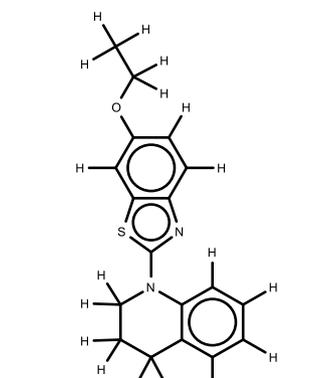
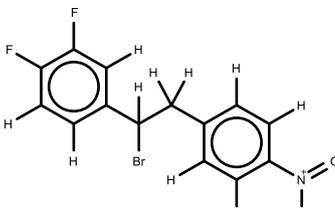
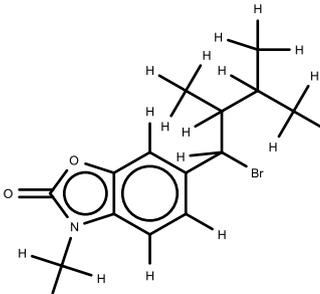
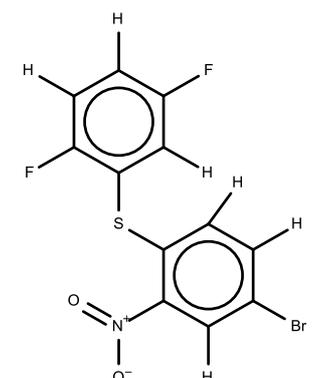
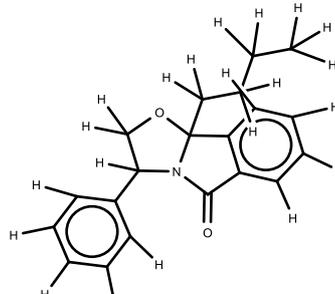
927

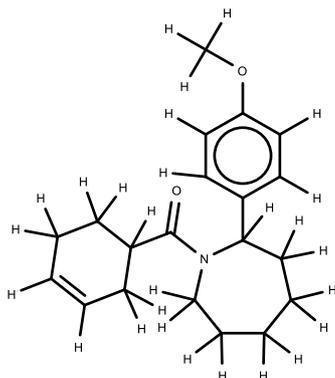
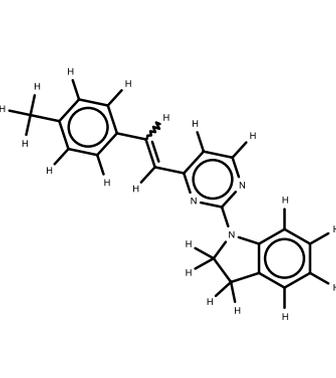
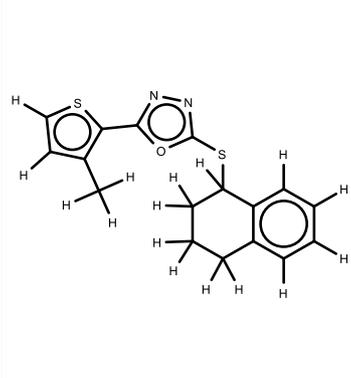
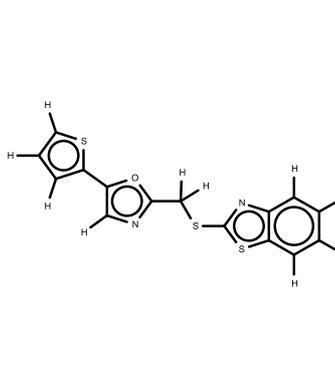
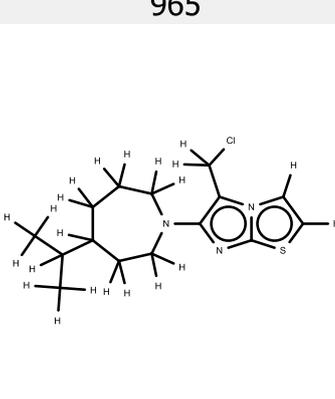
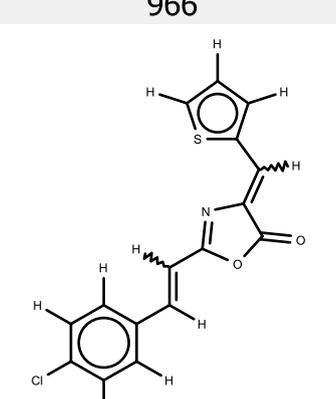
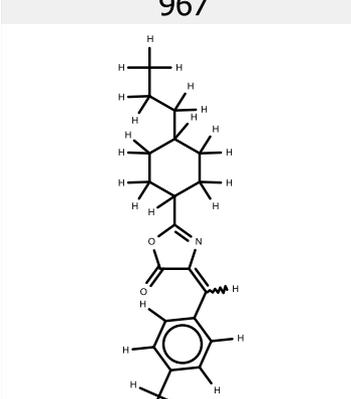
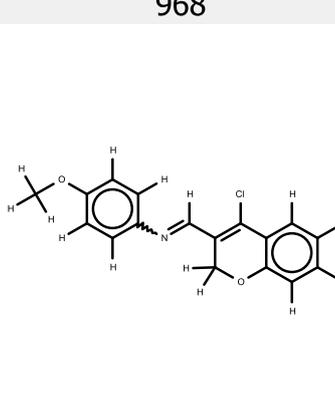
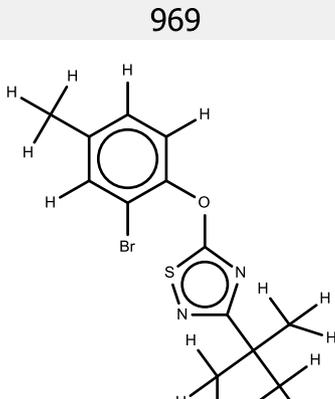
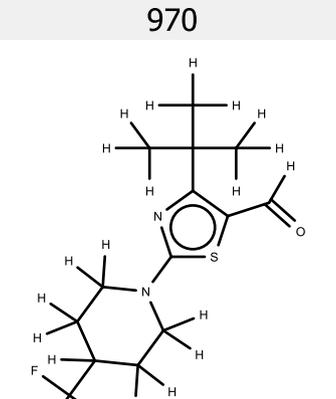
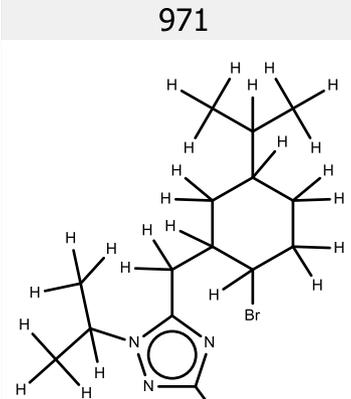
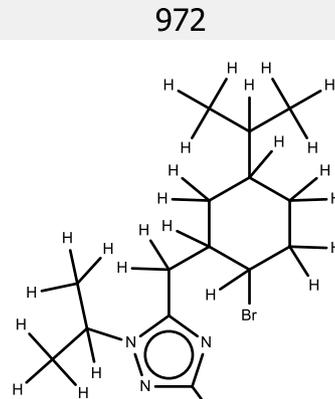
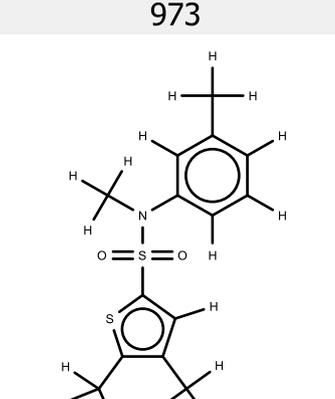
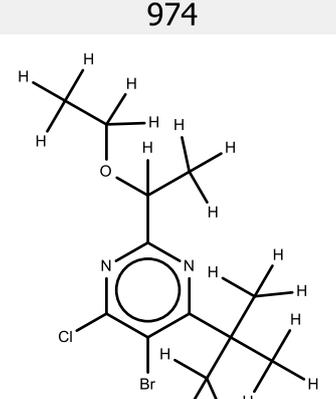
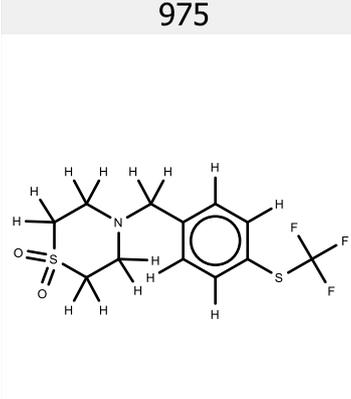
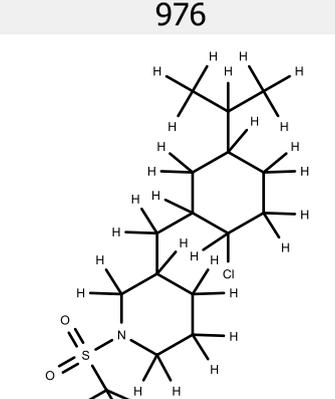


928

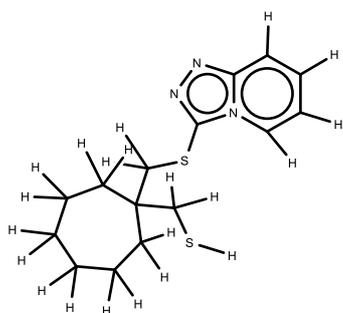


|  |   |  |   |
|--|---|--|---|
| <p style="text-align: center;"><b>929</b></p>    | <p style="text-align: center;"><b>930</b></p>    | <p style="text-align: center;"><b>931</b></p>    | <p style="text-align: center;"><b>932</b></p>    |
|  |   |  | C04268129   |
| <p style="text-align: center;"><b>933</b></p>   | <p style="text-align: center;"><b>934</b></p>   | <p style="text-align: center;"><b>935</b></p>   | <p style="text-align: center;"><b>936</b></p>   |
| C00508202  | C01164463   | C04907124  | C00545023   |
| <p style="text-align: center;"><b>937</b></p>  | <p style="text-align: center;"><b>938</b></p>  | <p style="text-align: center;"><b>939</b></p>  | <p style="text-align: center;"><b>940</b></p>  |
| C06315604  | C00475034   | C06793421  | C15575989   |
| <p style="text-align: center;"><b>941</b></p>  | <p style="text-align: center;"><b>942</b></p>  | <p style="text-align: center;"><b>943</b></p>  | <p style="text-align: center;"><b>944</b></p>  |
| C19848815  | C03527607   | C23042632  | C32530808   |

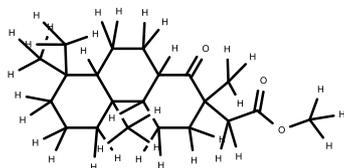
|   |  |   |  |
|---|--|---|--|
| <p style="text-align: center;"><b>945</b></p>  <p style="text-align: center;">C32854431</p>   | <p style="text-align: center;"><b>946</b></p>  <p style="text-align: center;">C00430422</p>   | <p style="text-align: center;"><b>947</b></p>  <p style="text-align: center;">C33329149</p>   | <p style="text-align: center;"><b>948</b></p>  <p style="text-align: center;">C33678745</p>   |
| <p style="text-align: center;"><b>949</b></p>  <p style="text-align: center;">C39440268</p>   | <p style="text-align: center;"><b>950</b></p>  <p style="text-align: center;">C39860186</p>   | <p style="text-align: center;"><b>951</b></p>  <p style="text-align: center;">C42203497</p>   | <p style="text-align: center;"><b>952</b></p>  <p style="text-align: center;">C50840503</p>   |
| <p style="text-align: center;"><b>953</b></p>  <p style="text-align: center;">C02954710</p> | <p style="text-align: center;"><b>954</b></p>  <p style="text-align: center;">C59382125</p> | <p style="text-align: center;"><b>955</b></p>  <p style="text-align: center;">C59518325</p> | <p style="text-align: center;"><b>956</b></p>  <p style="text-align: center;">C60244851</p> |
| <p style="text-align: center;"><b>957</b></p>  <p style="text-align: center;">C62167081</p> | <p style="text-align: center;"><b>958</b></p>  <p style="text-align: center;">C70126696</p> | <p style="text-align: center;"><b>959</b></p>  <p style="text-align: center;">C70139400</p> | <p style="text-align: center;"><b>960</b></p>  <p style="text-align: center;">C38881277</p> |

|  |   |  |   |
|--|---|--|---|
| <p>961</p>  <p>C52987730</p>   | <p>962</p>  <p>C72443623</p>   | <p>963</p>  <p>C78546456</p>   | <p>964</p>  <p>C78583039</p>   |
| <p>965</p>  <p>C79352133</p>   | <p>966</p>  <p>C00431095</p>   | <p>967</p>  <p>C00430081</p>   | <p>968</p>  <p>C00108480</p>   |
| <p>969</p>  <p>C87640102</p> | <p>970</p>  <p>C88438711</p> | <p>971</p>  <p>C88619222</p> | <p>972</p>  <p>C88619223</p> |
| <p>973</p>  <p>C88939374</p> | <p>974</p>  <p>C90748760</p> | <p>975</p>  <p>C93048007</p> | <p>976</p>  <p>C93701096</p> |

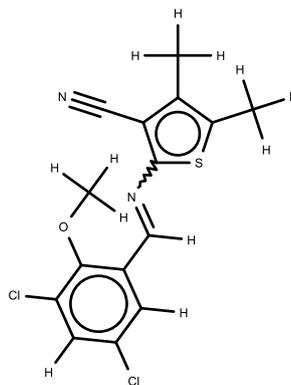
977



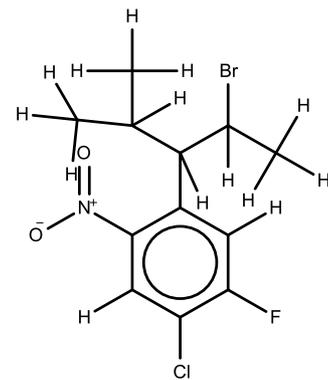
978



979



980



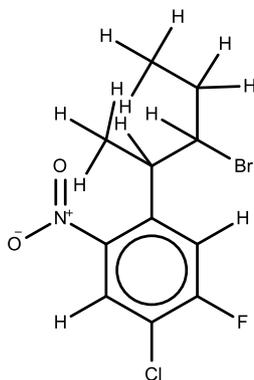
C94589508

C96346627

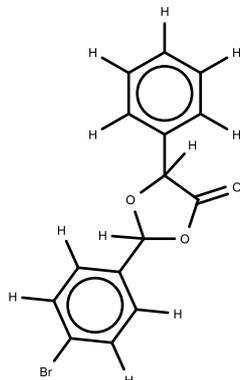
C13563737

C97362740

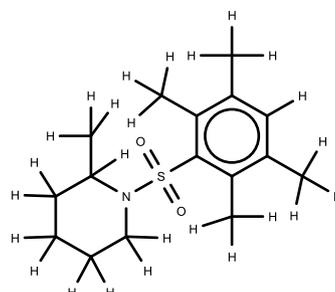
981



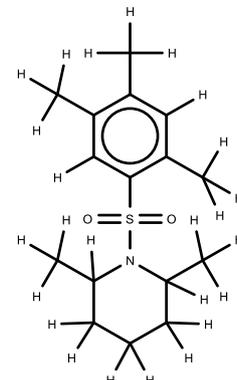
982



983

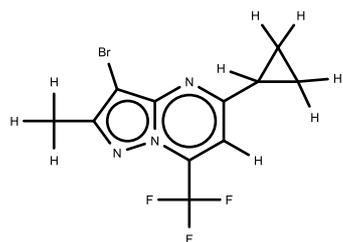


984

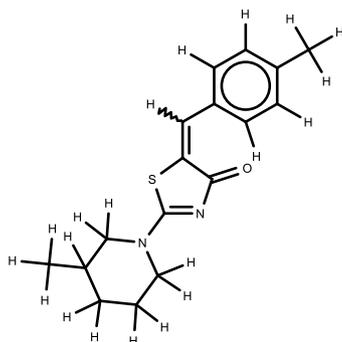


C97362898

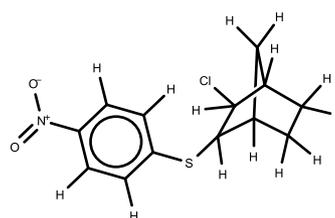
985



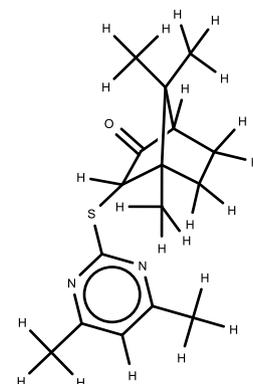
986



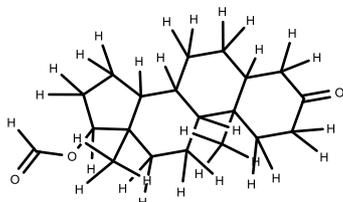
987



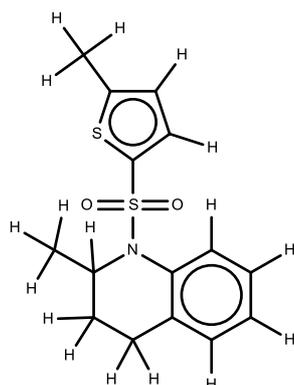
988



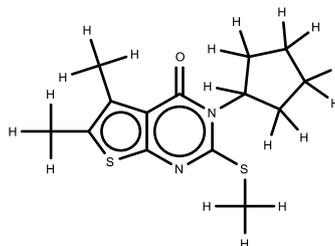
989



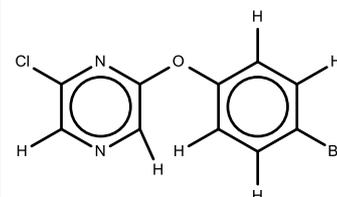
990



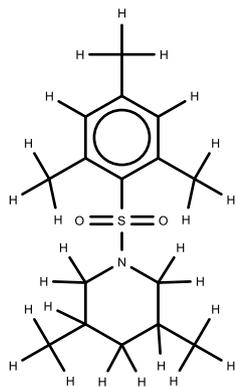
991



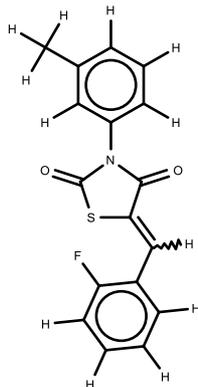
992



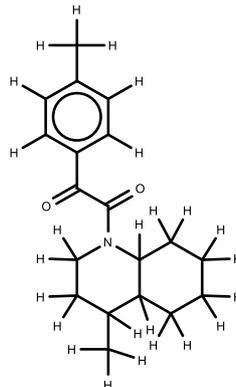
993



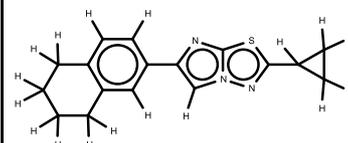
994



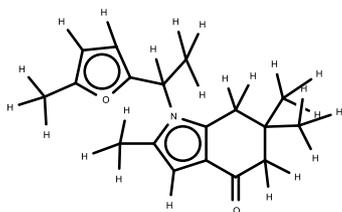
995



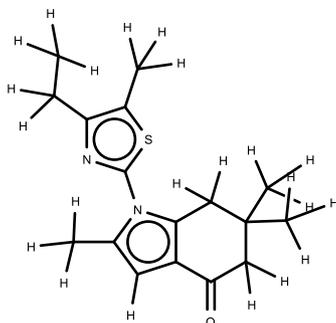
996



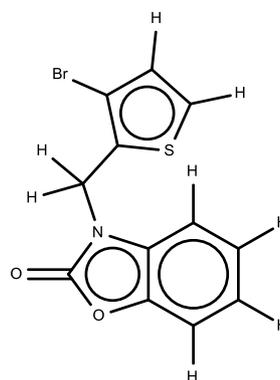
997



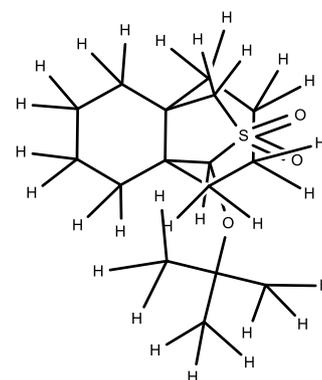
998



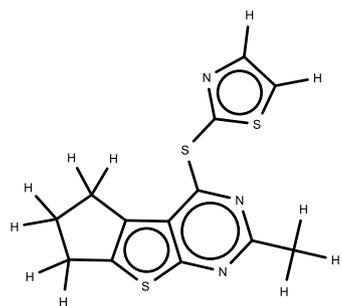
999



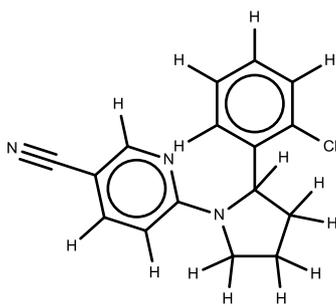
1000



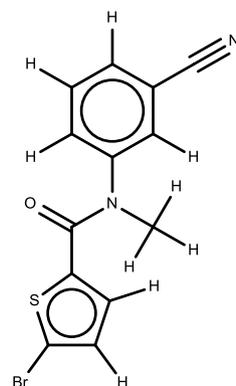
1001



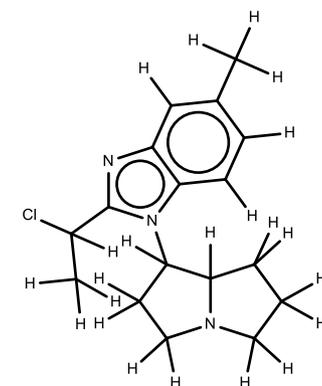
1002



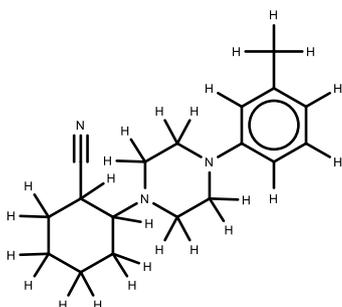
1003



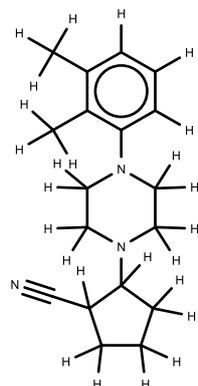
1004



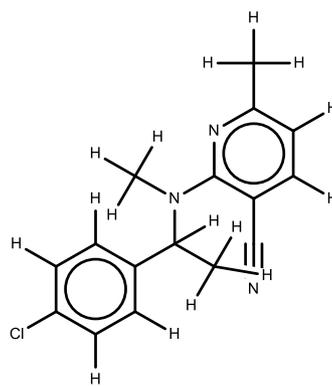
1005



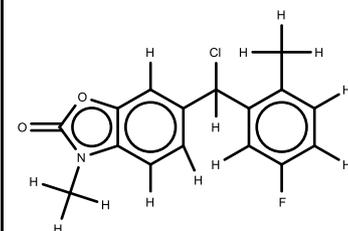
1006



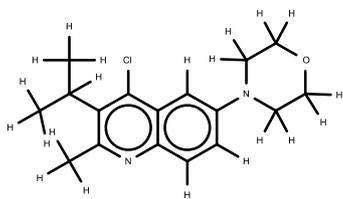
1007



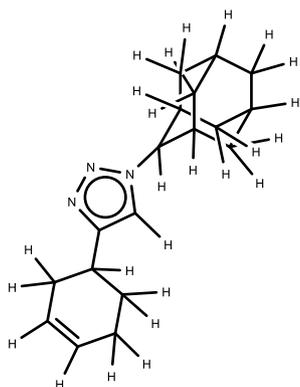
1008



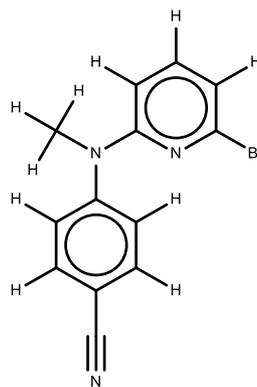
1009



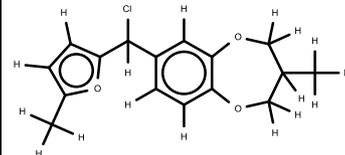
1010



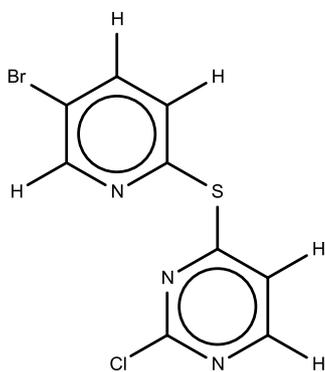
1011



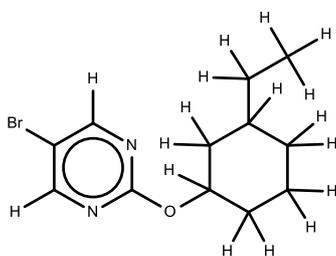
1012



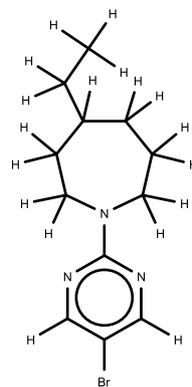
1013



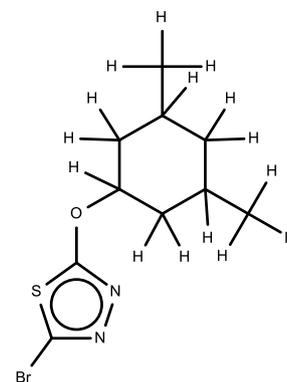
1014



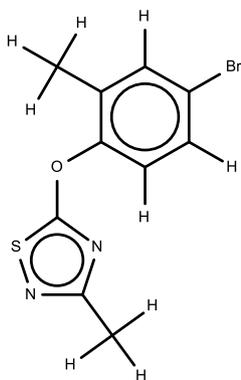
1015



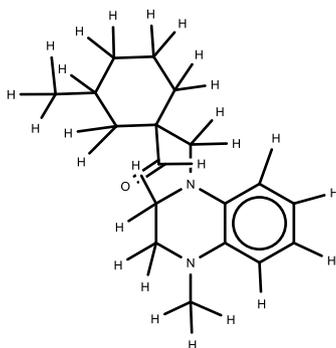
1016



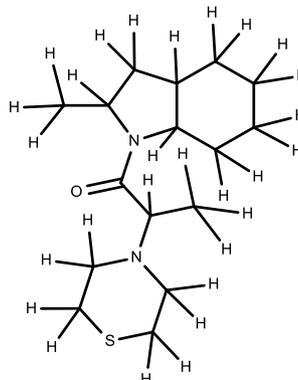
1017



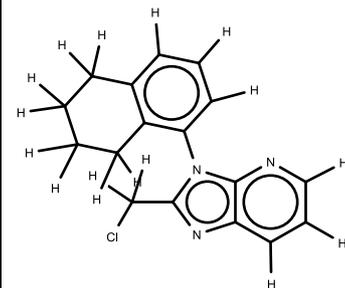
1018



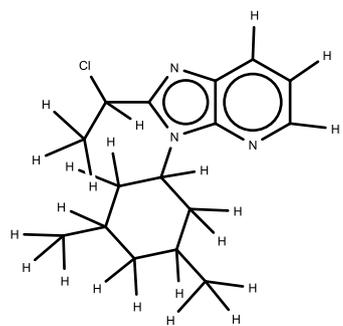
1019



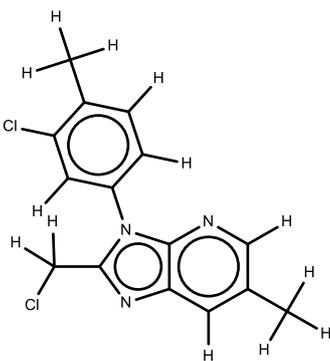
1020



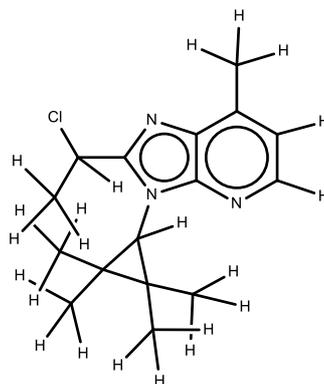
1021



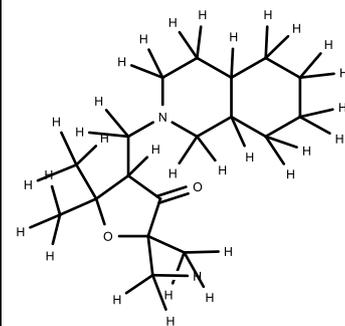
1022



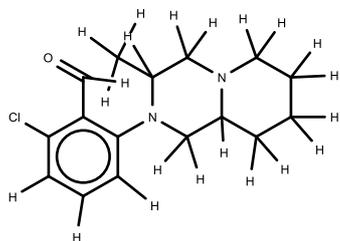
1023



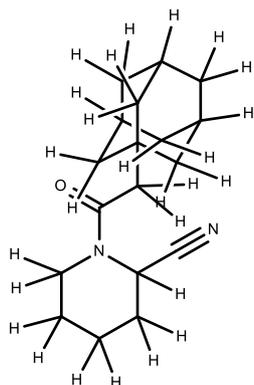
1024



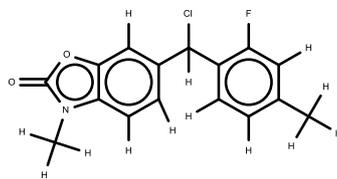
1025



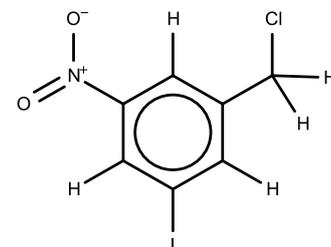
1026



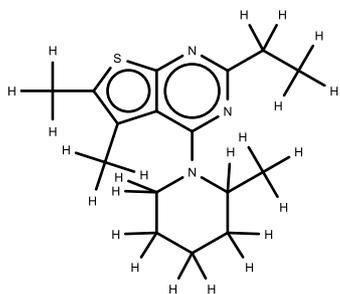
1027



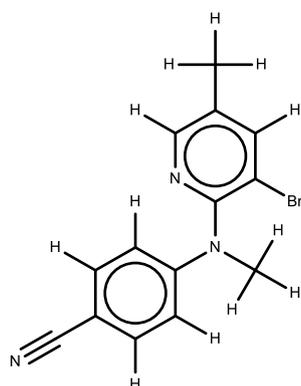
1028



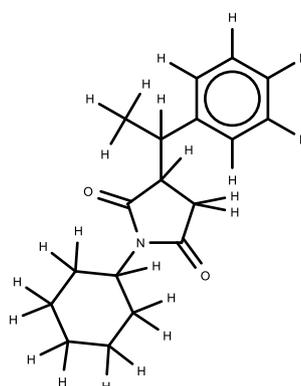
1029



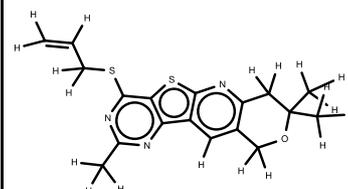
1030



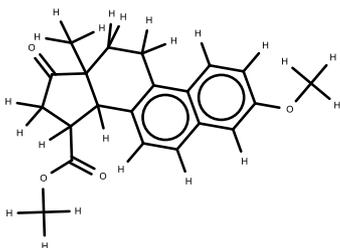
1031



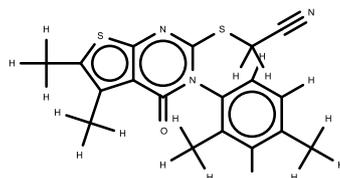
1032



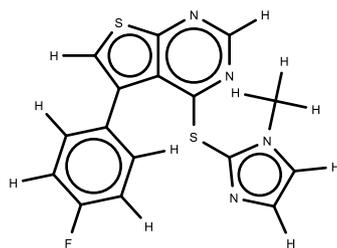
1033



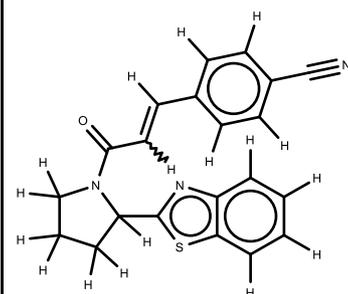
1034



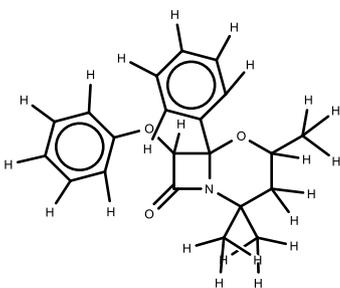
1035



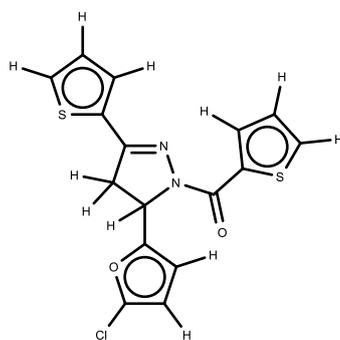
1036



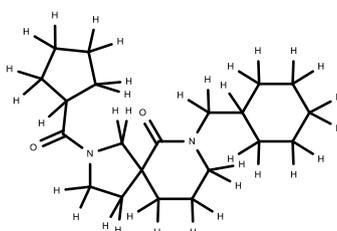
1037



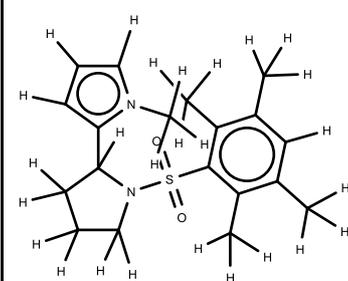
1038



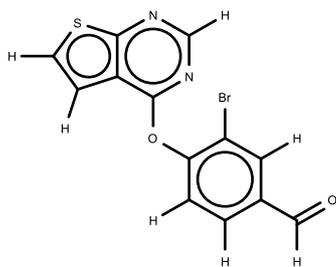
1039



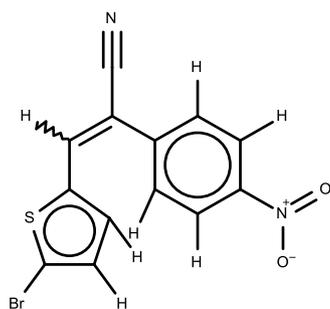
1040



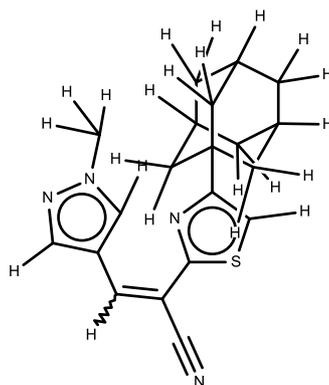
1041



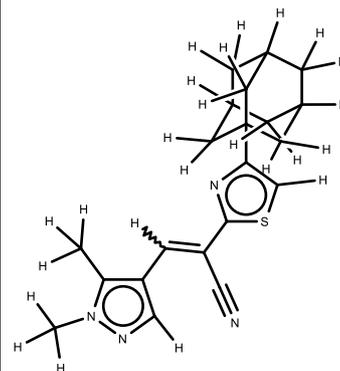
1042



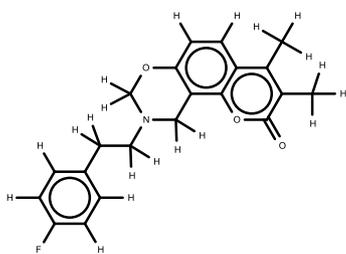
1043



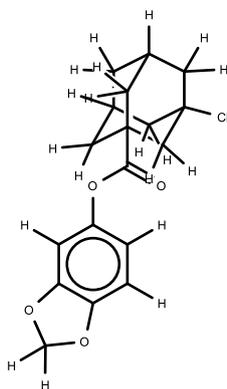
1044



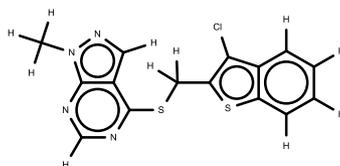
1045



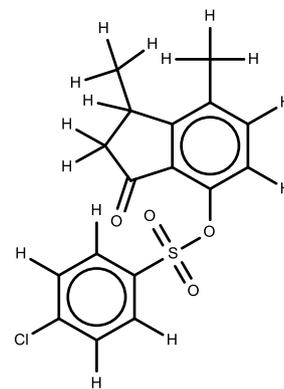
1046



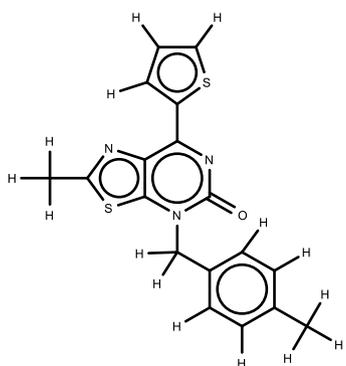
1047



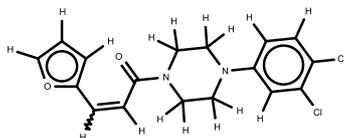
1048



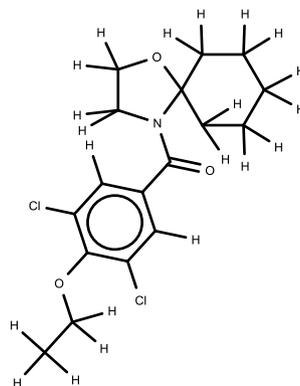
1049



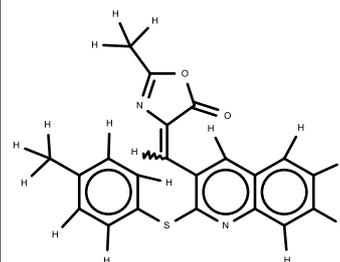
1050



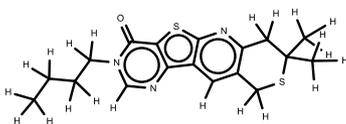
1051



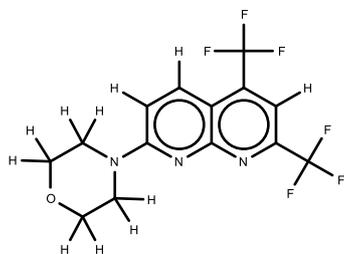
1052



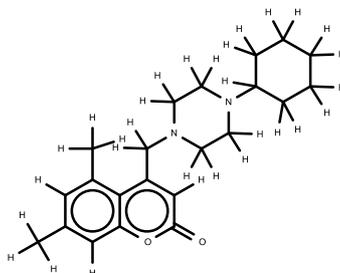
1053



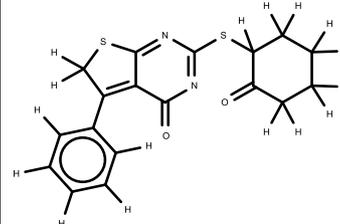
1054



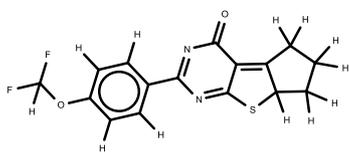
1055



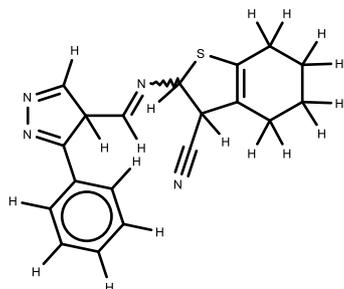
1056



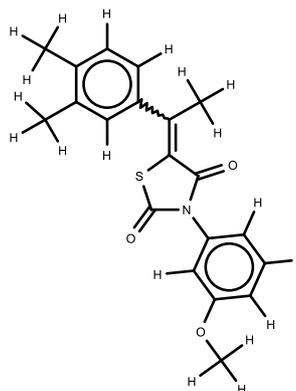
1057



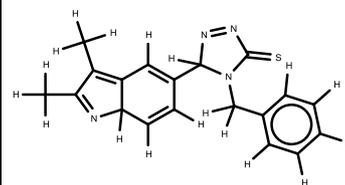
1058



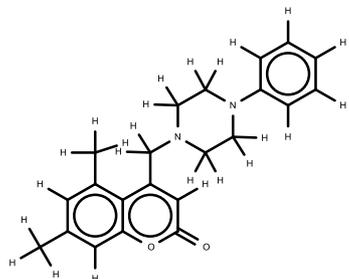
1059



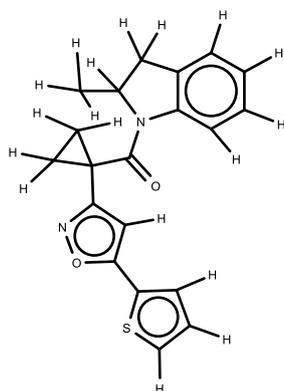
1060



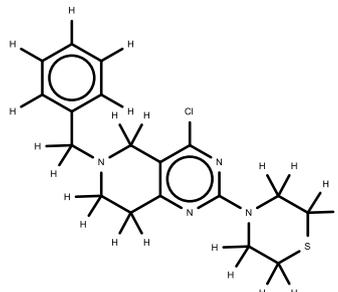
1061



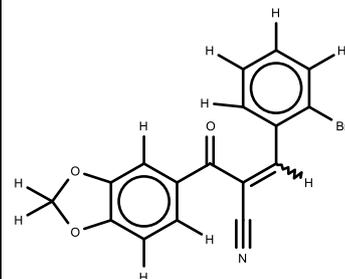
1062



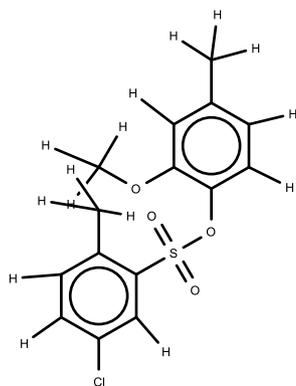
1063



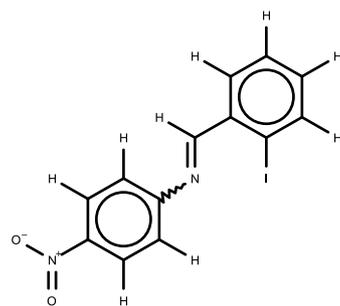
1064



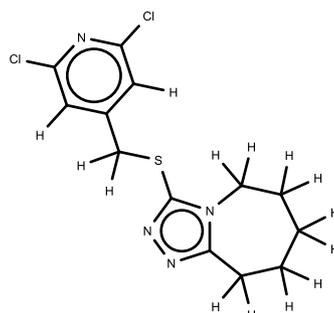
1065



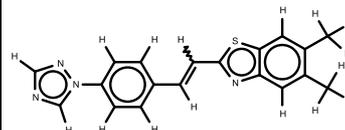
1066



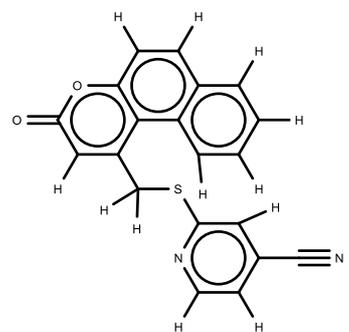
1067



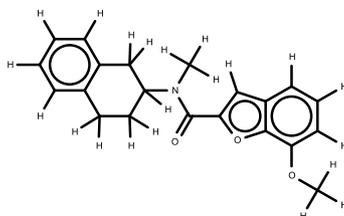
1068



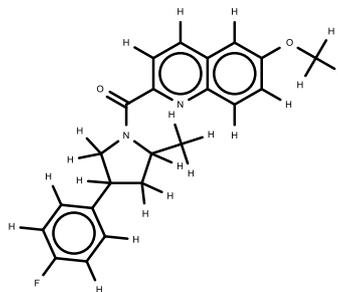
1069



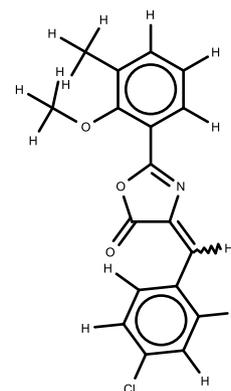
1070



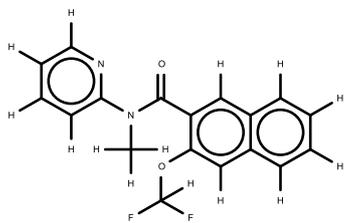
1071



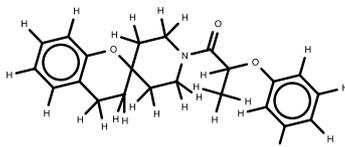
1072



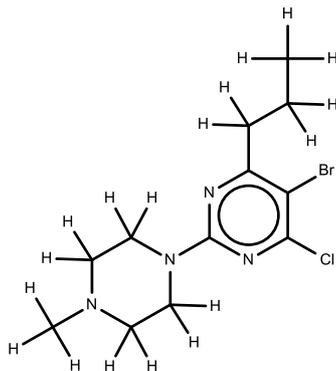
1073



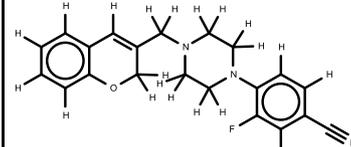
1074



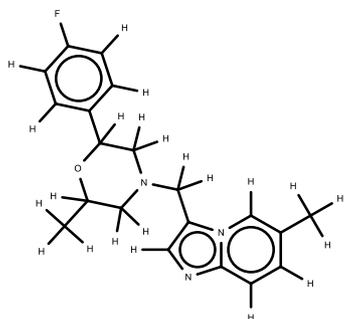
1075



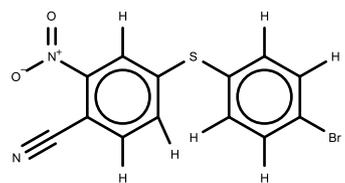
1076



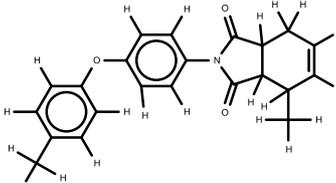
1077



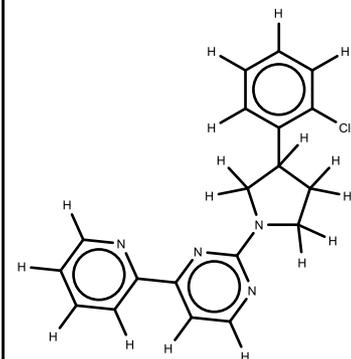
1078



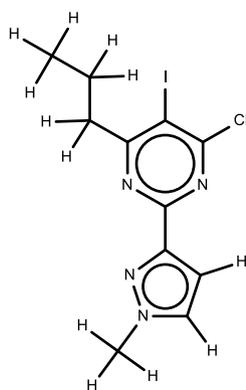
1079



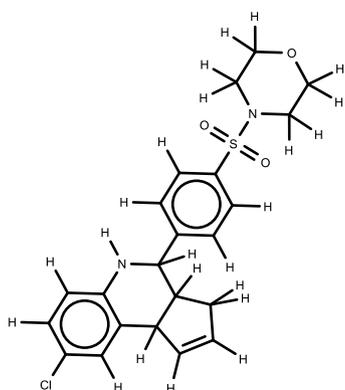
1080



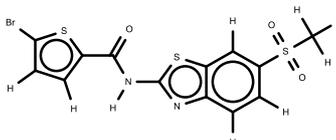
1081



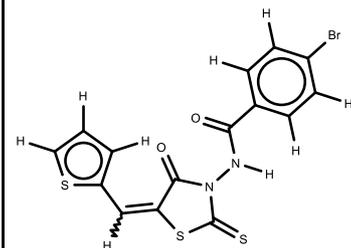
1082



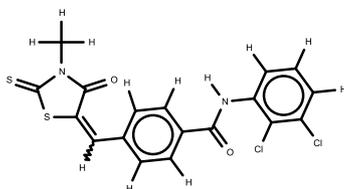
1083



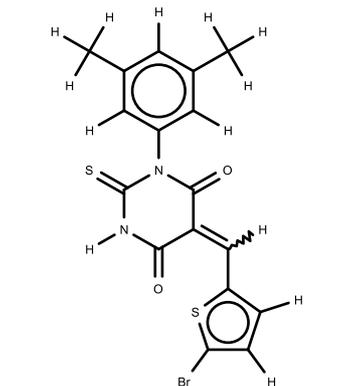
1084



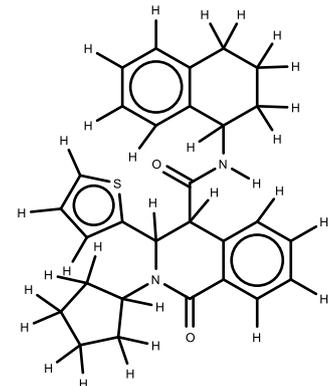
1085



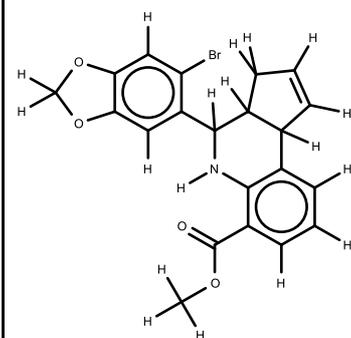
1086



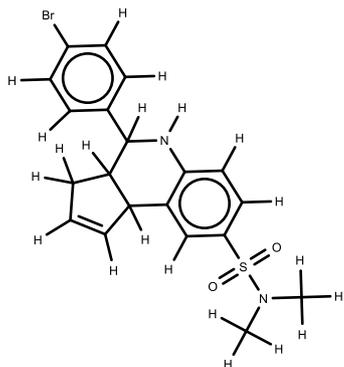
1087



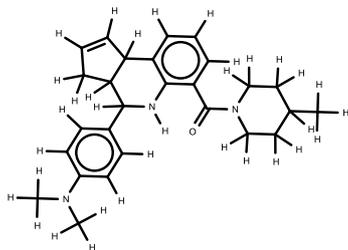
1088



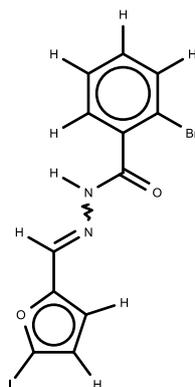
1089



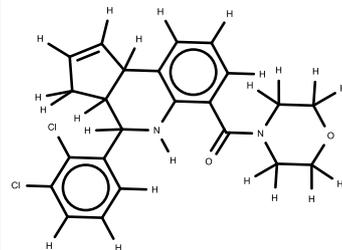
1090



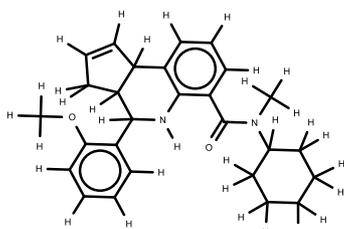
1091



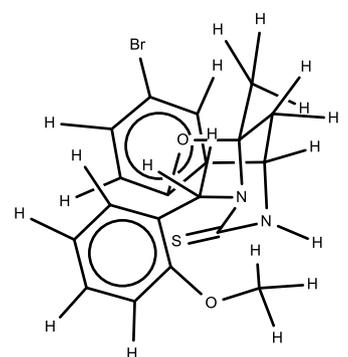
1092



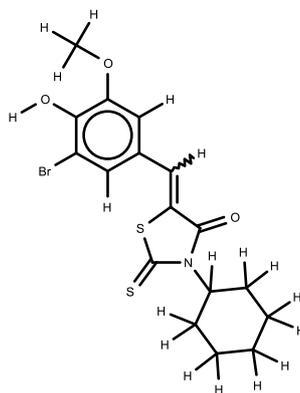
1093



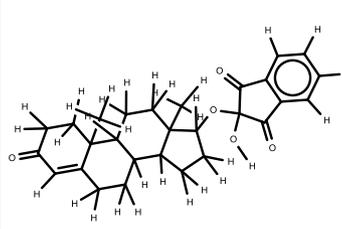
1094



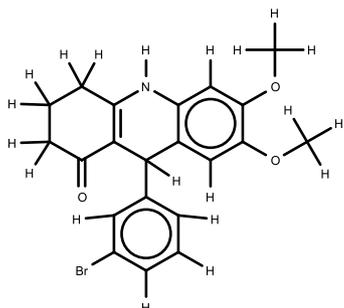
1095



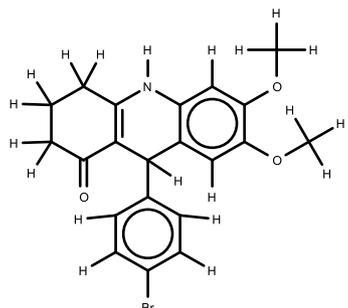
1096



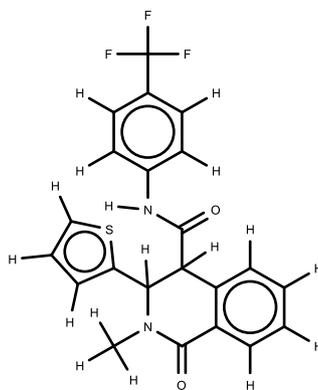
1097



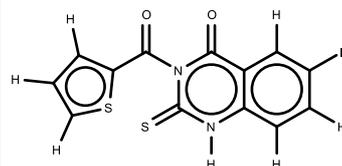
1098



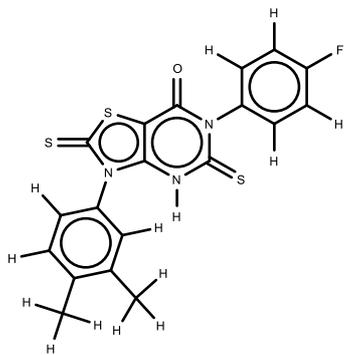
1099



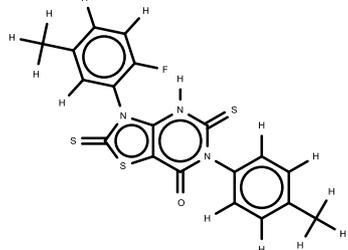
1100



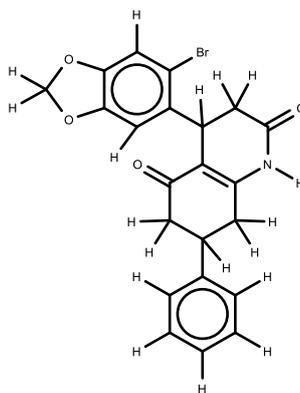
1101



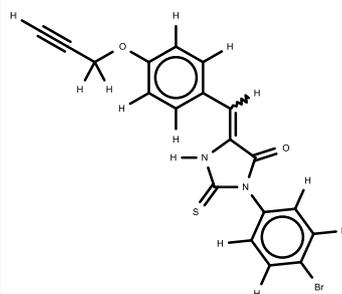
1102



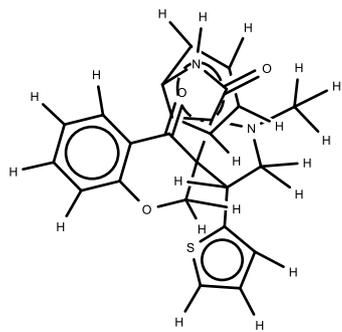
1103



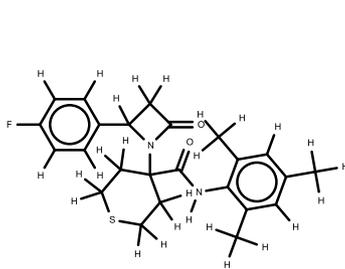
1104



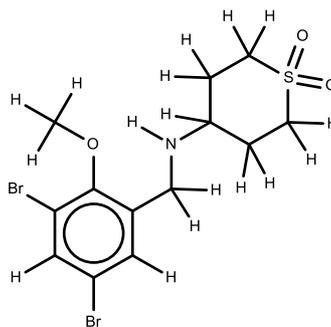
1105



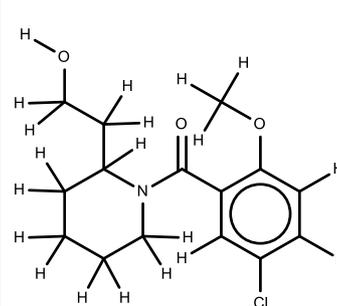
1106



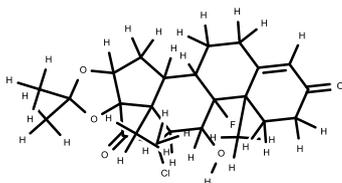
1107



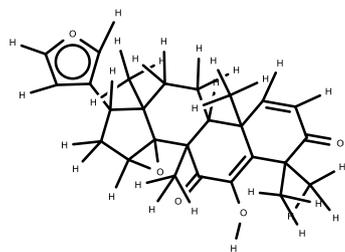
1108



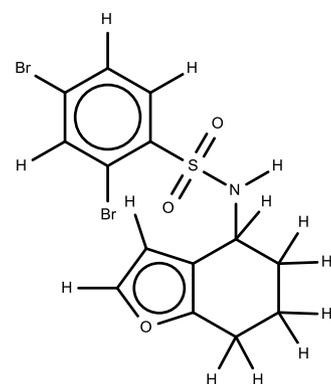
1109



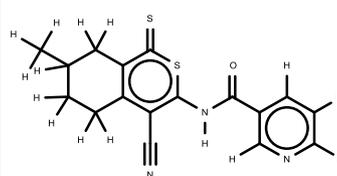
1110



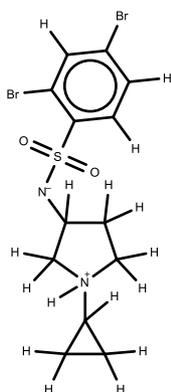
1111



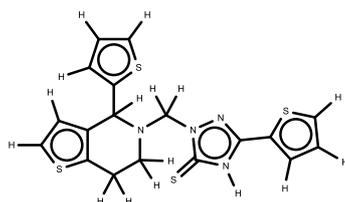
1112



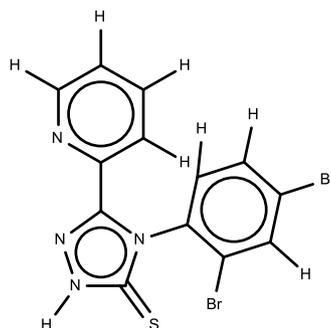
1113



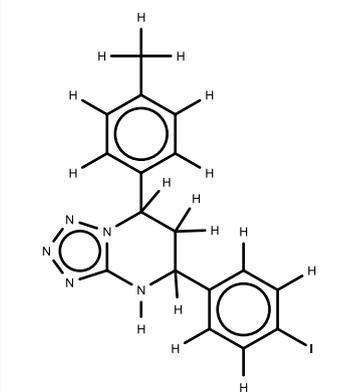
1114



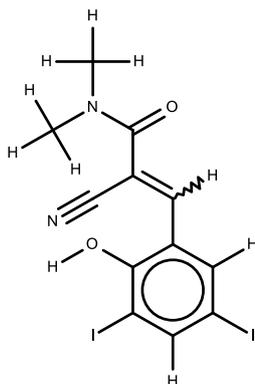
1115



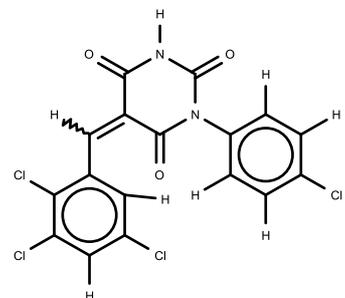
1116



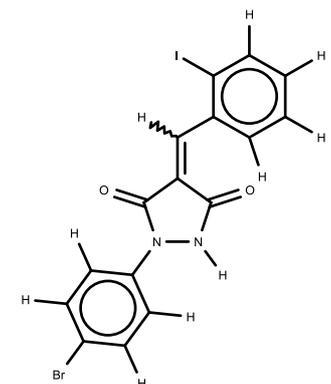
1117



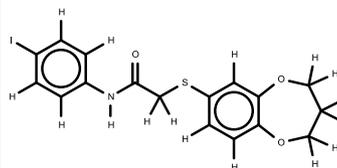
1118



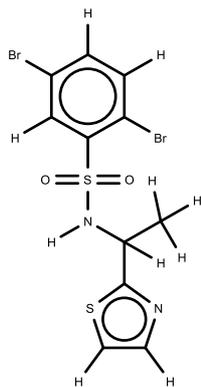
1119



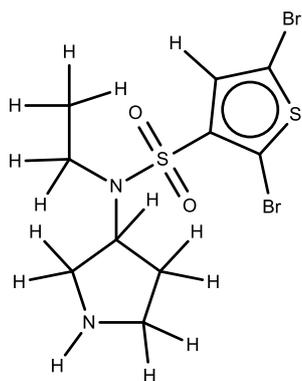
1120



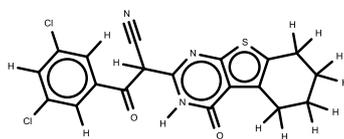
1121



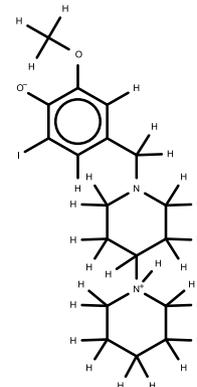
1122



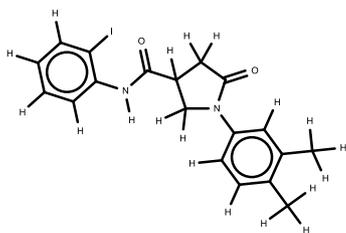
1123



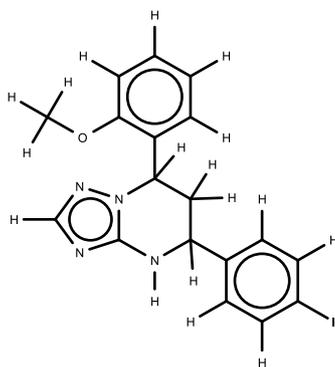
1124



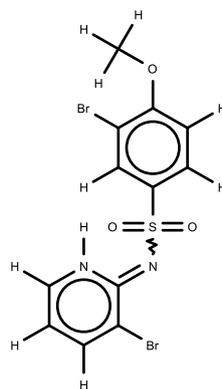
1125



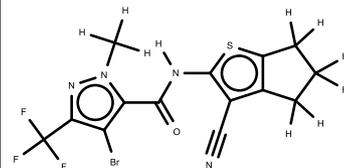
1126



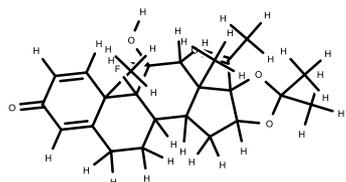
1127



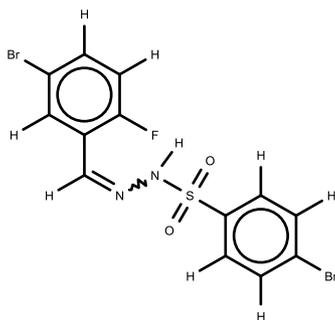
1128



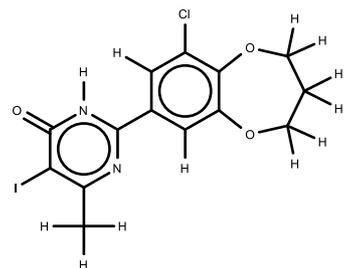
1129



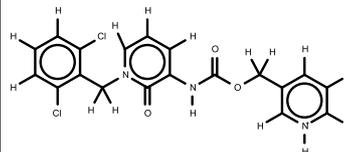
1130



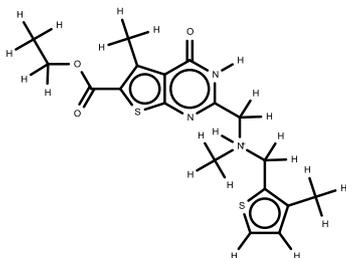
1131



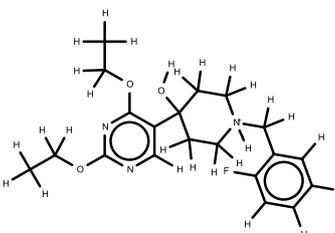
1132



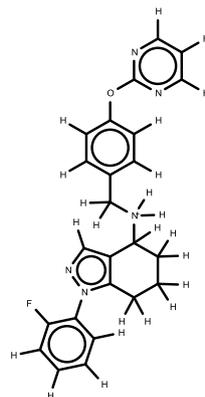
1133



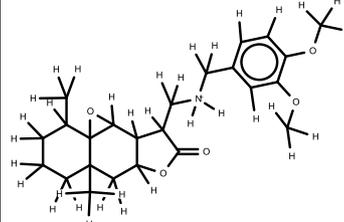
1134



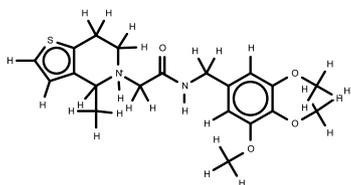
1135



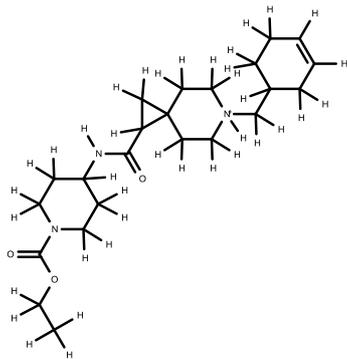
1136



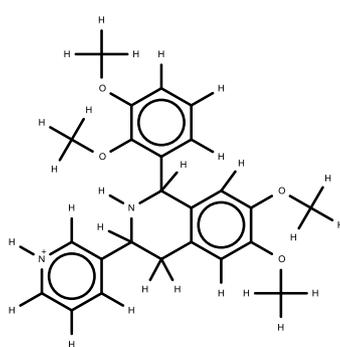
1137



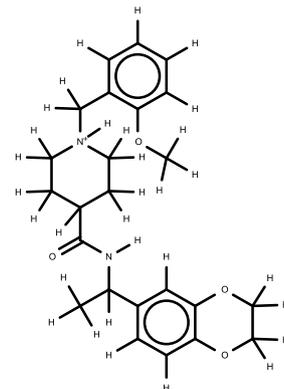
1138



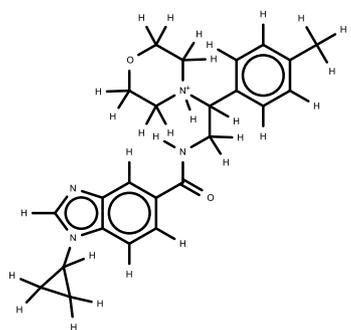
1139



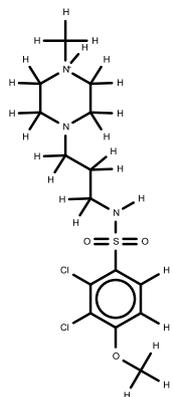
1140



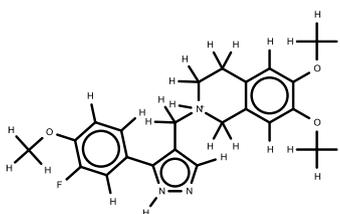
1141



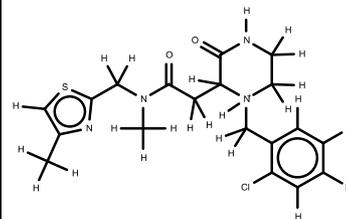
1142



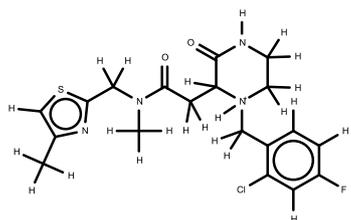
1143



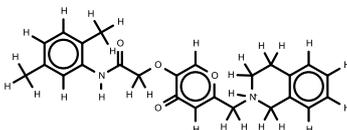
1144



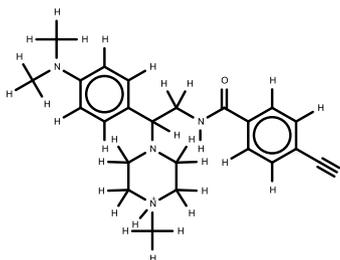
1145



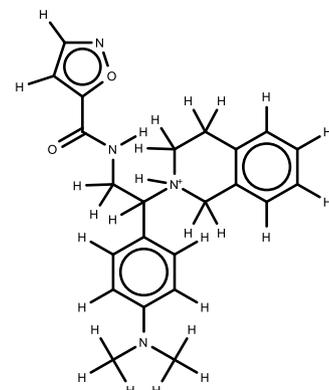
1146



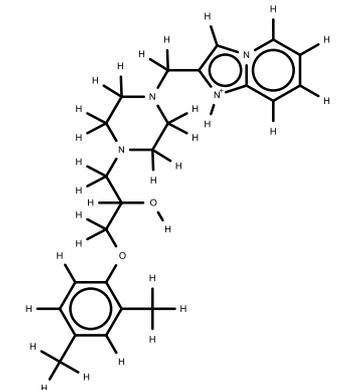
1147



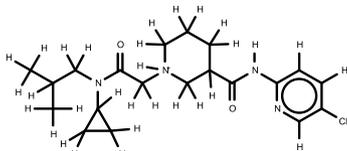
1148



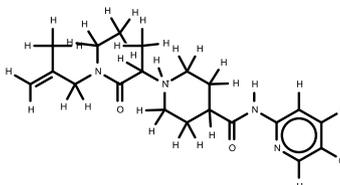
1149



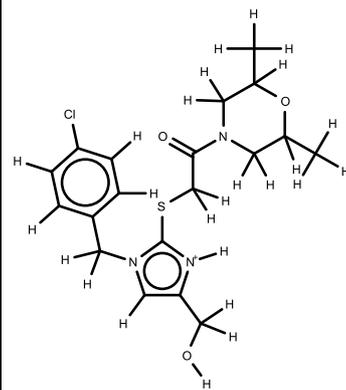
1150



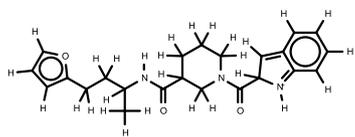
1151



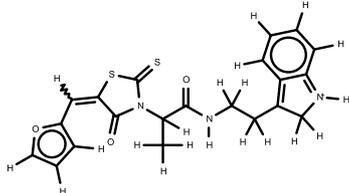
1152



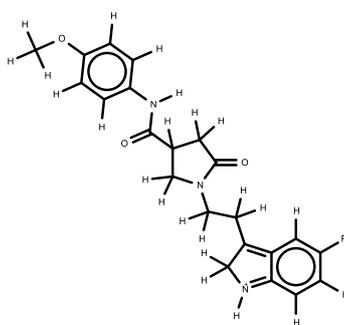
1153



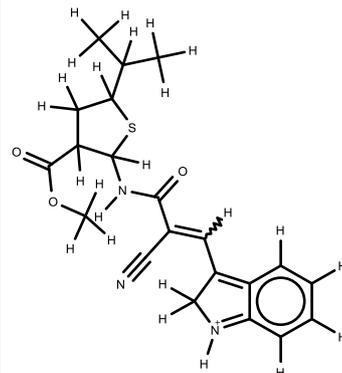
1154



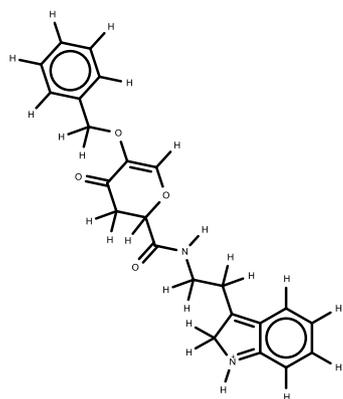
1155



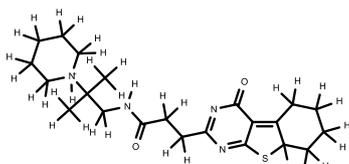
1156



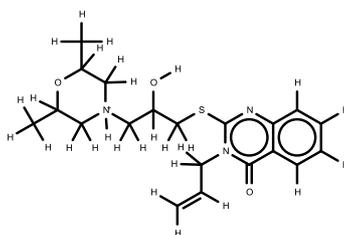
1157



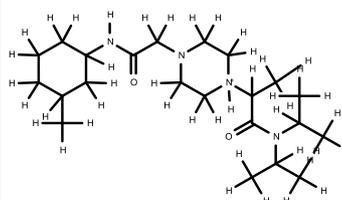
1158



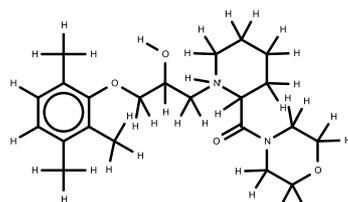
1159



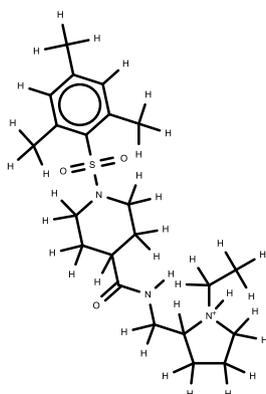
1160



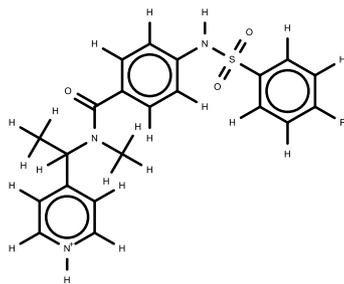
1161



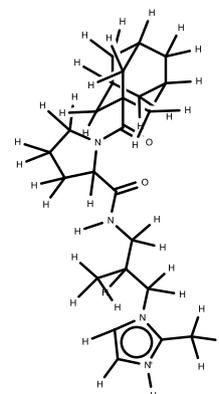
1162



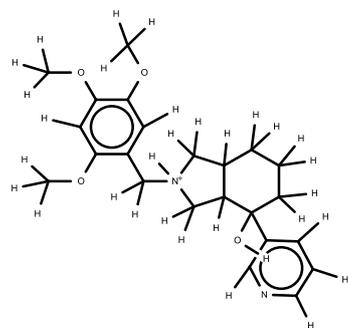
1163



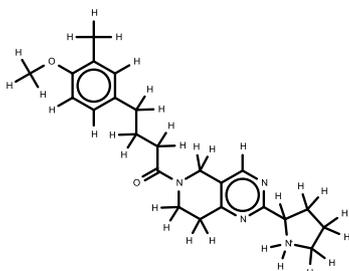
1164



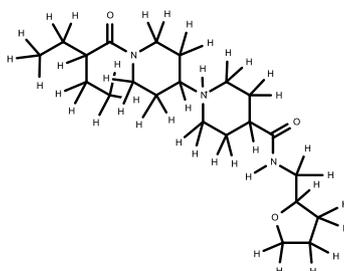
1165



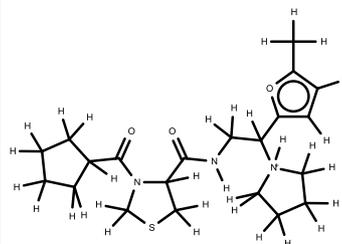
1166



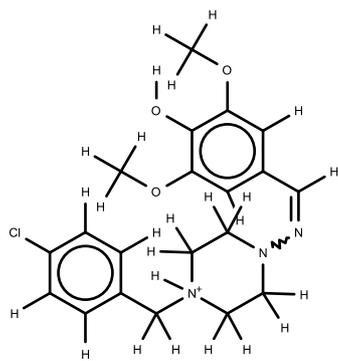
1167



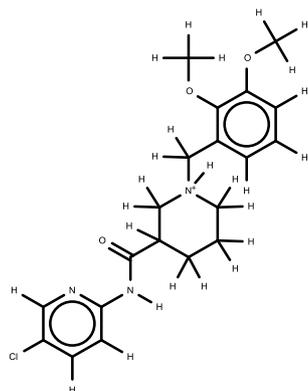
1168



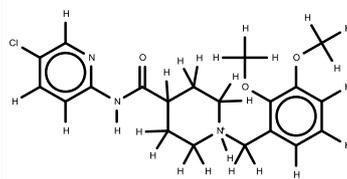
1169



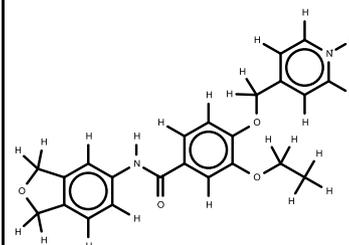
1170



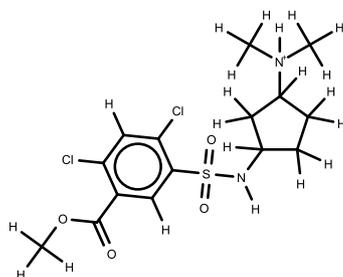
1171



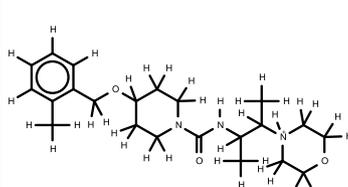
1172



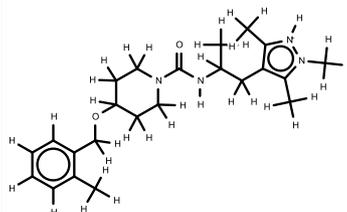
1173



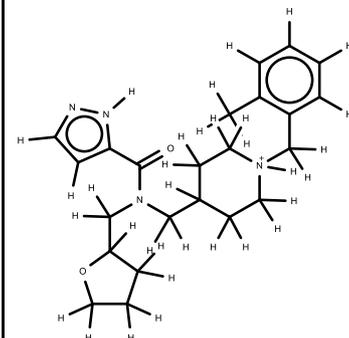
1174



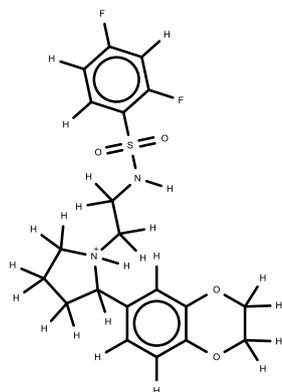
1175



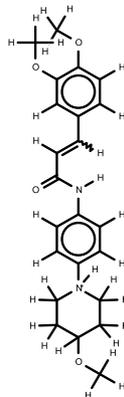
1176



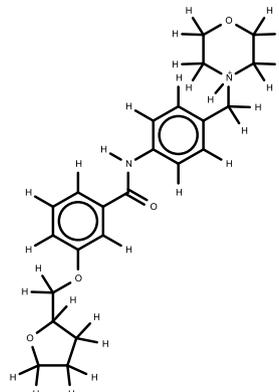
1177



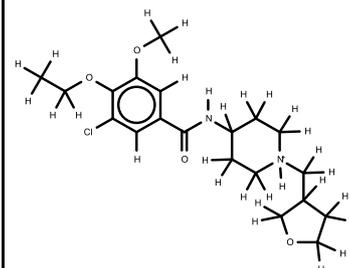
1178



1179



1180



1181

