



Universidad de Santiago de Compostela

Facultad de Química

Unidad de Resonancia Magnética Nuclear

Desarrollo de experimentos y aplicaciones de
RMN dirigidas al estudio de metabolitos y sus
interacciones supramoleculares

Marino Vega Vázquez

Santiago de Compostela, 2011

Manuel Martín Pastor, Técnico Superior de Investigación y responsable de la unidad de resonancia magnética de la RIAIDT, y Francisco Javier Sardina López, Catedrático del Departamento de Química Orgánica de la Universidad de Santiago de Compostela.

INFORMAN:

Que la presente Memoria, titulada “*Desarrollo de experimentos y aplicaciones de RMN dirigidas al estudio de metabolitos y sus interacciones supramoleculares*”, presentada por Marino Vega Vázquez para optar al grado de Doctor ha sido realizada bajo su supervisión en la Unidad de Resonancia Magnética de la RIAIDT.

Y considerando que la presente Memoria constituye trabajo de Tesis Doctoral, dan su aprobación para que sea presentada para su evaluación por la Comisión de Doctorado de la Universidad de Santiago de Compostela.

Y para que así conste, firman la presente en Santiago de Compostela, a 15 de Diciembre de 2011

Fdo: D. Manuel Martín Pastor

Fdo: D. Francisco Javier Sardina López

III

Agradecimientos

Con la presentación de esta Memoria concluye una etapa de mi formación académica, y en definitiva, de mi vida. A lo largo de estos años varias han sido las personas que, directa o indirectamente, en mayor o menor medida, han influido en mí y en mi trabajo. Por eso me gustaría tenerlas en cuenta y agradecerles en este rincón de mi Tesis todo lo que me han dado.

Agradezco al Doctor Manuel Martín Pastor la paciencia, la confianza y el esfuerzo dedicados durante todos estos años de dirección. Su dedicación y entusiasmo por la investigación han sido motores fundamentales en mi etapa como investigador y prácticamente todo lo que sé de RMN se lo debo a él. Agradezco al Doctor Javier Sardina López el haberme brindado, aún antes de haber acabado la Licenciatura, la posibilidad de acceder al fascinante mundo de la RMN. Agradezco asimismo a ambos el haber dirigido y supervisado esta Tesis Doctoral.

Agradezco a Nanetta y a Eva el haberme ayudado durante todos estos años en el laboratorio en todo cuanto les ha sido posible.

A los Doctores Mark Van Raaij y Antonio L. Llamas-Saiz, por darme la oportunidad de trabajar con ellos en el Proyecto Europeo “Artizymes” (*FP6-2003-NEST-B3 15471*) y por la financiación que eso supuso, fundamental para poder llevar a cabo todos estos años de investigación. Asimismo, me gustaría agradecer al Doctor Carlos Cobas por su inestimable colaboración en muchos apartados de esta Tesis y a la empresa MestreLab Research por el soporte y la financiación recibida.

En un plano mucho más personal, me gustaría agradecer a mis padres Marino e Isabel y a mis hermanos Jano, Isocha, Paula y Jorge por ser culpables en mucho de que sea como soy, y por todo el apoyo recibido de ellos en estos largos años. También a mi prima Alejandra, aunque al final me haya ganado y haya depositado antes que yo, por estos últimos meses de desahogo conjunto en los cafés mañaneros. Y como no, a mis cinco pequeños sobrinos, por ser un motivo de enorme alegría en mi vida.

Y en último lugar, y precisamente por eso, en el lugar más importante, a Beatriz. Por ser quien es, por estar siempre ahí, y en definitiva, por aguantarme. Porque si bien es complicado embarcarse en años de tesis, no lo es menos el compartir la vida con un doctorando. Por eso y otras muchas, mil gracias.

Índice:

Capítulo 1: Métodos de adquisición rápida de espectros de RMN multidimensionales.....	11
1.1 Introducción	13
1.2 Estrategias de espectroscopía nD-RMN rápida con muestreo reducido	15
1.2.1 Técnicas de muestreo reducido no-uniforme (aleatorio):.....	16
1.2.2 Técnica de muestreo reducido lineal optimizado: espectros plegados:	16
1.2.3 Técnica de muestreo reducido radial: espectros de dimensionalidad reducida:.....	18
1.2.4 Técnica de muestreo discreto de las frecuencias: espectroscopía Hadamard:	19
1.2.5 Técnica de codificación espacial en frecuencia: RMN ultrarrápida:.....	21
1.3 Estrategias de pulsado rápido para espectroscopía nD-RMN	23
1.3.1 Método de incremento de relajación longitudinal efectiva (LRE): SOFAST, BEST:.....	23
1.4 Estrategias de detección en paralelo	25
1.4.1 Experimentos en paralelo mediante evolución con tiempo compartido:	25
1.4.2 Experimentos en paralelo mediante tiempo de relajación entre scans compartido:	26
1.5 Referencias	27
Capítulo 2: Fast-PRE-NMR: Aceleración de espectros nD-RMN mediante incremento de relajación por adición de sustancias paramagnéticas	31
2.1 Introducción	33
2.2 Teoría	37
2.2.1 Relación señal-ruido por unidad de tiempo y la relajación longitudinal de ^1H :	37
2.2.2 Incremento de la relajación paramagnética (PRE):.....	38
2.3 Procedimiento experimental	41
2.3.1 Estudio de muestras Sacarosa: MnSO_4 :.....	41
Espectroscopía RMN.....	41
Medidas de tiempo de relajación longitudinal (T_1) de protón	42
Medidas de tiempo de relajación espín-espín (T_2) de protón	43
Medidas de tiempo de relajación longitudinal efectivo de protón en experimentos 2D-RMN (T_1^{eff})	43
2D g-COSY	44
2D TOCSY	44
2D g-HMQC	45
2D g-HMBC	45
2.3.2 Estudio de muestra N-(2-bromo-4,5-dimetoxifenetil)-N-(ciclopent-1-en-1-il)formamida: $\text{Fe}(\text{acac})_3$:.....	46
Espectroscopía RMN.....	46
Medidas de relajación longitudinal de espín (T_1) y de relajación espín-espín (T_2) de protón ..	47
2D g-COSY y 2D TOCSY.....	47
2D g-HMQC	47
2D g-HMBC	48
2.3.3 Estudio de muestras (E)-4-hidroxi-4-metil-1-(p-tolil)pent-1-en-3-ona: $\text{Fe}(\text{acac})_3$:.....	48
Espectroscopía RMN.....	48
2.4 Resultados y discusión.....	49
2.4.1 Optimización de PRE- T_1 y PRE- T_2 :	50
2.4.2 Fast-PRE-NMR con moléculas de tamaño mediano-pequeño:	54
2.4.3 Método rápido de ajuste de condiciones de fast-PRE-NMR:	60
2.4.4 Comprobación de la metodología Fast-PRE-NMR con otros solutos:.....	62
Estudio de la muestra 2 : $\text{Fe}(\text{acac})_3$ en DMSO	62

Estudio de la muestra $3:\text{Fe}(\text{acac})_3$ en CDCl_3	66
2.5 Conclusiones	69
2.6 Referencias	71
Capítulo 3: Adquisición rápida de experimentos multidimensionales de RMN en paralelo mediante espectroscopía localizada	73
3.1 Introducción	75
3.2 Teoría	77
3.2.1 Espectroscopía de RMN localizada espacialmente:.....	77
Efecto de un gradiente de campo.....	77
Método de pulsado en núcleos de una rodaja de muestra	78
3.2.2 Adquisición rápida de múltiples espectros mediante espectroscopía de RMN localizada: Método PALSYS:	80
3.3 Procedimiento experimental	83
3.3.1 Procesado de espectros 2D PALSYS:.....	83
3.3.2 Calibración del experimento MS-COSY:	84
3.4 Resultados y discusión	87
3.4.1 Calibraciones previas de los métodos PALSYS:	87
3.4.2 Parte 1 Experimentos rápidos PALSYS 2D COSY/2D DQF-COSY/ 2D TQF-COSY:	91
3.4.3 Parte 2: Combinación de espectros rápidos 2D PALSYS NOTOCO:	95
3.4.4 Parte 3: Combinación de espectros rápidos PALSYS COTOHMQC:	98
3.4.5 Parte 4: Experimento rápido 2D COSY con múltiple rodaja (2D MS-COSY):.....	100
Experimento rápido 2D MS-COSY.....	101
3.4.6 Parte 5: Experimento 1D de protón con selección de rodaja para el estudio de equilibrios de fase líquido-líquido:	104
3.5 Conclusiones	107
3.6 Referencias	109
Capítulo 4: Experimentos de RMN para el estudio de interacciones moleculares: Experimentos 1D STD con filtro de difusión reverso y 1D STD homodesacoplado	111
4.1 Introducción	113
4.2 Teoría	117
4.2.1 Experimento STD.....	117
4.2.2 Difusión	119
4.2.3 Espectros de RMN filtrados por difusión.....	120
4.2.4 Optimización del tiempo de difusión Δ en el experimento Drev	122
4.3 Procedimiento experimental	125
4.3.1 Experimentos con filtro de difusión reverso (Drev)	125
4.3.2 Experimento STD-Homodesacoplado	126
4.4 Resultados y discusión	129
4.4.1 Parte 1: Experimentos 1D con filtro de difusión reverso (Drev)	129
Experimento 1D-Drev:	129
Experimento 1D con filtro combinado de relajación transversal y filtro de difusión reverso; 1D CPMG-Drev:.....	132
Experimentos STD con filtro de difusión reverso; 1D STD-Drev y 1D-CPMG-Drev:	137

4.4.2 Parte 2 Experimento 1D-STD homodesacoplado.....	141
4.5 Conclusiones.....	147
4.6 Bibliografía.....	149
Capítulo 5: Métodos de clasificación automatizada para espectros de RMN. Clasificación automatizada de muestras de crudos basada en espectros 2D-HMBC.....	155
5.1 Introducción.....	157
5.2 Teoría.....	161
5.2.1 Quimiometría y análisis de datos:.....	161
5.2.2 Etapa previa de normalización o calibración de los datos experimentales:.....	162
5.2.3 Análisis de componentes principales (PCA):.....	163
5.2.4 Análisis de Clusters.....	166
1 Representación de los patrones.....	167
2 Definición de la medida de proximidad (similitud) de los objetos, apropiada a las características de los datos de las variables medidas en los objetos.....	167
3 Clusterizado o agrupado.....	168
5.2.5 Descripción del software específico realizado:.....	170
-Programa similarity.pas: Implementación de un algoritmo para calcular índices de similitud entre muestras.....	170
-Programa kmeans.pas: Implementación del algoritmo k-means para agrupamiento (clustering) automático de muestras definidas por coordenadas tridimensionales.....	177
5.3 Procedimiento experimental.....	183
5.3.1 Preparación de muestras:.....	183
5.3.2 Medidas de RMN:.....	183
5.3.3 Cálculos con espectros de RMN:.....	184
Integración en buckets de los espectros 2D HMBC.....	184
Cálculo del índice de similitud entre espectros.....	184
Análisis de componentes principales (PCA) sobre la matriz de similitud.....	185
Análisis de clusters k-means sobre resultados PCA.....	185
5.3.4 Cálculo de los buckets característicos de cluster:.....	186
5.4 Resultados y discusión.....	187
5.4.1 Integración en buckets:.....	189
5.4.2 Cálculo del índice de similitud entre espectros:.....	191
5.4.3 Análisis de componentes principales (PCA) sobre la matriz de similitud:.....	194
5.4.4 Análisis de clusters k-means guiado por PCA:.....	197
5.4.5 Efecto del nivel SN escogido en la clasificación:.....	199
5.4.6 Estudio de la evolución en el tiempo de las muestras de crudos:.....	202
5.4.7 Buckets exclusivos representativos de los clusters en los espectros 2D-HMBC de los crudos:.....	202
5.5 Conclusiones.....	207
5.6 Referencias.....	209
5.7 Apéndice.....	213
5.7.1 Detalles sobre el funcionamiento de los programas creados para este estudio:.....	213
Peakgeneratorclst.....	213
Similarityindex.....	214
Program_PCA.....	216
K-means8.....	216
Anexo.....	219

Lista de abreviaturas:

Aq	Acquisition
BEST	Band-selective Excitation Short-Transient
BPP-STE	STimulated Echo with BiPolar gradient Pulses
BUB	Índice Baroni-Urbani & Buser
CM	Confussion Matrix
CMI	Conditional Mutual Information
COR	Coeficiente Square CORrelation
COSY	CORrelation SpectroscopY
CPMG	Carr-Purcell-Meiboom-Gill
dB	deciBelio
Drev	REVerse Diffusion filter
EPI	Echo Planar-Imaging
EUC	Índice EUClideo
FID	Free Induction Decay
GARP	Globally Optimized Alternating Phase Rectangular Pulse
GDI	Global Discrimination Index
HMBC	Heteronuclear Multiple Bond Correlation
HMQC	Heteronuclear Multiple Quantum Correlation
Hz	Hertzio
INADEQUATE	Incredible Natural-Abundance Double-Quantum Transfer Experiment
LRE	Longitudinal Relaxation Enhancement
MCHORD	Modified CHORD Distance
MLEV-17	Malcolm LEVitt's Composite Pulse Decoupling sequence
MRI	Magnetic Resonance Imaging
NMR	Nuclear Magnetic Resonance
NOE	Nuclear Overhauser Effect
NRMSD	Normalized Root-Mean-Square Deviation
PALSY	PARallel Localized-SpectroscopY
PCA	Principal Components Analisis
PEA	Coeficiente PEArson de correlación
PFG	Pulse Field Gradients
PLS	Partial Least Square
PRE	Paramagnetic Relaxation Enhancement
RF	Radiofrecuencia
RMN	Resonancia Magnética Nuclear
RMSD	Root-Mean-Square Deviation
ROI	Regions Of Interest
SOFAST	Band-Selective Optimized Flip-Angle Short Transient
STD	Saturation Transfer Difference
TAN	Índice TANimoto
TOCSY	Total Correlated SpectroscopY
WaterLOGSY	Water-Ligand Observed via Gradient SpectroscopY
WATERGATE	Water suppression through GrAdient Tailored Excitation
WET	Water suppression Enhanced through T ₁ effects

Lista de anglicismos:

- Aliasing**.....Efecto sobre una señal procesada cuando el muestreo digital está por debajo del impuesto por el teorema de Nyquist.
- Array**.....Conjunto o serie.
- Back-projection**.....Retro-proyección.
- Bucket**.....Literalmente, cubo, en este caso se refiere a la región del espectro encerrada por un área de integración.
- Cluster**.....Agrupación.
- Dummy scans**.....Cada uno de los *Scans* en los que se pulsa completamente una secuencia de RMN pero no se acumulan datos.
- Dwell time**.....Es el tiempo transcurrido entre 2 puntos de adquisición de la FID. La FID es una onda analógica continua, pero se registra digitalmente, por lo tanto en discontinuo.
- Field-strength**.....Fuerza del campo generado por un imán, definido como B_0 .
- Filling factor**.....Factor de llenado, una medida de la relación geométrica entre la bobina receptora y el tamaño del tubo. Un mayor FF conlleva mejora en la relación SN.
- Folding**.....Plegamiento. Efecto que se produce al adquirir un espectro con un tamaño de ventana espectral insuficiente.
- Multi-slice**.....Múltiple rodaja.
- Outer-sphere relaxation**.....Mecanismo de relajación que ocurre cuando el Sólido no forma ningún complejo de solvatación con la especie paramagnética.
- Scan**.....Cada uno de los ciclos de pulsado completo mas adquisición y posterior acumulación de datos de una secuencia de RMN.
- Shimming**.....Corrección de las inhomogeneidades del campo magnético de un imán.

Introducción

La espectroscopía de resonancia magnética nuclear (RMN) ha probado ser una de las herramientas más versátiles en química y bioquímica. Permite conocer estructura, dinámica e interacciones moleculares tanto de pequeñas moléculas de origen sintético o metabólico como de biomoléculas de gran tamaño. Los estudios de RMN pueden realizarse en disolución (RMN de líquidos) en estado sólido o semi-sólido, e incluso, aunque es mucho menos común, en sistemas gaseosos.

Comparada con otras técnicas espectroscópicas, la RMN adolece de una inherente baja sensibilidad expresada como relación señal-ruido, por lo que gran parte del esfuerzo en el desarrollo instrumental y metodológico ha estado encaminado hacia el aumento de ésta a través de la fabricación de imanes de mayor campo, la creación de sondas específicas para cada núcleo y criosondas, secuencias de pulsos que utilizan detección inversa del heteronúcleo, etc. Gracias a las mejoras conseguidas, no es inusual conseguir muestras con sensibilidad por scan suficiente. La segunda limitación es el elevado tiempo que requiere el muestrear la señal hasta obtener la resolución necesaria en todas las dimensiones espectrales, problema que se hace patente en la espectroscopía multidimensional, donde los tiempos de adquisición son mucho mayores conforme aumentan el número de puntos a muestrear en las llamadas dimensiones indirectas.

Recientemente, se han propuesto varias estrategias para solventar esta segunda limitación. De manera muy general, éstas se pueden dividir en tres grandes grupos; experimentos con muestreo reducido, experimentos de pulsado rápido y experimentos con tiempo compartido (una ampliación sobre estas tres estrategias se expone en el capítulo 1). En esta Memoria se exploran dos de estos grupos, presentándose una

estrategia de aceleración de experimentos mediante el pulsado rápido (Fast-PRE-NMR, capítulo 2) y una estrategia de acelerado de adquisición de experimentos mediante la detección en paralelo, perteneciente al grupo de los experimentos con tiempo compartido (PALSY-NMR, capítulo 3).

Además de las limitaciones mencionadas sobre sensibilidad y tiempo requerido para obtener espectros multidimensionales, cuando se utiliza la RMN para estudiar mezclas complejas de compuestos de bajo peso molecular de origen diverso, biológico (metabolitos), alimentario, industrial, química combinatoria etc., o incluso en mezclas más simples de composición conocida, como las utilizadas en el estudio de las interacciones de pequeñas moléculas con grandes moléculas diana, es frecuente encontrarse con problemas de congestión de señales en los espectros, lo que introduce dificultades para su asignación y/o integración. Algunas de las soluciones planteadas para simplificar este tipo de espectros consisten en introducir, dentro de la secuencia de pulsos original del experimento, algún módulo adicional para el filtrado de señales. El filtrado opera de acuerdo con alguna propiedad que sea lo suficientemente diferente entre las señales de RMN de las moléculas de la mezcla, de forma que finalmente se consiga un espectro en el que tengan mayor preponderancia (intensidad) las señales de la/s molécula/s de interés. En esta Memoria se presenta un nuevo tipo de filtrado de señal basado en una propiedad como es el coeficiente de difusión molecular y que se ha denominado filtro de difusión reverso (filtro Drev, capítulo 4). Otra posibilidad que se plantea en esta Memoria para reducir el solapamiento de señales es extender una idea, originalmente propuesta para espectros de protón, al experimento de transferencia de saturación (STD-NMR), uno de los más utilizados para detectar interacciones intermoleculares. El método se basa en la transformación de espectros 2D J-resueltos a

1D-Homodesacoplados en los que la multiplicidad de las señales del espectro se simplifica a singletes (1D-STD-homodesacoplado, capítulo 4).

En el estudio por RMN de mezclas complejas como alguna de las mencionadas arriba, incluso con los diferentes métodos disponibles para separar señales o simplificar los espectros no suele ser suficiente. Normalmente la identificación de señal al nivel molecular requiere disponer de información adicional, proveniente de RMN y otras técnicas de análisis, y que idealmente puede venir sistematizada en forma de base de datos de los compuestos que son habituales para el tipo de muestra en cuestión. Frente a las dificultades que esto introduce, se ha planteado el término “huella dactilar de metabolitos”^{*}, con la intención de poder clasificar de forma rápida y efectiva entre diferentes muestras, en función de patrones de similitud que presentan sus espectros de RMN, identificando similitud en el patrón espectral con similitud en la composición de la mezcla. Los crudos son un ejemplo de este tipo de muestra compleja de origen geológico para la que es relevante su clasificación de cara a su posterior refinado. En el capítulo 5 de esta Memoria se exploran diversos métodos estadísticos y se exponen un conjunto de aplicaciones informáticas desarrolladas para la clasificación automática de espectros de RMN mono- o multidimensional. Se recogen los resultados obtenidos al clasificar por estos métodos una serie de muestras de crudo de diferente procedencia geográfica.

Esta Memoria está estructurada en cinco capítulos; el primero es una introducción general a los diferentes métodos de FAST-NMR, y cada uno de los siguientes corresponde al desarrollo de los temas comentados en esta introducción general. Se incluye en todos los casos una introducción al estado actual del tema a tratar.

^{*} Ellis, D.I.; Dunn, W.B.; Griffin, J.L.; Allwood, J.W.; Goodacre, R., “Metabolic fingerprinting as a diagnostic tool”, *Pharmacogenomics*, **2007**, 8, 1243-1266.

Capítulo 1:
Métodos de adquisición rápida de espectros de RMN
multidimensionales

1.1 Introducción

La RMN se ha convertido en una herramienta fundamental para el estudio de la estructura, dinámica e interacciones tanto de moléculas de bajo peso molecular como de biomoléculas de un tamaño relativamente grande. Los avances en la instrumentación han generalizado el uso de experimentos de RMN multidimensional (nD RMN).

El incremento de la sensibilidad de un espectro de RMN se consigue tradicionalmente mediante promediado de señal durante un número de repeticiones del experimento (*scans*). La sensibilidad se mejora también cuando esto se combina con algún método de transferencia de polarización homonuclear o heteronuclear. Los experimentos nD-RMN requieren un tiempo de medida considerable debido a que no sólo se requiere acumular *scans* hasta conseguir una sensibilidad suficiente, sino que además es necesario alcanzar una determinada resolución de señales en cada una de las denominadas dimensiones indirectas. Mientras que para 1D RMN la digitalización de las frecuencias de precesión de Larmor de todos espines se detecta directamente de su inducción eléctrica en la bobina de recepción de la sonda, en nD-RMN la digitalización de una dimensión indirecta se registra mediante un periodo de tiempo variable que clásicamente se va incrementando de forma lineal, y con un paso prefijado por el teorema de Nyquist, a lo largo de una serie de acumulaciones del experimento (Figura 1.1a). Como consecuencia de este proceso, el tiempo requerido para la adquisición de un experimento nD RMN se incrementa en aproximadamente 2 órdenes de magnitud por cada dimensión indirecta adicional. Por lo tanto, incluso cuando la sensibilidad por cada repetición del experimento (*scan*) sea la suficiente, el muestreo completo de toda la rejilla de puntos de una dimensión indirecta impone un tiempo mínimo de medida:

varios minutos para 2D, varias horas para 3D, varios días para 4D, y varios meses para 5D.

Durante la última década se han desarrollado una serie de métodos para acelerar la adquisición de experimentos nD-RMN (Atreya2005, Schanda2009, Felli2009, Atreya2010). Estos métodos no sólo son interesantes por el ahorro de tiempo que pueden aportar en su aplicación a experimentos multidimensionales de RMN (nD RMN), si no que pueden ser la única alternativa para muestras que son inestables en el tiempo y/o para muestras sobre las que se quiere estudiar su evolución en escalas de tiempo cortas o intermedias. A grandes rasgos, las estrategias utilizadas para conseguir reducir el tiempo de medida se basan bien en limitar el número de puntos adquiridos en las dimensiones espectrales indirectas (técnicas de muestreo reducido) o bien en reducir el tiempo necesario para cada repetición (*scan*) del experimento (técnicas de pulsado rápido). En ciertos experimentos la combinación de varias estrategias es factible y se puede obtener un ahorro de tiempo de medida acumulativo (Marion2010, Schanda2006a).

1.2 Estrategias de espectroscopía nD-RMN rápida con muestreo reducido

La forma más simple de reducir la duración total de un experimento de RMN es adquirir sólo un número reducido de puntos de la FID seleccionando puntos en las diferentes dimensiones, (t_1 , t_2 , ...) (Figura 1.1). A continuación se describen una serie de métodos de fast nD-RMN que utilizan diversas formas de realizar el muestreo reducido.

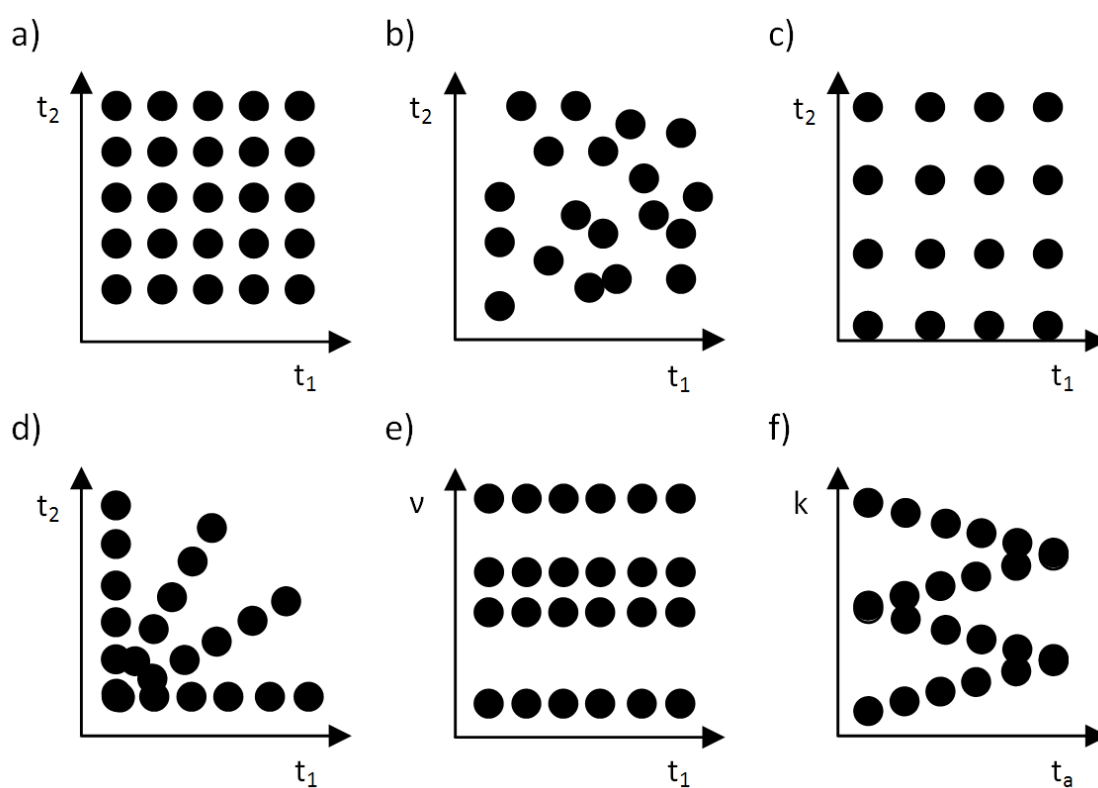


Figura 1.1: Experimentos de 2D RMN con a) muestreo lineal convencional, b) muestreo reducido aleatorio, c) muestreo lineal reducido, d) muestreo radial, e) muestreo selectivo con Hadamard, f) muestreo ultra-rápido.

1.2.1 Técnicas de muestreo reducido no-uniforme (aleatorio):

En este método se adquieren puntos en las diferentes dimensiones, (t_1, t_2, \dots) de forma aleatoria (Figura 1b) (Maciejewski2006). Normalmente, sólo se adquieren en torno al 10-30% de los puntos. Los experimentos con muestreo no-uniforme no pueden procesarse mediante la transformada rápida de Fourier, se requieren otros métodos de procesado para reconstruir el espectro nD-RMN. Algunos ejemplos del tipo de procesado que se requiere son la reconstrucción de máxima entropía (Hoch2001, Rovnyak2003), método de filtrado de diagonalización (Chen2004, Hu2000), descomposición multidimensional (Luan2005, Orekhov2003) o la integración de Fourier numérica (Kazimierczuk2006, Marion2005). Estos métodos mejoran la resolución espectral respecto al experimento adquirido en el mismo tiempo con muestreo lineal convencional. Sin embargo, un inconveniente importante de estos métodos es que en el procesado se utilizan en algún momento algoritmos de ajuste no lineal que no son por el momento tan robustos y fiables como la transformada de Fourier, y por ejemplo requieren del ajuste cuidadoso de algún parámetro, ya sea con un conocimiento previo del sistema, o bien mediante ensayo-error.

1.2.2 Técnica de muestreo reducido lineal optimizado: espectros plegados:

Esta técnica consiste en aumentar la separación de tiempo de evolución entre dos puntos contiguos a muestrear en una dimensión indirecta (aumento del paso de muestreo o *dwell time*) de forma que el salto entre los puntos sea más grande del que se necesitaría según la condición de muestreo de Nyquist (Figura 1.1c). El resultado de este déficit de muestreo en el dominio del tiempo es equivalente a una reducción aparente de la anchura espectral de esa dimensión indirecta. Se producen entonces fenómenos de plegamiento (*foldings* o *aliasing*) en el espectro (Figura 1.2), sin embargo,

la resolución espectral se mantiene porque el valor máximo de t_1 que se muestra (t_1^{\max}) no cambia. Un plegamiento intensivo del espectro en general no resulta práctico, ya que puede producir solapamiento accidental de los picos, lo que sería equivalente a una pérdida de resolución espectral. Sin embargo, dado que el fenómeno de plegamiento de señales es completamente previsible, cuando se conoce de antemano la posición de los picos por espectros previos, se puede utilizar esta información para optimizar convenientemente la reducción de la anchura espectral para no crear solapamiento de picos. De esta forma se han llegado a plantear plegamientos muy intensivos (Jeannerat2007, Maciejewski2006, Sakhaii2008). La sencillez del método lo ha hecho atractivo para muchas aplicaciones de rutina con algún conocimiento previo de los desplazamientos químicos.

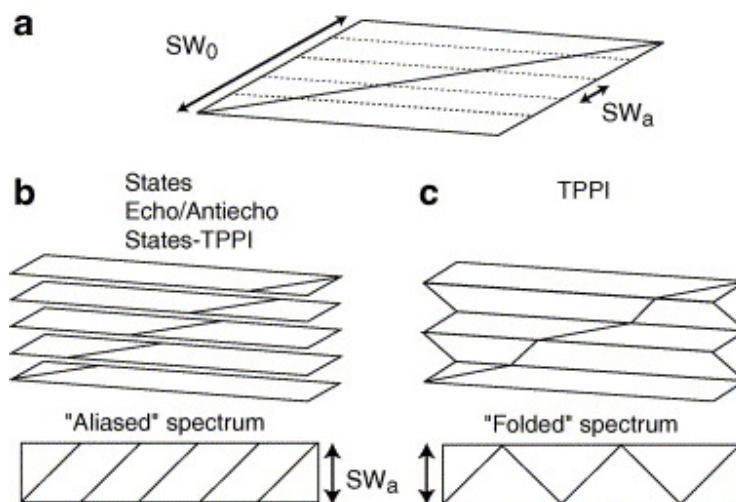


Figura 1.2: Cuando se reduce la anchura espectral de una dimensión indirecta a SW_a por debajo de la del espectro completo SW_0 las señales aparecen con aliasing (b) o plegadas (c) dependiendo del modo de adquisición en cuadratura utilizado. (Jeannerat2007, Figura 1)

1.2.3 Técnica de muestreo reducido radial: espectros de dimensionalidad reducida:

Esta técnica consiste en el incremento simultáneo de varios tiempos de evolución de dimensiones indirectas (por ej. t_1 y t_2 para el caso del 3D) (Figura 1.1d). Este tipo de espectros pueden ser procesados con la transformada de Fourier habitual si previamente se reorganizan las FIDs adquiridas. El espectro resultante de “dimensionalidad reducida” es conceptualmente idéntico a una proyección 2D en un determinado ángulo del espectro completo 3D (Brutscher1995, Kim2003, Kupče2003a Szyperski1993, Kupče2005, Kupče2004a). El ángulo de proyección del espectro 2D se puede controlar por medio del incremento de tiempo relativo que se elija para muestrear los puntos en t_1 y en t_2 . Normalmente lo que se hace es obtener una serie de espectros 2D con varios ángulos de proyección 2D (Figura 1.3) para posteriormente reconstruir por procesado el espectro completo de correlación 3D utilizando bien técnicas de retro-proyección (*back-projection*) (Kupče2004a, Coggins2004, Kupče2004b), o bien métodos de reconstrucción algebraica (Kim2003).

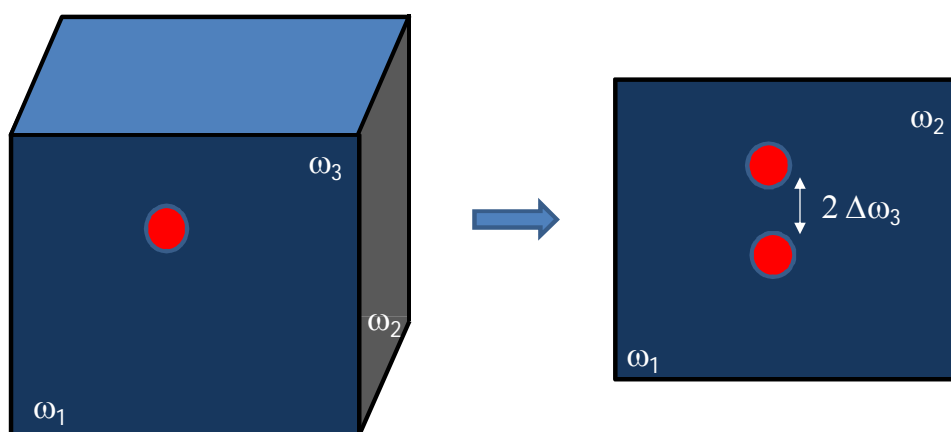


Figura 1.3: Método de muestreo radial de un espectro 3D para obtener un espectro 2D proyectado. Se adquiere con un cierto ángulo el plano ω_1/ω_2 que proyecta parcialmente el desplazamiento químico ω_3 en este plano. Esto puede permitir separar señales que originalmente solapaban en el plano ω_1/ω_3 .

Para una molécula desconocida no existe una forma fácil de predecir cuál es el número mínimo de espectros de dimensionalidad reducida que se debe adquirir y cuáles serán los mejores ángulos de proyección. El número de proyecciones depende un poco de la complejidad del espectro a tratar en cuanto a número de señales y su grado de solapamiento (Hiller2005). En general los ángulos de proyección deben ser suficientemente diferentes para evitar redundancias.

Con esta estrategia se consigue una reducción de varios órdenes de magnitud para espectros con dimensión $n \geq 4$. Se ha aplicado con éxito para espectros de correlación de hasta 7D. (Hiller2007, Staykova2008)

1.2.4 Técnica de muestreo discreto de las frecuencias: espectroscopía Hadamard:

Cuando los desplazamientos químicos se conocen de antemano y sólo interesa la información de un determinado número n de señales, es posible ahorrar mucho tiempo de medida si en lugar del experimento 2D, se adquieren experimentos 1D análogos de forma independiente para cada una de las n posiciones seleccionadas (Figura 1.1e). El método Hadamard que se describe a continuación es una alternativa más eficiente para realizar este tipo de espectros selectivos (Kupče2003b).

El método Hadamard sustituye el período de evolución de una dimensión indirecta de un experimento n D-RMN por una codificación en frecuencia mediante una serie de pulsos selectivos de excitación aplicados de forma simultánea a las frecuencias adecuadas (excitación multiplexada).

Consideremos un ejemplo con 4 señales diferentes que quieran ser editadas. En primer lugar se crearía una serie (*array*) de pulsos selectivos para aplicarlos de forma lo más simultánea posible (o consecutivos) para las 4 resonancias de interés. En un espectro así adquirido aparecerían las correlaciones de las señales seleccionadas y se plantearía la necesidad de cómo distinguir las que provienen de cada una de las 4 posiciones seleccionadas.

En espectroscopía Hadamard la forma de separar en el espectro final las correlaciones debidas a cada una de las señales se basa en el hecho de que la fase de cada pulso selectivo de excitación puede controlarse de forma independiente. Esto hace posible introducir a voluntad variaciones de fase de 180° que produzcan un cambio de signo en la correspondiente señal del espectro. Cuando se codifica el signo de cada señal elegida como positiva (fase 0°) o negativa (fase 180°) siguiendo el esquema de cada fila de una matriz de tipo Hadamard de dimensión $n \times n$ (donde n es igual o mayor que el número de frecuencias a distinguir, 4 en este caso), se obtienen n espectros con codificación Hadamard. Mediante combinaciones suma y resta de los n espectros Hadamard obtenidos, o más directamente mediante una transformada de Hadamard, se pueden obtener las correlaciones individuales de cada sitio seleccionado (Figura 1.4).

Dado que en la espectroscopía Hadamard se excitan y se detectan de forma simultánea las respuestas de todos los sitios a la vez (hay multiplexado de señales) y la relación señal-ruido de una señal es función del número de acumulaciones, se consigue una ganancia teórica de sensibilidad de \sqrt{n} donde n es el número de sitios seleccionados respecto a la adquisición de experimentos 1D independientes. Esta ganancia de sensibilidad es equivalente a la de la adquisición del propio experimento 2D.

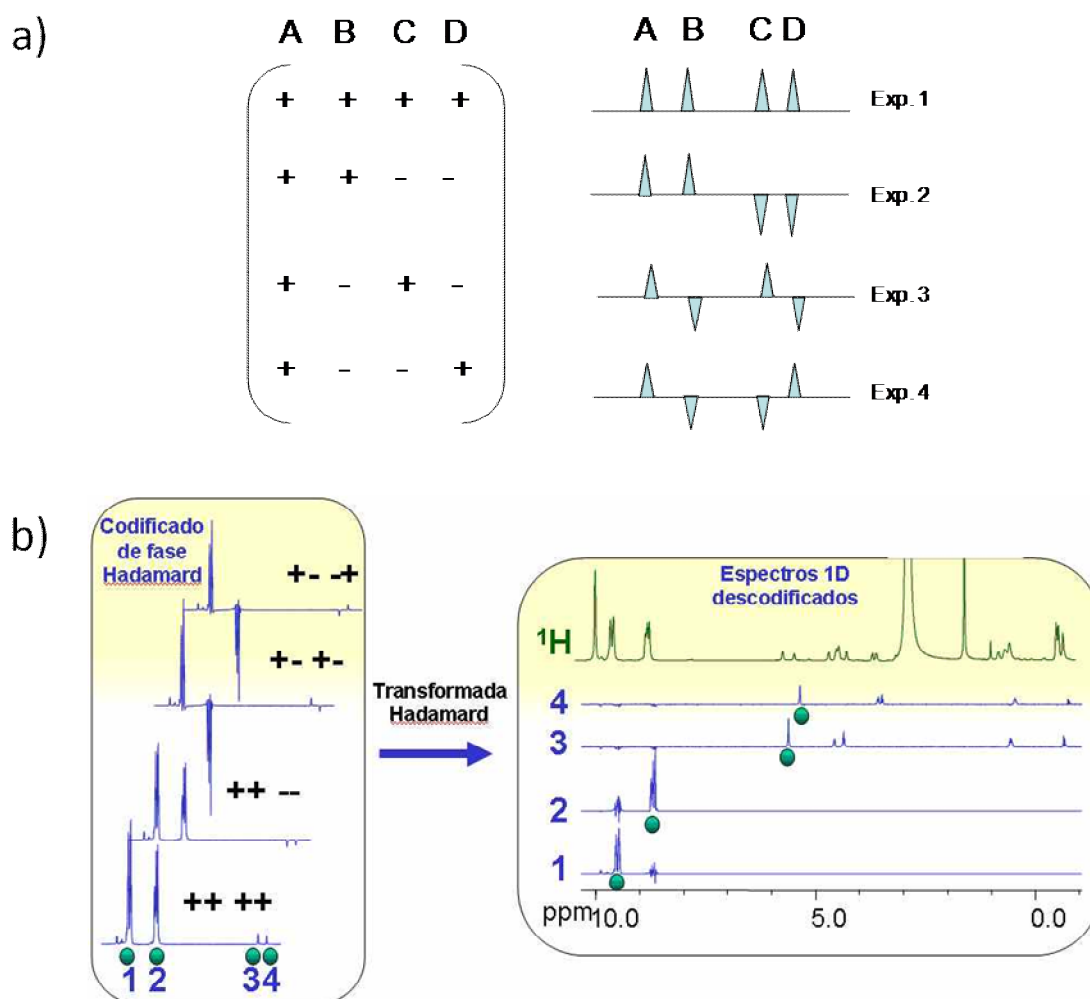


Figura 1.4: a) Matriz de Hadamard de orden 4 y esquema de excitación multiplexada para RMN. b) Aplicación del método Hadamard a la codificación en fase de cuatro señales de un espectro 1D Hadamard-TOCSY (izda.). La transformada de Hadamard proporciona los cuatro espectros individuales 1D TOCSY (dcha.).

1.2.5 Técnica de codificación espacial en frecuencia: RMN ultrarrápida:

Otra solución para el problema del muestreo en las dimensiones indirectas de la espectroscopía RMN se basa en realizar experimentos diferentes en partes diferentes de la muestra utilizando codificación espacial. El método de RMN ultra-rápida o de un solo *scan* (*single scan*) (Frydman2002, Shrot2003) codifica la evolución temporal de una dimensión indirecta a lo largo de una coordenada espacial (Figura 1.1f). Este método en principio permite obtener un espectro multidimensional en una sola

acumulación. Para la decodificación individual de cada tiempo de evolución de la información se reemplaza la detección convencional de la señal al final de la secuencia por una serie de pequeños gradientes bipolares de lectura de los ecos que se corresponden a los puntos de la señal en t_1 . El esquema es similar al utilizado en imagen rápida de resonancia con el experimento EPI o *Echo Planar-Imaging*.

En la RMN ultra-rápida existe un fuerte compromiso entre resolución y sensibilidad. La sensibilidad es proporcional aproximadamente a \sqrt{n} , donde n es el número de puntos a detectar en la dimensión indirecta. Por otro lado, existe cierta dificultad instrumental y metodológica para la implementación de esta técnica en los espectrómetros convencionales.

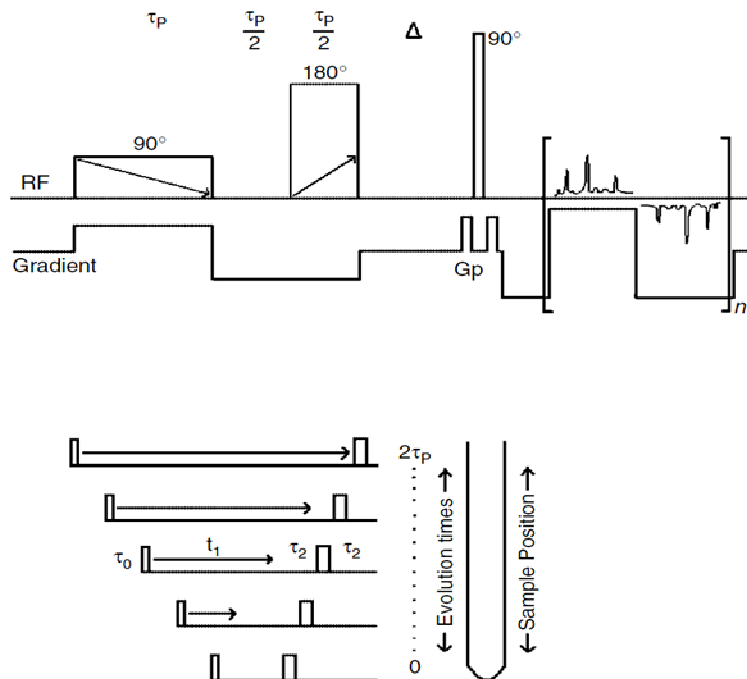


Figura 1.5: Esquema de pulsos de un experimento ultrafast-2D COSY (Andersen05). Debajo se muestra en esquema la codificación espacial de la dimensión indirecta t_1 a lo largo de una serie de rodajas de muestra paralelas a lo largo del tubo de RMN.

1.3 Estrategias de pulsado rápido para espectroscopía nD-RMN

Estas estrategias son un método alternativo de reducir los tiempos de medida en RMN. La idea principal es acortar el tiempo de espera de relajación entre *scans* sucesivos para conseguir mayores velocidades de repetición y así adquirir el mismo número de *scans* en un menor tiempo. El tiempo de relajación entre *scans* es el período que va desde el último pulso del *scan* anterior del experimento al primer pulso del *scan* actual, esto generalmente corresponde a la suma del tiempo de espera del *scan* (d_1) y del período que dura la adquisición de la FID (aq). Se requiere un tiempo mínimo de relajación entre *scans* para garantizar que los espines que fueron excitados en un *scan* anterior vuelvan total o parcialmente a la situación de equilibrio termodinámico y así alcancen un nivel suficiente de polarización antes del siguiente *scan* y por tanto se consiga la suficiente sensibilidad en cada adquisición después de la primera.

1.3.1 Método de incremento de relajación longitudinal efectiva (LRE):

SOFAST, BEST:

Una de los métodos de pulsado rápido se denomina Incremento de Relajación Longitudinal Efectiva (LRE: *Longitudinal Relaxation Enhancement*) (Pervushin2002). LRE se basa en el hecho de que la eficacia de la velocidad de relajación espín-red de ^1H se incrementa si los protones que estén cerca de uno de interés no son perturbados por la secuencia de pulsos. Esto permite que los espines no perturbados absorban parte de la energía de los espines perturbados mediante interacciones dipolo-dipolo (efecto NOE) o mediante procesos como el intercambio de hidrógeno. En la práctica LRE se realiza mediante manipulación con pulsos selectivos de un conjunto pequeño de espines de ^1H de interés, por ej. protones amida, de forma que el resto de protones que no participan

en los procesos de transferencia de coherencia del experimento de pulsos se mantengan inalterados. Esto consigue reducir los tiempos de relajación efectivos de protón desde algunos segundos a unos cientos de milisegundos lo que permite un pulsado mucho más rápido.

Para algunos experimentos como por ej. HMQC, la sensibilidad se puede incrementar aún más si se combina LRE con un ángulo de pulso de excitación ajustado al llamado ángulo de Ernst (Ernst1966) que proporciona la máxima sensibilidad por tiempo de medida. Ambas estrategias son utilizadas en los experimentos SOFAST (Schanda2005a, Schanda2005b) y BEST (Schanda2006b, Lescop2007) que proporcionan espectros 2D de correlación ^1H - ^{15}N o ^1H - ^{13}C de muestras marcadas de péptidos y proteínas en unos pocos segundos.

1.4 Estrategias de detección en paralelo

Se encuentran a medio camino entre el muestreo reducido no uniforme y el pulsado rápido.

1.4.1 Experimentos en paralelo mediante evolución con tiempo compartido:

Esta estrategia se refiere a la conversión de dos experimentos diferentes pero complementarios en un único experimento donde el tiempo de evolución variable de una dimensión indirecta (ej. t_1) es compartido y evoluciona de forma simultánea para ambos. Así por ej. se pueden combinar las dimensiones indirectas de dos experimentos 2D HMQC de varias especies nucleares diferentes en uno sólo tal como se muestra en la siguiente figura. (Figura 1.6) (Parella2010, Malmodin2005, Kupče2006)

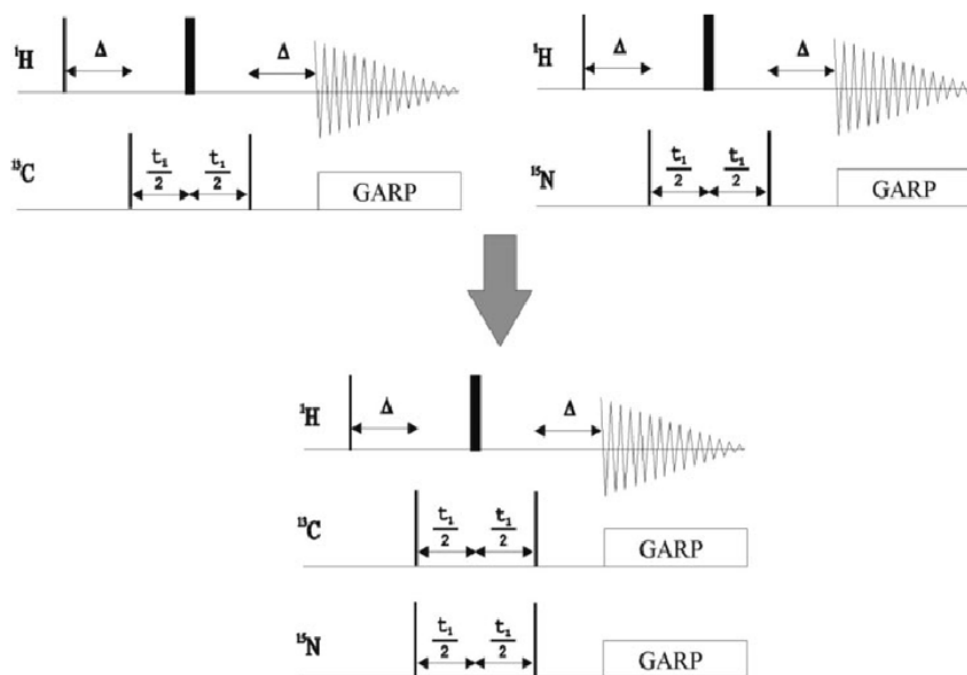


Figura 1.6: Conversión de dos experimentos independientes HMQC en uno de tiempo compartido con un tiempo de evolución t_1 simultáneo.

1.4.2 Experimentos en paralelo mediante tiempo de relajación entre scans compartido:

El tiempo de relajación entre *scans* define el tiempo de repetición y la velocidad de pulsado. Se pueden plantear diversas combinaciones de experimentos de RMN aprovechando el tiempo de espera de relajación de uno para el pulsado y adquisición del otro de forma que se ahorre tiempo de medida. (Kupče2008, Kupče2006, Kupče2007)

Un ejemplo de este tipo de estrategia se muestra en la Figura 1.7.

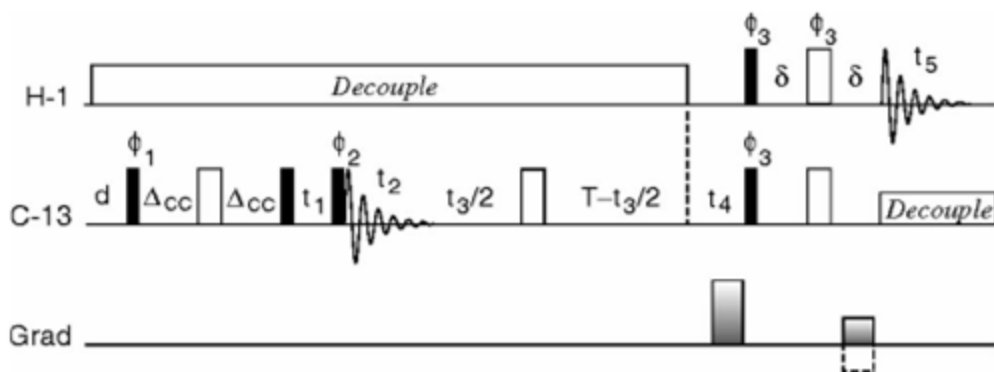


Figura 1.7: Secuencia de pulsos para la observación de un espectro de ^{13}C desacoplado, experimento ^{13}C - ^{13}C INADEQUATE, ^1H - ^{13}C HSQC editado y 3D ^1H - ^{13}C J-HMBC con adquisición en paralelo (Kupče2008). Nótese que hay distintos períodos de detección en la secuencia.

1.5 Referencias

- Andersen, S.N.; Köckenberger, W., “A simple approach for phase-modulated single-scan 2D NMR spectroscopy”, *Magn. Reson. Chem.*, **2005**, *43*, 795–797.
- Atreya, H.S.; Szyperski, T., “Rapid NMR Data Collection”, *Methods Enzymol.*, **2005**, *394*, 78-108.
- Atreya, H.S., “NMR methods for fast data acquisition”, *J. of the Indian Inst. of Sci.*, **2010**, *90*, 87-104.
- Brutscher, B.; Morelle, N.; Cordier, F.; Marion, D., “Determination of an Initial Set of NOE-Derived Distance Constraints for the Structure Determination of $^{15}\text{N}/^{13}\text{C}$ -Labeled Proteins”, *J. Magn. Reson. B*, **1995**, *109*, 238-242.
- Chen, J.H.; Nietlispach, D.; Shaka, A.J.; Mandelshtam, V.A., “Ultra-high resolution 3D NMR spectra from limited-size data sets”, *J. Magn. Reson.*, **2004**, *169*, 215-224.
- Coggins, B.E.; Venters, R.A.; Zhou, P., “Generalized Reconstruction of n-D NMR Spectra from Multiple Projections: Application to the 5-D HACACONH Spectrum of Protein G B1 Domain”, *J. Am. Chem. Soc.*, **2004**, *126*, 1000-1001.
- Ernst, R.R.; Anderson, W.A., “Application of Fourier Transform Spectroscopy to Magnetic Resonance”, *Rev. Sci. Instrum.*, **1966**, *37*, 93-103.
- Felli, I.C.; Brutscher, B., “Recent Advances in Solution NMR: Fast Methods and Heteronuclear Direct Detection”, *ChemPhysChem*, **2009**, *10*, 1356-1368.
- Frydman, L.; Scherf, T.; Lupulescu, A., “The acquisition of multidimensional NMR spectra within a single scan”, *PNAS*, **2002**, *99*, 15858-15862.
- Hiller, S.; Fiorito, F.; Wüthrich, K.; Wider, G., “Automated projection spectroscopy (APSY)”, *PNAS*, **2005**, *102*, 10876-10881.
- Hiller, S.; Wasmer, C.; Wider, G.; Wüthrich, K., “Sequence-Specific Resonance Assignment of Soluble Nonglobular Proteins by 7D APSY-NMR Spectroscopy”, *J. Am. Chem. Soc.*, **2007**, *129*, 10823-10828.
- Hoch, J.C.; Stern, A.S., “Maximum Entropy Reconstruction, Spectrum Analysis and Deconvolution in Multidimensional Nuclear Magnetic Resonance”, *Meth. Enzym.*, **2001**, *338*, 159-178.

- Hu, H.; De Angelis, A.A.; Mandelshtam, V.A.; Shaka, A.J., "The Multidimensional Filter Diagonalization Method: II. Application to 2D Projections of 2D, 3D, and 4D NMR Experiments", *J. Magn. Reson.*, **2000**, *144*, 357-366.

- Jeannerat, D., "Computer optimized spectral aliasing in the indirect dimension of ^1H - ^{13}C heteronuclear 2D NMR experiments. A new algorithm and examples of applications to small molecules", *J. Magn. Reson.*, **2007**, *186*, 112-122.

- Kazimierczuk, K.; Koźmiński, W.; Zhukov, I., "Two-dimensional Fourier transform of arbitrarily sampled NMR data sets", *J. Magn. Reson.*, **2006**, *179*, 323-328.

- Kim, S.; Szyperski, T., "GFT NMR, a New Approach To Rapidly Obtain Precise High-Dimensional NMR Spectral Information", *J. Am. Chem. Soc.*, **2003**, *125*, 1385-1393.

- Kupče, E.; Freeman, R., "Projection-Reconstruction of Three-Dimensional NMR Spectra", *J. Am. Chem. Soc.*, **2003**, *125*, 13958-13959.

- Kupče, E.; Freeman, R., "Frequency domain Hadamard spectroscopy", *J. Magn. Reson.*, **2003**, *162*, 158-165.

- Kupče, E.; Freeman, R., "Projection-reconstruction technique for speeding up multidimensional NMR spectroscopy", *J. Am. Chem. Soc.*, **2004**, *126*, 6429-6440.

- Kupče, E.; Freeman, R., "The Radon transform: A new scheme for fast multidimensional NMR", *Concepts in Magn. Reson. part A*, **2004**, *22A*, 4-11.

- Kupče, E.; Freeman, R., "Resolving ambiguities in two-dimensional NMR spectra: the 'TILT' experiment", *J. Magn. Reson.*, **2005**, *172*, 329-332.

- Kupče, E.; Freeman, R.; John, B.K., "Parallel acquisition of two-dimensional NMR spectra of several nuclear species", *J. Am. Chem. Soc.*, **2006**, *128*, 9606-9607.

- Kupče, E.; Cheatham, S.; Freeman, R., "Two-dimensional spectroscopy with parallel acquisition of ^1H -X and ^{19}F -X correlations", *Magn. Reson. in Chem.* **2007**, *45*, 378-380.

- Kupče, E.; Freeman, R., "Molecular Structure from a Single NMR Experiment", *J. Am. Chem. Soc.*, **2008**, *130*, 10788-10792.

- Lescop, E.; Schanda, P.; Brutscher, B., "A set of BEST triple-resonance experiments for time-optimized protein resonance assignment", *J. Magn. Reson.*, **2007**, *187*, 163-169.

-
- Luan, T.; Jaravine, V.; Yee, A.; Arrowsmith, C.H.; Orekhov, V.Y., “Optimization of resolution and sensitivity of 4D NOESY using Multi-dimensional Decomposition”, *J. Biomol. NMR*, **2005**, *33*, 1-14.

 - Maciejewski, M.W.; Stern, A.S.; King, G.F.; Hoch, J.C., “Nonuniform Sampling in Biomolecular NMR” En *Modern Magnetic Resonance*; Webb G.A., Ed.; Springer: Dordrecht, The Netherland, **2006**; Vol. 2, pp. 1305-1311.

 - Malmodin, D.; Billeter, M., “Multiway Decomposition of NMR Spectra with Coupled Evolution Periods”, *J. Am. Chem. Soc.*, **2005**, *127*, 13486-13487.

 - Marion, D., “Fast acquisition of NMR spectra using Fourier transform of non-equispaced data”, *J. Biomol. NMR*, **2005**, *32*, 141-150.

 - Marion, D., “Combining methods for speeding up multi-dimensional acquisition. Sparse sampling and fast pulsing methods for unfolded proteins”, *J. Magn. Reson.*, **2010**, *206*, 81-87.

 - Orekhov, V.Y.; Ibraghimov, I.; Billeter, M., “Optimizing resolution in multidimensional NMR by three-way decomposition”, *J. Biomol. NMR*, **2003**, *27*, 165-173.

 - Parella, T.; Nolis, P., “Time-Shared NMR Experiments”, *Concepts in Magn. Reson. part A*, **2010**, *36A*, 1-23.

 - Pervushin, K.; Vögeli, B.; Eletsky, A., “Longitudinal ^1H Relaxation Optimization in TROSY NMR Spectroscopy”, *J. Am. Chem. Soc.*, **2002**, *124*, 12898-12902.

 - Rovnyak, D.; Filip, C.; Itin, B.; Stern, A.S.; Wagner, G.; Griffin, R.G.; Hoch, J.C., “Multiple-quantum magic-angle spinning spectroscopy using nonlinear sampling”, *J. Magn. Reson.*, **2003**, *161*, 43-55.

 - Schanda P.; Brutscher, B., “Very fast two-dimensional NMR spectroscopy for real-time investigation of dynamic events in proteins on the time scale of seconds”, *J. Am. Chem. Soc.*, **2005**, *127*, 8014-8015.

 - Schanda, P.; Kupče, E.; Brutscher, B., “SOFAS-T-HMQC experiments for recording two-dimensional heteronuclear correlation spectra of proteins within a few seconds”, *J. Biomol. NMR*, **2005**, *33*, 199-211.

 - Schanda, P.; Brutscher, B., “Hadamard frequency encoded SOFAST-HMQC for ultrafast two-dimensional protein NMR”, *J Magn. Reson.*, **2006**, *178*, 334-339.

- Schanda, P.; Van Melckebeke, H.; Brutscher, B., “Speeding Up Three-Dimensional Protein NMR Experiments to a Few Minutes”, *J. Am. Chem. Soc.*, **2006**, *128*, 9042-9043

- Schanda, P., “Fast-pulsing longitudinal relaxation optimized techniques: Enriching the toolbox of fast biomolecular NMR spectroscopy”, *Prog. NMR Spectrosc.*, **2009**, *55*, 238-265.

- Shrot, Y.; Frydman, L., “Single-Scan NMR Spectroscopy at Arbitrary Dimensions”, *J. Am. Chem. Soc.*, **2003**, *125*, 11385-11396.

- Sakhaii, P.; Haase, B.; Bermel, W., “An alternative approach for recording of multidimensional NMR data based on frequency dependent folding mechanism”, *J. Magn. Reson.*, **2008**, *191*, 291-303.

- Staykova, D.K.; Fredriksson, J.; Bermel, W.; Billeter, M., “Assignment of protein NMR spectra based on projections, multi-way decomposition and a fast correlation approach”, *J. Biomol. NMR*, **2008**, *42*, 87-97.

- Szyperski, T.; Wider, G.; Bushweller, J.H.; Wüthrich, K., “Reduced Dimensionality in Triple-Resonance NMR Experiment”, *J. Am. Chem. Soc.*, **1993**, *115*, 9307-9308.

Capítulo 2:
Fast-PRE-NMR: Aceleración de espectros nD-RMN
mediante incremento de relajación por adición de
sustancias paramagnéticas

2.1 Introducción

Los avances técnicos y metodológicos producidos en los últimos años han incrementado notablemente la sensibilidad de los experimentos de RMN. Por citar algunos, las mejoras en las consolas y en la digitalización de la señal de RMN, criosondas que reducen el ruido térmico de detección, microsondas que reducen la distancia entre la bobina receptora y la muestra (*filling factor*), técnicas de marcaje isotópico en núcleos activos como ^{13}C y/o ^{15}N , etc. hasta el punto de que, a las concentraciones de muestra accesibles, puede plantearse la situación de que la sensibilidad por *scan* deje de ser una limitación. La segunda limitación es el tiempo total necesario para muestrear con la resolución necesaria las dimensiones indirectas de un experimento nD-RMN. Entre las diferentes estrategias que se han desarrollado para acelerar la adquisición de espectros nD-RMN se encuentran las técnicas de pulsado rápido. En este capítulo se describe un método de pulsado rápido para nD-RMN que se basa en el incremento de la relajación por adición de una sustancia paramagnética a la muestra del soluto que se quiere estudiar, PRE (PRE: *Paramagnetic Relaxation Enhancement*).

En un experimento de RMN el tiempo de espera (d_1) entre acumulaciones sucesivas (*scans*) normalmente supone una parte considerable de la duración total del experimento. Éste debe ser al menos 5 veces el máximo valor de tiempo de relajación longitudinal, T_1^{max} , de los núcleos a estudiar, para garantizar que todas las resonancias afectadas por los pulsos de RF en el anterior *scan* han regresado a su situación de equilibrio con el campo magnético (> 99 %). Si no se cumple esta condición, se produce saturación parcial o total de la señal que disminuye su intensidad (pérdida de

cuantitatividad del espectro) y se puede producir ruido t_1 y otros artefactos en los espectros nD (Kupče2007).

Las interacciones intermoleculares entre un soluto e iones paramagnéticos presentes en una disolución pueden producir en el soluto una disminución de los tiempos de relajación longitudinal (T_1) y transversal (T_2) de sus espines nucleares. La base física del efecto PRE que es útil a nuestro objetivo es la interacción dipolar o de pseudo-contacto entre el soluto y el ión paramagnético. Esta interacción, que se produce entre los momentos magnéticos del electrón y el núcleo, se transmite a través del espacio y depende de la distancia entre ambos. Su contribución al efecto PRE puede ser modulada ajustando la concentración del paramagnético (Bakhmutov2004). Este efecto PRE ha sido usado para el estudio de estructura y propiedades dinámicas de moléculas (Bertini1996), y también históricamente para incrementar la velocidad de adquisición de espectros 1D de ^{13}C así como para acelerar la relajación de la señal correspondiente al H_2O . Recientemente la estrategia de pulsado rápido mediante PRE se ha usado con éxito para acelerar la adquisición de experimentos nD-RMN de macromoléculas (Cai2006).

El efecto PRE en el caso de moléculas pequeñas y medianas es teóricamente menos favorable que para macromoléculas (Cai2006), la razón es que además de acortarse los tiempos T_1 también pueden reducirse significativamente los T_2 , algo no deseable ya que provoca ensanchamiento de señal y cierta pérdida de sensibilidad en los espectros. Sin embargo, para experimentos nD-RMN de moléculas pequeñas-medianas en muestras no limitadas por la sensibilidad, estos dos problemas no siempre deben suponer una limitación para este tipo de estrategia. Hay que tener en cuenta que los experimentos nD-RMN ya incorporan algunas limitaciones que afectan a la resolución

de señales (se limita el tiempo de medida de las dimensiones indirectas, el número de puntos en la dimensión de detección cuando hay desacoplamiento de alta potencia, el tamaño de la matriz de datos, etc.), y de forma práctica es habitual que los puntos se digitalicen muy por debajo de lo necesario de acuerdo con la anchura real de las señales. El objetivo de este capítulo será pues poner de manifiesto que escogiendo cuidadosamente las condiciones experimentales, el efecto PRE permite acelerar experimentos nD-RMN de moléculas medianas-pequeñas con mínimas pérdidas de sensibilidad y resolución.

En este capítulo se ha considerado el uso de PRE para acelerar experimentos de rutina 2D RMN que proporcionan correlaciones a través del acoplamiento escalar ^1H - ^1H y ^1H - ^{13}C . En concreto, se ha utilizado PRE para los experimentos 2D g-COSY, 2D TOCSY, 2D g-HMQC, y 2D g-HMBC para varias moléculas de tamaño mediano-pequeño, un método al que nos referiremos como fast-PRE-NMR. Mediante el ajuste de la concentración de la sustancia paramagnética en la muestra para producir el efecto PRE óptimo en cada uno de estos experimentos, en concordancia con la duración optimizada de los eventos de la propia secuencia de pulsos, se puede obtener la mejor relación de sensibilidad y velocidad posible del experimento. Finalmente, se presenta un protocolo sencillo para la rápida optimización de las condiciones de fast-PRE-NMR, en el conjunto de los experimentos considerados, para cualquier tipo de soluto.

2.2 Teoría

2.2.1 Relación señal-ruido por unidad de tiempo y la relajación longitudinal de 1H:

La velocidad de repetición de un experimento de RMN es función del tiempo de espera d_1 situado entre el último pulso de un *scan* y el primero del siguiente. La relación señal-ruido en el tiempo de medida (SN_t), referido como la sensibilidad del experimento, viene dada por la Ecuación 2.1 (Ernst1987, Pervushin2002)

$$SN_t = c(1 - \exp(-\frac{1}{T_1} d_1)) / (\sqrt{d_0 + d_1}) \quad \text{Ecuación 2.1}$$

donde T_1 es el tiempo de relajación longitudinal, d_0 es la duración de la secuencia de pulsos y el tiempo de adquisición de la FID*, d_1 es el tiempo de espera entre *scans* y c es un coeficiente de proporcionalidad. Para la relajación monoexponencial, la sensibilidad maximizada (Pervushin2002)

$SN_t^{max} = c \cdot 0.71 / \sqrt{d_1^{opt}} \sim c \cdot 0.64 \sqrt{1/T_1}$ se consigue con un tiempo óptimo de espera dado por la Ecuación 2.2:

$$d_1^{opt} = 1.25 \cdot T_1 - d_0 \quad \text{Ecuación 2.2}$$

Por lo tanto la velocidad de la relajación longitudinal (el inverso de T_1) del espín excitado afecta a la sensibilidad de la señal considerada, para un *scan* del experimento de RMN en cuestión.

* En la publicación original (Pervushin2002) d_0 solo comprende la duración de la secuencia de pulsos.

2.2.2 Incremento de la relajación paramagnética (PRE):

Los tiempos de relajación de espín observados en un soluto al añadir paramagnéticos son la suma de las contribuciones diamagnética y paramagnética a la relajación. Dado que la concentración no es muy alta, el PRE es proporcional a la concentración del paramagnético (Ecuación 2.3) (Cai2006)

$$\frac{1}{T_i^{obs}} = \frac{1}{T_i^{dia}} + c \cdot \frac{1}{T_i^{para}} \quad i = 1, 2 \quad \text{Ecuación 2.3}$$

donde T_i se refiere a las relajaciones longitudinal y transversal, T_1 y T_2 respectivamente, c es el coeficiente de concentración de la especie paramagnética que también tiene en cuenta el intercambio. Para moléculas de soluto pequeñas y medianas, todos los protones son igualmente accesibles para el disolvente y por lo tanto, los cambios esperados en T_1^{obs} o T_2^{obs} serán básicamente iguales para todos ellos con una dependencia lineal de la concentración de acuerdo con la Ecuación 2.3. Desviaciones de esa linealidad pueden ser útiles en estudios estructurales como forma de refutar la presunción inicial de la ausencia de interacciones específicas con el soluto (formación de aductos), pero en cualquier caso esta desviación no es necesariamente perjudicial para el método de fast-PRE-NMR que se propone en este capítulo.

Dado que en principio no se produce reacción química del soluto con el ión paramagnético, puede descartarse la contribución del mecanismo de contacto a través de enlaces químicos, y por tanto sólo la interacción de pseudo-contacto o dipolar y, si acaso, el mecanismo de Curie entre un núcleo y electrones desapareados pueden dar lugar a PRE significativo (relajación “*outer-sphere*”) (Freed1978, Ayant1975). En campos magnéticos altos (frecuencia de $^1\text{H} \geq 400$ MHz) las velocidades de relajación paramagnética dipolar $R_1^{para, dd}$ y $R_2^{para, dd}$ (que vienen a ser las inversas de los tiempos

de relajación $1/T_1^{\text{para, dd}}$ y $1/T_2^{\text{para, dd}}$, respectivamente) vienen dadas por la teoría de Solomon-Bloembergen (Ecuaciones 2.4 y 2.5) (Solomon1955, Bloembergen1961):

$$R_1^{\text{para, dd}} = \frac{2}{15} \left(\frac{\mu_0}{4\pi} \right)^2 \gamma_I^2 g^2 \mu_B^2 S(S+1) r^{-6} \left\{ \frac{3\tau_c}{1+(\omega_H\tau_c)^2} + \frac{7\tau_c}{1+(\omega_S\tau_c)^2} \right\} \quad \text{Ecuación 2.4}$$

$$R_2^{\text{para, dd}} = \frac{1}{15} \left(\frac{\mu_0}{4\pi} \right)^2 \gamma_I^2 g^2 \mu_B^2 S(S+1) r^{-6} \left\{ 4\tau_c + \frac{3\tau_c}{1+(\omega_H\tau_c)^2} \right\} \quad \text{Ecuación 2.5}$$

donde r es la distancia entre el centro paramagnético y el núcleo observado, μ_0 la permeabilidad del vacío, γ_I la constante giromagnética nuclear, g el factor electrónico, μ_B el magnetón de Bohr, S el número cuántico de espín electrónico ($S = 5/2$ para Mn^{2+} y Fe^{3+} , $S = 1/2$ para Cu^{2+}), ω_H y ω_S las frecuencias de Larmor de ^1H y del electrón, respectivamente, y $\omega_S = 658 \cdot \omega_H$. El tiempo de correlación τ_c en las Ecuaciones 2.4 y 2.5 viene dado por la Ecuación 2.6:

$$\frac{1}{\tau_c} = \frac{1}{T_{1e}} + \frac{1}{\tau_M} + \frac{1}{\tau_R} \quad \text{Ecuación 2.6}$$

donde T_{1e} es el tiempo de relajación electrónico, τ_R es el tiempo de correlación rotacional del soluto; y τ_M es el tiempo de residencia del paramagnético en la superficie del soluto. El proceso más rápido será el que tenga un mayor peso en τ_c .

T_{1e} depende del ion metálico paramagnético, con unos valores situados en el rango de 10^{-12} a 10^{-9} s. τ_R suele tener un valor que va desde 10^{-10} a 10^{-8} en moléculas pequeñas y medianas (peso molecular < 2000). τ_M ha sido calculado (Pintacuda2002) usando $\tau_M = x^2/6D$, donde D es el cociente de difusión y x la distancia requerida por el paramagnético para salir de la interacción del complejo (por ej. para Mn^{2+} la distancia x es 34 \AA) (Iwahara2006) y para distancias mayores que este valor el efecto del paramagnético se vuelve indetectable.

A medida que la fuerza del campo magnético es mayor, $\omega_{S, H} \tau_c \gg 1$ en las Ecuaciones 2.4 y 2.5 y la función de densidad espectral $J(\omega)$ disminuye, así pues, para la misma concentración de paramagnético, el PRE disminuye con el aumento de la fuerza del campo. Por otro lado, una mayor fuerza de campo incrementa la relación $R_1^{\text{para,dd}}/R_2^{\text{para,dd}}$ especialmente en casos de mayor τ_c y por lo tanto son favorables para el método fast-PRE-NMR aquí propuesto (Figura 2.1). Para moléculas de pequeño y mediano tamaño en solventes no viscosos τ_M es mucho mayor que τ_R y por lo tanto no contribuye al valor de τ_c en la Ecuación 2.6. Dado que el más rápido de ambos procesos T_{1e} o τ_R será el que más influya en τ_c , la única condición que se le requiere al paramagnético para obtener la mayor relación $R_1^{\text{para,dd}}/R_2^{\text{para,dd}}$ es que $T_{1e} \geq \tau_R$. Muchos de los paramagnéticos más sencillos de obtener satisfacen esta condición; por ejemplo, Mn^{2+} y Fe^{3+} tienen un T_{1e} en el rango de 100 a 1000 ps, lo cual está cerca del τ_R típico de moléculas pequeñas y medianas.

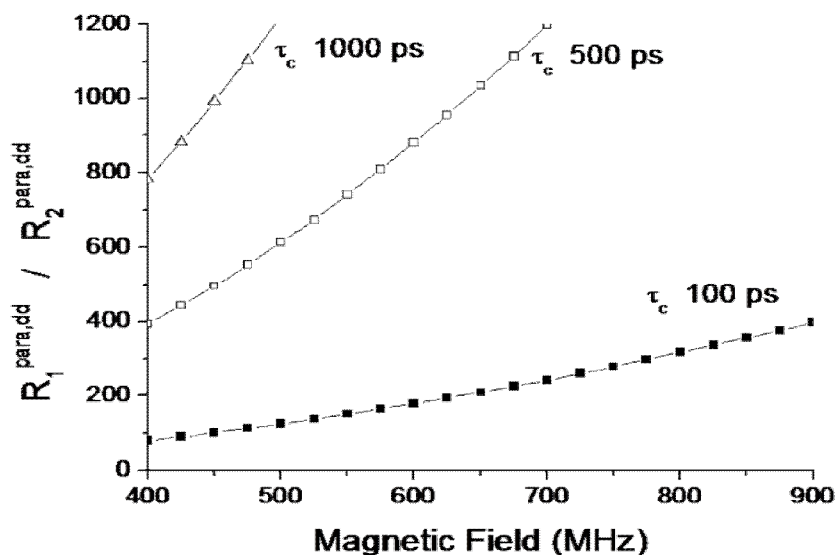
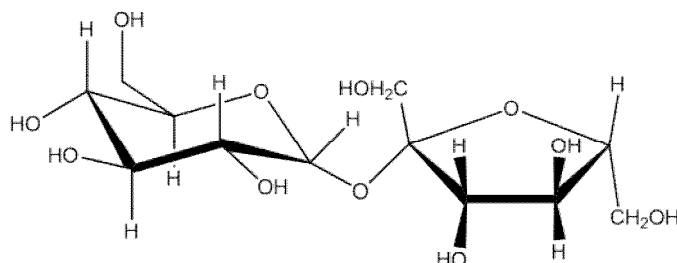


Figura 2.1: Dependencia teórica de la relación $R_1^{\text{para,dd}}/R_2^{\text{para,dd}}$ con el campo magnético. Las curvas representadas corresponden a tiempos de correlación τ_c 1000 ps (triángulos), 500 ps (cuadrados sin relleno) y 100 ps (cuadrados con relleno).

2.3 Procedimiento experimental

2.3.1 Estudio de muestras Sacarosa: $MnSO_4$:



Sacarosa (**1**)

Para la adquisición de los experimentos COSY, TOCSY, HMQC y HMBC se prepararon cinco muestras, con diferentes concentraciones de Mn^{2+} y con la misma concentración de (**1**). Las concentraciones de Mn^{2+} fueron 0, 0.5, 1, 1.5 y 2 mM. La concentración de (**1**) fue de 9.75 mM en todas las muestras. Para la medición de los tiempos de relajación T_1 y T_2 se prepararon muestras adicionales con diferentes concentraciones de Mn^{2+} , con valores de 0.1, 0.2, 0.3, 0.4, 0.6, 0.7, 0.8 y 0.9 mM.

Espectroscopía RMN

Todos los espectros se adquirieron a 300K, en un espectrómetro BRUKER DRX-500. Salvo en los casos especificados, todos los espectros fueron adquiridos con 32 *dummy scans*, que consiste en realizar la secuencia de pulsos correspondiente al experimento pero sin registrar la FID, para así asegurarnos el estado estacionario de espín antes de la adquisición de los datos. Además, en los experimentos 2D se insertó un bloque pulso 90° (1H)-*gradiente* en cada secuencia, justo antes del tiempo de espera

de relajación entre *scans* d_1 , para purgar la magnetización residual del *scan* anterior, con el objetivo de reducir la presencia de artefactos en los espectros (Kupče2007).

Los datos fueron procesados y analizados con el programa MestRe-C v4.9 (*Mestrelab Research*). El ajuste de los datos obtenidos de intensidad/integral a funciones matemáticas y su representación gráfica fueron realizados con Origin v 6.1 (*Originlab Corporation*).

La asignación de las señales de ^1H y ^{13}C de **(1)** se realizó con los espectros adquiridos para las muestras preparadas, de acuerdo con la asignación bibliográfica (Bock1984). Para referirse a las señales de RMN de la sacarosa se usaron las letras g y f para los residuos de glucosa y fructosa, respectivamente.

Medidas de tiempo de relajación longitudinal (T_1) de protón

El valor de T_1 de los diferentes protones se determinó a partir de 14 muestras de concentración constante de sacarosa y de concentraciones de Mn^{2+} entre 0 y 2 mM. Estas medidas se realizaron con la secuencia de saturación progresiva (Mitrovic2001) ($d_1 - 90^\circ - \text{aq}$). Para cada una de las muestras de sacarosa, se realizaron 16 experimentos, variando en cada uno el tiempo de espera entre *scans* (d_1). Los valores d_1 utilizados fueron $4 \cdot 10^{-6}$, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1, 1.1, 1.2, 1.3, 1.4 y 1.5 s. Los experimentos se adquirieron con 128 *dummy scans* y 32 *scans*. Las FIDs fueron adquiridas con 1024 puntos complejos y el tiempo de adquisición (aq) fue de 0.225 s en todos los experimentos. Para cada señal no solapada de **(1)**, el tiempo de relajación longitudinal T_1 se determinó por ajuste no lineal de los valores obtenidos al integrar la señal correspondiente en cada uno de los experimentos, utilizando la siguiente ecuación

$$\frac{I(\tau)}{I^{\max}} = 1 - \exp(-\tau/T_1) \quad \text{Ecuación 2.7}$$

donde $\tau = d_1 + aq$, $I(\tau)$ es la integral de la señal medida en cada espectro al valor correspondiente de τ e I^{\max} es la integral de la señal correspondiente a experimento con mayor valor de τ .

Medidas de tiempo de relajación espín-espín (T_2) de protón

Los T_2 de los protones fueron determinados a partir de las mismas 14 muestras anteriores. Las medidas se hicieron usando la secuencia clásica de CPMG (Meiboom1958) [$d_1-90^\circ-(\delta -180^\circ-\delta)_n-aq$]. El número de ciclos n del bloque CPMG se varió linealmente entre 50 puntos para obtener duraciones del bloque entre 3 y 700 ms. Los espectros fueron adquiridos con 16 *dummy scans* y 16 *scans*. Las FIDs fueron detectadas con 2048 puntos. Para cada señal no solapada de (1), el tiempo de relajación longitudinal T_2 se determinó por ajuste no lineal de los valores obtenidos al integrar la señal correspondiente en cada uno de los experimentos, utilizando la siguiente ecuación:

$$\frac{I(\tau)}{I_0} = \exp(-\tau/T_2) \quad \text{Ecuación 2.8}$$

donde $\tau = 2*n*\delta$ es la duración del bloque CPMG, $I(\tau)$ es la integral de la señal al valor correspondiente de τ e I_0 es la integral de la señal correspondiente al experimento con mínimo valor de τ .

Medidas de tiempo de relajación longitudinal efectivo de protón en experimentos 2D-RMN (T_1^{eff})

Los tiempos de relajación T_1^{eff} fueron medidos para picos de cruce aislados de la sacarosa en los espectros 2D g-COSY, 2D TOCSY, 2D g-HMQC y 2D g-HMBC,

usando el método de saturación progresiva (Mitrovic2001). El T_1^{eff} se determinó mediante un ajuste no lineal a la Ecuación 2.7 de la intensidad normalizada (I/I_{max}) para una serie de valores de d_1 .

2D g-COSY

Estos experimentos fueron adquiridos en modo magnitud, usando la secuencia COSY que emplea los gradientes para la selección de coherencias (la secuencia *cosygp* de la biblioteca de Bruker). Para cada una de las muestras de sacarosa: Mn^{2+} indicadas anteriormente se determinó el T_1^{eff} , adquiriendo 15 experimentos con valores de d_1 en un rango desde $4 \cdot 10^{-6}$ a 1.4 s. La única excepción es la muestra con Mn^{2+} 2 mM en la que se usó un rango de d_1 desde $4 \cdot 10^{-6}$ a 0.7 s. Los espectros se adquirieron con 8 *scans* y el tiempo de adquisición fue de 0.255 s. El número de puntos detectados fue de 128 y 1024 en las dimensiones t_1 y t_2 , respectivamente.

2D TOCSY

Estos experimentos se adquirieron en modo sensible a la fase (la secuencia *mlevst* de la biblioteca Bruker). Para cada una de las muestras de sacarosa: Mn^{2+} indicadas anteriormente se determinó el T_1^{eff} adquiriendo 15 experimentos con valores de d_1 en un rango desde $4 \cdot 10^{-6}$ a 1.4 s. La única excepción es la muestra con Mn^{2+} 2mM en la que se usó un rango de d_1 desde $4 \cdot 10^{-6}$ a 0.7 s. Los espectros se adquirieron con 8 *scans* y el tiempo de adquisición fue de 0.255 s. Se usó el esquema MLEV-17 (Bax1985), para el periodo de bloqueo de espines el tiempo de mezcla fue de 80 ms. El número de puntos detectados fue de 128 y 1024 en las dimensiones t_1 y t_2 , respectivamente.

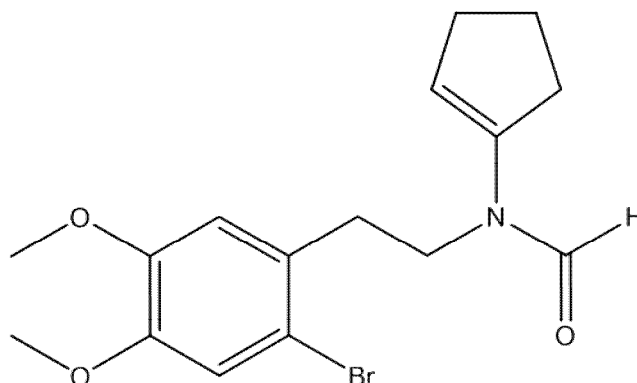
2D g-HMQC

Estos experimentos se adquirieron en modo magnitud usando la secuencia HMQC que emplea los gradientes para la selección de coherencias (la secuencia *inv4gp* en la biblioteca de Bruker). Para cada muestra de sacarosa:Mn²⁺ se determinó el T₁^{eff} adquiriendo 15 experimentos con valores de d₁ en un rango desde 4.10⁻⁶ a 1.4 s. excepto en el caso de las muestras con Mn²⁺ 1.5 y 2 mM en las que se usó un rango de d₁ desde 4.10⁻⁶ a 0.72 s. Los espectros se adquirieron con 16 *scans* y los tiempos de espera en la secuencia se ajustaron a un valor nominal de ¹J_{CH} de 145 Hz. Se utilizó desacoplamiento de banda ancha de carbono durante el tiempo de adquisición (0.12 s) con el esquema GARP (Shaka1985). El número de puntos detectados fue de 128 y 384 en las dimensiones t₁ y t₂, respectivamente.

2D g-HMBC

Estos experimentos se adquirieron en modo magnitud usando la secuencia HMBC que emplea los gradientes para la selección de coherencias (*inv4gplplrnd* en la biblioteca de Bruker). Para cada muestra de sacarosa:Mn²⁺ se determinó el T₁^{eff} adquiriendo 15 experimentos con valores de d₁ en un rango desde 4.10⁻⁶ a 1.4 s para las muestras con Mn²⁺ 0 y 0.5 mM mientras que para las muestras con Mn²⁺ 1, 1.5 y 2 mM se usó un rango de d₁ desde 4.10⁻⁶ a 0.72 s. Los espectros se adquirieron con 16 *scans* y un tiempo de adquisición de 0.255 s. Los tiempos de espera en la secuencia se ajustaron a unos valores nominales de ¹J_{CH} de 145 Hz y ⁿJ_{CH} of 7.7 Hz. El número de puntos detectados fue de 200 y 1024 en las dimensiones t₁ y t₂, respectivamente.

2.3.2 Estudio de muestra *N*-(2-bromo-4,5-dimetoxifenil)-*N*-(ciclopent-1-en-1-il)formamida:*Fe*(*acac*)₃:



(2)

Se prepararon cinco muestras (2):*Fe*(*acac*)₃ en DMSO, variando la concentración de Fe^{3+} mediante adiciones progresivas sobre una muestra inicial. La concentración de la amida (2) se mantuvo prácticamente constante en 10 mM. Las concentraciones de Fe^{3+} fueron 0, 0.157, 0.314, 0.584 y 1.025 mM.

Espectroscopía RMN

Los experimentos de RMN adquiridos para las muestras (2):*Fe*(*acac*)₃ se adquirieron con las mismas condiciones vistas anteriormente para las muestras (1):*MnSO*₄. Se señalan únicamente los cambios introducidos en algunos parámetros. El procesado de los espectros y el tratamiento de los datos se realizó con los mismos programas utilizados para la muestra (1):*MnSO*₄.

Medidas de relajación longitudinal de espín (T_1) y de relajación espín-espín (T_2) de protón

Para la medida de los tiempos T_1 se realizaron 16 experimentos de saturación progresiva variando el d_1 . Para las muestras con $[\text{Fe}^{3+}]$ entre 0 y 0.157 mM se exploraron 16 valores de d_1 entre 10^{-4} y 2.2 s. Para la muestra con $[\text{Fe}^{3+}]$ 0.314 mM se exploraron 16 valores entre 10^{-4} y 1.5 s y para las muestras con $[\text{Fe}^{3+}]$ entre 0.584 y 1.025 mM se exploraron 16 valores entre 10^{-4} y 0.9 s. Para cada señal no solapada de (2) se determinó el tiempo de relajación longitudinal T_1 .

Las medidas de tiempos T_2 se realizaron únicamente para dos muestras, la muestra sin Fe^{3+} y la de máxima concentración de Fe^{3+} . En cada muestra se realizaron 16 experimentos, con duraciones del bloque CPMG comprendidas entre 140 y 4200 ms. Para cada señal sin solapamiento de (2), se determinó el tiempo de relajación espín-espín T_2 .

2D g-COSY y 2D TOCSY

Se adquirieron espectros 2D g-COSY y 2D TOCSY para la validación del método. Para la muestra sin Fe^{3+} se adquirieron ambos espectros por duplicado, con valores de d_1 de 1.5 s y 0.2 s. Se adquirieron también ambos espectros para la muestra de mayor concentración de Fe^{3+} con un d_1 de 0.2 s. Cada espectro se adquirió con 8 *scans* y un tiempo de adquisición de 0.182 s.

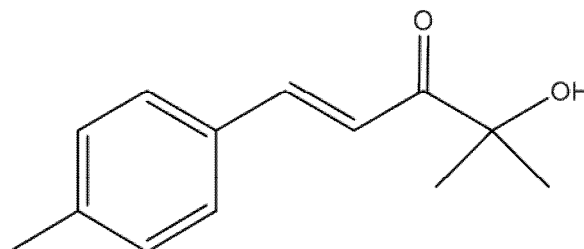
2D g-HMQC

Se adquirieron tres espectros 2D g-HMQC para la validación del método. Se adquirieron dos espectros de la muestra sin Fe^{3+} con d_1 de 1.5 s y 0.3 s. El tercer espectro se adquirió para la muestra de mayor concentración de Fe^{3+} , con d_1 de 0.3 s. Los espectros se adquirieron con 8 *scans* y el tiempo de adquisición fue de 0.184 s.

2D g-HMBC

Se adquirieron tres espectros 2D g-HMBC para la validación del método. Se adquirieron dos espectros de la muestra sin Fe^{3+} , con d_1 de 1.5 s y de 0.3 s. El tercer espectro se adquirió con la muestra de mayor concentración de Fe^{3+} , con d_1 de 0.3 s. Los espectros se adquirieron con 8 *scans* y el tiempo de adquisición fue de 0.184 s.

2.3.3 Estudio de muestras (E)-4-hidroxi-4-metil-1-(p-tolil)pent-1-en-3-ona: $\text{Fe}(\text{acac})_3$:



(3)

Se prepararon dos disoluciones en CDCl_3 del compuesto (3) de concentración 100 mM de (3) una sin paramagnético y otra añadiendo $\text{Fe}(\text{acac})_3$ a concentración de Fe^{3+} de 1 mM.

Espectroscopia RMN

Los experimentos de RMN adquiridos para las muestras (3): $\text{Fe}(\text{acac})_3$ se adquirieron con las mismas condiciones vistas anteriormente para las muestras (1): MnSO_4 . Para las dos muestras se midieron tiempos de relajación T_1 y T_2 y los experimentos 2D g-COSY, 2D TOCSY, 2D-g HMQC y 2D-gHMBC.

2.4 Resultados y discusión

Para moléculas de pequeño y mediano tamaño a frecuencias de resonancia entre 400 y 500 MHz, el efecto PRE, calculado como

$$PRE_{T_i} = T_i^{dia} - T_i^{exp} \quad \text{para } i = 1, 1^{eff}, 2 \quad \text{Ecuación 2.9}$$

depende principalmente de las propiedades de la sustancia paramagnética (espín electrónico, tipo de coordinación del metal y tiempo de relajación electrónica T_{1e}) y de su concentración en la muestra. El efecto PRE introduce una disminución en los tiempos de relajación T_1 's y T_2 's en los núcleos del soluto (ver sección de teoría). Mientras que un efecto PRE en T_1 es útil para la reducción del tiempo de espera entre *scans* de los experimentos de RMN (permite reducir d_1^{opt} en Ecuación 2.2), el efecto PRE en T_2 causa ensanchamiento de señales y pérdida de sensibilidad en los espectros. Las condiciones más favorables para fast-PRE-NMR pueden encontrarse experimentalmente de forma que se produzca la menor relación PRE- T_1 /PRE- T_2 para la mayoría de los protones de moléculas de tamaño mediano-pequeño. Esto requiere fundamentalmente el ajuste de la concentración del paramagnético a su valor idóneo en la muestra y para el experimento de RMN en cuestión.

La forma práctica de optimizar fast-PRE-NMR para una sustancia paramagnética que se utiliza por primera vez, es determinar su concentración óptima con una muestra de referencia de un soluto de tamaño medio-pequeño (datos de referencia PRE) bajo las condiciones experimentales a utilizar (campo magnético, temperatura, disolvente, etc.). Es esperable que esta concentración óptima sea

fácilmente extrapolable a otros solutos de similares características y que presumiblemente vayan a interactuar de forma similar con el paramagnético.

En lo que sigue se presenta un análisis completo de efecto PRE con una muestra de referencia de sacarosa en D₂O utilizando como ión paramagnético Mn²⁺. Basándonos en estos resultados de referencia, finalmente se propone un protocolo simplificado para la determinación rápida de la concentración óptima de esta especie paramagnética en solutos de características similares, y se prueba con diferentes muestras.

2.4.1 Optimización de PRE-T₁ y PRE-T₂:

El efecto PRE-T₁ y PRE-T₂ de un paramagnético se determina en una muestra de referencia a través un estudio de valoración con el paramagnético elegido. En este caso se obtuvieron datos de referencia PRE con una muestra de sacarosa en D₂O y el paramagnético Mn²⁺. A cada concentración de paramagnético se midieron los tiempos de relajación T₁ y T₂ de los protones no solapados de la sacarosa. Estos resultados se dan en la Tabla 2.1 y se representan en las Figuras 2.2a y 2.2b.

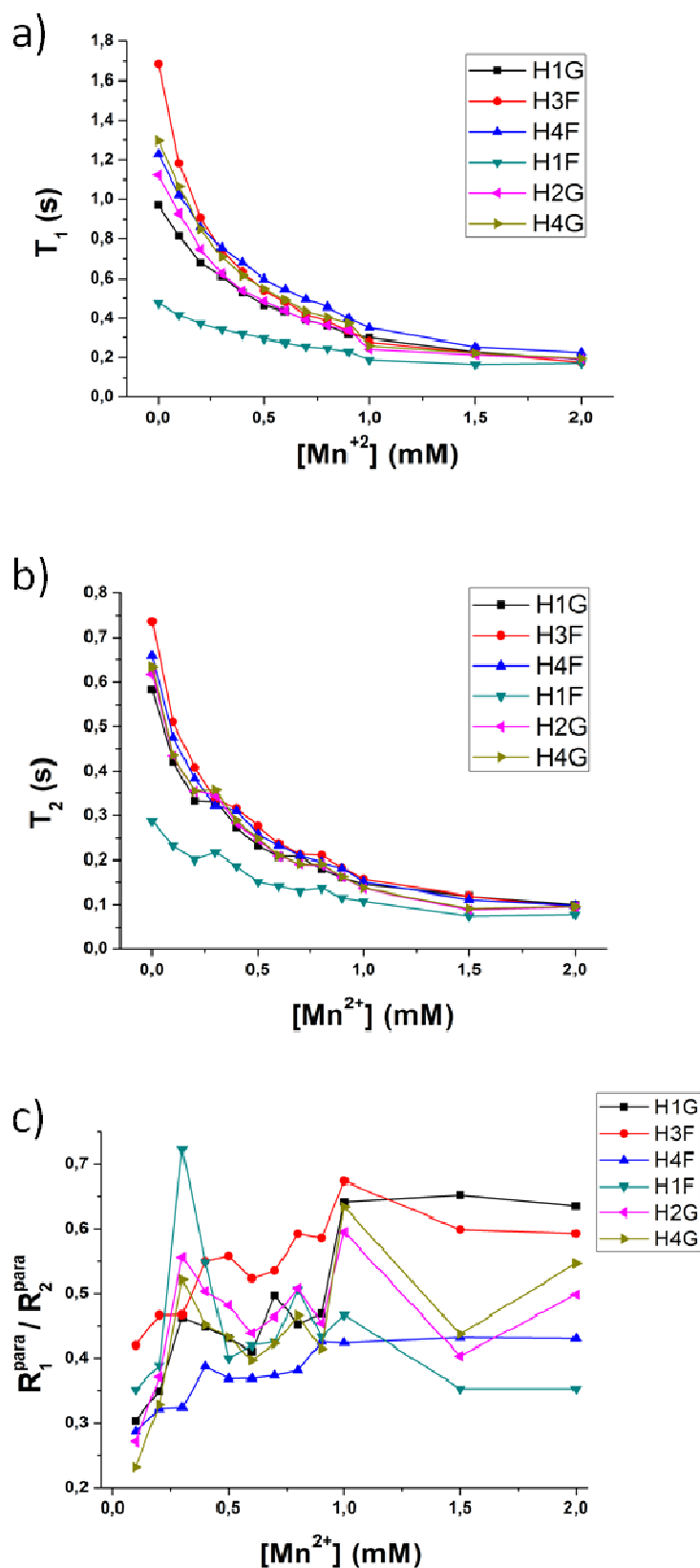


Figura 2.2: Curvas de valoración para los protones no solapados de (1) en función de la concentración de Mn^{2+} (mM). a) Tiempos de relajación longitudinal T_1 (s), b) Tiempos de relajación transversal T_2 (s) y c) relación experimental R_1^{para}/R_2^{para} .

Mn²⁺ (mM)	T₁(s) H-1g	T₁(s) H-2g	T₁(s) H-4g	T₁(s) H-1f	T₁(s) H-3f	T₁(s) H-4f
0.0	0.970	1.122	1.296	0.474	1.684	1.230
0.1	0.814	0.928	1.066	0.417	1.182	1.021
0.2	0.677	0.748	0.849	0.372	0.908	0.860
0.3	0.614	0.626	0.710	0.341	0.743	0.756
0.4	0.527	0.538	0.615	0.318	0.633	0.680
0.5	0.465	0.483	0.548	0.296	0.538	0.597
0.6	0.435	0.437	0.490	0.275	0.480	0.544
0.7	0.392	0.389	0.433	0.256	0.419	0.494
0.8	0.361	0.362	0.405	0.246	0.385	0.456
0.9	0.317	0.336	0.375	0.229	0.335	0.399
1.0	0.299	0.242	0.260	0.187	0.279	0.351
1.5	0.229	0.213	0.222	0.163	0.224	0.255
2.0	0.189	0.193	0.195	0.170	0.174	0.226
Mn²⁺ (mM)	T₂(s) H1g	T₂(s) H2g	T₂(s) H4g	T₂(s) H1f	T₂(s) H3f	T₂(s) H4f
0.0	0.583	0.618	0.634	0.288	0.736	0.658
0.1	0.422	0.434	0.436	0.232	0.510	0.476
0.2	0.334	0.355	0.355	0.201	0.409	0.383
0.3	0.332	0.346	0.357	0.217	0.337	0.323
0.4	0.274	0.283	0.289	0.186	0.317	0.311
0.5	0.232	0.246	0.249	0.150	0.276	0.259
0.6	0.208	0.208	0.210	0.140	0.237	0.232
0.7	0.209	0.191	0.192	0.130	0.212	0.210
0.8	0.180	0.189	0.192	0.136	0.211	0.195
0.9	0.160	0.161	0.163	0.115	0.183	0.182
1.0	0.145	0.136	0.137	0.108	0.157	0.151
1.5	0.119	0.088	0.091	0.073	0.119	0.111
2.0	0.100	0.095	0.096	0.077	0.094	0.098

Tabla 2.1 Tiempos de relajación de protón T₁ (s) y T₂ (s) determinados a 500 MHz para las señales seleccionadas de la sacarosa a diferentes concentraciones de Mn²⁺.

Para cada protón analizado en la Figuras 2.2a ó 2.2b, la correspondiente curva claramente se separa de la linealidad predicha por la Ecuación 2.3, sugiriendo la formación de aductos (Slappendel1982). Cada curva puede descomponerse en dos zonas con cierto carácter lineal, una entre 0 a 0.3 mM que sugiere la formación de aducto 1:1, y otra en el rango 1 a 2 mM que sugiere un aducto 1:2.

En las Figuras 2.2a y 2.2b puede verse que la mayor caída de los tiempos T_1 y T_2 se produce en el rango de $[Mn^{2+}]$ entre 0 a 1.0 mM. A concentraciones superiores, el decaimiento es mucho más moderado y asintótico. Es interesante apuntar que la magnitud del PRE es mayor cuanto mayor es el T_1 (o T_2) original en ausencia de paramagnético. Como consecuencia de esto, el incremento de $[Mn^{2+}]$ tiende a reducir la dispersión en los T_1 and T_2 observados de los diferentes protones de la sacarosa considerados.

La contribución paramagnética a la relajación R_1^{para} y R_2^{para} se calculó para los protones de la sacarosa a cada concentración de Mn^{2+} usando los datos de la Tabla 2.1 con la Ecuación 2.4 y la Ecuación 2.5, respectivamente y la relación experimental R_1^{para}/R_2^{para} está representada en la Figura 2.2c. Para la mayor parte de los protones de la Figura 2.2c hay una tendencia significativa en el cociente R_1^{para}/R_2^{para} a aumentar con la concentración del Mn^{2+} , indicando que el paramagnético causa una mayor caída en el T_1 que en el T_2 de los protones. La única excepción es la señal H1f que tiene un comportamiento errático; aumenta rápidamente en las muestras con concentración de Mn^{2+} 0 a 0.25 mM y entonces disminuye para concentraciones más altas. La interacción paramagnética dipolar entre Mn^{2+} y la sacarosa por sí sola no explica las curvas representadas en la Figura 2.2c, dado que la teoría de Solomon-Bloembergen (Ecuaciones 2.4 y 2.5) predice un cociente R_1^{para}/R_2^{para} constante a concentraciones diferentes de Mn^{2+} . La posibilidad de que la contribución a la relajación por el

mecanismo de Curie sea relevante en este caso puede en principio descartarse, ya que normalmente el resultado de este mecanismo cuando sus efectos son apreciables es un aumento considerable de R_2^{para} y poco efecto sobre R_1^{para} , lo que significa una reducción de la relación $R_1^{\text{para}}/R_2^{\text{para}}$, que es lo contrario de lo que se observa en la Figura 2.2c. Estos resultados deben reflejar pues cierta influencia de la interacción específica entre el paramagnético y el soluto (formación de aductos) que contribuye a la relación PRE- T_1 /PRE- T_2 .

2.4.2 Fast-PRE-NMR con moléculas de tamaño mediano-pequeño:

Para el desarrollo de un protocolo general de fast-PRE-NMR para acelerar los experimentos 2D g-COSY, 2D TOCSY, 2D g-HMQC y 2D g-HMBC, en paralelo con el estudio de PRE- T_1 y PRE- T_2 descrito antes, se hizo una valoración con el paramagnético Mn^{2+} para determinar el efecto PRE en los tiempos de relajación longitudinal efectivos, PRE- T_1^{eff} de estos experimentos 2D. Se obtuvieron PRE- T_1^{eff} para aquellos picos no solapados utilizando el método de saturación progresiva (Mitrovic2001) aplicado al propio experimento 2D (parte experimental).

El PRE- T_1^{eff} determina cómo de rápido un pico de cruce recupera su máxima intensidad de equilibrio en función del tiempo de espera entre *scans* d_1 para una concentración dada de paramagnético. El PRE- T_1^{eff} depende del tiempo de relajación de todas las coherencias generadas durante la secuencia de pulsos, no solo aquellas detectadas, y por lo tanto es más apropiado que el T_1 para definir el tiempo óptimo de relajación entre *scans* en un experimento 2D (Pervushin2002).

La Figura 2.3 muestra dos curvas de saturación progresiva de 2D g-COSY para el pico de cruce H1G-H2G de **(1)** a concentraciones de Mn^{2+} 0 y 1.0 mM. Para $d_1 = 0$ s, la cantidad de señal recuperada en la primera muestra es solo de ~30% de la máxima

intensidad, mientras que es ~75 % en el segundo caso. Los valores correspondientes de $\text{PRE-T}_1^{\text{eff}}$ 1.84s, y 0.25 s para Mn^{2+} 0 y 1.0 mM, respectivamente, son consistentes con estas observaciones.

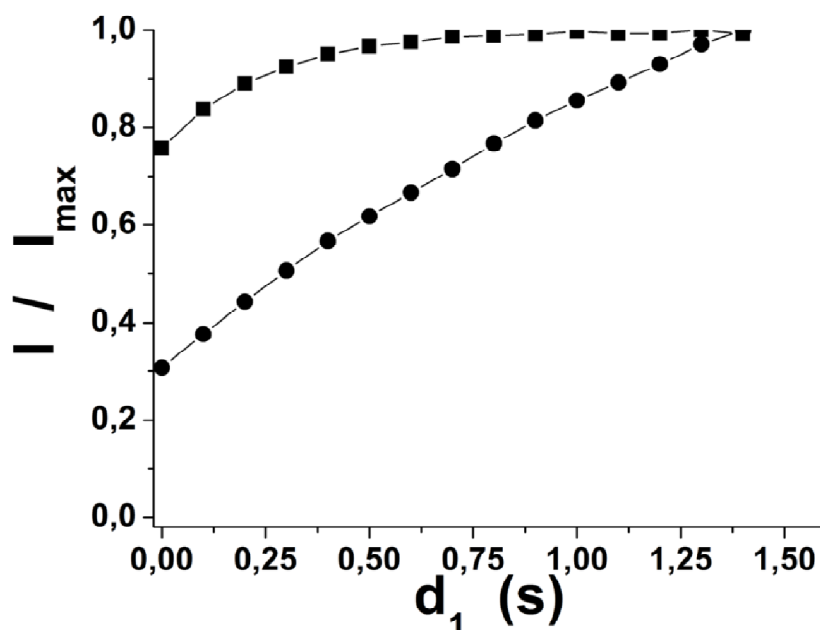


Figura 2.3: Grafica de la intensidad de señal normalizada (%) vs. tiempo entre *scans* d_1 (s) obtenido para el pico de cruce H1G-H2G de (1) con el experimento 2D g-COSY. Los círculos y las cajas corresponden a Mn^{2+} 0 mM y 1.0 mM, respectivamente.

Los valores de $\text{PRE-T}_1^{\text{eff}}$ determinados para diversos picos de cruce de la sacarosa en los espectros 2D g-COSY, 2D TOCSY, 2D g-HMQC y 2D g-HMBC están representados en las Figuras 2.4a, 2.4b, 2.4c, y 2.4d, respectivamente. En estas figuras se puede observar que el mayor cambio en $\text{PRE-T}_1^{\text{eff}}$ ocurre a concentraciones de Mn^{2+} situadas en el rango de 0 a 1 mM en todos los espectros. Para concentraciones por encima de 1.0 mM los $\text{PRE-T}_1^{\text{eff}}$ siguen disminuyendo pero su caída es menos acusada y por lo tanto no causan un gran impacto en la cantidad de señal recuperada en esos experimentos, por el contrario, puede amplificar efectos adversos para la sensibilidad

como incrementar PRE- T_2 e incluso disminuir la sensibilidad de la sonda de RMN en el caso de sondas de detección inversa con un elevado factor-Q y poca tolerancia salina.

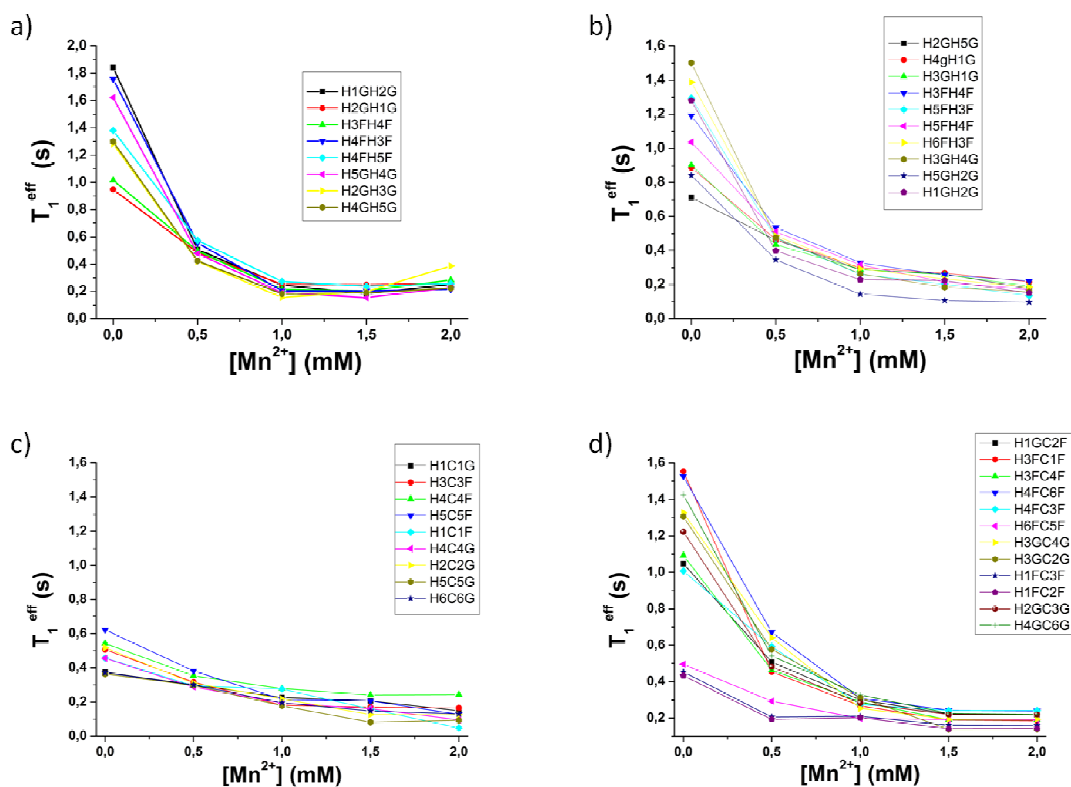


Figura 2.4: Tiempo efectivo de relajación longitudinal (T_1^{eff}) obtenido de los picos de cruce de (1) a diferentes concentraciones de Mn^{2+} a) progressive-saturation 2D g-COSY b) progressive-saturation 2D TOCSY c) progressive-saturation 2D g-HMQC d) progressive-saturation 2D g-HMBC.

Los resultados de PRE- T_1^{eff} de la Figura 2.4 indican que si la duración se hace similar en los diferentes tipos de espectros 2D considerados ($d_0 =$ tiempo de pulsos + adquisición FID ~ 0.3 s), una concentración de Mn^{2+} entre 1.0 y 1.5 mM resulta la mejor elección para maximizar la cantidad de señal recuperada en los diferentes picos de cruce analizados al mínimo valor de d_1 .

Cuando se introducen en la Ecuación 2.2 los datos de PRE- T_1^{eff} obtenidos para sacarosa a concentración Mn^{2+} 1.5 mM, se observa que el d_1^{opt} es 0 para los cuatro

experimentos. Comparando con los valores generalmente asumidos de d_1 comprendidos entre 1.2 y 1.6 s para este tipo de experimentos se pone de manifiesto el considerable ahorro de tiempo de espectrómetro obtenido al utilizar el Fast-PRE-NMR.

Para el caso del experimento 2D-g-HMQC, dado que se utiliza desacoplamiento de banda ancha heteronuclear de alta potencia, en vez del d_1^{opt} de 0 s, se utilizó un valor de 0.3 s para prevenir daños en la sonda como resultado de los ciclos de trabajo excesivamente rápidos. Existen algunas versiones más recientes del experimento HMQC como 2D IPAP-HMQC (Kern2008) que no emplean desacoplamiento heteronuclear y por lo tanto no sería necesario aumentar el d_1^{opt} .

Los espectros de la Figura 2.5 ilustran los resultados obtenidos con el método fast-PRE-NMR para la muestra de sacarosa. El espectro 2D-gHMBC de la Figura 2.5a se adquirió en 17 minutos sin la adición de Mn^{2+} y sufre los efectos perniciosos de utilizar un d_1 muy corto (d_1 0.181 s). La baja calidad de este espectro se refleja en las muchas bandas de ruido t_1 y la pérdida evidente de sensibilidad cuando se le compara con el espectro convencional (Figura 2.5b) adquirido en 60 min. (d_1 1.4 s). Sin embargo, cuando a la muestra se le añade Mn^{2+} 1.5 mM, el espectro puede adquirirse en 11 min. (d_1 0 s) (Figura 2.5c) obteniéndose picos perfectamente resueltos y con una calidad similar a la del espectro convencional y sin ningún tipo de artefactos.

La Tabla 2.2 resume los resultados experimentales del método fast-PRE-NMR aplicado a 2D-g-COSY, 2D-TOCSY, 2D-g-HMQC y 2D-g-HMBC con la muestra de sacarosa a una concentración de Mn^{2+} 1.5 mM. A esa concentración de paramagnético, el tiempo de experimento se reduce, según el caso, en un factor entre 3 y 6, sin la aparición de artefactos y con una resolución de pico de cruce igual a la de los experimentos convencionales.

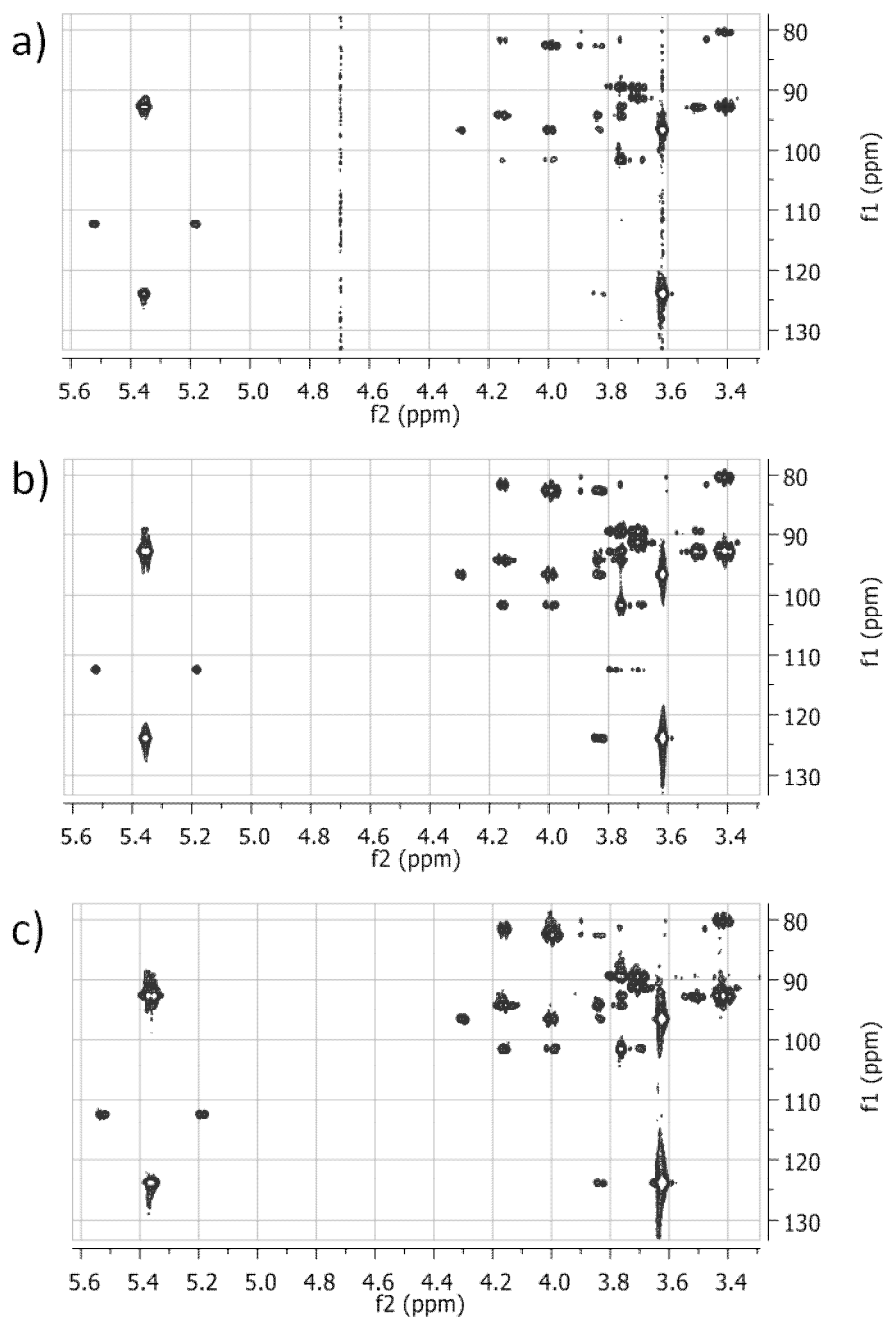


Figura 2.5: Comparación de espectro 2D g-HMBC de (**1**). a) Experimento convencional con d_1 0.2 s en una muestra sin Mn^{2+} adquirido en 17 minutos b) Experimento convencional con d_1 1.4 s en una muestra sin Mn^{2+} adquirido en 60 minutos y c) Experimento Fast-PRE-NMR con d_1 0 s en una muestra con $[Mn^{2+}]$ 1.5 mM adquirido en 11 minutos. El resto de condiciones experimentales y de procesamiento fue igual en los 3 espectros y los contornos se representaron al mismo nivel.

Experimento	Fast-PRE-NMR $d_1^{\text{opt (a)}}$ (s)	Fast-PRE-NMR Exp. time ^(b) (min.)	RMN Convencional Exp. time ^(c) (min.)	Factor reducción tiempo	Fast-PRE-NMR Sensibilidad ^(f) (%)
2D g-COSY	0	5	30	6	56 to 80%
2D TOCSY ^(d)	0	12	60	5	45 to 64%
2D g-HMQC	0.300(e)	8	30	3.75	56 to 87%
2D g-HMBC	0	11	60	5.45	20 to 40%

(a) d_1^{opt} está calculado con la Ecuación 2.2 para un tiempo T_1^{efectivo} nominal de protón de 0.225 s.

(b) 8 *scans* para 2D g-COSY, 2D TOCSY y 2D g-HMQC, 16 *scans* para 2D g-HMBC. Numero de incrementos en t_1 128. Tiempo de adquisición 0.255 en todos los experimentos menos 2D g-HMQC en que fue 0.12 s.

(c) Los experimentos convencionales utilizaron las mismas condiciones que los fast-PRE-NMR excepto el d_1 , que fue de 1.4 s.

(d) Experimento sensible a la fase.

(e) El d_1^{opt} teórico fue 0 s pero el valor final utilizado fue 0.300 s debido a los requisitos del ciclo de trabajo del desacoplamiento heteronuclear.

(f) Calculado como una media de la relación señal-ruido de los picos de en fast-PRE-NMR respecto de los experimentos convencionales.

Tabla 2.2 Resultados experimentales para Fast-PRE-NMR de 2D-COSY, 2D-TOCSY, 2D-HMBC y 2D-HMQC para la muestra de sacarosa y Mn^{2+} 1.5 mM. Los resultados se comparan con el experimento análogo de la muestra sin paramagnético bajo las mismas condiciones exceptuando que el tiempo de espera entre *scans* d_1 es 1.4 s.

Un aspecto importante del experimento fast-PRE-NMR se refiere a la sensibilidad absoluta respecto al experimento convencional. El efecto PRE- T_2 asociado al uso de paramagnéticos inevitablemente introduce pérdidas de sensibilidad en los espectros. Los resultados de la Tabla 2.2 muestran que los experimentos fast-PRE-NMR 2D g-COSY, 2D-TOCSY y 2D-g HMQC tienen una sensibilidad siempre por encima del 45% en las muestras de oligosacáridos estudiados. La peor situación ocurre para el fast-PRE 2D g-HMBC en el cual la sensibilidad de ciertos picos se reduce hasta un 20% respecto del experimento convencional (Tabla 2.2). La razón de esa baja sensibilidad está relacionada con que las coherencias heteronucleares ^1H - ^{13}C de múltiple cuanto emplean un tiempo relativamente mayor en este experimento (65 ms) expuestas a los procesos de relajación transversal incrementados por el PRE- T_2 .

En conjunto, la reducción de la sensibilidad obtenida que se muestra en la Tabla 2.2 para los experimentos fast-PRE-NMR indica que el método es más apropiado para estudiar muestras que no estén limitadas por la sensibilidad del espectrómetro.

2.4.3 Método rápido de ajuste de condiciones de fast-PRE-NMR:

Los resultados obtenidos de PRE-T₁, PRE-T₂ y PRE-T₁^{eff} con la muestra de sacarosa y el paramagnético Mn²⁺ (Figuras 2.2 y 2.4) proporcionan la base para conseguir resultados óptimos con el método fast-PRE-NMR en otras moléculas de tamaño mediano-pequeño utilizando este ión paramagnético. Básicamente, lo que se requiere es ajustar la concentración de paramagnético para que el efecto PRE-T₁ se sitúe aproximadamente en el valor óptimo PRE-T₁^{opt} ~ 0.225 s (valor medio de PRE-T₁ en Figura 2.2a para Mn²⁺ 1.5 mM) que se ha comprobado que es adecuado para el conjunto de experimentos 2D considerados aquí.

En la Figura 2.6 se representa en forma esquemática un protocolo para poder ajustar rápidamente la concentración necesaria de paramagnético en un soluto cualquiera para conseguir PRE-T₁^{opt}. La concentración de paramagnético se incrementa en pequeños pasos y su efecto PRE-T₁ se determina con el experimento rápido de inversión-recuperación (180°-τ-90°-aq). Este ajuste puede hacerse en un solo paso mediante el método conocido como anulación de señal T₁ (*T₁-null*). Para conseguir PRE-T₁^{opt} ≤ 0.225 s, la condición es que se anule o cambie de negativa a positiva la señal del experimento de inversión-recuperación al utilizar un valor $\tau_{\text{null}} = T_1^{\text{opt}} * \ln(2) = 0.156 \text{ s}$.

Dado que PRE-T₂ no se ha considerado para este protocolo de ajuste rápido de la condición de fast-PRE-NMR, es aconsejable comprobar que la calidad del propio

espectro de protón (anchuras de línea y sensibilidad) es aceptable para estas condiciones.

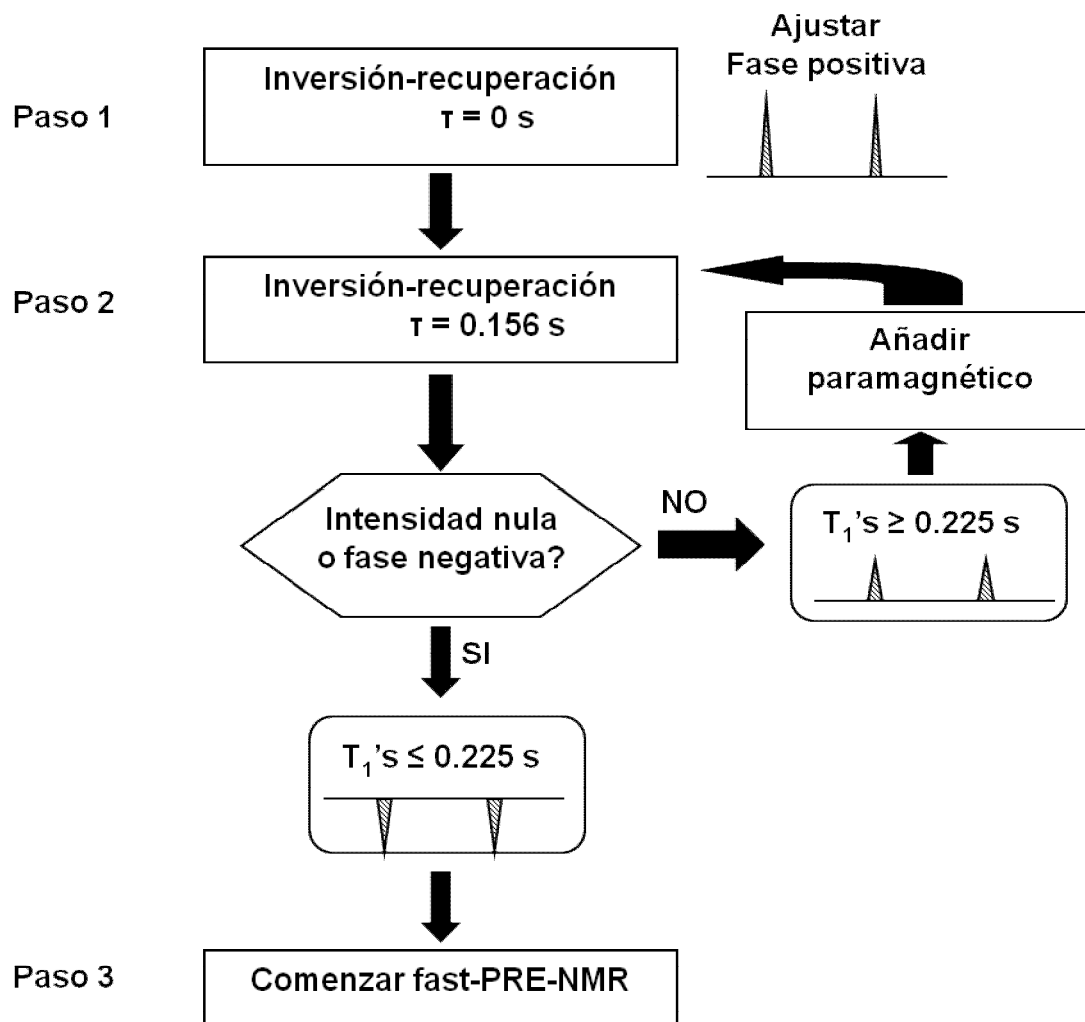


Figura 2.6: Protocolo propuesto para el ajuste rápido de la concentración de paramagnético en experimentos fat-PRE-NMR. El experimento de inversión-recuperación se utiliza con el método de T_1 -null. La concentración de paramagnético se aumenta en pasos pequeños hasta que el T_1^{\max} del soluto es igual o inferior a 0.225 s.

2.4.4 Comprobación de la metodología Fast-PRE-NMR con otros solutos:

Para comprobar la viabilidad y/o generalidad del método Fast-PRE-NMR para otro tipo de solutos, se comprobó su funcionamiento para diferentes condiciones experimentales en cuanto a soluto, paramagnético y disolvente. En los experimentos 2D considerados se ajustó el nivel de efecto PRE para conseguir valores óptimos de $T_1^{\text{opt}} \sim 0.225$ s. Este valor óptimo es precisamente el que se obtuvo anteriormente en el estudio del sistema **1**:Mn²⁺.

*Estudio de la muestra **2**:Fe(acac)₃ en DMSO*

Para la molécula (**2**) la valoración con el paramagnético Fe³⁺ se llevó a cabo variando la concentración de Fe³⁺ desde cero hasta 1.025 mM. Las muestras de concentración intermedia se realizaron para seguir las variaciones de T_1 . Tras cada adición no se usó el método rápido T_1 -null para obtener un dato aproximado del PRE- T_1 , si no que se realizó el experimento de saturación-progresiva para la medida exacta de los T_1 de todos los protones no solapados de (**2**). En la Tabla 2.3 se recoge la evolución del T_1 y puede verse que a la concentración 1.025 M de paramagnético se obtienen valores adecuados para fast-PRE-NMR con valores de $T_1^{\text{exp}} \sim 0.225$ s. Con esta muestra se adquirieron los cuatro experimentos fast-PRE-NMR, 2D g-COSY, 2D-TOCSY 2D-g HMBC y 2D-g HMQC. Para realizar la comparación se repitieron en idénticas condiciones los espectros rápidos de esta muestra sin el paramagnético y también se repitieron con el pulsado a la velocidad habitual. Las Figuras 2.7, 2.8, 2.9 y 2.10 son comparaciones entre los distintos experimentos convencionales y sus variantes fast-PRE. Se han omitido los espectros obtenidos en condiciones fast, pero sin adición

de metal, ya que son equivalentes a lo visto en la Figura 2.5a (baja calidad de espectro por aparición de ruido y artefactos)

Fe³⁺ (mM)	T₁(s) H-1	T₁(s) H-9	T₁(s) H-8	T₁(s) H-2	T₁(s) H-7	T₁(s) H-5	T₁(s) H-3	T₁(s) H-4
0	1,2	1,37	0,92	0,96	0,46	0,55	0,65	0,75
0,157	0,82	0,92	0,69	0,72	0,38	0,45	0,5	0,58
0,314	0,59	0,64	0,53	0,55	0,33	0,37	0,4	0,45
0,584	0,39	0,41	0,37	0,38	0,26	0,28	0,3	0,32
1,025	0,23	0,24	0,24	0,24	0,19	0,18	0,2	0,22

Tabla 2.3: Tiempos de relajación de protón T_1 (s) determinados a 500 MHz para las señales seleccionadas de (2) a diferentes concentraciones de Fe^{3+} .

Igualmente, se puede observar en las Figuras 2.7 a 2.10 que, al igual que en el caso anterior, la aplicación del método fast-PRE tiene como resultado experimentos con calidad similar al espectro original y carentes de artefactos, si bien obtenidos en un tiempo sustancialmente menor.

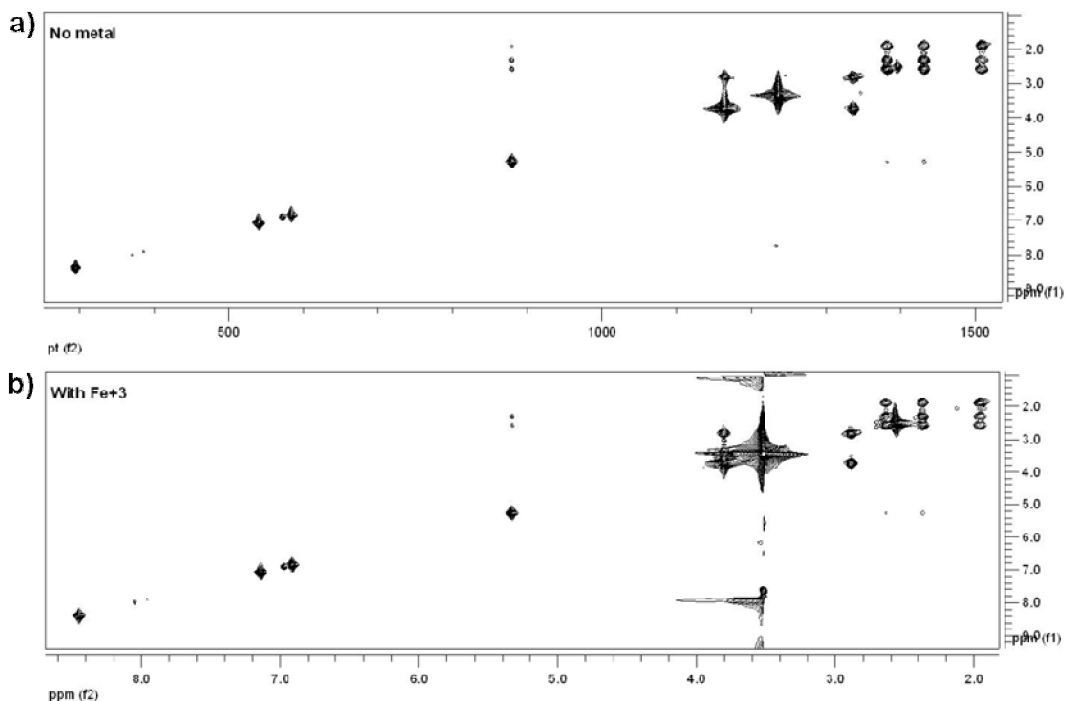


Figura 2.7: Comparación de espectros COSY de (2) a) en condiciones normales y ausencia de paramagnético y b) en condiciones de fast-PRE con concentración de 1,025 mM de Fe^{3+} . Ambos espectros están representados a la misma escala de contornos. La duración del espectro a) fue de 28 minutos, mientras que la del b) fue de 7 minutos.

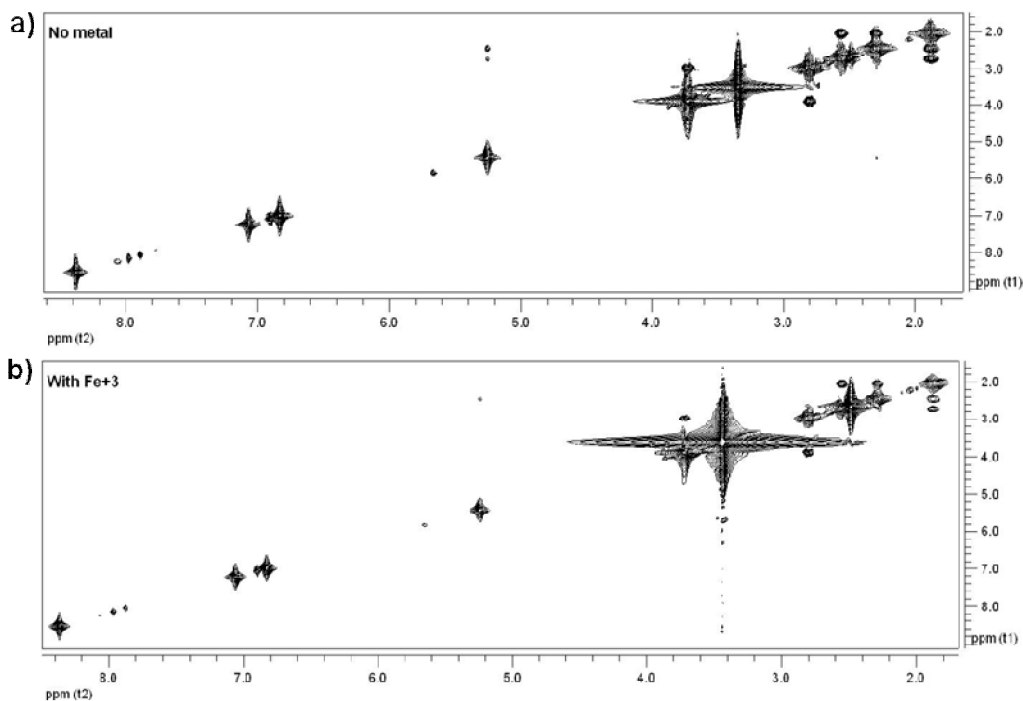


Figura 2.8: Comparación de espectros TOCSY de (2) a) en condiciones normales y ausencia de paramagnético y b) en condiciones de fast-PRE con concentración de 1,025 mM de Fe^{3+} . Ambos espectros están representados a la misma escala de contornos. La acumulación del espectro a) fue de 56 minutos, mientras que la del b) fue de 13 minutos.

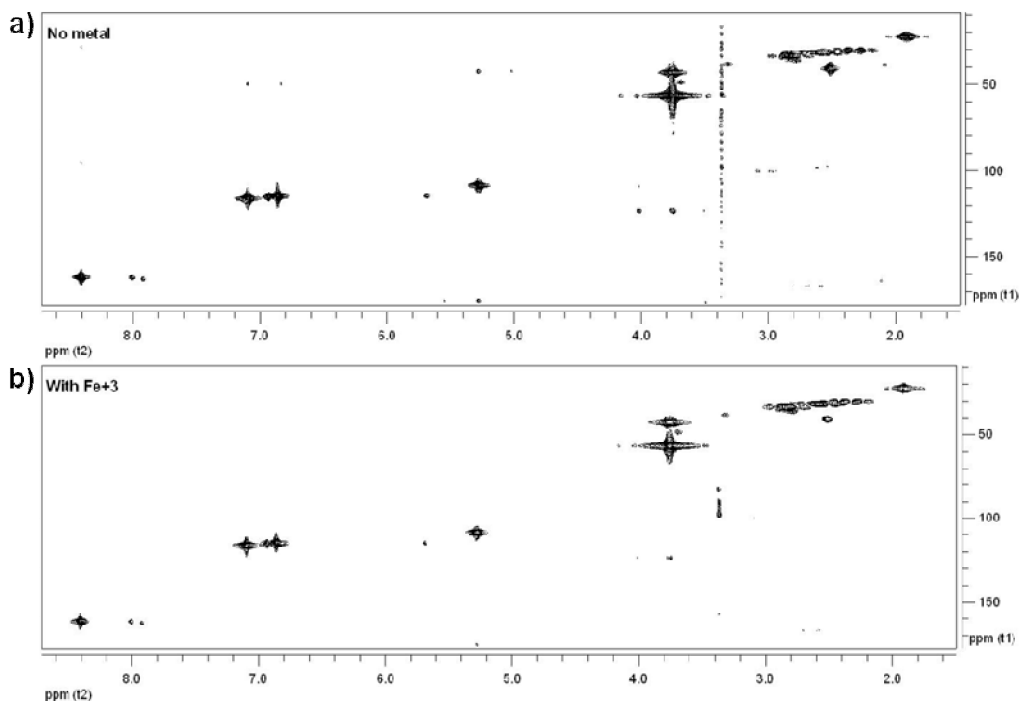


Figura 2.9: Comparación de espectros HMQC de (2) a) en condiciones normales y ausencia de paramagnético y b) en condiciones de fast-PRE con concentración de 1,025 mM de Fe^{3+} . Ambos espectros están representados a la misma escala de contornos. La acumulación del espectro a) fue de 34 minutos, mientras que la del b) fue de 13 minutos.

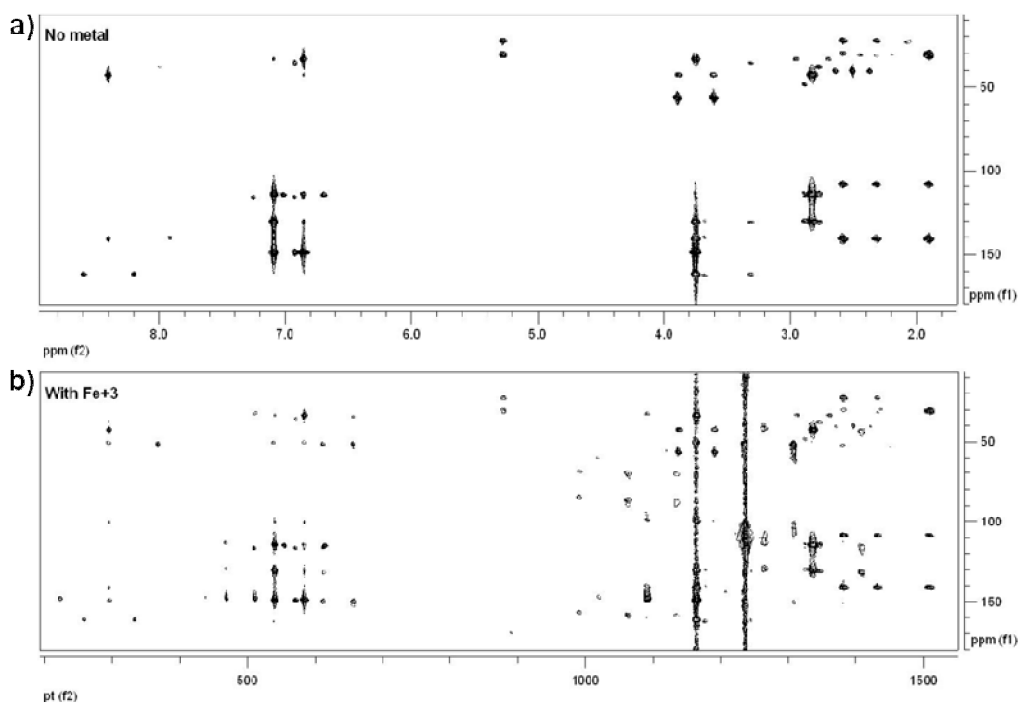


Figura 2.10: Comparación de espectros HMBC de (2) a) en condiciones normales y ausencia de paramagnético y b) en condiciones de fast-PRE con concentración de 1,025 mM de Fe^{3+} . Ambos espectros están representados a la misma escala de contornos. La acumulación del espectro a) fue de 32 minutos, mientras que la del b) fue de 12 minutos.

Estudio de la muestra 3: Fe(acac)₃ en CDCl₃

Para la molécula (**3**) la valoración con el paramagnético Fe³⁺ se realizó entre concentraciones de Fe³⁺ desde cero hasta 1 mM, de acuerdo con el método rápido T₁-null de la Figura 2.6. Con una concentración de Fe³⁺ de 1 mM se obtienen valores de T₁^{exp} ~0.225 s, aptos para la aplicación de los experimentos desarrollados. A esta concentración de paramagnético se realizaron los experimentos fast-PRE-NMR 2D g-COSY, 2D-TOCSY y 2D-g HMQC. En las Figuras 2.11, 2.12, 2.13 y 2.14 se pueden ver los resultados.

En este caso, se aplicó rigurosamente el protocolo de acción propuesto en la Figura 2.6, midiendo solamente T₁ tras cada adición, y adquiriendo los espectros fast-PRE-NMR solamente tras la última adición. Además, no se realizaron los espectros fast en ausencia de paramagnético como en los casos anteriores. Las comparaciones entre los espectros convencionales en ausencia de paramagnético y las versiones fast-PRE-NMR con la cantidad apropiada de metal nos permiten comprobar que también en este caso obtenemos espectros de calidad muy similar y en mucho menos tiempo.

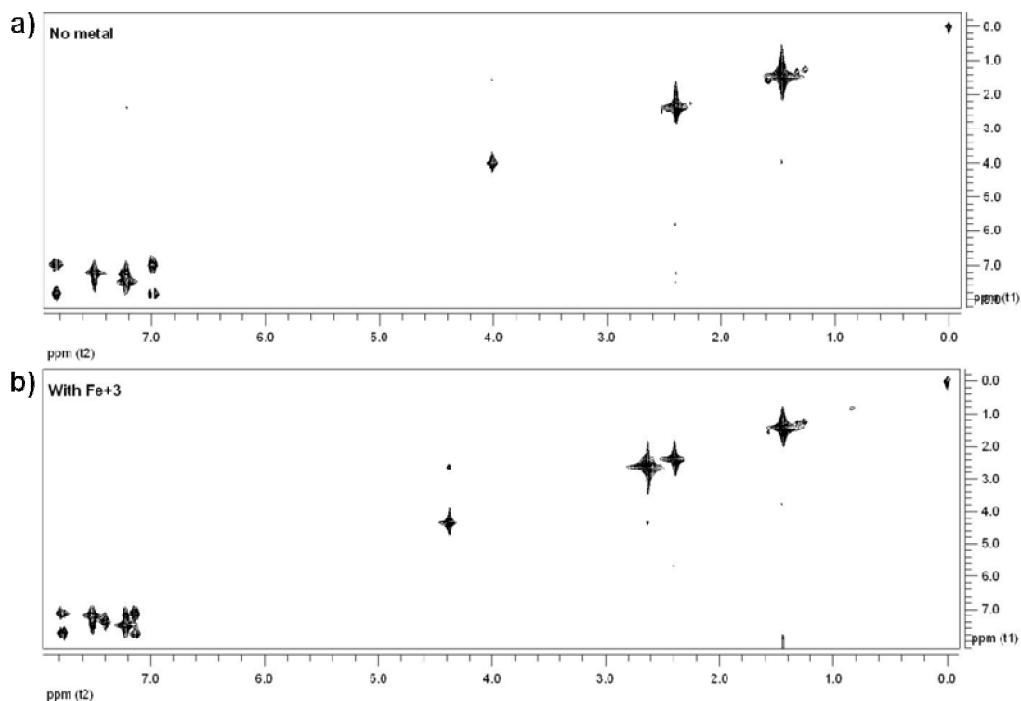


Figura 2.11: Comparación de espectros COSY de **(3)** a) en condiciones normales y ausencia de paramagnético y b) en condiciones de fast-PRE con concentración de 1 mM de Fe^{3+} . Ambos espectros están representados a la misma escala de contornos. La duración del espectro a) fue de 28 minutos, mientras que la del b) fue de 7 minutos.

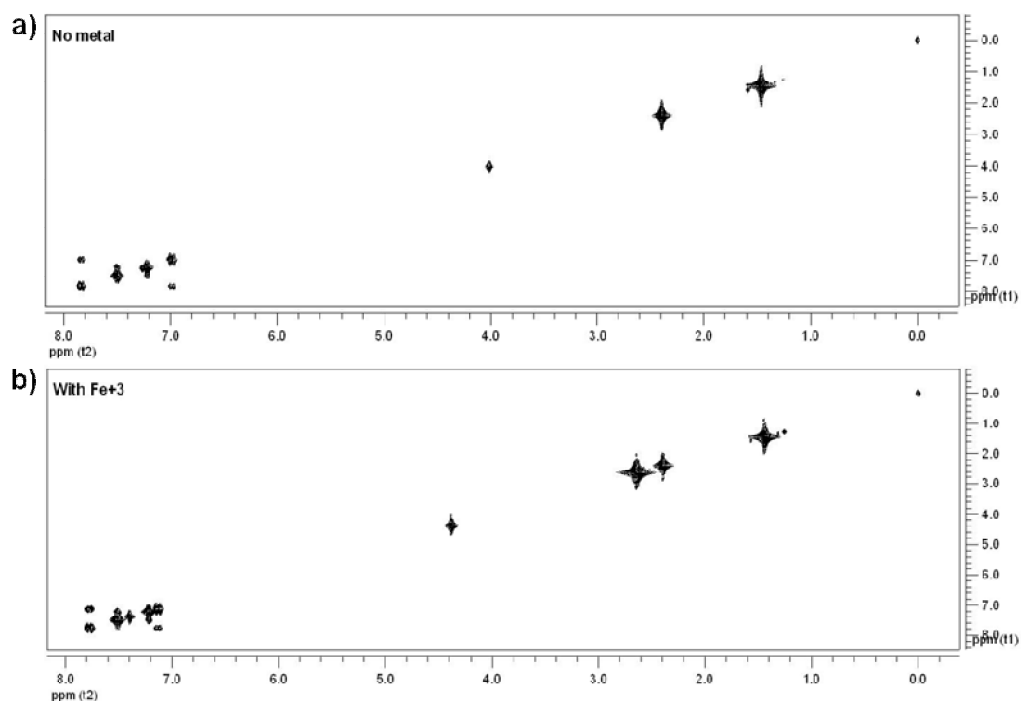


Figura 2.12: Comparación de espectros TOCSY de **(3)** a) en condiciones normales y ausencia de paramagnético y b) en condiciones de fast-PRE con concentración de 1 mM de Fe^{3+} . Ambos espectros están representados a la misma escala de contornos. La acumulación del espectro a) fue de 56 minutos, mientras que la del b) fue de 13 minutos.

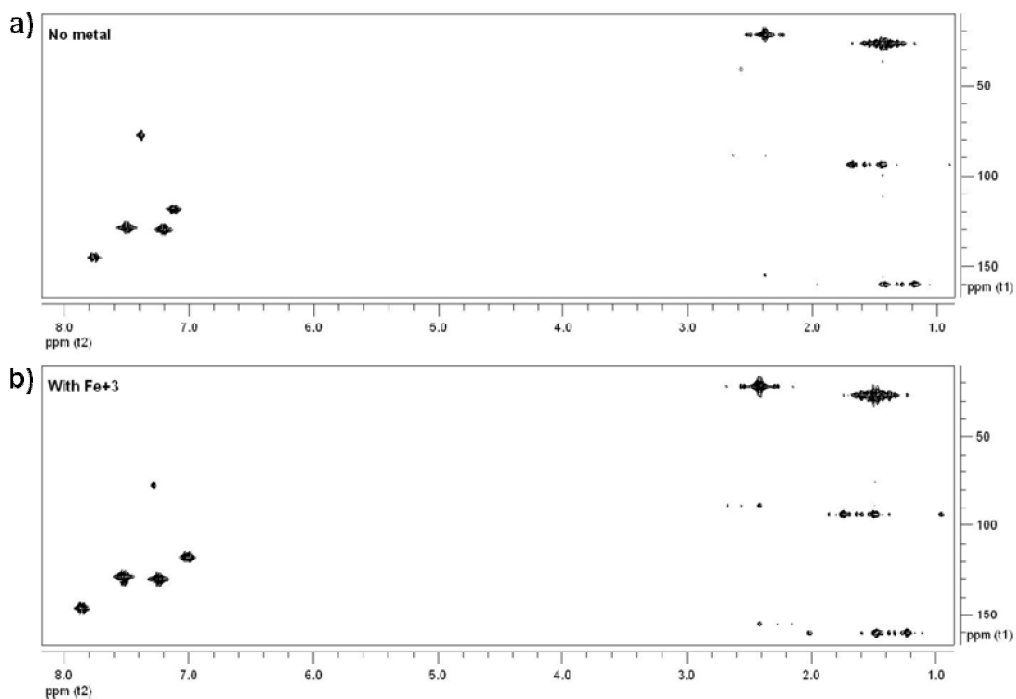


Figura 2.13: Comparación de espectros HMQC de (3) a) en condiciones normales y ausencia de paramagnético y b) en condiciones de fast-PRE con concentración de 1 mM de Fe^{3+} . Ambos espectros están representados a la misma escala de contornos. La acumulación del espectro a) fue de 34 minutos, mientras que la del b) fue de 13 minutos.

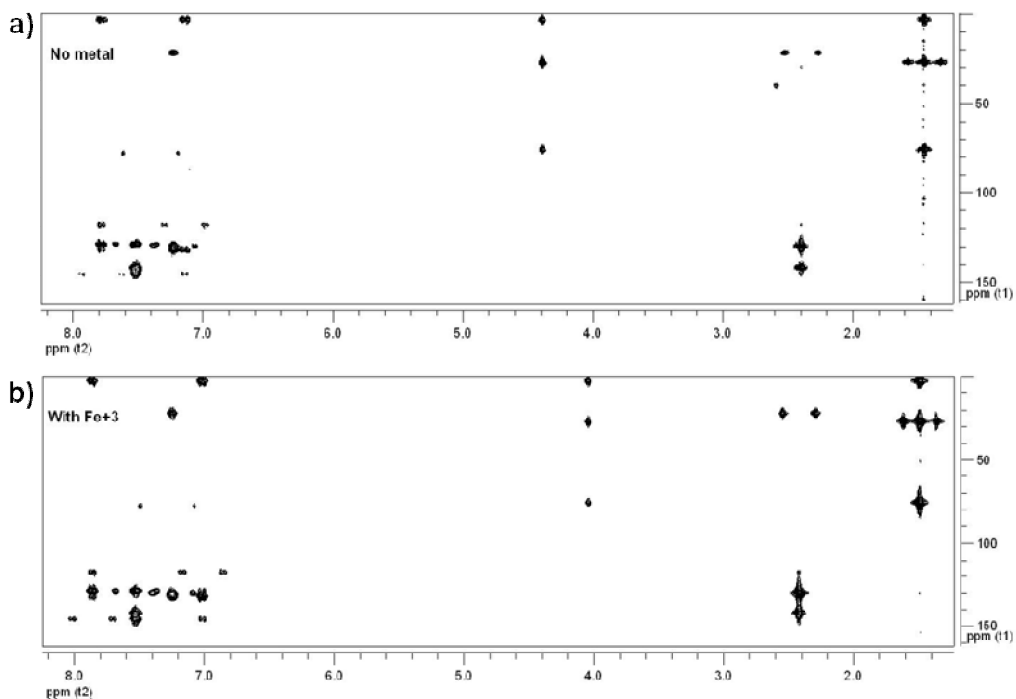


Figura 2.14: Comparación de espectros HMBC de (3) a) en condiciones normales y ausencia de paramagnético y b) en condiciones de fast-PRE con concentración de 1 mM de Fe^{3+} . Ambos espectros están representados a la misma escala de contornos. La acumulación del espectro a) fue de 32 minutos, mientras que la del b) fue de 12 minutos.

2.5 Conclusiones

La adición de sustancias paramagnéticas para producir efecto PRE en solutos de tamaño mediano-pequeño puede ser un método conveniente para disminuir o incluso suprimir el tiempo de espera entre *scans* (d_1) en los experimentos nD-RMN que establecen correlaciones a través del acoplamiento escalar. Utilizando esta estrategia, se ha propuesto el método fast-PRE-NMR que permite adquirir experimentos nD-RMN con un ahorro considerable de tiempo de espectrómetro evaluado en un factor entre 3 y 6. El método se ha probado para los experimentos fast-PRE-NMR 2D g-COSY, 2D TOCSY, 2D g-HMQC and 2D g-HMBC obteniéndose espectros de una calidad similar a los convencionales.

La simplicidad del método Fast-PRE-NMR, que no requiere secuencias de pulsos o procesamiento de datos especiales, y el hecho de que se puedan utilizar especies paramagnéticas de bajo coste (ej. Mn^{2+} o Fe^{3+}) hace que pueda tener aplicaciones para la aceleración de análisis rutinarios de nD-RMN de otros solutos de tamaño mediano-pequeño por ej. para acelerar determinados protocolos de control de calidad basados en RMN en los que se conoce de antemano cual es el soluto.

El protocolo propuesto para el cálculo de la cantidad de sustancia paramagnética es de muy sencilla aplicación, y permite conocer la cantidad necesaria de paramagnético en muy pocos minutos.

2.6 Referencias

- Ayant, Y.; Belorizky, E.; Alizon, J.; Gallice, J., “Calcul des densités spectrales résultant d'un mouvement aléatoire de translation en relaxation par interaction dipolaire magnétique dans les liquides”, *Le Journal de Physique*, **1975**, *36*, 991-1004.
- Bakhmutov, V.I., “Paramagnetic NMR Relaxation” En *Practical NMR Relaxation for Chemists*, **2004**, John Wiley & Sons: Chichester, UK. pp 179-195.
- Bax, A.; Davis, G.D., “MLEV-17-Based Two-Dimensional Homonuclear Magnetization Transfer Spectroscopy”, *J. Magn. Reson.*, **1985**, *65*, 355-360.
- Bertini, I.; Luchinat, C., “NMR of paramagnetic substances”, *Coord. Chem. Rev.*, **1996**, *150*, 1-300.
- Bloembergen, N.; Morgan, L.O., “Proton Relaxation Times in Paramagnetic Solutions. Effects of Electron Spin Relaxation”, *J. Chem. Phys.*, **1961**, *34*, 842–850.
- Bock, K.; Pedersen, C.; Pedersen, H., “Carbon-13 Nuclear Magnetic Resonance Data for Oligosaccharides”, *Adv. Carbohydr. Chem. Biochem.*, **1984**, *42*, 193-225.
- Cai, S.; Seu, C.; Kovacs, Z.; Sherry, A.D.; Chen, Y., “Sensitivity Enhancement of Multidimensional NMR Experiments by Paramagnetic Relaxation Effects”, *J. Am. Chem. Soc.*, **2006**, *128*, 13474-13478.
- Ernst, R.R.; Bodenhausen, G.; Wokaum, A., “Principles of Nuclear Magnetic Resonance in One and Two Dimensions”, Oxford University Press, Ed.; Clarendon Press: New York, **1987**.
- Freed, J.H., “Dynamic effects of pair correlation functions on spin relaxation by translational diffusion in liquids. II. Finite jumps and independent T_1 processes”, *J. Chem. Phys.*, **1978**, *68*, 4034-4037.
- Iwahara, I.; Clore, M.G., “Detecting transient intermediates in macromolecular binding by paramagnetic NMR”, *Nature*, **2006**, *440*, 1227-1230.
- Kern, T.; Schanda, P.; Brutscher, B., “Sensitivity-enhanced IPAP-SOFAST-HMQC for fast-pulsing 2D NMR with reduced radiofrequency load”, *J. Magn. Reson.*, **2008**, *190*, 333-338.
- Kupče, E.; Freeman, R., “Fast multidimensional NMR by polarization sharing”, *Magn. Reson. Chem.*, **2007**, *45*, 2-4.

- Meiboom, S.; Gill, D., “Modified Spin-Echo Method for Measuring Nuclear Relaxation Times”, *Rev. Sci. Instr.*, **1958**, 29, 688-691.

- Mitrovic, V.F.; Sigmund, E.E.; Halperin, W.P., “Progressive saturation NMR relaxation”, *Phys. Review B.*, **2001**, 64, 24520-24527.

- Pervushin, K.; Vögeli, B.; Eletsky, A., “Longitudinal ¹H Relaxation Optimization in TROSY NMR Spectroscopy”, *J. Am. Chem. Soc.*, **2002**, 124, 12898-12902.

- Pintacuda, G.; Otting, G., “Identification of Protein Surfaces by NMR Measurements with a Paramagnetic Gd(III) Chelate”, *J. Am. Chem. Soc.*, **2002**, 124, 372-373.

- Shaka, A.J.; Barker, P.B.; Freeman, R., “Computer-optimized decoupling scheme for wideband applications and low-level operation”, *J. Magn. Reson.*, **1985**, 64, 547-552.

- Slappendel, S.; Aasa, R.; Falk, K.E.; Malmström, B.G.; Vänngård, T.; Veldink, G.A.; Vliegthart, J.F.G., “¹H-NMR spectroscopic study on the binding of alcohols to soybean lipoxygenase-1”, *Biochim. et Biophys. Acta.*, **1982**, 708, 266-271.

- Solomon, I., “Relaxation Processes in a System of Two Spins”, *Phys. Rev.*, **1955**, 99, 559-565.

Capítulo 3:
Adquisición rápida de experimentos
multidimensionales de RMN en paralelo mediante
espectroscopía localizada

3.1 Introducción

En el anterior capítulo se trató la posibilidad de hacer experimentos rápidos de RMN multidimensional mediante la adición de alguna sustancia paramagnética a la muestra, de forma que se pueda reducir el tiempo de espera necesario entre acumulaciones sucesivas (d_1) y por lo tanto la duración total del experimento. En este capítulo se plantea una posibilidad diferente para reducir d_1 que no requiere ninguna adición química a la muestra. La nueva estrategia que se propone se basa en el uso de espectroscopía localizada espacialmente. Esta metodología ya se ha utilizado para acelerar la medida de tiempos de relajación longitudinal T_1 (Loening2003) y también para la adquisición a velocidad convencional de espectros multidimensionales localizados en estudios de MRI (Prescot2005). Sin embargo, nunca se había aplicado esta estrategia como forma general de acelerar la adquisición de un conjunto de experimentos multidimensionales, un método al que se ha denominado PALSYP (PARallel Localized-SpectroscopY). Una versión inicial, menos perfeccionada, de este tipo de metodología consistió en el experimento 2D MS COSY que se describe en la parte 4 de este capítulo.

3.2 Teoría

3.2.1 Espectroscopía de RMN localizada espacialmente:

Es conveniente hacer una introducción breve de los principios de espectroscopía localizada, aplicados en el desarrollo de las metodologías expuestas en este capítulo. Un desarrollo más completo de la teoría puede encontrarse en bibliografía específica sobre MRI (Callaghan1991). Para simplificar la explicación se considerará el caso de una muestra con un solo tipo de señal de RMN, por ej. la señal de H₂O, y se obviarán efectos de desplazamiento químico, acoplamientos y relajación.

Efecto de un gradiente de campo

Cuando se aplica un gradiente de campo que varía linealmente a lo largo de un eje espacial arbitrario en la muestra, por ej. un gradiente G_z , el campo neto que sienten los espines nucleares en cada punto de la muestra es la suma del campo principal B_0 y el gradiente G_z (Figura 3.1).

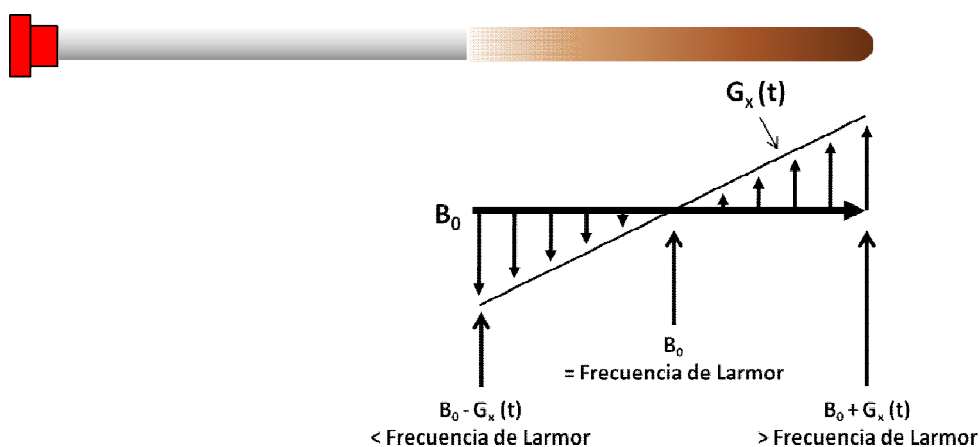


Figura 3.1: Efecto de un gradiente de campo G_z en la frecuencia de resonancia de Larmor. Solamente en el llamado isocentro de la muestra la frecuencia de Larmor es igual a la existente en ausencia del gradiente.

Cuando la forma del gradiente es lineal, el campo neto a lo largo del eje de aplicación del gradiente varía también linealmente con la posición, y con ello también la correspondiente frecuencia de resonancia de Larmor (ω) de los espines (Figura 3.1). Para espines situados a una misma coordenada respecto al eje del gradiente aplicado, esto es, en una misma rodaja de la muestra, les corresponderá un mismo valor de su frecuencia de Larmor (Ecuaciones 3.1 y 3.2).

$$\text{Sin gradiente:} \quad \omega = \gamma \cdot B_0 \quad \text{Ecuación 3.1}$$

$$\text{Con gradiente } G_z: \quad \omega(z) = \gamma \cdot (B_0 + G_z) \quad \text{Ecuación 3.2}$$

Método de pulsado en núcleos de una rodaja de muestra

Como se vio antes, al aplicar un gradiente de campo, la frecuencia de resonancia de Larmor de los espines se hace dependiente de su posición en la muestra. Por otra parte, en los espectrómetros modernos es posible aplicar un pulso de RF selectivo a una determinada frecuencia de Larmor y que cubra una determinada anchura de banda del espectro. Si se aplican simultáneamente el gradiente y el pulso selectivo, el pulso sólo producirá rotación de la magnetización para los espines localizados espacialmente en una determinada rodaja de la muestra para los que se cumpla exactamente la condición de Larmor (Figura 3.2).

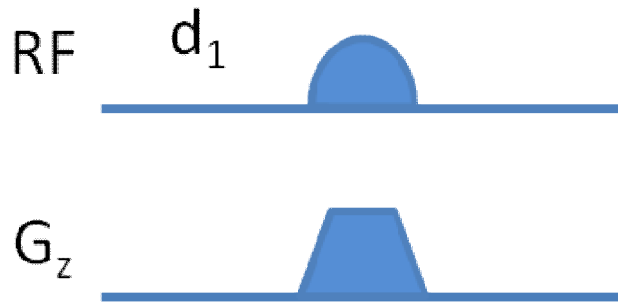


Figura 3.2: Experimento de RMN para excitación espacialmente selectiva de espines en una sección de la muestra a una determinada altura en el eje Z.

Cuando se aplica un gradiente en la dirección Z, las siguientes ecuaciones permiten conocer la posición de la sección de la muestra dentro de la cual los espines se verán afectados por el pulso selectivo:

Frecuencia offset pulso $(Hz) = \gamma \cdot z \cdot G_z$ Ecuación 3.3

Anchura de banda pulso $(Hz) = \gamma \cdot Thk \cdot G_z$ Ecuación 3.4

Donde G_z es la potencia del gradiente (G/cm), z es la coordenada de altura (positiva o negativa) respecto al isocentro de la muestra (cm) y Thk (*thickness*) se refiere a la anchura de la rodaja (cm). La Figura 3.3 resume la relación entre las características de frecuencia del pulso y la localización espacial que se produce.

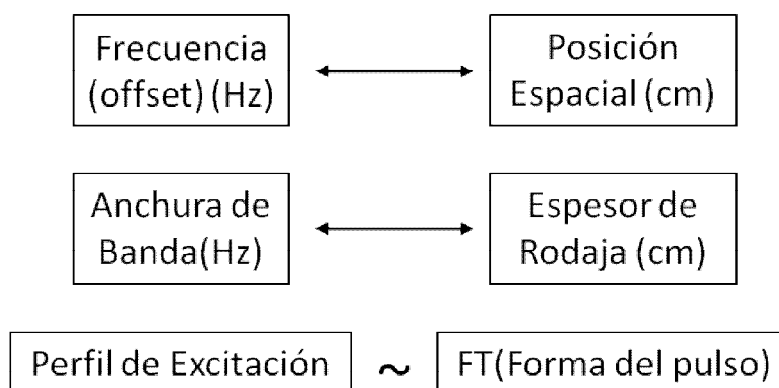


Figura 3.3: Relación entre las propiedades del pulso y las coordenadas espaciales de los espines que son afectados en presencia del gradiente.

3.2.2 Adquisición rápida de múltiples espectros mediante espectroscopía de RMN localizada: Método PALSy:

El tiempo de espera entre pulsos en acumulaciones sucesivas (d_1) puede acortarse si se divide virtualmente la zona de excitación y adquisición del tubo de RMN en una serie de rodajas virtuales, no solapadas entre sí, y sobre las que podemos adquirir diferentes espectros n-dimensionales de forma independiente (Figura 3.4). Así, durante la adquisición de puntos del experimento (p. ej. dimensión t_1), mientras se pulsa sobre una determinada rodaja, los espines en el resto de rodajas no se verían afectados y bien se mantendrán en la situación de equilibrio (si ésta ya se alcanzó o porque no se ha pulsado sobre ellas todavía), o bien continuarán su retorno hacia el equilibrio (si se había pulsado sobre ellas con anterioridad). Si la rodaja seleccionada se va alternando ordenadamente durante *scans* sucesivos del experimento y el tiempo de pulsado de cada una es aproximadamente el mismo para las n rodajas (Figura 3.4), el resultado es que se puede reducir el tiempo de espera d_1 individual para el pulsado de un espectro en un factor n, donde n es el número de rodajas utilizado.

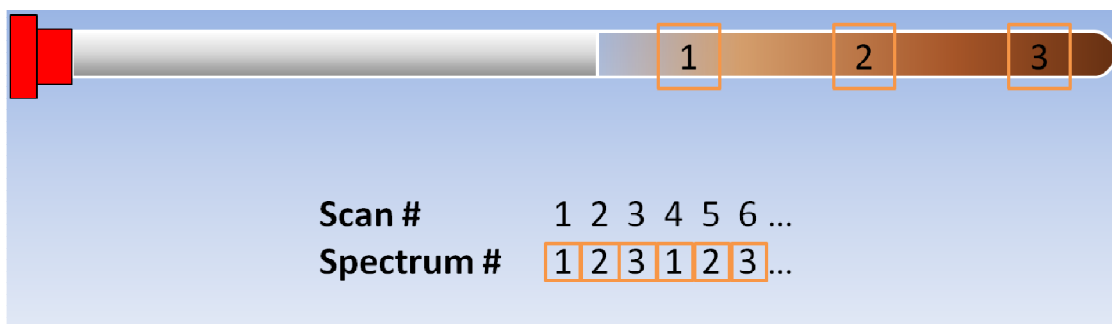


Figura 3.4: Esquema del método PALSy para la adquisición de diferentes espectros en tres rodajas de muestra a lo largo del tubo de RMN. Durante el experimento la rodaja sobre la que se pulsa se alterna entre *scans* sucesivos.

La mejor relación entre la sensibilidad y el ahorro de tiempo se obtiene cuando el número de rodajas es el mínimo imprescindible para poder reducir a cero el d_1

necesario en cada espectro y cuando la anchura que cubre cada rodaja sea la máxima posible y cubra al máximo el volumen de muestra efectivo de la sonda (Figura 3.4). En experimentos multidimensionales de muestras con moléculas de pequeño o mediano tamaño, el d_1 habitual que se requiere se encuentra en el rango de 1 a 1.5 s, y el tiempo de pulsado y adquisición de la FID está en torno a los 0.3 s, así que para muestras orgánicas con los tiempos de relajación habituales, la situación óptima ocurre cuando se seleccionan entre 3 y 4 rodajas a lo largo del tubo de muestra.

En principio, el método PALSYP permite la adquisición en paralelo de espectros que impliquen tanto los mismos como diferentes tipos de núcleos en cada dimensión indirecta, sin la necesidad de hardware especial, como puede ser múltiples receptores (Kupče2006, Kupče2007) o sondas con varias bobinas receptoras (Wang2005). El método solo requiere un espectrómetro equipado con gradientes de campo apantallados, lo cual hoy en día es habitual en cualquier aparato. Las secuencias de pulsos de RMN que usan gradientes de campo se pueden beneficiar del uso de sondas con gradientes de triple eje, ya que es más sencillo en su caso el prevenir el reenfoque de coherencias no deseadas, ya provengan del propio espectro o de diferentes rodajas de muestra. En todo caso, el método PALSYP podría aplicarse también en sondas que posean solamente gradiente en el eje Z, mediante el uso de los valores apropiados de los gradientes que prevengan este tipo de reenfoques.

A pesar de las pérdidas en sensibilidad inherentes al método de espectroscopía localizada que implementa PALSYP, como se verá, se puede obtener un ahorro considerable en tiempo total de adquisición cuando se quiere adquirir un número de espectros nD para una muestra que no esté muy limitada por la concentración.

3.3 Procedimiento experimental

Los experimentos de RMN se realizaron con una muestra de sacarosa 500 mM en D₂O, en un espectrómetro Varian INOVA-750 (Agilent Inc., Varian Inc.) operando con software de adquisición VNMRJ 2.2D y equipado con una sonda de detección inversa HCP y con gradientes de triple eje. Los espectros RMN fueron procesados con software MestRe-C (Cobas2003) y MestreNova (*Mestrelab. S.L.*). Las secuencias de pulsos desarrolladas y los detalles de los experimentos se dan en la sección de resultados.

3.3.1 *Procesado de espectros 2D PALSYP*:

En la implementación realizada del método PALSYP, un conjunto de espectros 2D PALSYP se almacenan en un sólo fichero “fid”. Para su procesado se requiere una operación de reorganización de las FIDs individuales desde el orden en que fueron adquiridas para repartirlas y conseguir cada uno de los espectros 2D individuales en el dominio del tiempo. Después de esta operación ya se puede utilizar un procesado convencional del espectro 2D RMN aplicando las opciones de apodización, rellenado de ceros, predicción lineal etc. que sean más convenientes para cada espectro. Se ha implementado una función dentro del software de procesado MestRe-C/MestreNova para facilitar la extracción automática de los espectros individuales PALSYP o MS. El proceso es el siguiente:

- 1) Se importa el fichero “fid” del espectrómetro, que contiene todos los espectros PALSYP adquiridos.

2) Desde la ventana del editor de comandos (script editor) se ejecuta el comando “fc”, que extrae las FIDs de cada espectro individual de forma que genera cada uno de los espectros 2D individuales en el dominio del tiempo.

3) Se procesa cada espectro 2D individual en la forma habitual.

3.3.2 Calibración del experimento MS-COSY:

Para este experimento inicialmente se calibró con el experimento de calibración de rodajas de la Figura 3.5a la selección de 4 rodajas de muestra no solapadas y del mismo espesor. La calibración de cada rodaja se realizó en un solo *scan* con una muestra de D₂O al 99.9% mediante la detección de la señal de protón del HDO residual. El pulso espacialmente selectivo de 90° tuvo una duración de 1 ms. La fuerza del gradiente de selección de rodaja G_{ss} fue de 4 G/cm y la fuerza del gradiente de lectura G_{RO} fue de 6 G/cm. La anchura espectral fue de 60 kHz y el Δ fue de 14.5 ms. La FID se detectó con 3500 puntos complejos en un tiempo de adquisición aproximado de 29 ms y en presencia del gradiente G_{RO} . La posición de la rodaja es el único parámetro variado en cada una de las adquisiciones, y ésta es controlada por el offset de frecuencia del pulso de selección espacial respecto a la frecuencia de la portadora en el centro del espectro. Las cuatro posiciones seleccionadas fueron 0.5, 0.2 -0.2 y -0.5 cm respecto al isocentro de la sonda y los correspondientes offset de frecuencia fueron 8514, 4704, -4704 y -8514 Hz respectivamente. Los cuatro experimentos se repitieron con los mismos parámetros pero usando pulsos del tipo *eburp-1* y *sinc* para la selección espacial.

Para la obtención del espectro 2D MS-COSY se utilizó la secuencia de la Figura 3.15 aplicando la calibración de rodajas obtenida. El primer pulso de 90° fue del tipo

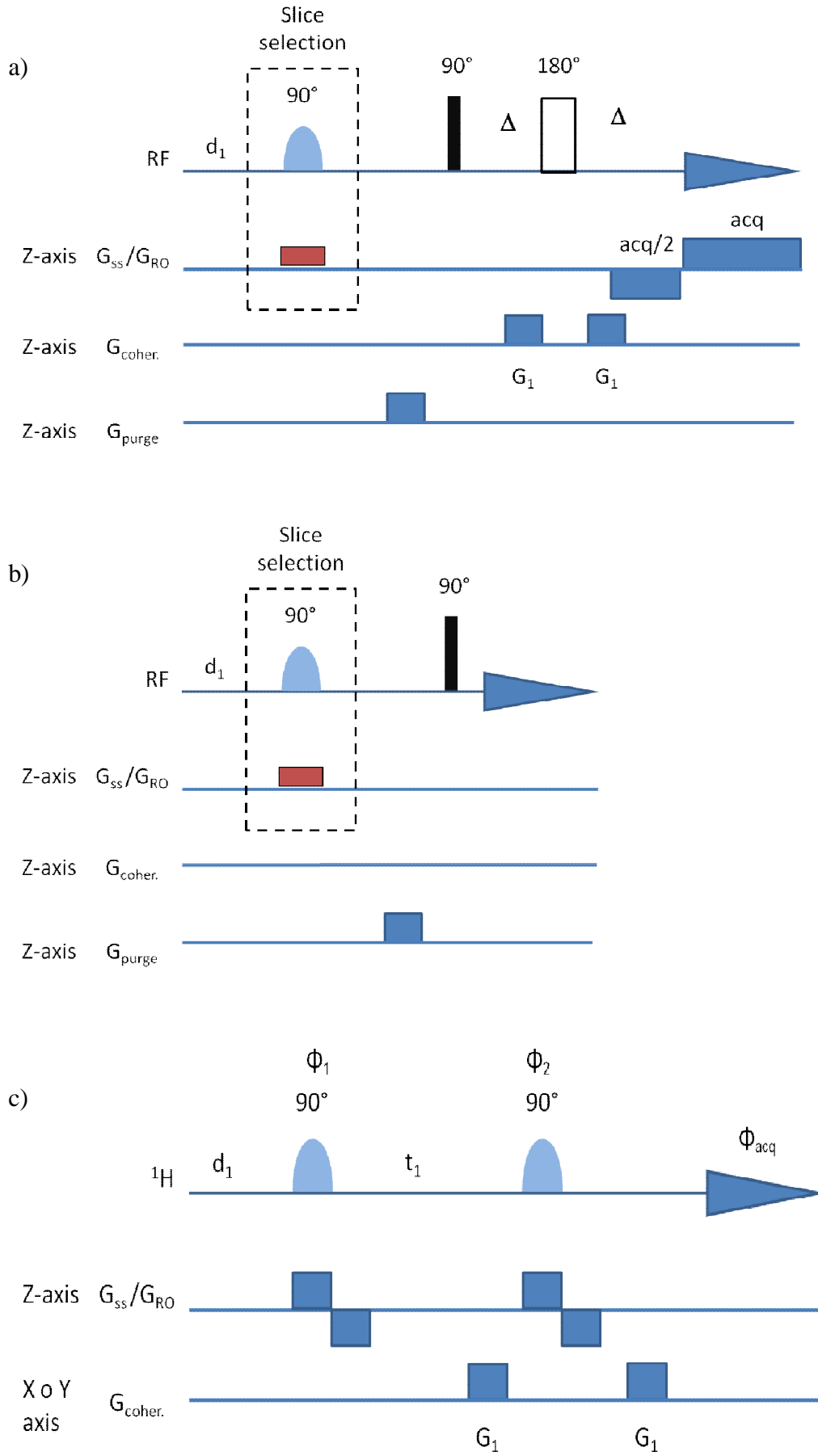
eburp-1 y el segundo del tipo *sinc*. El experimento se adquirió con un solo *scan* por incremento en t_1 . Se aplicó un término de corrección ϕ_{rodaja} en ambos pulsos para compensar la acumulación de errores de fase. Se adquirieron 128 y 4069 puntos reales en F1 y F2 respectivamente. El d_1 fue de 0 s y el tiempo de adquisición de cada FID fue de 0.3 s. La duración total del experimento fue de 49 s. El procesado del espectro obtenido fue el habitual para este experimento.

3.4 Resultados y discusión

3.4.1 Calibraciones previas de los métodos PALSy:

El punto de partida del método PALSy es la calibración de los pulsos y gradientes para seleccionar convenientemente los espines nucleares en un número de rodajas no superpuestas a lo largo del eje principal del tubo de muestra (eje Z). Se utilizó el experimento de mapeo de rodajas 1D (Figura 3.5a) para seleccionar experimentalmente la posición y anchura de 3 rodajas no superpuestas a lo largo del tubo de muestra.

El funcionamiento de la secuencia de mapeo de rodajas (Figura 3.5a) es el siguiente: la magnetización excitada por el primer pulso espacialmente selectivo es inmediatamente destruida (desfasada) por el gradiente de purga que se aplica a continuación a lo largo del eje Y. La siguiente parte de la secuencia consiste en un eco de espín que utiliza pulsos duros para la excitación de todo el volumen de la muestra, lo que es seguido por la detección de la FID en presencia de un eco de gradiente de lectura (G_{RO}) que proporciona una imagen de distribución de densidad de protones, en este caso a lo largo del eje Z. La parte de la magnetización afectada por el pulso espacialmente selectivo que fue destruida por el gradiente de purga no aparecerá en esta imagen, creándose así una imagen que refleja el perfil de excitación espacial que produce el pulso inicial.



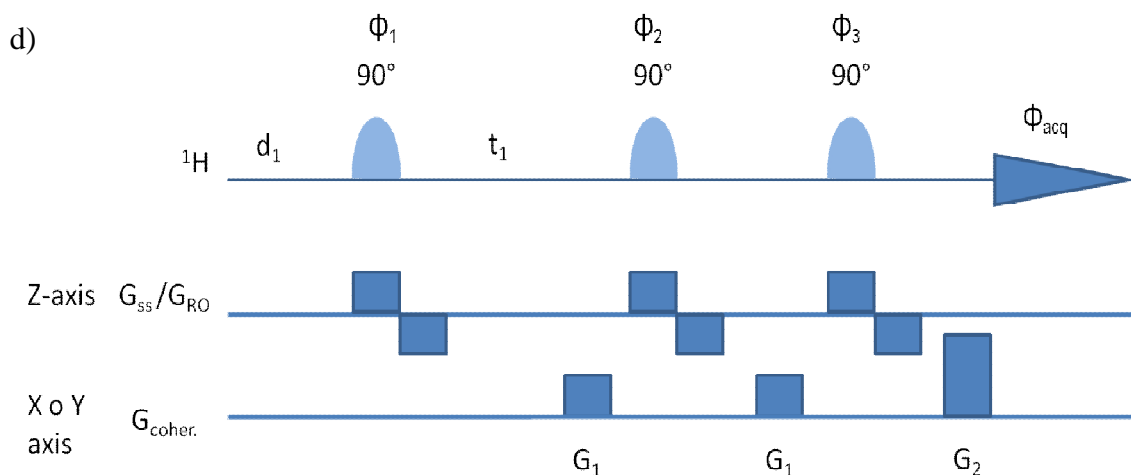


Figura 3.5: a) Secuencia 1D Slice-profile mapping. Los gradientes de selección de rodaja y los de lectura (G_{ss} y G_{RO} respectivamente) son aplicados ambos en la dirección Z. G_{purge} es un gradiente de 1ms y de 10 G/cm. El eco de espera Δ se reguló a 15 ms. b) Secuencia para obtener un espectro cuantitativo 1D selectivo de rodaja con fase de absorción pura. El espectro 1D de referencia con el volumen total de la muestra se obtiene cuando el elemento de selección de rodaja de la muestra es desactivado. El espectro cuantitativo 1D selectivo de rodaja se obtiene de la resta entre el espectro ON y el OFF en idénticas condiciones experimentales. c) 2D selectivo en rodaja COSY en magnitud. El ciclo de fases es ϕ_1 : 0, 2, 0, 2; ϕ_2 : 0, 0, 2, 2, ϕ_{aq} = 0, 2, 0, 2 y d) 2D selectivo en rodaja MQF COSY en magnitud. El ciclo de fases es ϕ_1 : 1, 3; ϕ_2 : 0; ϕ_3 : 1, ϕ_{aq} = 0, 2. En todas las secuencias los pulsos rectangulares corresponden a pulsos duros no selectivos, la forma de los gradientes es rectangular y el pulso de selección de rodaja es de tipo *sinc* con una anchura de 500 μ s. Se aplica simultáneamente con un gradiente de selección de rodaja G_{ss} de 5 G/cm a lo largo del eje Z. La duración de los gradientes G_1 y G_2 es de 1 ms y la fuerza de G_1 es de 10 G/cm. Se aplicó la misma fase a todos los pulsos y al receptor. En las secuencias b) y c) tras el pulso de 90° de selección de rodaja se aplica un gradiente de signo opuesto y mitad de área que G_{ss} durante 1 ms para el reenfoque espacial. La relación de gradientes G_1 : G_2 es 1:2 y 1:3 para DQF-COSY y para TQF-COSY, respectivamente.

Usando la secuencia de la Figura 3.5a se calibraron de forma interactiva tres rodajas no solapadas directamente con la muestra de control (Figura 3.6). Para esta calibración se requiere optimizar los siguientes parámetros: 1) La potencia del gradiente G_{ss} , que controla la anchura de la rodaja, 2) la frecuencia de offset del pulso de selección, que controla la posición espacial de la rodaja a lo largo del eje Z, y 3) la potencia del pulso de selección, se calibra para maximizar la profundidad de la depresión que se crea en el perfil de la muestra.

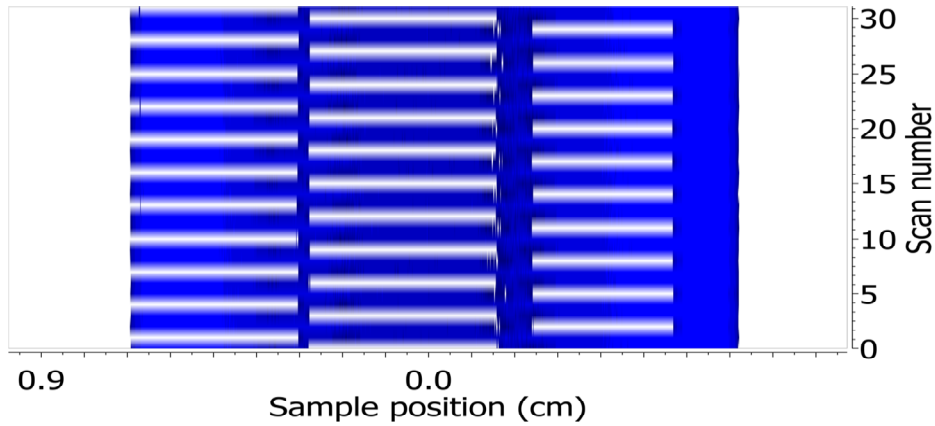


Figura 3.6: Calibración experimental de los tres pulsos de selección de rodaja usados para el método PALSy con la secuencia de pulsos de la Figura 3.5a y una muestra de sacarosa en D_2O . La figura representa en blanco los perfiles de los volúmenes seleccionados por cada rodaja frente al número de *scan*, y el fondo representa el conjunto del volumen activo de muestra de la sonda. El gradiente G_{ss} aplicado a lo largo del eje Z fue de 5 g/cm, el pulso fue del tipo *sinc* y la duración de 500 μs . En cada *scan* la frecuencia de offset del pulso se alternó para excitar los spines en las posiciones 0.07, 0.050, y -0.42 cm respecto del isocentro del gradiente de la sonda.

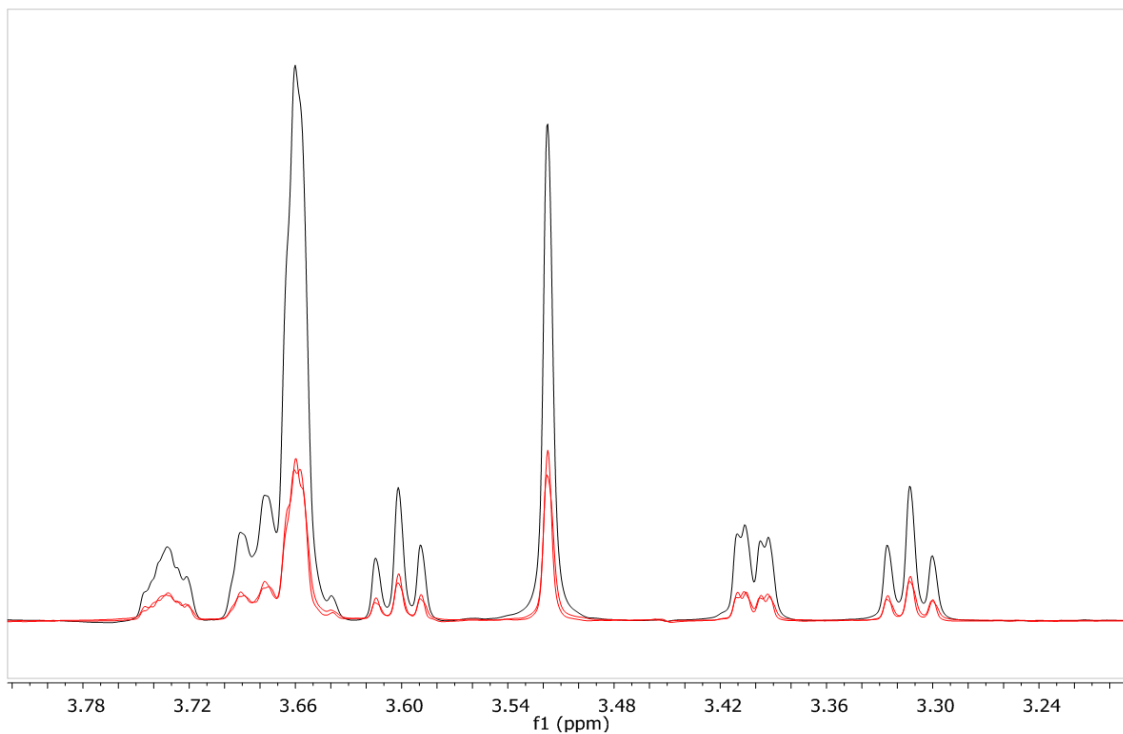


Figura 3.7: Superposición de tres espectros de protón 1D selectivos de rodaja obtenidos con la secuencia de pulsos de la Figura 3.5b (en rojo) y el espectro 1D de protón de todo el volumen activo de la sonda (en negro). Los espectros se adquirieron con la muestra de calibración de sacarosa bajo condiciones cuantitativas y están superpuestos con el mismo escalado vertical. Los parámetros usados para la selección de las tres rodajas fueron los mismos que en la Figura 3.6.

La sensibilidad de la señal de RMN de protón obtenida en la selección de cada una de estas tres rodajas en la muestra se midió con el experimento de la Figura 3.5b que proporciona un espectro 1D cuantitativo selectivo de rodaja y con fase pura (Figura 3.7). La sensibilidad absoluta conseguida en cada rodaja, referida al espectro de referencia con todo el volumen activo de la sonda, fue 27.9%, 29.2% y 30.9%. Estos resultados están satisfactoriamente cerca de los máximos teóricos de sensibilidad por rodaja del método PALSYS, esta se calcula simplemente dividiendo la sensibilidad para todo el volumen activo (100 %) dividida entre el número de rodajas, es decir, 33.3%. En la calibración efectuada arriba, con las tres rodajas se recupera en total el 88% del total de la intensidad de la muestra.

3.4.2 Parte 1 Experimentos rápidos PALSYS 2D COSY/2D DQF-COSY/ 2D TQF-COSY:

El método PALSYS se probó para obtener en paralelo tres experimentos: 2D COSY/ 2D DQF-COSY/2D TQF-COSY (Cohen1989, Shaw1996, Prescott2005) con la muestra de sacarosa. El primer paso fue adaptar las secuencias de pulsos de estos tres experimentos a espectroscopía localizada y se dan en las Figuras 3.5b y 3.5c. (Prescott2005, Loening2003, Thomas2001, Delmas2001) A cada experimento se le asigna a una rodaja de la muestra, de forma que en cualquier *scan* los pulsos espacialmente selectivos de un determinado espectro sean siempre aplicados a la misma rodaja. Esto último resulta imprescindible para conseguir una perfecta coherencia de fase e intensidad de la FID a lo largo de la dimensión t_1 de cada espectro. Los tres espectros localizados se adquirieron con la calibración previa de rodajas de las Figuras 3.6 y 3.7 y dedicando toda la duración de cada *scan* exclusivamente al pulsado y a la adquisición de los puntos de la FID (esto es, haciendo $d_1 = 0$ s). Se utilizó un *scan* por

cada incremento en la dimensión t_1 y un total de 128 incrementos. El tiempo total de adquisición de los tres experimentos fue de tan solo 3 minutos, con una duración aproximada de 0.3 s por *scan*.

La calidad de los espectros es comparable a los mismos espectros adquiridos de manera convencional, con buena sensibilidad y con una total ausencia de artefactos (Figura 3.8). La Figura 3.9 compara un 2D TQF-COSY convencional, usando pulsos duros de 90° y adquiriendo sobre el volumen total de la muestra, con el 2D PALSY TQF-COSY localizado en rodaja de 33% del volumen total de muestra, ambos adquiridos en 1 minuto. Es evidente que en el espectro 2D TQF-COSY convencional (Figura 3.9a) aparecen un conjunto de artefactos que incluyen una doble diagonal de doble cuanto generada por el reenfoque de coherencias no deseadas del *scan* anterior debido al tiempo insuficiente de relajación entre *scans* (ver desarrollo en Figura 3.10). En el espectro 2D PALSY TQF-COSY de la Figura 3.9b adquirido con el mismo tiempo de relajación entre *scans* no encontramos este tipo de artefacto.

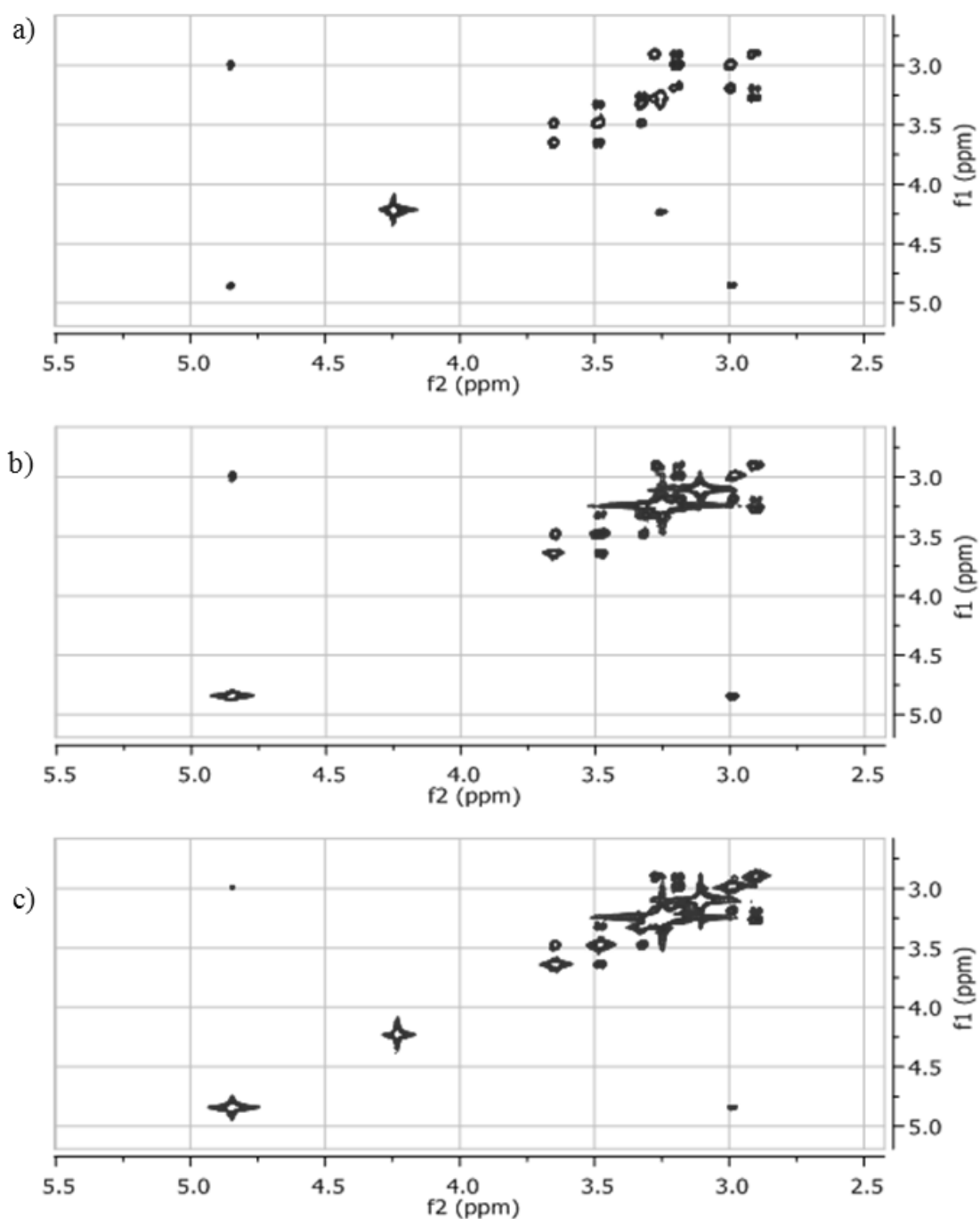


Figura 3.8: Espectros RMN de la sacarosa obtenidos mediante el método PALSY a) 2D PALS TQF-COSY, b) 2D PALS DQF-COSY y c) 2D PALS g-COSY. Los tres espectros se obtuvieron en 3 minutos, con 128 puntos en t_1 , un solo *scan* por incremento en t_1 y una duración aproximada de 0.3 s por *scan*.

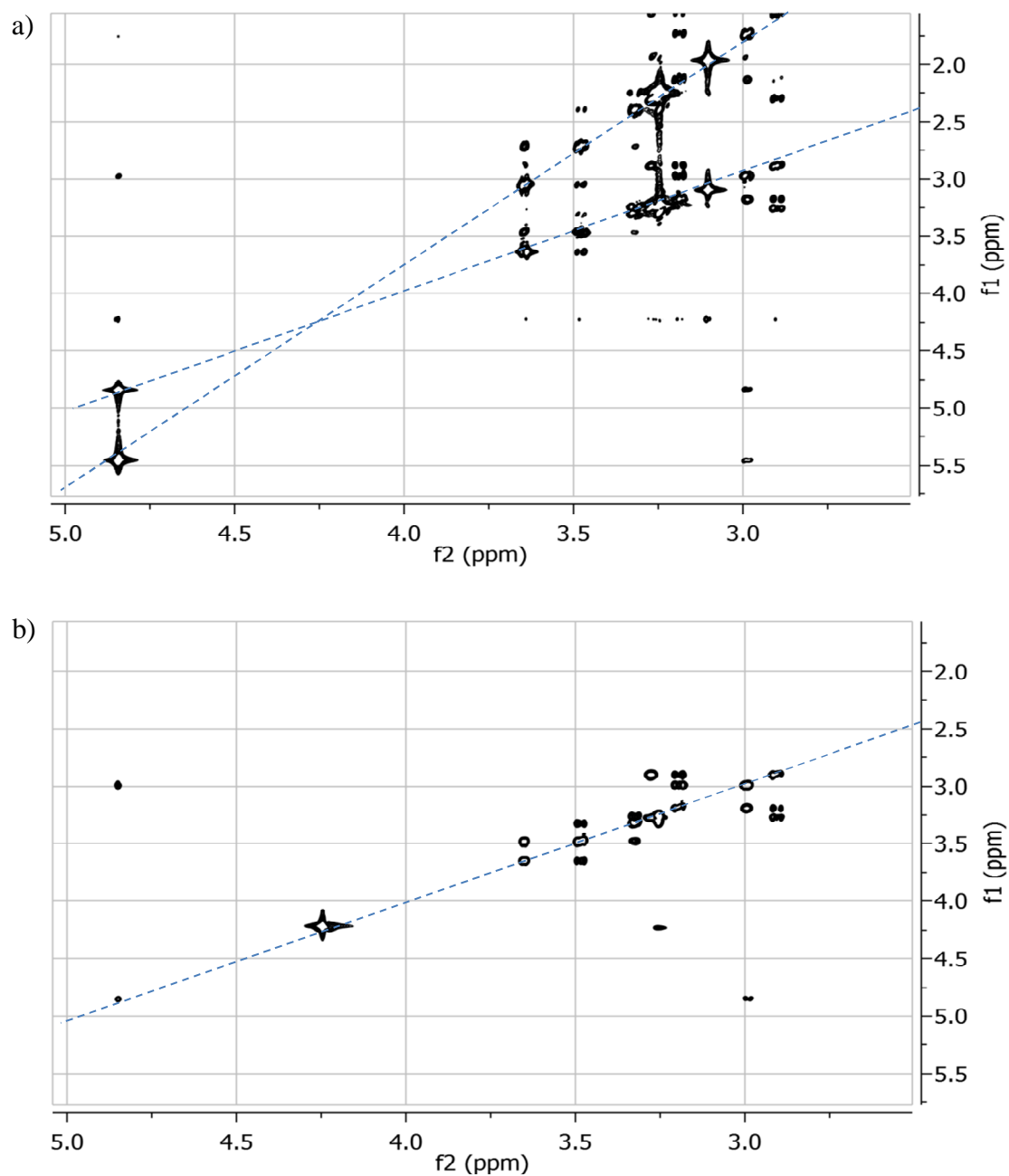


Figura 3.9: a) 2D TQF-COSY convencional usando pulsos duros de 90° sobre el total del volumen de la muestra y con una adquisición estándar y b) 2D PALS Y TQF-COSY (selectivo de rodaja). Cada espectro se adquirió en 1 minuto, con aproximadamente 0.3 s por *scan*, 128 puntos en t_1 y un solo *scan* por incremento en t_1 . Están procesados de idéntica manera, y representados con el mismo nivel de contornos.

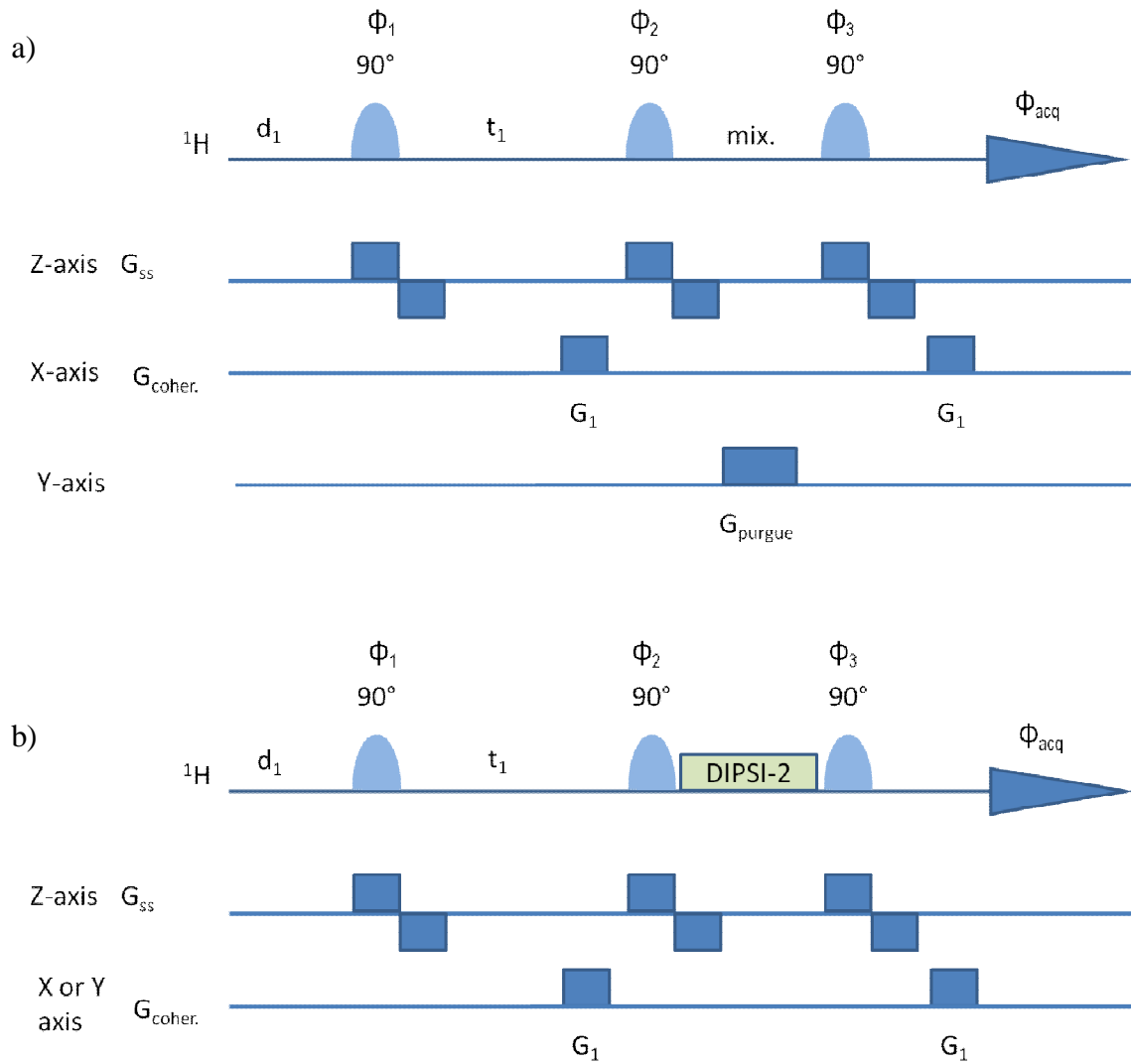


Figura 3.11: Secuencias de pulsos de los experimentos a) 2D PALS-NOESY y b) 2D PALS-DIPSI-TOCSY. Se probó el experimento 2D PALS NOTOCO con la muestra de sacarosa.

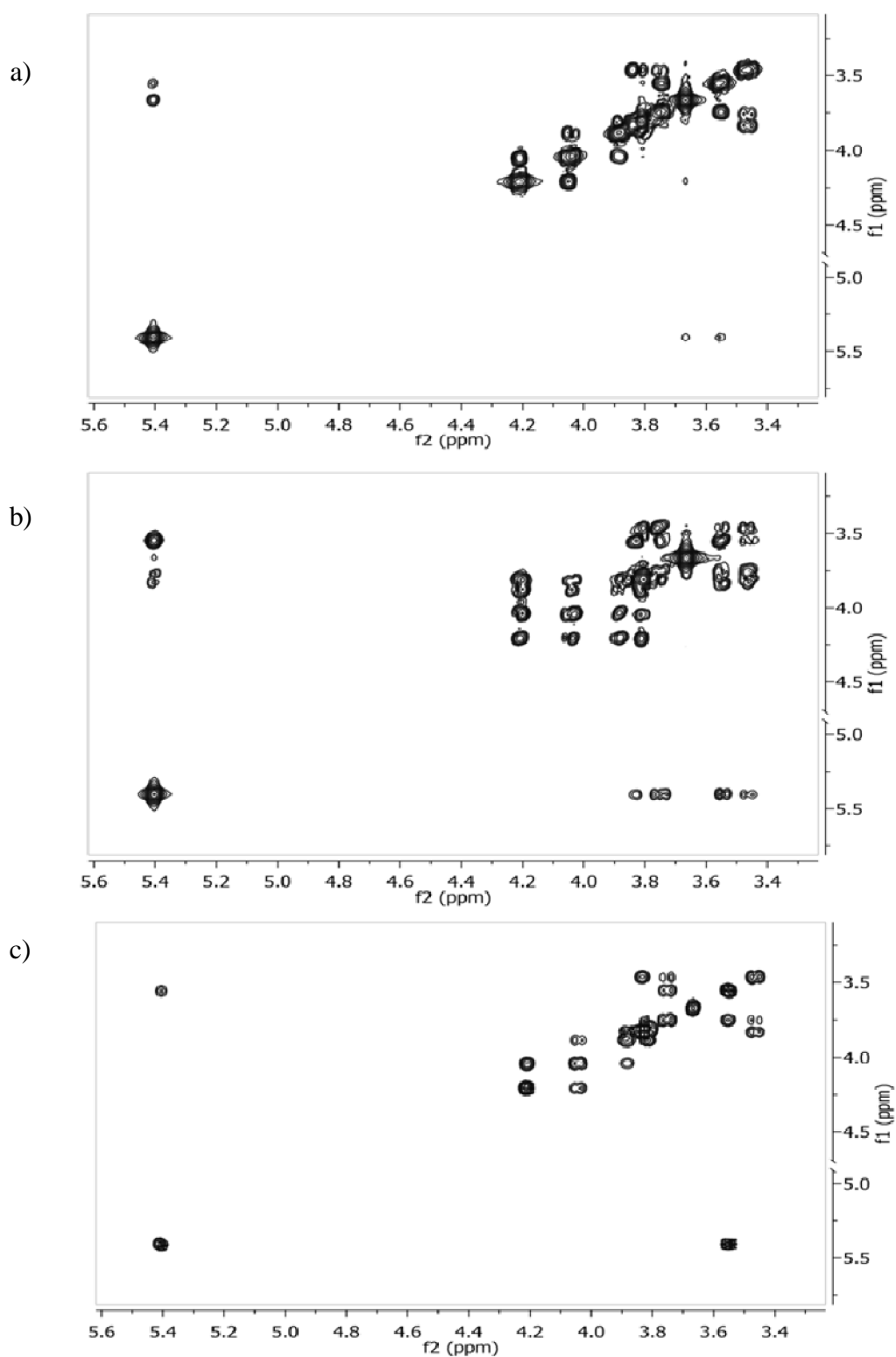


Figura 3.12: Espectros PALSY-NOTOCO de una muestra de sacarosa. a) 2D PALSY-NOESY, b) 2D PALSY-TOCSY, c) 2D PALSY DQF-COSY. Los tres espectros se adquirieron en 3 minutos, con 128 puntos en t_1 , un solo *scan* por incremento en t_1 y una duración aproximada de 0.3 s por *scan*.

3.4.4 Parte 3: Combinación de espectros rápidos PALS Y COTOHMQC:

Se probó la metodología PALS Y para la adquisición rápida en paralelo de una combinación de experimentos homonucleares y heteronucleares. Se probó la combinación espectros 2D DQF-COSY/2D Z-DIPSI-TOCSY/2D-HMQC, una combinación a la que se denominó 2D PALS Y COTOHMQC.

Las secuencias de pulsos convencionales se adaptaron para espectroscopía localizada. La secuencia del experimento 2D PALS Y-DQF-COSY se da en Figura 3.5d, la de 2D PALS Y DIPSI-TOCSY se da en la Figura 3.11b y la del experimento 2D PALS Y HMQC se da en la Figura 3.13.

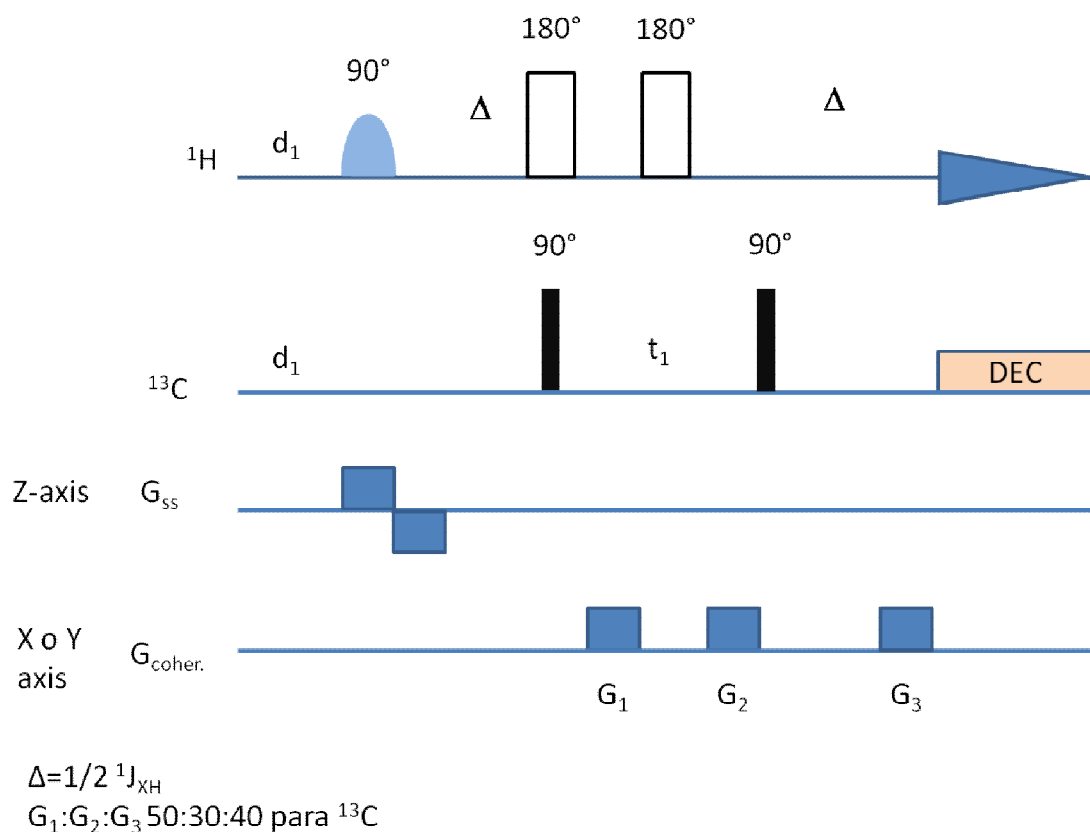


Figura 3.13: Secuencia de pulsos del experimento 2D PALS Y-HMQC ^1H - ^{13}C . Nótese que se aplican dos pulsos duros de 180° en protón en lugar de uno sólo con el fin de reestablecer la magnetización en el eje Z de los espines que están fuera de la rodaja seleccionada por el experimento.

Se probó el experimento 2D PALSYS COTOHMQC con la muestra de sacarosa.

Los espectros obtenidos se muestran en la Figura 3.14.

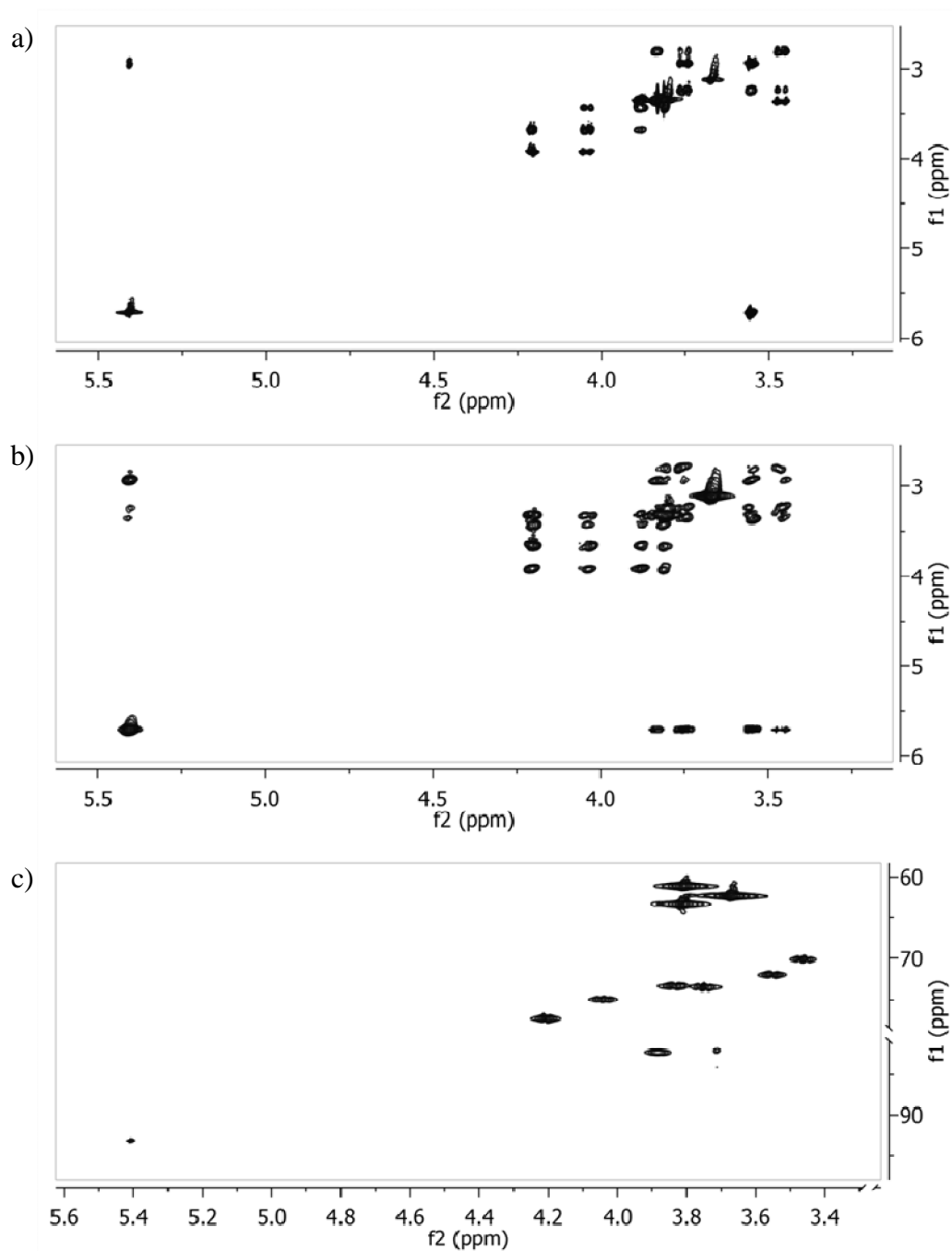


Figura 3.14: Espectros PALSYS-COTOHMQC de una muestra de sacarosa. a) 2D PALSYS-DQF-COSY, b) 2D PALSYS-DIPSI-TOCSY, y c) 2D PALSYS HMQC. Los tres espectros se adquirieron en 3 minutos con 128 puntos en t_1 , un solo *scan* por incremento en t_1 y una duración aproximada de 0.3 s por *scan*.

3.4.5 Parte 4: Experimento rápido 2D COSY con múltiple rodaja (2D MS-COSY):

Se probó la posibilidad de utilizar espectroscopía localizada para adquirir de forma rápida los puntos de una dimensión indirecta (por ej. t_1) de un único espectro 2D en lugar de varios espectros. Se trata pues de un método diferente al método PALSYS que se describió anteriormente al que se llamó método de múltiple rodaja o MS (abreviatura de *Multi-Slice*).

En el método MS al igual que en PALSYS se emplea espectroscopía localizada para dividir virtualmente el tubo de RMN en una serie de rodajas virtuales no solapadas entre sí. Asimismo, mientras en el experimento se pulsa sobre una determinada rodaja, los espines en el resto de rodajas continúan sin perturbar su relajación hacia el equilibrio e igualmente la rodaja seleccionada se va alternando ordenadamente durante *scans* sucesivos del experimento. La diferencia fundamental con PALSYS es que para el experimento MS cada punto de t_1 se adquiere de forma consecutiva con una rodaja de muestra diferente, entre aquellas que se hayan calibrado con anterioridad.

Mediante este método se puede también en principio reducir el tiempo de espera entre *scans* consecutivos, d_1 , de forma que pueda acelerarse la adquisición de los puntos de la dimensión indirecta del experimento. Teóricamente se puede conseguir una reducción máxima del tiempo total del experimento hasta un factor $1/n$, donde n es el número de rodajas seleccionadas.

Experimento rápido 2D MS-COSY

Para probar la metodología MS se calibró primero la selección espacial de cuatro rodajas no solapadas a lo largo del eje Z de la muestra. Se utilizó el mismo experimento de calibración de la Figura 3.5a descrito anteriormente para PALSYS. Los detalles de esta calibración se dan en la parte experimental. La secuencia de pulsos 2D COSY con gradientes en modo magnitud se adaptó para espectroscopía localizada *multi-slice* y se denominó 2D MS-COSY (Figura 3.15).

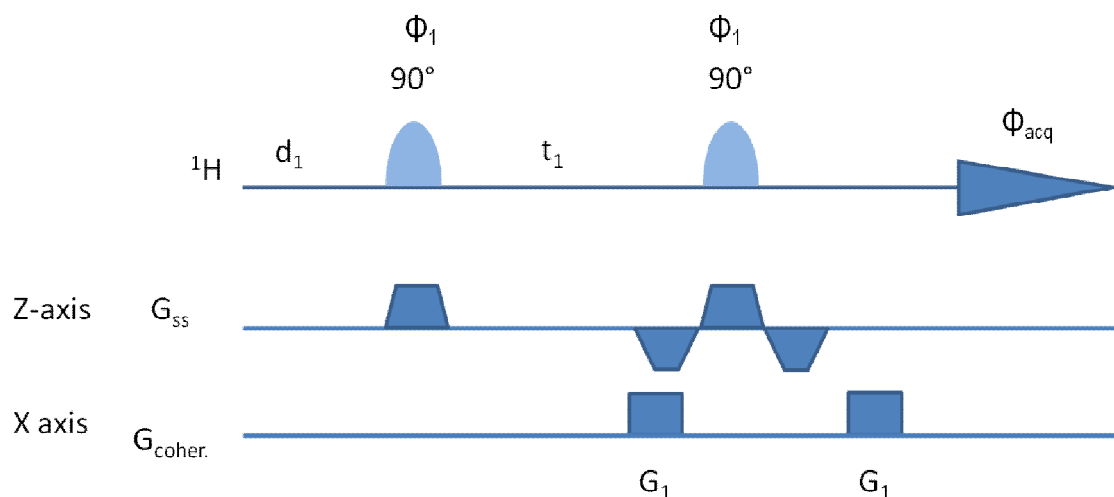


Figura 3.15: Secuencia de pulsos del experimento de múltiple rodaja 2D-MS-COSY. G_1 es el gradiente de selección de coherencias, de una fuerza de 10 G/cm. El primer pulso de selección es de tipo *eburp-1* y el segundo del tipo *sinc*. Se aplican simultáneamente con un gradiente trapezoidal G_{ss} de 4 G/cm a lo largo del eje Z. La fase de ambos pulsos es $\phi_1 = x + \phi_{rodaja}$, donde ϕ_{rodaja} es un término de corrección de fase para compensar errores de fase acumulados por la selección de rodaja y consigue coherencia de fase a lo largo de la dimensión t_1 . La fase del receptor es $\phi_{aq} = x$.

Se adquirió el experimento de la Figura 3.15 con la muestra de sacarosa utilizando la calibración previa de cuatro rodajas de muestra no solapadas.

El espectro 2D MS-COSY obtenido para una muestra de sacarosa se muestra en la Figura 3.16. En este espectro además de las señales de la diagonal y los picos de

correlación esperables, puede verse que también aparecen una serie de artefactos que discurren en paralelo a la diagonal principal a diferentes posiciones en la dimensión F1 (Figura 3.16a). La existencia de estos artefactos se atribuyó a una serie de efectos perniciosos que son i) la inhomogeneidad residual de campo B_0 que queda incluso después del ajuste cuidadoso de los shims, ii) la inhomogeneidad de la radiofrecuencia de la propia sonda a lo largo del volumen de muestra efectivo y iii) imperfección en la linealidad del gradiente de selección de rodaja a lo largo del volumen de muestra efectivo. La suma de estos efectos causa incoherencias de fase y/o de intensidad en las rodajas creadas para puntos de la dimensión t_1 del experimento que fueron adquiridos con diferente rodaja. Como en el método MS las rodajas se van seleccionando de forma alternada, estas incoherencias de fase e intensidad se repiten con una periodicidad n en el dominio de tiempo t_1 , donde n es el número de rodajas.

Como forma de evitar estos artefactos inherentes al propio método MS utilizado, se probó la posibilidad de reducir la incoherencia de fase e intensidad entre rodajas diferentes mediante la introducción de una pequeña corrección de fase extra denominada ϕ_{rodaja} en los pulsos de RF que se aplican a una determinada rodaja (Figura 3.15). Esto consiguió aliviar el problema en buena medida pero aún así no consiguió eliminarlo por completo. Experimentalmente se consiguió reducir la intensidad de los picos artefacto a menos de un 10% de la intensidad que tendrían sin este factor de corrección. Es importante señalar que en PALSY estas incoherencias de fase e intensidad nunca ocurren debido a que la misma rodaja es asignada consistentemente para adquirir el dominio de tiempo t_1 de un mismo experimento 2D.

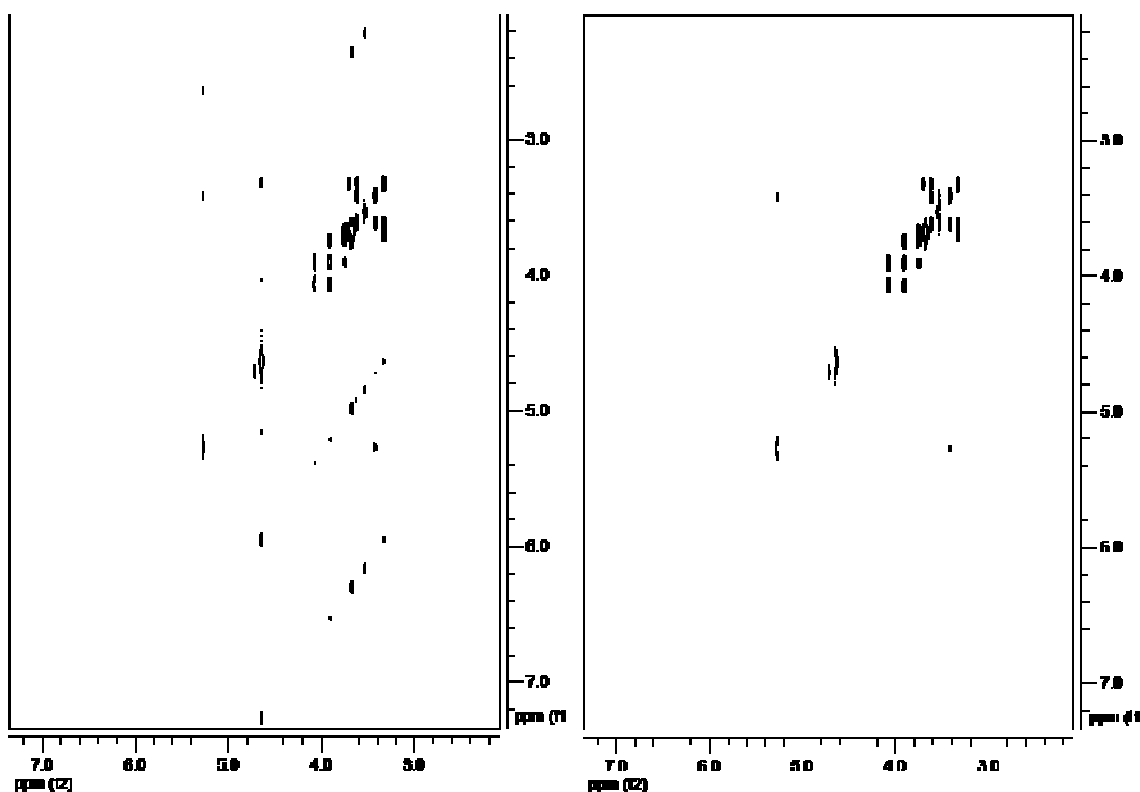


Figura 3.16: 2D-MS COSY con selección de 4 rodajas obtenido con una muestra de sacarosa 20 mM en 49 s. a) Procesado convencional 2D. Aparecen una serie de picos artefacto discurriendo paralelos a la diagonal principal. b) El mismo espectro tras aplicar postprocesado con el método de diferencia de variaciones en F1. El valor asignado a los tres factores de peso del algoritmo de corrección fue de 0.1. El postprocesado elimina correctamente los artefactos.

Debido a que los artefactos comentados del experimento 2D MS-COSY son periódicos en el dominio del tiempo t_1 , estos aparecen en posiciones del espectro predecibles tras aplicar la transformada de Fourier. La cantidad de diagonales artefacto que aparecen es la misma que el número de rodajas utilizadas en el experimento, su posición a lo largo de la dimensión F1 viene dada por la ecuación siguiente:

$$F_1(i) = F_1^{original} + (i/ns) * SW_1 \quad \text{Ecuación 3.5}$$

donde $F_1^{original}$ es la posición en F1 del pico en el COSY convencional, SW_1 es la amplitud espectral en la dimensión F1 e i es el número de pico desde 1 hasta ns . Por ejemplo, en el espectro que se muestra en la Figura 3.8a adquirido con 4 rodajas, para

cada pico del COSY convencional se generan 4 picos con coordenadas en F1 de $+1/4SW_1$, $+1/2SW_1$, $+3/4SW_1$ y $+SW_1$. Solo el último pico corresponde al COSY convencional. Se procedió a su eliminación mediante un algoritmo de postprocesado en el dominio de la frecuencia. El algoritmo de post-procesado de 2D MS-COSY opera a partir del espectro procesado normalmente. Se crean a partir de éste n espectros (donde n es el número de rodajas utilizadas) con un desplazamiento de los datos a lo largo del eje F1 en una cantidad de SW_1/n . El algoritmo calcula un nuevo espectro mediante la sustracción al espectro original de cada uno de los espectros desplazados en F1, multiplicados por un factor de pesado que se ajusta convenientemente al mínimo valor requerido para eliminar cada uno de los artefactos de simetría diagonal. Una vez que se calcula el espectro diferencia, el algoritmo procede a la anulación de la parte negativa del espectro para eliminar totalmente las señales negativas residuales. El resultado de la utilización de este algoritmo puede verse en la Figura 3.16b. Este método de eliminación de los picos artefacto de 2D MS-COSY se ha implementado en el software de procesado MestRe-C de forma que es posible el ajuste interactivo de los factores de pesado. En el espectro de la Figura 3.16b. se utilizó un valor de 0.1 para los tres factores de peso.

3.4.6 Parte 5: Experimento 1D de protón con selección de rodaja para el estudio de equilibrios de fase líquido-líquido:

El uso de espectroscopía localizada puede resultar de utilidad para el estudio de muestras heterogéneas. Las aplicaciones cubren un rango muy amplio que van desde el estudio de equilibrios líquido-líquido con fines cuantitativos hasta incluso la simplificación de los espectros de RMN de metabonómica en muestras complejas tales como tejidos *ex-vivo*.

Las secuencias de pulsos de las Figuras 3.5a y 3.5b pueden utilizarse respectivamente para calibrar la localización espacial en la muestra y para obtener espectros 1D cuantitativos localizados con fase pura. Se utilizaron estas secuencias para calibrar la selección de un total de 64 rodajas a lo largo del tubo de RMN, utilizando para ello una muestra de sacarosa en D₂O (Figura 3.17).

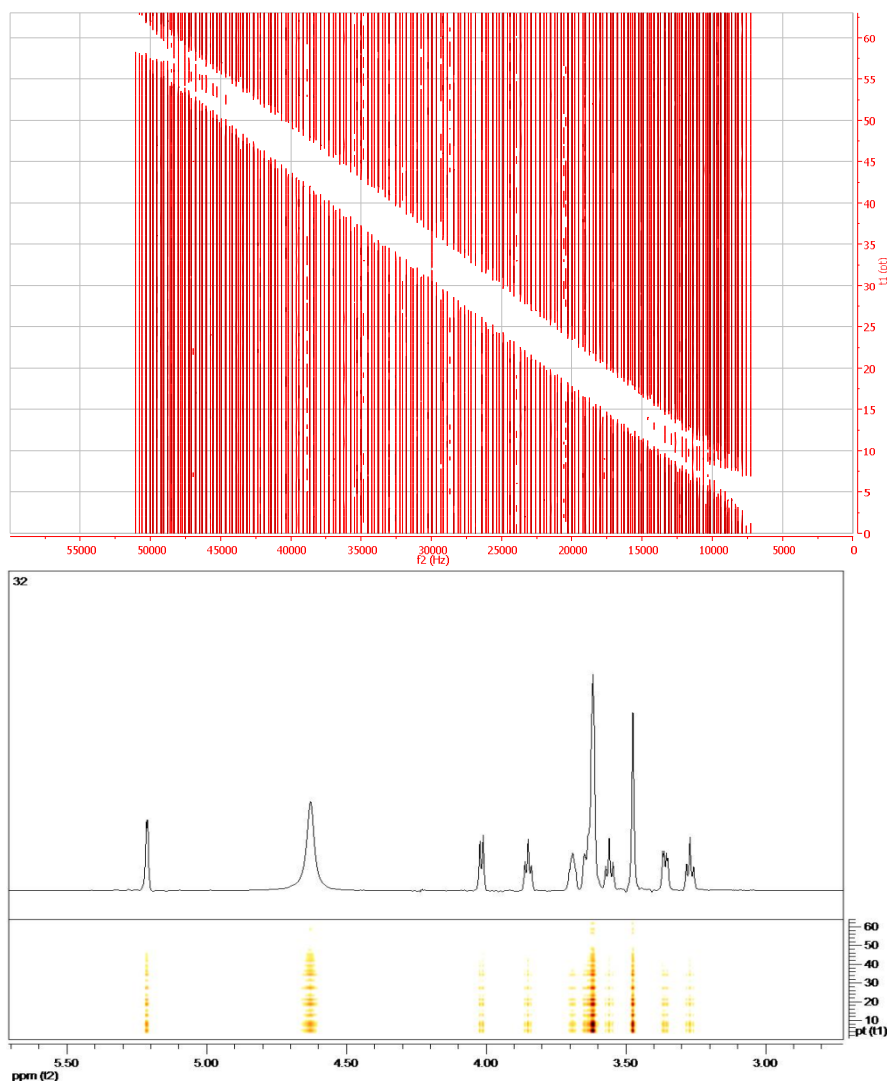


Figura 3.17: Experimentos de RMN localizados obtenidos para una muestra de sacarosa en D₂O. a) Experimento 1D Slice-profile (secuencia Figura 3.5a) utilizado para calibrar 64 rodajas a lo largo del tubo de RMN (eje Z). b) Representación apilada de los correspondientes 64 espectros 1D localizados de protón obtenidos (secuencia Figura 3.5b). El pulso de RF fue de tipo *sinc*, de duración 500 μ s y potencia 58 dB. El gradiente de selección de rodaja G_{ss} fue de 6 G/cm. La frecuencia del pulso espacialmente selectivo se va incrementando en pasos de 1200 Hz para cada una de las 64 rodajas de la muestra.

Una vez calibrados los experimentos se probó su funcionamiento para una muestra heterogénea, consistente en una mezcla de dos fases inmiscibles de H₂O y acetato de metilo entre las que se puede repartir el líquido iónico 2HEAA (2 hidroxietanolamina). Los 64 espectros 1D localizados obtenidos se dan en la Figura 3.18. Puede verse que la interfase entre ambos disolventes se produce en torno a la rodaja número 32. El líquido iónico se distribuye mayoritariamente por la fase acuosa y en menor medida en posiciones próximas a la interfase. Su presencia en la fase de acetato de metilo es indetectable en los espectros obtenidos.

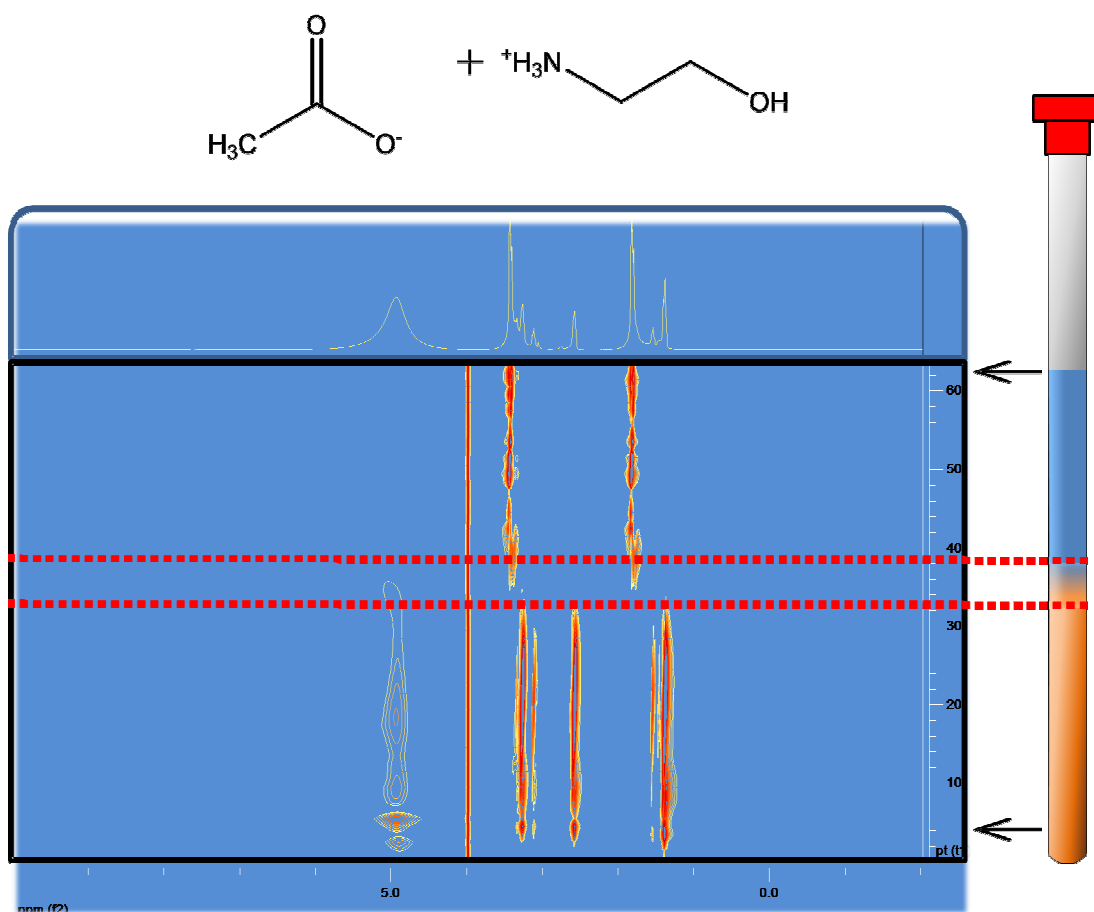


Figura 3.18: Espectros localizados 1D 1H de una muestra consistente en una mezcla de dos fases inmiscibles de H₂O y acetato de metilo entre los que se puede repartir el líquido iónico 2HEAA (2 hidroxietanolamina). Se han representado 64 espectros apilados de protón obtenidos a distinta altura en el tubo de muestra que cubren un espesor de 0.3 mm cada uno. La traza de arriba muestra el espectro 1D de protón convencional de todo el volumen de muestra.

3.5 Conclusiones

El método de adquisición en paralelo PALSYS permite un ahorro de tiempo considerable en la adquisición de varios espectros nD, sin que su uso implique aparición de ruido o artefactos en ellos. El método es válido para cualquier tipo de muestra, pero es más recomendable para muestras no muy limitadas en concentración y para experimentos con suficiente sensibilidad por *scan*, ya que en cada experimento PALSYS el volumen de muestra que se detecta es inferior (normalmente en un factor $1/n$, donde n es la cantidad de experimentos diferentes introducidos en PALSYS) respecto al volumen del espectro convencional. Esto se traduce en una reducción efectiva de la sensibilidad proporcional a n . En los experimentos realizados, optimizando el equiparticionado del volumen de rodaja, la penalización a la sensibilidad fue de un factor aproximado de 3 para cada uno de los experimentos en una serie de tres.

Por otro lado, algunas ventajas excepcionales de este método de adquisición rápida son: 1) no se limita la cantidad de puntos adquiridos en cada dimensión, 2) no conlleva por su mecánica la aparición de ningún artefacto y 3) el procesado se lleva a cabo por métodos convencionales. El método además es compatible con esquemas de supresión de señal/es de disolvente/s que dependan de pulsos selectivos convencionales (es decir, que no sean selectivos espacialmente) como puede ser WET (Smallcombe1995).

El método de adquisición en múltiples rodajas 2D MS-COSY proporciona una reducción considerable de la duración total del experimento. Sin embargo hay una serie de factores experimentales que introducen artefactos y perjudican seriamente su aplicación. La inhomogeneidad residual de campo B_0 incluso después del ajuste cuidadoso de los shims, unida a la inhomogeneidad de la propia sonda a lo largo del

volumen de muestra efectivo tanto en lo que se refiere a la generación de los pulsos de RF, como a los defectos de linealidad de los gradientes de campo, finalmente introducen incoherencias de fase y/o de intensidad en las rodajas creadas para la adquisición de la dimensión indirecta t_1 . Estos efectos no son siempre fáciles de compensar hasta el nivel necesario y causan artefactos en el espectro MS-COSY. Con todo, dado que estos artefactos aparecen en posiciones predecibles fue posible su eliminación mediante un algoritmo de postprocesado.

Por último, cabe destacar que los métodos de espectroscopía localizada como los que se describen en este capítulo pueden tener aplicaciones para estudiar propiedades físico-químicas en muestras con cierto grado de heterogeneidad, como es el caso de los equilibrios de fase. En el futuro se puede plantear introducir el esquema básico de localización de la Figura 3.5b para realizar otros experimentos de RMN localizados, por ejemplo para estudiar los tiempos de relajación u otros parámetros medibles por RMN como coeficientes de difusión en este tipo de muestras.

3.6 Referencias

- Callaghan, P.T., “Principles of Nuclear Magnetic Resonance Microscopy”, Oxford University Press, Ed.; Clarendon Press: New York, **1991**.
- Cobas, J.C.; Sardina, F.J., “Nuclear magnetic resonance data processing. MestReC: A software package for desktop computers”, *Concepts Magn. Reson.*, **2003**, *19A*, 80-96.
- Cohen, Y.; Chan, L.H.; Litt, L.; James, T.L., “Spatially localized COSY spectra from a surface coil using phase-encoding magnetic field gradients”, *J. Magn. Reson.*, **1989**, *85*, 203–208.
- Delmas, F.; Beloeil, J-C.; Van der Sanden, B.P.J.; Nicolay, K.; Gillet, B., “Two-Voxel Localization Sequence for in Vivo Two-Dimensional Homonuclear Correlation Spectroscopy”, *J. Magn. Reson.*, **2001**, *149*, 119–125.
- Kupče, E.; Freeman, R.; John, B.K., “Parallel acquisition of two-dimensional NMR spectra of several nuclear species”, *J. Am. Chem. Soc.*, **2006**, *128*, 9606-9607.
- Kupče, E.; Cheatham S.; Freeman, R., “Two-dimensional spectroscopy with parallel acquisition of $^1\text{H-X}$ and $^{19}\text{F-X}$ correlations”, *Magn. Reson. Chem.*, **2007**, *45*, 103-105.
- Loening, N.M.; Thrippleton, M.J.; Keeler, J.; Griffin; R.G., “Single-scan longitudinal relaxation measurements in high-resolution NMR spectroscopy”, *J. Magn. Reson.*, **2003**, *164*, 321–328.
- Prescott, A.P.; Dzik-Jurasz, A.S.K.; Leach, M.O.; Sirohi, B.; Powles, R.; Collins, D.J., “Localized COSY and DQF-COSY $^1\text{H-MRS}$ Sequences for Investigating Human Tibial Bone Marrow In Vivo and Initial Application to Patients With Acute Leukemia”, *J. Magn. Reson. Imaging*, **2005**, *22*, 541-548.
- Shaw, A.A.; Salaun, C.; Dauphin, J.F.; Ancian, B., “Artifact-Free PFG-Enhanced Double-Quantum-Filtered COSY Experiments”, *J. Magn. Reson.*, **1996**, *120*, 110–115.
- Smallcombe, S.H.; Patt, S.L.; Keifer, P.A., “WET Solvent Suppression and Its Applications to LC NMR and High-Resolution NMR Spectroscopy”, *J. Magn. Reson., Ser. A*, **1995**, *117*, 295-303.
- Thomas, M.A.; Yue, K.; Binesh, N.; Davanzo, P.; Kumar, A.; Siegel, B.; Frye, M.; Curran, J.; Lufkin, R.; Martin, P.; Guze, B., “Localized two-dimensional shift correlated MR spectroscopy of human brain”, *Magn. Reson in Med.*, **2001**, *46*, 58–67.
- Wang, H.; Ciobanu, L.; Webb, A., “Reduced data acquisition time in multi-dimensional NMR spectroscopy using multiple-coil probes”, *J. Magn. Reson.*, **2005**, *173*, 134–139.

Capítulo 4:
Experimentos de RMN para el estudio de
interacciones moleculares: Experimentos 1D STD
con filtro de difusión reverso y 1D STD
homodesacoplado

4.1 Introducción

La RMN es una herramienta excelente para lograr la identificación de compuestos de pequeño tamaño que interaccionan con macromoléculas diana. Dependiendo de la cinética del proceso de asociación se pueden plantear experimentos de valoración para determinar parámetros termodinámicos como la constante del equilibrio de asociación (K_a) (Fielding2000) y/o experimentos como STD (*Saturation Transfer Difference*) (Mayer1999) para conocer las interacciones intermoleculares y determinar el llamado epítipo de unión del ligando, que son las zonas en la molécula del ligando que interaccionan directamente con el receptor macromolecular.

Para este tipo de estudios y en otros en los que utilizan mezclas de compuestos, es habitual que se produzca solapamiento de señales de moléculas diferentes, lo que puede plantear problemas en el análisis de los espectros. Para poder observar en los espectros exclusivamente las señales de RMN de las moléculas pequeñas se suele recurrir a experimentos que incluyen en su secuencia de pulsos algún bloque destinado a lograr el filtrado de señales basándose en la diferencia de magnitud entre ambas moléculas de alguna/s de las propiedades medibles por RMN (Figura 4.1). Las dos estrategias más habituales son utilizar o bien alguno de los denominados filtros de relajación, normalmente de relajación transversal (Rastrelli2009) (aunque también se han descrito filtros de relajación longitudinal) (Ma2001, Borlototti2010) o bien los filtros de difusión (Esturau2006, Khera2010, Hajduk1997).

Los filtros de relajación y de difusión simplifican el espectro introduciendo una atenuación diferente (y modulable mediante algún parámetro) para las señales de una mezcla de forma que se atenúen selectivamente las señales de alguno de los componentes, llegando incluso a cancelarse.

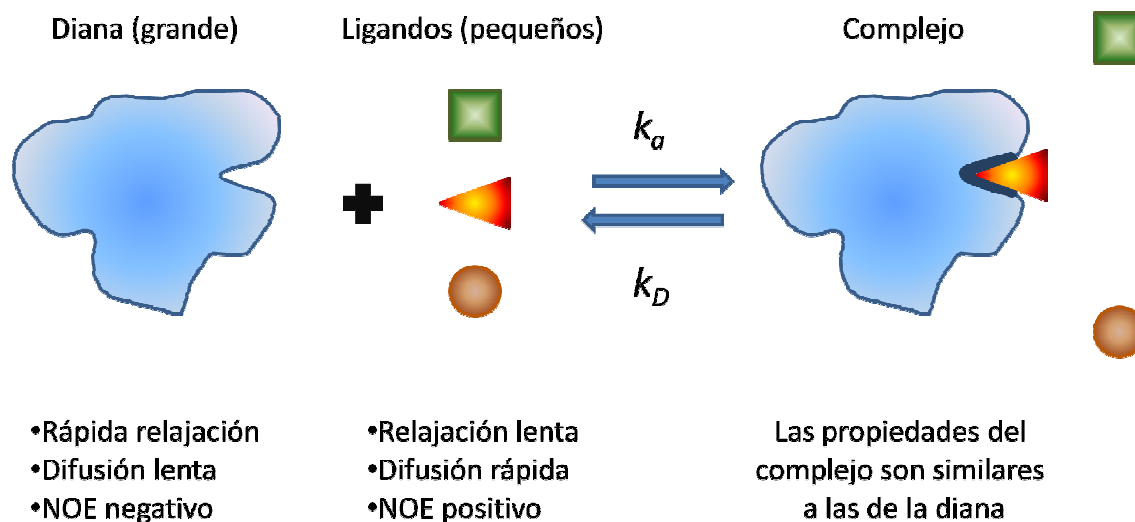


Figura 4.1: Moléculas con tamaño molecular muy diferente generalmente tienen propiedades de RMN muy diferentes en cuanto a su coeficiente de difusión traslacional (D), velocidad de relajación transversal (R_2) y NOE.

Estos métodos de filtrado resultan muy convenientes, ya que no suponen un aumento de la dimensionalidad del espectro ni con ello un aumento significativo del tiempo de experimento y tampoco suelen producir una merma importante en la sensibilidad. Además de esto, pueden incluirse sin gran dificultad dentro de las secuencias de pulsos de muchos de los experimentos de RMN convencionales para obtener su correspondiente versión filtrada.

El filtro de difusión se basa en el diferente valor del coeficiente de difusión traslacional de las moléculas de la mezcla. El parámetro que permite controlar el grado de atenuación es el denominado tiempo de difusión (Δ). La atenuación de la intensidad de señal es mayor para las señales de compuestos con coeficiente de difusión alto (moléculas pequeñas) que para las señales de los compuestos con coeficiente de difusión bajo (moléculas grandes). El grado de atenuación es muy similar para todas las señales de una misma molécula (salvo para señales en intercambio químico o en el caso poco habitual de que alguna de éstas presente un tiempo de relajación longitudinal

extremadamente corto) y por tanto éstas mantienen en gran medida las integrales relativas originales.

El experimento de filtro de relajación transversal se basa en la diferencia de velocidad de relajación transversal entre las distintas señales de la mezcla. Experimentalmente estos filtros consisten en introducir en el experimento de RMN un período donde la magnetización macroscópica se mantiene durante un tiempo relevante en el plano transversal antes de la detección de la FID, si bien evitando los efectos adversos de evolución con el desplazamiento químico y acoplamiento/s que esto lleva asociado. Hay dos maneras de realizar esta operación: Una de ellas es el denominado *filtro- T_2* basado en secuencias como la Carr-Purcell-Meiboom-Gill (CPMG) (Meiboom1958) y otras análogas, como la XY-16 (Dierks2009), en las que se aplica un ciclo de ecos de espín con pulsos duros de 180° que producen reenfoque en el plano transversal de los desplazamientos químicos, con tiempos de eco muy cortos para evitar la evolución con los acoplamientos escalares. La otra es el denominado *filtro- $T_{1\rho}$* que consiste en un único pulso largo y continuo de baja potencia (*field-strength* entre 2000-3000 Hz) que confina los espines en torno a un eje del plano rotatorio, de forma que opere la velocidad de relajación longitudinal del marco rotatorio de cada señal ($R_{1\rho}$, el inverso de $T_{1\rho}$) como forma de filtrado. La duración de este pulso puede ajustarse para controlar el grado de atenuación de intensidad entre las diferentes señales de la mezcla. En general, las señales de las moléculas de elevado tamaño molecular tienen velocidades de relajación transversal, R_2 y $R_{1\rho}$, mucho mayores que las de pequeño tamaño. Esto permite que incluso con una duración relativamente corta del pulso del filtro de relajación transversal se consiga que las primeras se vean atenuadas en mucha mayor medida que las segundas. La atenuación de señales que produce un filtro T_2 y un filtro $T_{1\rho}$ es teóricamente algo diferente debido a su distinta dependencia con la

dinámica molecular (Kowalewski2006), sin embargo en la práctica, para moléculas de tamaño molecular muy diferente y duraciones similares del filtro, el espectro filtrado suele resultar bastante similar. Es interesante destacar que el filtro de relajación sólo distingue, *grosso modo*, por tamaño de molécula, ya que en realidad es la relajación transversal de la señal el parámetro relevante. Las señales de una misma molécula en general tienen valores de R_2 (o $R_{1\rho}$) diferentes, y su diferente grado de atenuación por este tipo de filtrado lleva a que el espectro filtrado no necesariamente mantenga la proporción original de las áreas relativas de sus señales.

Los filtros de difusión y de relajación transversal son pues experimentos complementarios entre sí en lo que se refiere al filtrado de señales. En esta sección se expone una nueva herramienta de filtrado que se ha desarrollado, denominada filtro de difusión reverso (filtro *Drev*), basada en espectroscopía diferencia. Se describen algunas de las aplicaciones de este filtro tanto por sí sólo como en su aplicación combinada con experimentos STD (Meyer2003).

El solapamiento de señales en un espectro 1D-STD puede introducir problemas de ambigüedad en su asignación. Es por ello que desde su aparición se han desarrollado varias modificaciones sobre el experimento base del STD, como son el 2D-STD-TOCSY (Mayer1999, Vogtherr2000, Möller2002) o la versión selectiva 1D-STD-selective TOCSY (Johnson2002). En este capítulo se expone también otra secuencia de experimento diseñada para la simplificación de espectros STD, el experimento 1D-STD-Homodesacoplado, que proporciona un espectro en el que las respuestas STD aparecen como singletes, lo que alivia el solapamiento de señales y facilita su estudio. Este experimento está basado en la espectroscopía 2D J-resuelta (Aue1976, Morris2010).

4.2 Teoría

4.2.1 Experimento STD

El experimento STD (Mayer1999) (Figura 4.2) se basa en saturar selectivamente y de forma continuada una pequeña región del espectro de protón donde aparecen de forma exclusiva la/s señal/es de una macromolécula. Durante el tiempo que dura la saturación, los protones seleccionados transfieren parte de su saturación hacia otros protones espacialmente próximos de la propia macromolécula (transferencia de saturación mediante el mecanismo de relajación cruzada dipolar). Este proceso es sumamente eficaz en macromoléculas, y por ello, si la saturación dura el tiempo suficiente, puede transmitir rápidamente por toda su red de protones y llegar a afectar a muchos protones del receptor que estaban alejados del lugar inicial de saturación. En estas condiciones, si un pequeño ligando se une débilmente con la macromolécula, su movilidad y propiedades de RMN se vuelven transitoriamente similares a ésta, y puede ocurrir que ciertos protones del ligando queden próximos a protones saturados de la macromolécula y reciban también eficazmente saturación. Posteriormente, si el ligando se disocia con la suficiente rapidez, éste adquiere instantáneamente sus propiedades originales de RMN, sin embargo, la fracción de saturación que recibieron algunos de sus protones puede permanecer el tiempo suficiente como para ser detectada antes de que se desvanezca por relajación.

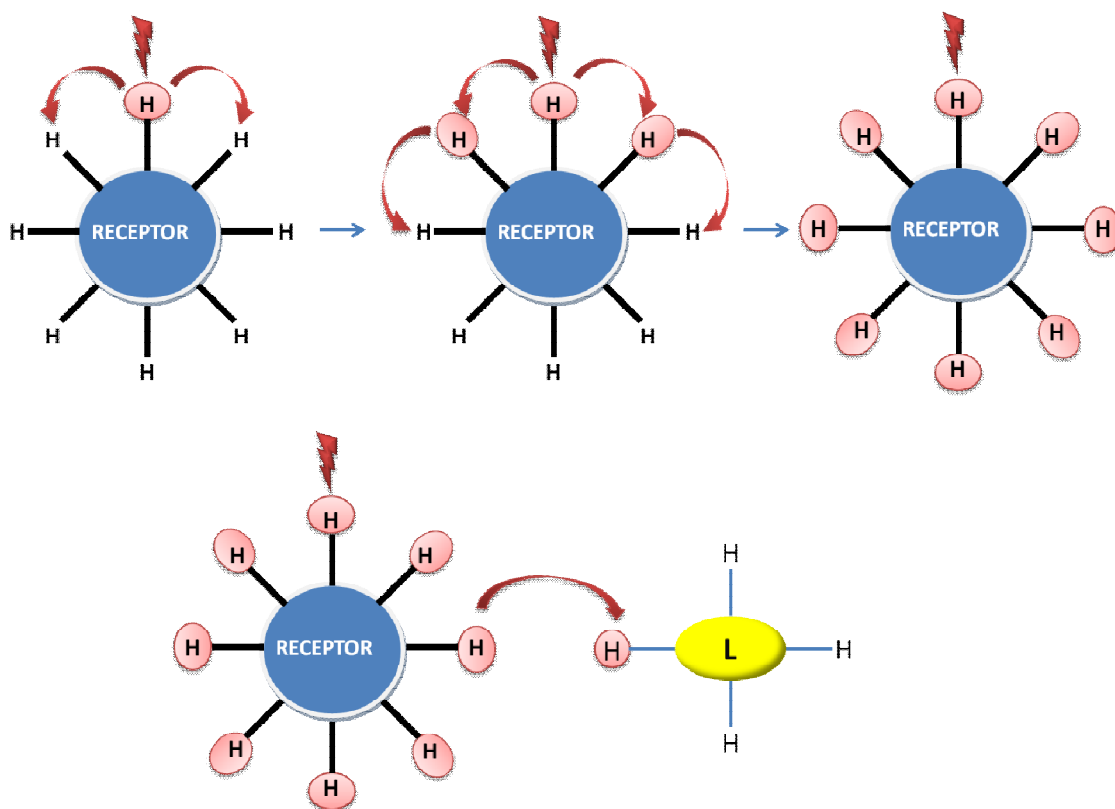


Figura 4.2: Evolución de la magnetización intra e intermolecular en el tiempo, tras la saturación selectiva de una señal o señales del receptor. La magnetización evoluciona rápidamente entre los protones del receptor debido al mecanismo de difusión de spin, y puede pasar a protones de otras moléculas que se sitúen lo suficientemente próximas en el espacio por el mecanismo de relajación cruzada.

El experimento STD trata precisamente de detectar esta pequeña saturación de los protones del ligando en el estado libre en disolución que delata que hubo unión con la macromolécula diana. Se adquieren dos espectros diferentes, uno situando la frecuencia del pulso de saturación selectivo en una zona donde aparezcan (o sean esperables) en exclusiva las señales de la macromolécula por ej. en la zona entre 0 y 1 ppm (llamado espectro *on-resonance*) y otro situando el mismo pulso de saturación a una frecuencia a la que no aparezca (ni sean esperables) protones de la muestra (equivalente a un espectro sin saturación selectiva, y llamado espectro *off-resonance*). La resta entre los espectros on- y off-resonance dará lugar a un espectro diferencia, que muestra las señales del ligando que recibieron saturación por proximidad con protones

de la macromolécula (Figura 4.3). Inicialmente (Mayer1999) esta técnica se presentó como una herramienta de cribado de ligandos capaces de interactuar con proteínas, y con el tiempo se ha convertido en una herramienta fundamental en la búsqueda y desarrollo de fármacos.

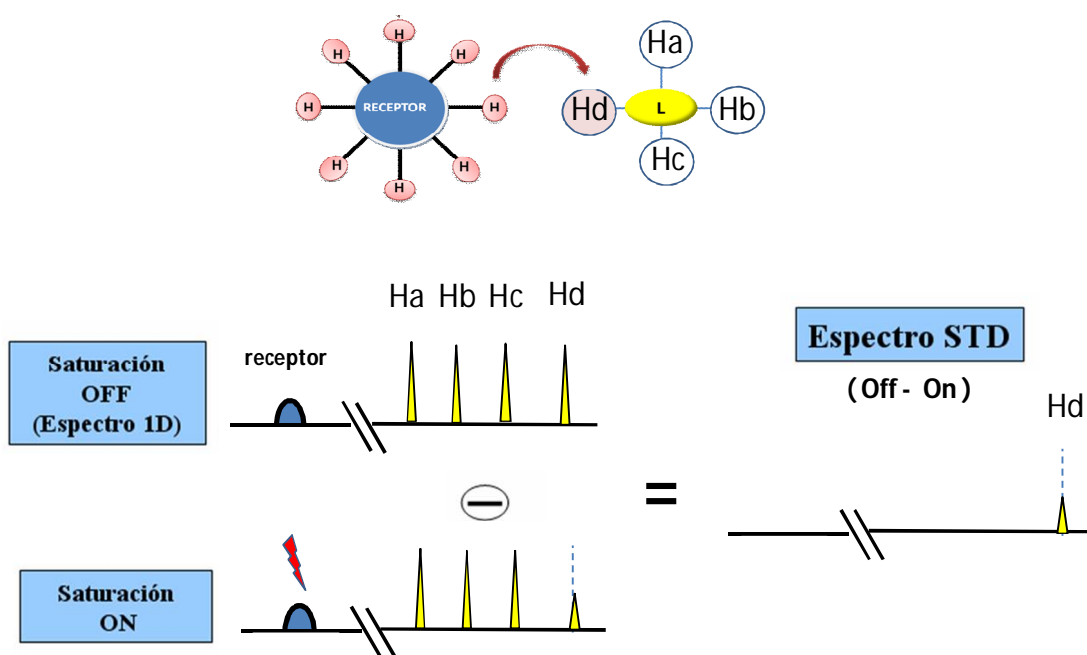


Figura 4.3: Ilustración del experimento STD en la interacción transitoria con un ligando con una macromolécula. En el experimento STD-off las señales del ligando aparecen con su intensidad original. En el experimento STD-on se saturan los protones del receptor macromolecular, lo que tiene el efecto de perturbar la intensidad de algunas de las señales del ligando. A la derecha se muestra el espectro diferencia STD-off – STD-on. Las señales en este espectro prueban que existe interacción intermolecular y cuáles son los protones implicado en el epítipo de unión del ligando.

4.2.2 Difusión

La difusión se refiere a la velocidad del movimiento aleatorio de las partículas en un líquido o un gas por el efecto térmico en su energía cinética. La velocidad a que difunde la partícula depende no solo de la temperatura sino también de la viscosidad del fluido y del tamaño (masa) de la partícula. El coeficiente de difusión (D) mide la velocidad del movimiento traslacional aleatorio de las partículas por el efecto térmico.

En un medio homogéneo su valor es función de la distancia radial promedio que alcanza la partícula que difunde en recorrido aleatorio en un tiempo dado (Ecuación 4.1) (Chen1984, Espinosa1987).

$$\langle r^2 \rangle = n \cdot D \cdot t \quad \forall n = 2, 4, 6 \quad \text{Ecuación 4.1}$$

La medida de coeficientes de difusión D mediante RMN u otra técnica reviste interés porque puede relacionarse con parámetros estructurales de la partícula como su tamaño (masa) a través de la teoría de Debye-Einstein (Ecuación 4.2) (Johnson1999):

$$D = \frac{k_B T}{f_T} \quad \text{Ecuación 4.2}$$

donde k_B es la constante de Boltzmann, T es la temperatura absoluta y f_T es el factor de fricción que depende de la forma y tamaño de la partícula y de la viscosidad de la disolución que es un parámetro que engloba aspectos como la concentración tipo de disolvente, etc. Se han deducido expresiones de f_T para distintas formas moleculares (ej. esférica, cilíndrica o elipsoidal), si bien, la ecuación para la forma esférica (ecuación de Stokes-Einstein) es una de las más utilizadas debido a su simplicidad (Ecuación 4.3):

$$D = \frac{k_B T}{6\pi\eta r_H} \quad \text{Ecuación 4.3}$$

donde η es la viscosidad y r_H el radio hidrodinámico.

4.2.3 Espectros de RMN filtrados por difusión

La secuencia de pulsos de eco estimulado BPP-STE (Figura 4.4) se utiliza en RMN para la medida de coeficientes y/o el filtrado de señales por difusión. A grandes

rasgos, tras la excitación inicial de los espines de la muestra, se utilizan gradientes de campo magnético (PFG *Pulse Field Gradients*) para la codificación/ descodificación de la posición espacial instantánea que ocupa cada uno los espines en la muestra, a lo largo de la dirección en que se aplican gradientes. Entre los períodos de codificación y descodificación espacial se deja un período de tiempo ajustable Δ para que ocurra la difusión molecular.

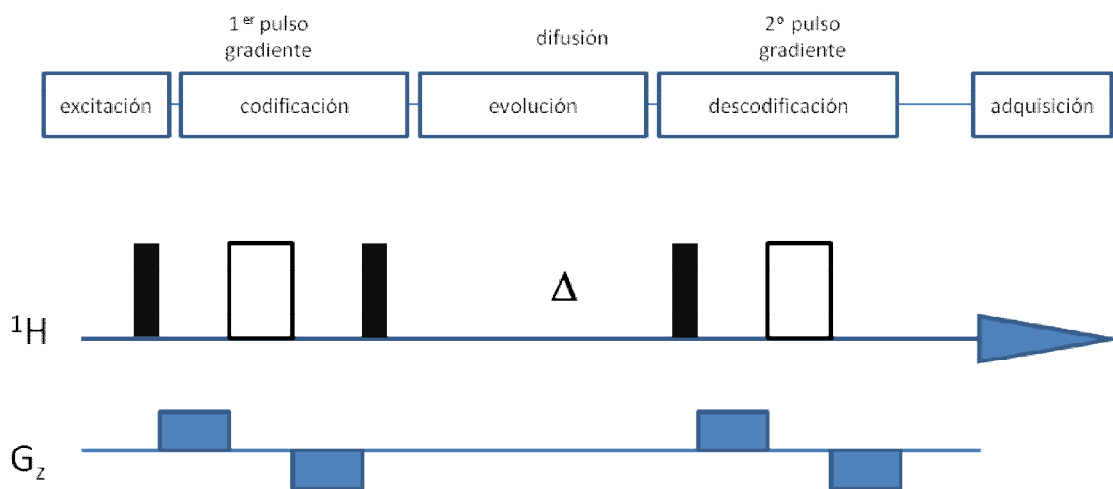


Figura 4.4: Experimento de eco estimulado con gradientes bipolares (BPP-STE) mostrando arriba el esquema de funcionamiento. En el canal ^1H rectángulos estrechos y anchos corresponden a pulsos duros de 90° y 180° , respectivamente.

La difusión traslacional permite que una cierta fracción de moléculas cambie de posición espacial durante el periodo Δ , y con ello, que la magnetización correspondiente a sus espines nucleares no se descodifique y adquiera un desfase, lo que produce una cierta atenuación en la intensidad de la señal de RMN que se detecta. La ecuación de Stejskal-Tanner que se deduce del desarrollo teórico del experimento (Price1998, Johnson1999 y referencias en ellos) permite relacionar el grado de atenuación con el coeficiente de difusión (D) y otros parámetros del experimento (Figura 4.5).



	Gradiente rectangular	$\ln \frac{I(G)}{I(0)} = -\gamma^2 G^2 D \delta^2 \left(\Delta - \frac{\delta}{3} \right)$
	Gradiente sinusoidal	$\ln \frac{I(G)}{I(0)} = -\gamma^2 G^2 D \delta^2 \left(\frac{4\Delta - \delta}{\pi^2} \right)$

Figura 4.5: Ecuaciones de Stejskal-Tanner para determinar la difusión molecular para las versiones del experimento BPP-STE que utilizan gradientes bipolares con forma sinusoidal y rectangular, respectivamente. $I(0)$ es la intensidad inicial, γ es la constante magnetogrfica, que para el ^1H es 26751 rad. G-1 s-1, G el potencial de gradiente (G/cm), D la constante de difusión, en $\text{cm}^2 \text{s}^{-1}$, δ la duración de los gradientes (aprox. 0.001 s) y Δ la duración del eco estimulado.

La relación entre el coeficiente de difusión traslacional y el grado de atenuación de la señal en el experimento BPP-STE permite también utilizar el experimento en mezclas de compuestos utilizándolo como filtro de difusión, para obtener un espectro 1D con una mayor atenuación en las señales de moléculas que difundan más rápidamente (por ej. señales de disolventes) de forma que se simplifique el espectro

4.2.4 Optimización del tiempo de difusión Δ en el experimento *Drev*

En el experimento *Drev* propuesto la fuerza del gradiente se alterna, en *scans* sucesivos, entre la mínima y la máxima disponible en el espectrómetro. El único parámetro que se debe optimizar para conseguir los mejores resultados, esto es, atenuación máxima para las señales de macromoléculas y mínima para las moléculas pequeñas, es el periodo de evolución por difusión Δ .

Utilizando la ecuación de Stejskal-Tanner puede deducirse teóricamente el período de evolución por difusión óptimo (Δ_{opt}) del experimento *Drev* para que se

consiga el espectro con mayor atenuación relativa posible de las señales de la molécula de mayor tamaño, para dos moléculas con coeficientes de difusión cualesquiera. Las ecuaciones que se dedujeron para el experimento *Drev* con gradientes rectangulares y sinusoidales se dan en Ecuaciones 4.4 y 4.5, respectivamente:

$$\Delta_{opt} = \frac{\ln(D_s/D_f)}{-\gamma^2 G_{max}^2 \delta^2 (D_f - D_s)} + \frac{\delta}{3} \quad \text{Ecuación 4.4}$$

$$\Delta_{opt} = \frac{4 \cdot \ln(D_s/D_f)}{-\gamma^2 G_{max}^2 \delta^2 (D_f - D_s)} + \frac{4 \cdot \delta}{\pi^2} \quad \text{Ecuación 4.5}$$

Donde Δ_{opt} es el período de evolución por difusión óptimo, D_s y D_f son los cocientes de difusión de la molécula “lenta” y la molécula “rápida” respectivamente, γ es la constante magnetogírica, G es la potencia de gradiente y δ la duración del gradiente. La forma de la curva Δ_{opt} se muestra en Figura 4.6.

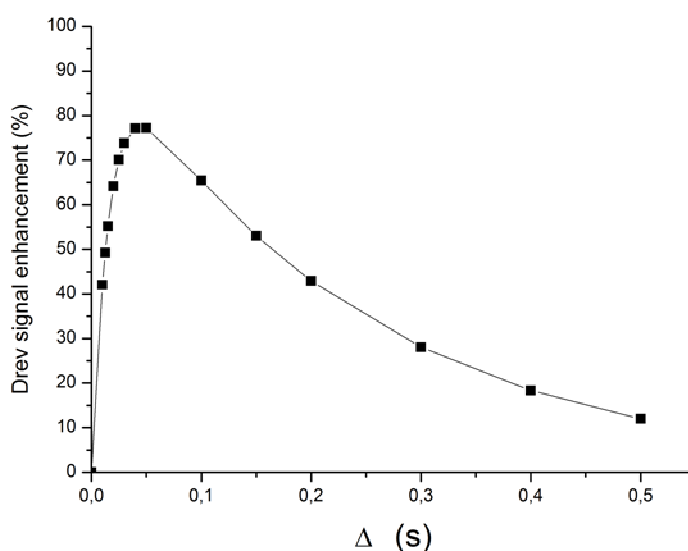


Figura 4.6: Simulación teórica de la atenuación de las señales de dos moléculas con diferente coeficiente de difusión en el experimento *Drev*. Los parámetros utilizados en la simulación son: D_f y D_s 5.45×10^{-6} y $3.51 \times 10^{-7} \text{ cm}^2 \text{ s}^{-1}$, respectivamente. La fuerza de los gradientes con forma rectangular y duración 2 ms fue de G_{min} y G_{max} 0 y 65 G cm^{-1} , respectivamente. En estas condiciones el Δ óptimo que maximiza la diferencia de intensidades entre las dos moléculas está en torno a 45.1 ms en la que se consigue un factor de atenuación de $\sim 77.5\%$ para las señales de la macromolécula.

En las Ecuaciones 4.4 y 4.5 se han obviado los efectos de atenuación debidos a la relajación T_1 , que operaría durante el tiempo de difusión Δ_{opt} , y la relajación T_2 , que operaría durante el tiempo de los gradientes bipolares. Para todos los parámetros iguales, el uso de gradientes sinusoidales requiere incrementar el Δ_{opt} en un factor 4, aproximadamente respecto al mismo experimento con gradientes rectangulares.

4.3 Procedimiento experimental

4.3.1 Experimentos con filtro de difusión reverso (*Drev*)

Los experimentos de RMN se adquirieron a 300 K en un espectrómetro Varian Inova-750 equipado con sonda HCP con gradientes en tres ejes x, y, z y operando con software VNMRJ. 2.2D. Los espectros se procesaron con software MestreNova (*Mestrelab Research inc*). Para comprobar el funcionamiento de los experimentos con filtro *Drev* (ver abajo) se preparó en un tubo de RMN una muestra de 7.15 mM de indometacina (IND) (catálogo Guinama) y 0.15 mM de proteína zeína del maíz (catálogo Aldrich) en la mezcla de disolventes deuterados CD₃OD: D₂O 90:10 v/v. La proporción entre ligando IND y la proteína fue de ~ 40:1.

Se implementaron los siguientes experimentos: 1D CPMG, 1D *Drev*, 1D CPMG-*Drev*, 1D STD-CPMG, 1D STD-*Drev*, y 1D-STD-CPMG-*Drev*. Para los experimentos STD se implementó una versión para obtener directamente el espectro resta STD (ON-OFF) y otra para obtener solamente STD-OFF.

En los experimentos que utilizan *Drev* la duración de cada gradiente fue de 1 ms de duración. La potencia del gradiente que codifica la difusión G_z on/off se alternó entre 65 y $\sim 0.0 \text{ G cm}^{-1}$, respectivamente. El periodo de difusión Δ en el eco estimulado fue de 50 ms. En los experimentos que utilizan CPMG el período entre pulsos de 180° fue de 600 μs y el número de ciclos (m en las secuencias de pulsos mostradas) se ajustó al número par más próximo para una duración nominal del filtro de 50 ms.

En los experimentos que utilizan STD el número de *scans* fue de 1024. El tiempo de saturación fue de 3 s. Para la saturación se utilizó un ciclo de pulsos

gaussianos de 1 ms de baja potencia (4 dB). Para los *scans* correspondientes a STD-ON se aplicó la saturación en la zona aromática de la proteína zeína a 7.26 ppm. Para los *scans* correspondientes a STD-OFF la saturación se aplicó a 20 ppm.

4.3.2 Experimento STD-Homodesacoplado

Los experimentos RMN se adquirieron a 298 K en un espectrómetro Varian INOVA 750 MHz equipado con una sonda HCP con gradientes en tres ejes x, y, z y operando con software VNMRJ. Los espectros fueron procesados con el software MestRe-C v4.9.

Para comprobar el funcionamiento del experimento STD-homodesacoplado se prepararon en tubos de RMN independientes una disolución de 10 mg del tetrasacárido lacto-n-tetraosa (β -D-Gal-(1 \rightarrow 3)- β -D-GlcNAc-(1 \rightarrow 3)- β -D-Gal-(1 \rightarrow 4)-D-Glc) y otra de 12 mg de lactosa, disueltos cada uno en 3 mg de proteína galectina de pollo (CG-14) en 600 μ l de buffer NaCl/P_i (concentración de buffer 20 mM, H₂O/D₂O = 9:1, 0.05% NaN₃) a pH = 7. La proporción entre oligosacárido y proteína fue ~ 200:1.

Se implementó el experimento 2D-STD-T_{1ρ}-J-resolved-watergate. El tiempo total de saturación fue de 1 s (con d₁=1 s) para el espectro on-resonance y de 2 s (con d₁=0 s) para el espectro off-resonance. La saturación selectiva on/off-resonance se situó a 7.5 ppm y 20 ppm. La saturación se consiguió mediante una serie de pulsos gaussianos de 50 ms de duración con una fuerza de campo (*field strength*) de 32 Hz, separados por un tiempo de espera de 1 ms. El *filtro-T_{1ρ}* consistió en un pulso continuo de bloqueo de espines de 60 ms de duración y 2 kHz. La anchura espectral de adquisición y de la dimensión J-resuelta fue de 9500 y 25 Hz, respectivamente. El

espectro se adquirió con 128 *scans* y 48 incrementos en la dimensión J-resuelta. El tiempo total de adquisición fue de aprox. 4 horas.

La forma de procesar el experimento 2D-STD- $T_{1\rho}$ -Jres-wg para obtener un espectro homodesacoplado se basa en el método propuesto por Gueneau (Gueneau1999). Primero, si es necesario, se elimina por procesado la señal residual de disolvente aplicándole a la dimensión de adquisición de la FID un filtro de paso alto a la frecuencia del disolvente. Después, se aplica la transformada de Fourier a lo largo de la dimensión de adquisición y de la dimensión J-resuelta con una apodización seno desplazada 45° y 90° respectivamente en F2 y F1, y se aplica un llenado con ceros para extender los datos hasta los 4096 y 128 puntos complejos en F2 y F1, respectivamente. Tras la transformada, el espectro se gira 45° a lo largo de la dimensión de adquisición, y se aplica un algoritmo de derivada en F1 para realzar la resolución en esa dimensión. En este punto, en vez de usar el método derivativo discreto habitual, se utilizó un método de onda continua (Nie2002), que dio lugar a una mejora en la relación señal-ruido. El espectro obtenido como resultado de este procesado se representa en modo potencia.

Mediante la suma de las proyecciones de todas las trazas de la dimensión F2 del espectro 2D-STD- $T_{1\rho}$ -Jres-wg procesado de esta forma se obtiene el espectro 1D STD homodesacoplado (Guennenau1999, Cobas 2004). Se adquirió un experimento 1D-STD- $T_{1\rho}$ -wg (Wang2004) en condiciones análogas a las descritas para el experimento 2D-STD- $T_{1\rho}$ -Jres-wg. El filtrado $T_{1\rho}$ de este experimento se calibra para eliminar las señales de la galectina, de forma que el espectro resultante muestre solamente las señales del oligosacárido. Para lactosa se adquirió un espectro 2D-HSQC-wg para una correcta asignación de las señales de la lactosa bajo las condiciones experimentales, basándonos en asignaciones completas previamente publicadas (Platzer1989;

Breg1988). La asignación La asignación de señales de protón de lacto-n-tetraosa se tomó de la bibliografía (Strecker89).

El resultado de aplicar el filtro *Drev* es un espectro en el que se atenúan en mayor medida las señales de compuestos que menos cambian de intensidad con la fuerza del gradiente, es decir, la de los agregados de tamaño macromolecular. Cuando se utilizan gradientes de una duración habitual, en torno a 1-2 ms, y su potencia se alterna en el experimento entre la máxima y la mínima disponible, el único parámetro que hay que optimizar es el período de evolución por difusión Δ del experimento. De acuerdo con los cálculos teóricos representados en la Figura 4.6 y para la potencia máxima de gradiente accesible en los espectrómetros, en general se deben utilizar valores de Δ pequeños (ej. 10-100 ms) para que el cambio de intensidad producido por el gradiente sea máximo en las señales de moléculas pequeñas y por tanto den señales más intensas y suceda lo contrario para las señales de macromoléculas y agregados moleculares de gran tamaño.

Se comprobó el funcionamiento del experimento *ID Drev* con una mezcla de proteína zeína (Figura 4.8a) y del ligando Indometacina (IND) (Figura 4.8b) preparada en CD₃OD en relación molar IND:zeína 40:1. La zeína es una proteína de 24 kDa que se obtiene de los cuerpos proteicos del endosperma del maíz que ha despertado interés para la preparación de biomateriales (Wang2003, Subramanian2007). Es una proteína insoluble en D₂O que contiene un número relevante de aminoácidos apolares y su estructura presenta diez segmentos helicoidales pegados en disposición antiparalela estabilizados por enlaces de hidrógeno y unidas por lazos ricos en glutamina (Figura 4.8a) (Matsushima1997).

Estudios previos en nuestro grupo de investigación habían mostrado la afinidad de IND en estas condiciones para unirse a zeína y determinado la constante de disociación del complejo, K_D 0.5±1.0 mM (Sousa2011, en preparación).

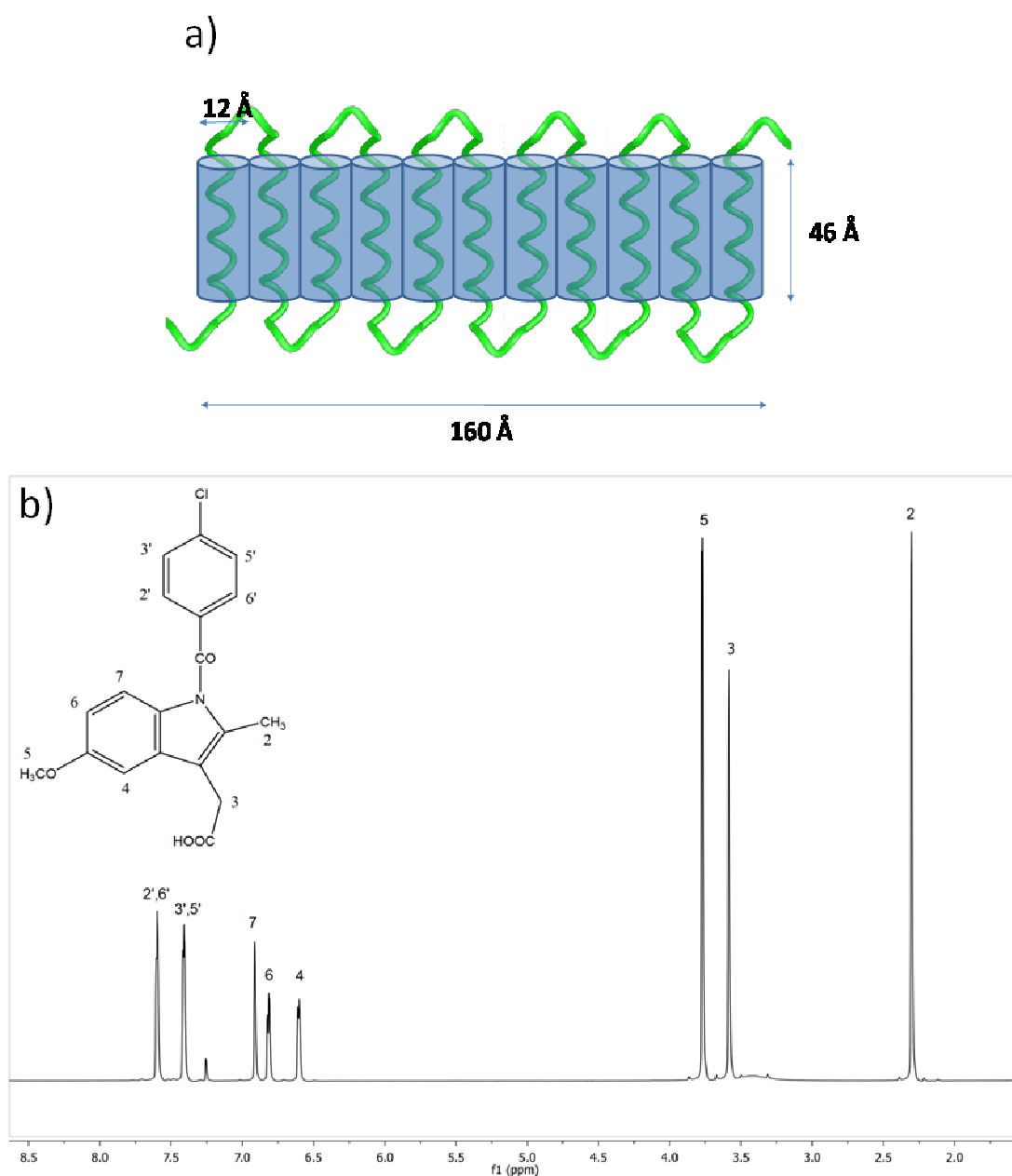


Figura 4.8: a) Estructura de la unidad monomérica de proteína zeína. b) Estructura y espectro 1D ^1H de IND puro mostrando la asignación de sus señales de protón.

Los espectros de la Figura 4.9 permiten comparar el resultado de aplicar el filtro *Drev* a la muestra IND:zeína 40:1. En el espectro 1D ^1H (Figura 4.9a), las señales de zeína aparecen como un fondo de señales claramente identificables por su mayor anchura respecto a las de IND. Si se toma como referencia la intensidad de la señal aromática 2', 6' de la molécula de interés IND, las señales alifáticas de zeína a ~ 0.92

ppm se ven atenuadas en el espectro *1D Drev* (Figura 4.9c) en un factor ~ 9.0 respecto a las mismas señales en el espectro $1D\ ^1\text{H}$ original. Se comprueba además que las señales de IND en el espectro *1D Drev* conservan en buena medida las integrales relativas originales. Por comparación, en el experimento de filtro- T_2 $1D$ CPMG (Figura 4.9b) con un tiempo de filtro de 50 ms solamente consigue una atenuación de las señales de proteína en un factor 2.9. Se observan además en éste ciertas imperfecciones en la fase de las señales de IND que son generadas por componentes antifase que no pueden corregirse por un ajuste de fase convencional. No obstante, estas componentes antifase se podrían eliminar si se incorpora un filtro-Z antes de la detección (Rastrelli2009).

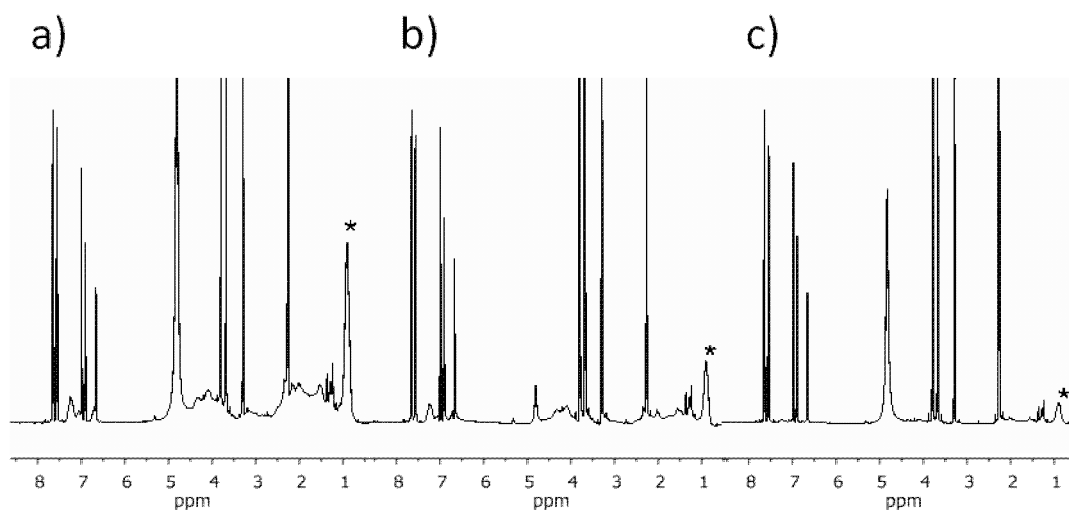


Figura 4.9: Diferentes espectros de una mezcla IND:zeína en proporción 40:1 a) $1D\ ^1\text{H}$, b) $1D$ CPMG 50 ms y c) *1D Drev* 50 ms. Los espectros están representados a la misma escala vertical respecto a la señal aromática 2', 6' de IND. La señal alifática de zeína utilizada para la cuantificación de la atenuación está indicada con un asterisco.

Experimento 1D con filtro combinado de relajación transversal y filtro de difusión reverso; 1D CPMG-Drev:

Se puede potenciar el efecto de filtrado de señales de agregados de gran tamaño o macromoléculas si se combinan secuencialmente en el mismo experimento un filtro de

relajación transversal CPMG con el filtro *Drev*. Esta combinación de filtros se muestra en la secuencia de pulsos *ID CPMG-Drev* de la Figura 4.10.

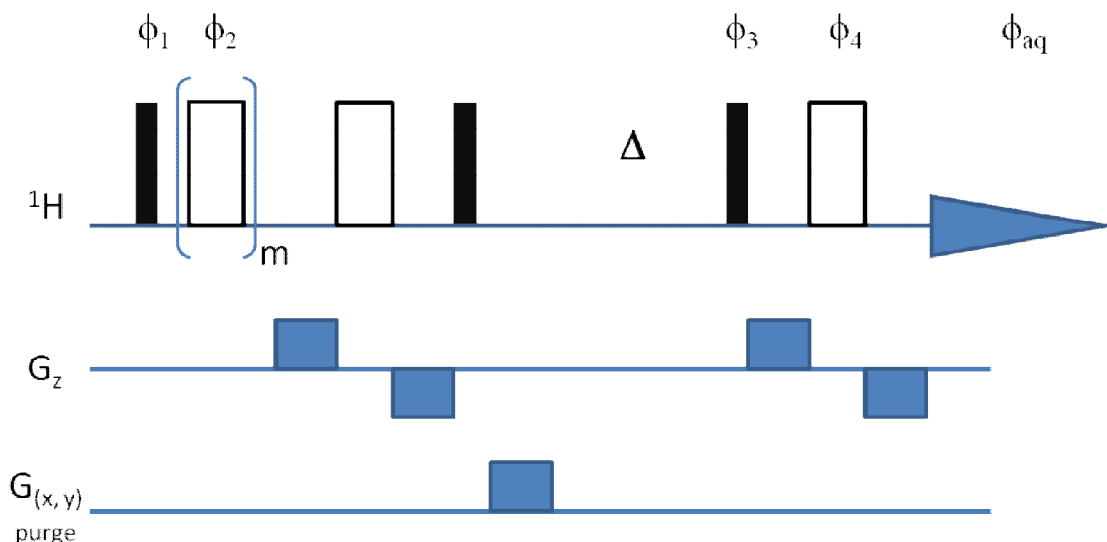


Figura 4.10: Secuencia de pulsos del experimento *ID CPMG-Drev*. En el canal ^1H rectángulos estrechos y anchos corresponden a pulsos duros de 90° y 180° , respectivamente. El ciclo de fases es: ϕ_1 : 0, 2, 2, 0, 1, 3, 3, 1; ϕ_2 : 1, 1, 1, 1, 2, 2, 2, 2; ϕ_3 : 0, 0, 2, 2, 0, 0, 2, 2, 1, 1, 3, 3, 1, 1, 3, 3, 2, 2, 0, 0, 2, 2, 0, 0, 3, 3, 1, 1, 3, 3, 1, 1; ϕ_4 : 0^{16} , 2^{16} ; ϕ_{aq} : 2, 0, 0, 2, 3, 1, 1, 3, 1, 3, 3, 1, 2, 0, 0, 2, 0, 2, 2, 0, 1, 3, 3, 1, 3, 1, 1, 3, 0, 2, 2, 0, 0, 2, 2, 0, 1, 3, 3, 1, 3, 1, 1, 3, 0, 2, 2, 0, 2, 0, 0, 2, 3, 1, 1, 3, 3, 1, 2, 0, 0, 2. G_z on/off: 1, 1, 0, 0.

Los espectros de la Figura 4.11 ilustran el funcionamiento del filtro *ID CPMG-Drev* para la muestra de IND:zeína 40:1. Se obtiene un factor de atenuación de 19.9 para la señal alifática de zeína a 0.92 ppm mientras que para el espectro 1D CPMG con 50 y 100 ms de filtro las atenuaciones fueron de 2.9 y 8.4, respectivamente. En la ampliación que se muestra de esta zona en los espectros de la Figura 4.11d, e, f puede observarse como al utilizar *ID CPMG-Drev* comienzan a apreciarse debajo de la señal alifática ancha de zeína algunas señales estrechas que no podía detectarse con el espectro 1D ^1H . Se trata, con toda probabilidad, de señales debidas a los protones de impurezas de la zeína (compuestos de bajo peso molecular).

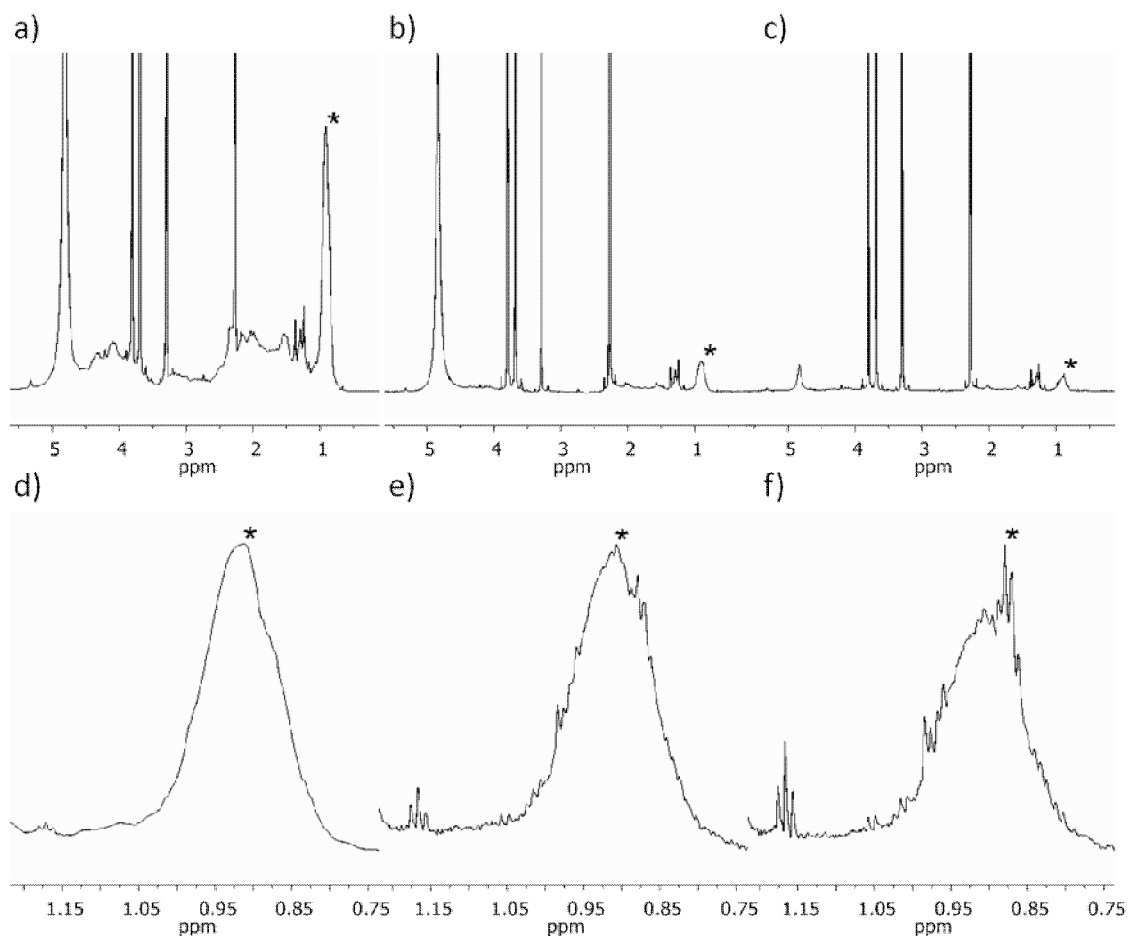


Figura 4.11: Espectros de la mezcla IND:zeína 40:1. a) $1D\ ^1H$, b) $1D\ Drev$ y c) $1D\ CPMG-Drev$. En los espectros la duración de los filtros CPMG y *Drev* fue cada uno de 50 ms. Los espectros a, b y c están representados a la misma escala vertical respecto a la señal aromática 2', 6' de IND. Los espectros d), e) y f) muestran una ampliación de la zona alifática correspondiente a los espectros a), b) y c), respectivamente. La señal alifática de zeína utilizada para la cuantificación de la atenuación está indicada con un asterisco.

Los espectros de la Figura 4.12 permiten comparar para la misma muestra el espectro $1D\ CPMG-Drev$ y $1D\ CPMG$, ambos con la misma duración total de los filtros de 100 ms. La señal alifática de la proteína a 0.92 ppm (marcada con asterisco * en Figura 4.12) se reduce en un factor 2.1 en el primer espectro. Por el contrario la señal de un compuesto de pequeño tamaño (indicada con almohadilla # en Figura 4.12) se ve incrementada en un factor 1.4 en el primer espectro.

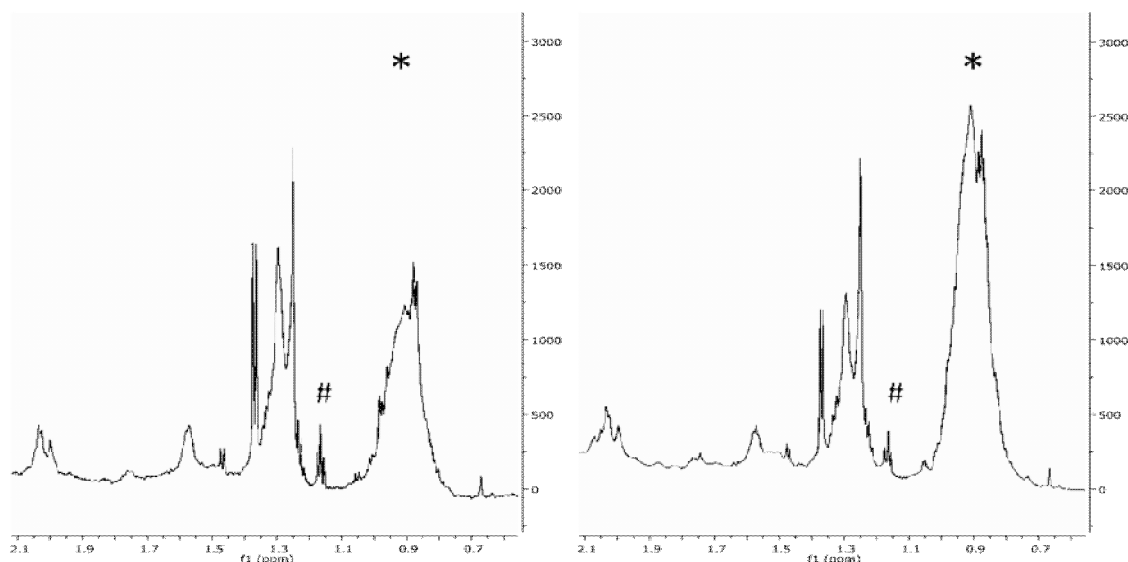


Figura 4.12: Espectros de la mezcla IND:zeína 40:1. a) *1D CPMG-Drev* con filtro total de 100 ms (50 ms CPMG +50 ms *Drev*) y b) *1D CPMG* con filtro de 100 ms. Los espectros están representados a la misma escala vertical respecto a la señal aromática 2', 6' de IND. Los espectros muestran una ampliación de la zona alifática. Se hace referencia en el texto a las señales indicadas con el asterisco y almohadilla.

Los resultados cuantitativos obtenidos para la muestra IND:zeína en la Figura 4.13 demuestran que con los filtros *Drev* y *CPMG-Drev* pueden conseguirse factores de supresión de señales macromoléculas bastante mayores que con un filtro convencional CPMG de la misma duración, llegando a ser por ej. hasta un orden de magnitud mayores. Esto puede conseguirse incluso para una duración del filtro *Drev* de tan sólo 10 ms. La Figura 4.13 muestra además que en el experimento *CPMG-Drev* el factor de supresión que se consigue a cada duración de filtro es aproximadamente el producto de ambos filtros individuales.

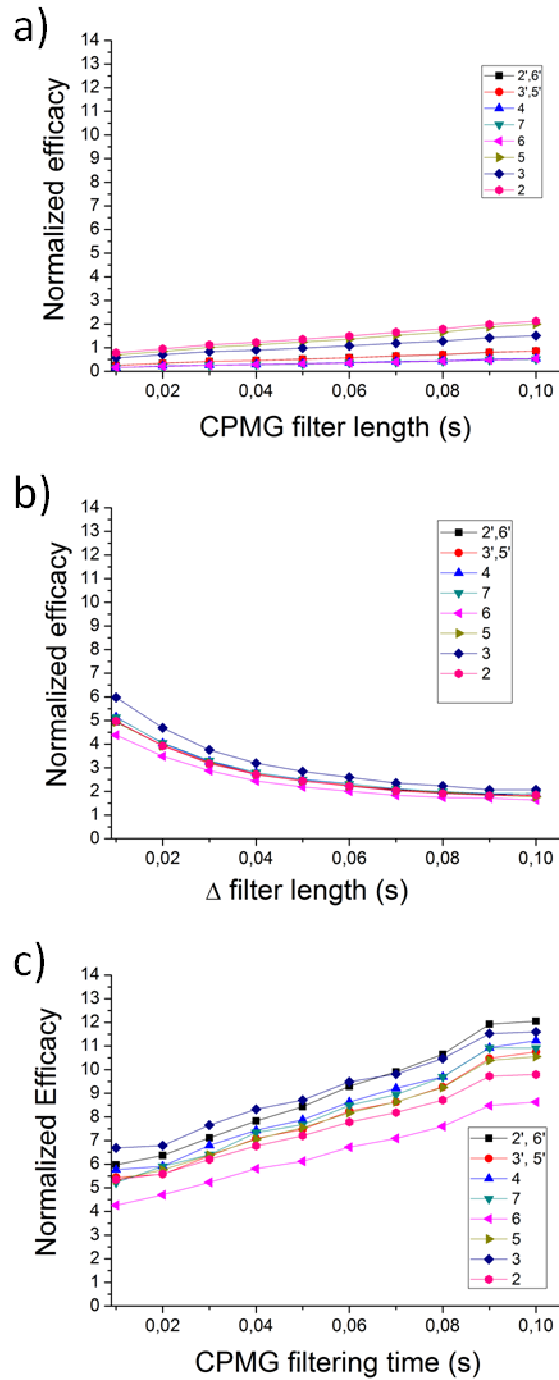


Figura 4.13: Curvas experimentales de la eficacia normalizada de los filtros CPMG, *Drev*, y *CPMG-Drev* para la suppression de señales de la macromolécula en la muestra IND:zéina 40:1 frente al tiempo de filtrado utilizado. a) 1D CPMG (filtro T2 convencional), b) 1D *Drev* (adquirido con $\delta = 3$ ms, G_z -on = 65 G/cm⁻¹), y c) 1D *CPMG-Drev* (adquirido con $\Delta = 10$ ms, $\delta = 3$ ms y G_z -on = 65 G/cm⁻¹). Cada punto representa la integral de la señal de IND normalizada por el número de protones que generan la señal, y dividido por la integral de la señal de zéina a ~ 0.9 ppm.

Experimentos STD con filtro de difusión reverso; 1D STD-Drev y 1D-CPMG-Drev:

Para el cribado por RMN de compuestos activos es importante disponer de experimentos capaces de detectar e identificar limpiamente las señales de un ligando con la afinidad suficiente para formar complejo con una proteína diana, a la vez que suprimir las señales de esta última del espectro que pueden ocultar algunas de las señales de interés. Esto es especialmente útil para estudios en los que interesa realizar medidas de intensidad cuantitativa para determinar el epítipo de unión del ligando. Algunos de los experimentos disponibles para detectar ligandos que forman complejos con macromoléculas son: *STD- $T_{1\rho}$* (*saturation transfer difference* que incorpora un filtro $T_{1\rho}$) (Mayer2001), *WaterLOGSY* (Dalvit2000, Dalvit2001), *NOE-pumping* (Chen1998) y *reverse NOE-pumping experiment* (Chen2000).

En el experimento STD la forma convencional de eliminar las señales del receptor macromolecular es utilizar un filtro de relajación transversal como el filtro $T_{1\rho}$ (Mayer2001) o bien un filtro CPMG (Figura 4.14a) (esta última versión es menos habitual aunque se ha comprobado que funciona mejor en los espectrómetros Varian de nuestro laboratorio). En este apartado se ha probado la introducción de los filtros *Drev* y *CPMG-Drev*, descritos anteriormente, al experimento STD como una alternativa a los métodos de filtrado mencionados. Las secuencias de pulsos desarrolladas para los experimentos *1D STD-Drev* y *1D STD-CPMG-Drev* se muestran en las Figuras 4.14b y 4.14c, respectivamente.

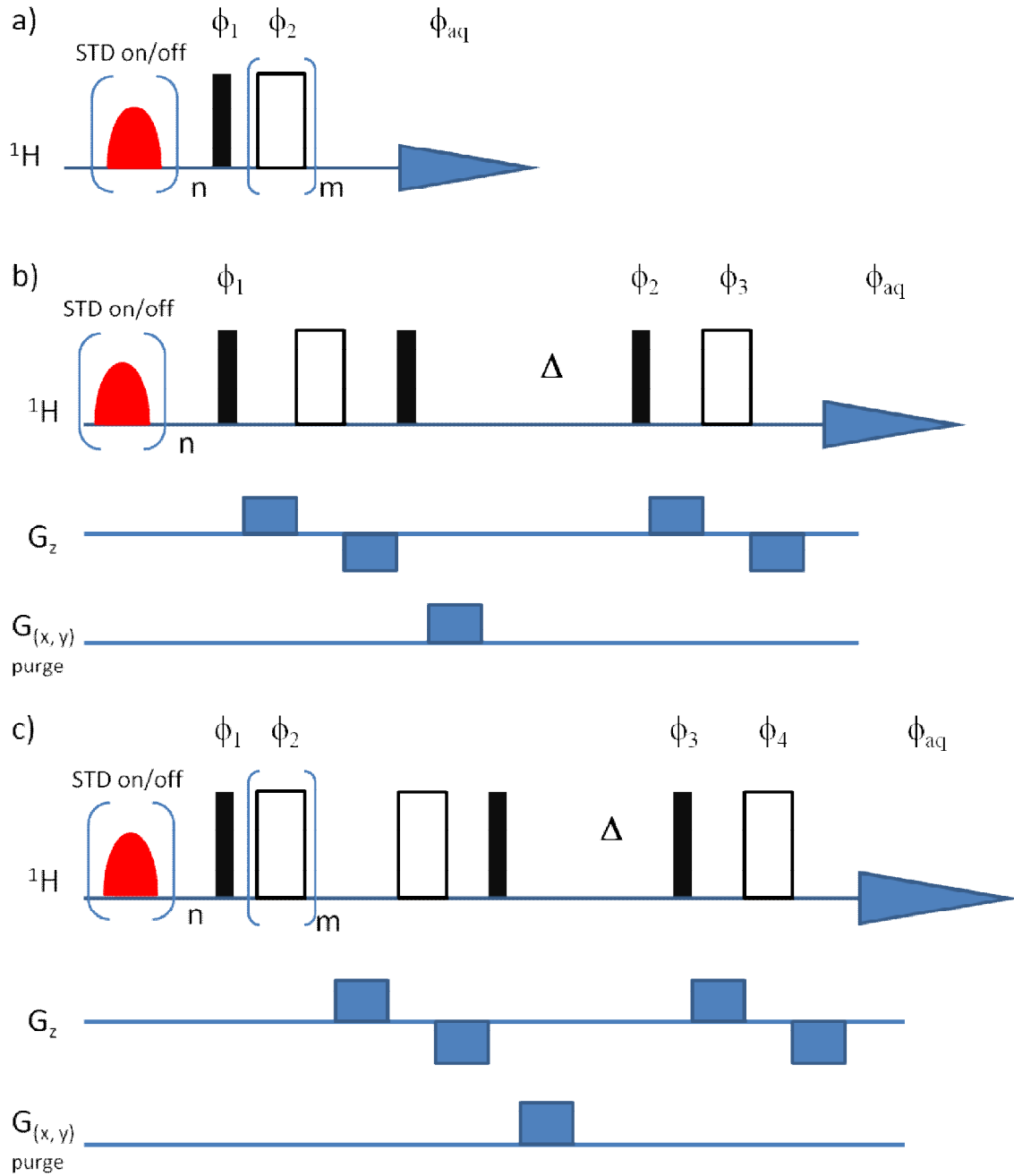


Figura 4.14: Secuencia de pulsos de varios experimentos STD utilizados. En el canal ^1H rectángulos estrechos y anchos corresponden a pulsos duros de 90° y 180° , respectivamente. a) Experimento *1D STD-CPMG*. El ciclo de fases que proporciona el espectro STD on-off es: ϕ_1 : 0, 0, 2, 2, 1, 1, 3, 3; ϕ_2 : 1, 1, 1, 1, 0, 0, 0, 0; ϕ_{aq} : 0, 2, 2, 0, 1, 3, 3, 1. El ciclo de fases que proporciona el espectro STD Off es idéntico salvo ϕ_{aq} : 0, 0, 2, 2, 1, 1, 3, 3 y STD on/off: 0, 0. b) Experimento *1D STD-Drev*. El ciclo de fases que proporciona el espectro STD on-off es: ϕ_1 : 0, 0, 2, 2, 1, 1, 3, 3; ϕ_2 : 0, 0, 2, 2, 0, 0, 2, 2, 1, 1, 3, 3, 1, 1, 3, 3, 2, 2, 0, 0, 2, 2, 0, 0, 3, 3, 1, 1, 3, 3, 1, 1; ϕ_3 : $0^{16}, 2^{16}$; ϕ_{aq} : 2, 0, 0, 2, 3, 1, 1, 3, 1, 3, 3, 1, 2, 0, 0, 2, 0, 2, 2, 0, 1, 3, 3, 1, 3, 1, 1, 3, 0, 2, 2, 0, 0, 2, 2, 0, 1, 3, 3, 1, 3, 1, 1, 3, 0, 2, 2, 0, 2, 0, 0, 2, 3, 1, 1, 3, 1, 3, 3, 1, 2, 0, 0, 2. G_z on/off: 1, 1, 0, 0; STD on/off: 1, 0. El ciclo de fases que proporciona el espectro STD off es idéntico salvo: ϕ_1 : 0, 2, 2, 0, 1, 3, 3, 1 y STD on/off: 0, 0. c) Experimento *1D STD-*

CPMG-Drev. El ciclo de fases que proporciona el espectro STD on-off es: ϕ_1 : 0, 0, 2, 2, 1, 1, 3, 3; ϕ_2 : 1, 1, 1, 2, 2, 2, 2; ϕ_3 : 0, 0, 2, 2, 0, 0, 2, 2, 1, 1, 3, 3, 1, 1, 3, 3, 2, 2, 0, 0, 2, 2, 0, 0, 3, 3, 1, 1, 3, 3, 1, 1; ϕ_4 : 0^{16} , 2^{16} ; ϕ_{aq} : 2, 0, 0, 2, 3, 1, 1, 3, 1, 3, 3, 1, 2, 0, 0, 2, 0, 2, 2, 0, 1, 3, 3, 1, 3, 1, 1, 3, 0, 2, 2, 0, 0, 2, 2, 0, 1, 3, 3, 1, 3, 1, 1, 3, 0, 2, 2, 0, 2, 0, 0, 2, 3, 1, 1, 3, 1, 3, 3, 1, 2, 0, 0, 2. G_z on/off: 1, 1, 0, 0; STD on/off: 1, 0. El ciclo de fases que proporciona el espectro STD off es idéntico salvo: ϕ_1 : 0, 2, 2, 0, 1, 3, 3, 1 y STD on/off: 0, 0.

Se utilizó la mezcla IND:zeína 40:1 para probar el funcionamiento de los experimentos *ID STD-Drev* y *ID-CPMG-Drev* y se compararon sus resultados con los de los experimentos más habituales *ID STD* y *ID STD-CPMG* (Figura 4.15). En cualquiera de los espectros de Figura 4.15 se observa que únicamente las señales aromáticas de IND dan respuestas STD y por tanto forman parte del epítipo de unión a la proteína zeína, además puede verse la influencia del tipo de filtrado en la eliminación de las señales de la proteína. El filtrado es claramente superior en el experimento *ID STD-Drev* y mejora aún más en *ID STD-CPMG-Drev*. Cabe destacar que aunque la sensibilidad de estos dos últimos experimentos es teóricamente inferior y por tanto se requiere mayor tiempo de acumulación para conseguir el mismo grado de sensibilidad que al utilizar las secuencias sin el filtro *Drev*, pueden ser una buena alternativa para determinar señales del epítipo de unión donde la eficacia del filtrado CPMG no sea suficiente.

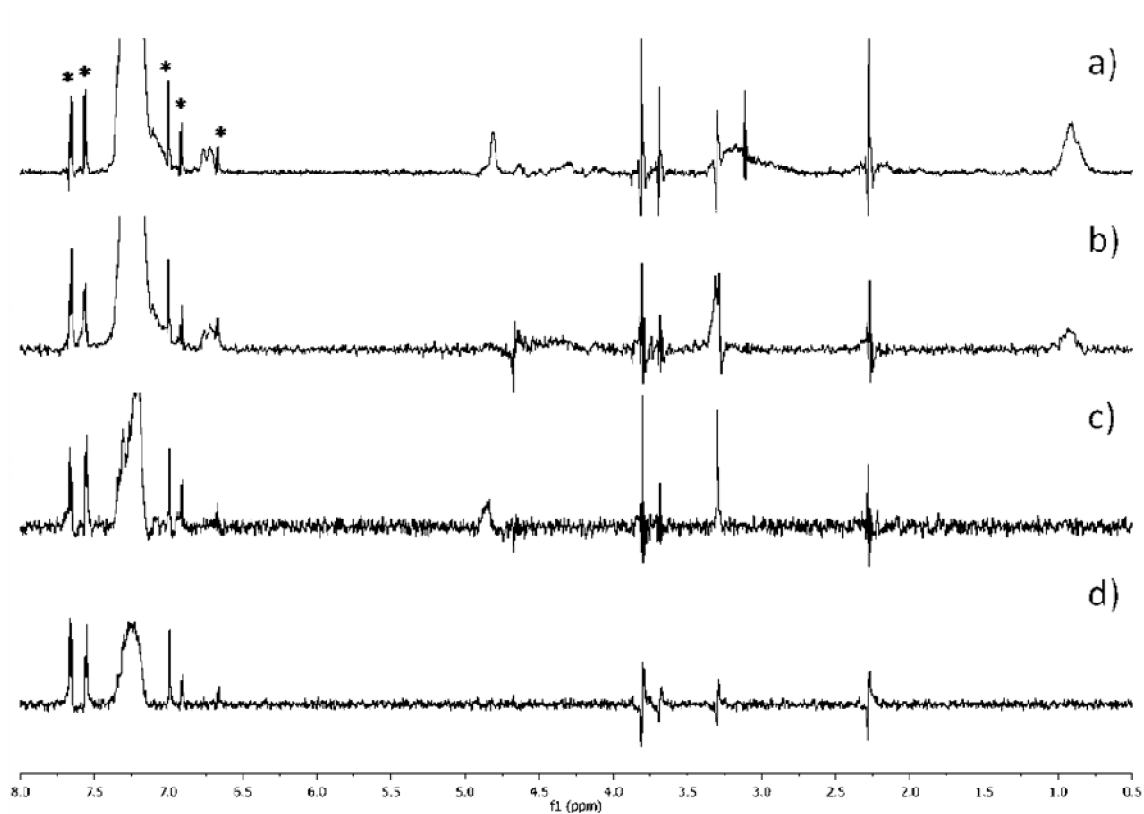


Figura 4.15: Espectros de la mezcla IND:zeína 40:1. a) 1D STD, b) 1D STD-CPMG con duración de filtro 50 ms c) *1D STD-Drev* con duración de filtro 50 ms, y d) *1D-STD-CPMG-Drev* con duración del filtro CPMG y *Drev* de 50 ms cada uno. En el espectro a) se han marcado con un asterisco las señales STD de IND, el resto son señales de la proteína zeína o artefactos de resta.

4.4.2 Parte 2 Experimento 1D-STD homodesacoplado

La espectroscopía J-resuelta es uno de los experimentos 2D más sensibles y rápidos de obtener por los métodos habituales de adquisición. Guenneau y otros (Guenneau1999) desarrollaron un método de procesado de espectros J-resueltos que lo transforma en el correspondiente espectro 1D de protón homodesacoplado, es decir, donde todas las señales aparecen como singletes. Este tipo de procesado del espectro reduce el solapamiento espectral a la vez que reduce la dimensionalidad del espectro que resulta.

En nuestro laboratorio se había explorado esta metodología de homodesacoplamiento a experimentos DOSY (Cobas 2004). En este apartado se explora su aplicación a los experimentos STD como forma de facilitar la asignación de las señales de ligandos que interaccionan con biomoléculas diana. La secuencia de pulsos que se propone para obtener un espectro STD-homodesacoplado se denominó 2D-STD- $T_{1\rho}$ -Jres-wg (Figura 4.16). Este experimento consta de varios módulos cuyo efecto se detalla a continuación.

La secuencia comienza con un período de saturación en el que se alterna la frecuencia de la saturación selectiva entre on- y off-resonance en *scans* consecutivos y de forma sincronizada con el ciclo de fases para obtener, cada dos *scans* consecutivos, una FID con auto-substracción de los espectros STDoff y STDOn. Esta implementación permite minimizar los errores de resta que aparecen a veces en los espectros STD debidos a fluctuaciones de temperatura o por agregación en la muestra durante el experimento. Tras la saturación, le sigue el primer pulso excitación y un filtro $T_{1\rho}$ cuya duración puede ajustarse a conveniencia para eliminar las señales del receptor macromolecular (Hajduk1997).

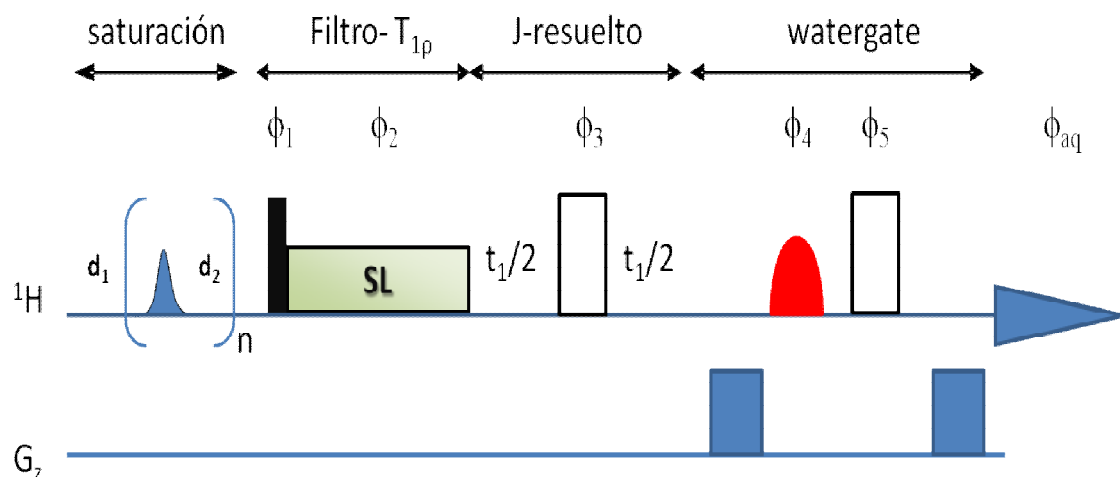


Figura 4.16: Secuencia detallada del experimento 2D-STD- $T_{1\rho}$ -Jres-wg. En el canal ^1H rectángulos estrechos y anchos corresponden a pulsos duros de 90° y 180° , respectivamente. d_1 es un tiempo de espera de relajación previo a la secuencia opcional. La saturación selectiva de la proteína se consigue mediante una serie de n pulsos gaussianos blandos separados por un tiempo de espera d_2 . El filtro- $T_{1\rho}$ se consigue mediante un pulso continuo de bloqueo de espines. En el bloque watergate, la fuerza del gradiente y el pulso selectivo de 180° deben calibrarse para suprimir las señales del disolvente. En *scans* alternos, la saturación se calibra fuera del rango de señales de la proteína, y la resta se consigue mediante el ciclo de fases: ϕ_1 : x, x, -x, -x, y, y, -y, -y, ϕ_2 : y, y, -y, -y, -x, -x, x, x, ϕ_3 : x, x, x, x, y, y, y, y, ϕ_4 : x, -x, -x, x, y, -y, -y, y, ϕ_5 : -x, x, x, -x, -y, y, y, -y, y ϕ_{aq} : x, -x, -x, x, y, -y, -y, y.

A continuación va un módulo J-resuelto que introduce la dimensión indirecta t_1 en la que las respuestas STD se desdoblan con sus acoplamientos homonucleares. No es necesario adquirir un gran número de puntos en esta dimensión (por ej. 64 puntos en modo magnitud con anchura espectral a digitalizar de aprox. 25 Hz), lo que contribuye a reducir drásticamente la duración total del experimento en comparación con otros experimentos 2D-STD, que normalmente requieren un mayor número de puntos en F1 (Mayer1999, Vogtherr000, Möller2002). En las condiciones mencionadas, el tiempo máximo de evolución de t_1 es relativamente corto y por ello no es esperable que se produzcan durante este periodo pérdidas relevantes de la saturación STD en las señales del ligando. En cualquier caso, cabe la posibilidad de incluir también saturación selectiva durante el periodo de evolución t_1 . Finalmente, la secuencia incorpora un

módulo *soft-watergate* (Piotto1992) para la supresión fuerte de la señal del disolvente, un método adecuado por ej. para mezclas H₂O:D₂O 90:10.

Para comprobar el funcionamiento de este experimento se estudió la interacción de oligosacáridos con una lectina animal perteneciente al grupo de las galectinas (lectinas β -galactósido-específicas, Leonidas1998). En concreto, la lectina escogida fue CG-14 (*Chicken Galectine-14*) aislada de pollo. Se trata de una de las galectinas monoméricas que han sido utilizadas en ensayos funcionales sobre proliferación tumoral y diferenciación neuronal (André2005, Kopitz2004). CG-14 es capaz de unirse específicamente a unidades de galactosa de configuración beta que estén próximas al extremo no reductor y con un orden de preferencia Gal β (1-4) > Gal β (1-3) (Wu2007). Se escogieron dos oligosacáridos que incorporan este tipo de unidad que son lacto-N-tetraosa (LNT) y lactosa. La constante de disociación K_d de las galectinas con este tipo de oligosacáridos está normalmente en el rango de micro a milimolar, por lo que se satisfacen las condiciones de aplicabilidad del experimento STD, e incluso en ciertos casos las de otro experimento que proporciona información adicional sobre el modo de interacción, el experimento NOE transferido (tr-NOE) (Meyer2003, Dam2005).

Los espectros 2D-STD- $T_{1\rho}$ -Jres-wg obtenidos para la mezcla LNT:CG-14 se muestran en la Figura 4.17. La Figura 4.17a corresponde al espectro de referencia STD off que muestra todas las señales de LNT. El espectro de la Figura 4.17b corresponde al espectro diferencia STDoff – STDOn en el que la saturación STDOn es en la región aromática de la galectina. En este último espectro sólo se observan las señales STD del residuo β -D-Gal-(1 \rightarrow 4) de LNT generadas por su interacción dinámica con la lectina. En ambos espectros, al desplazamiento químico de cada señal (dimensión F2), se observa que los picos se desdoblán en F1 con sus acoplamientos escalares

homonucleares. Esto ayuda a detectar el solapamiento de señales y contribuye a la fiabilidad de la asignación de las señales STD.

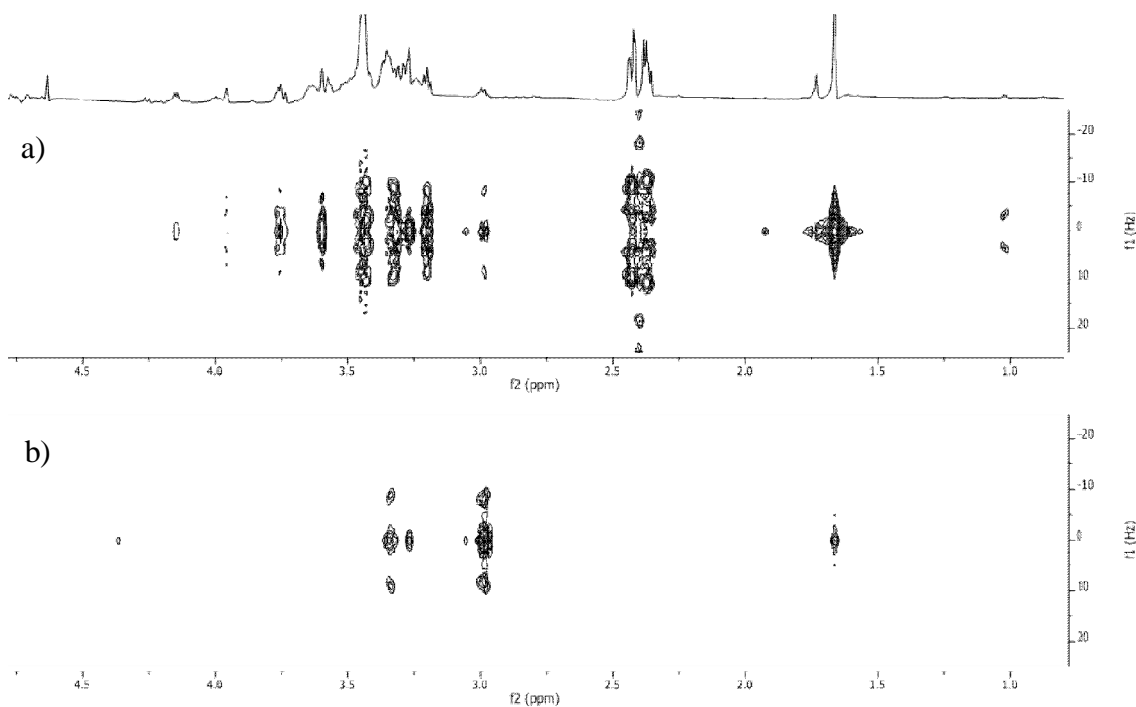


Figura 4.17: Espectros de la mezcla de LNT y galectina CG-14 en relación molar $\sim 200:1$. a) Espectro 2D- STD-Jres-wg de referencia (off), b) Espectro 2D-STD- $T_{1\rho}$ -Jres-wg (off-on con saturación-on de 2 s en la región aromática de la galectina). Sobre ambos espectros se da el espectro 1D ^1H - $T_{1\rho}$ watergate de la muestra.

Aunque en principio es viable realizar la asignación del espectro 2D STD J-resuelto por este método, resulta interesante para ganar sensibilidad utilizar un procesamiento especial de este espectro, descrito en la parte experimental, para convertirlo en el espectro 1D-STD-homodesacoplado. Un ejemplo del resultado que puede obtenerse con este tipo de procesamiento se da en la Figura 4.18 con los espectros 2D STD J-resueltos obtenidos con la mezcla de lactosa y CG-14. En la Figura 4.18a se da el espectro 1D STD convencional off-on obtenido con la secuencia 1D-STD- $T_{1\rho}$ -wg (512 *scans*) con 2 s de saturación. Este espectro es comparable con el espectro 1D-STD-Homodesacoplado (128 *scans*) de la Figura 4.18b. En este último espectro las señales

simplifican sus desdoblamientos sin afectar por ello a la anchura de línea conseguida que resultó similar a la observada en el espectro 1D- ^1H original de la muestra. Se observa una importante ganancia de sensibilidad en el espectro de la Figura 4.18b que está de acuerdo con su mayor tiempo de adquisición y con el hecho de que este espectro se construye como la suma de las trazas 1D de la dimensión J-resuelta.

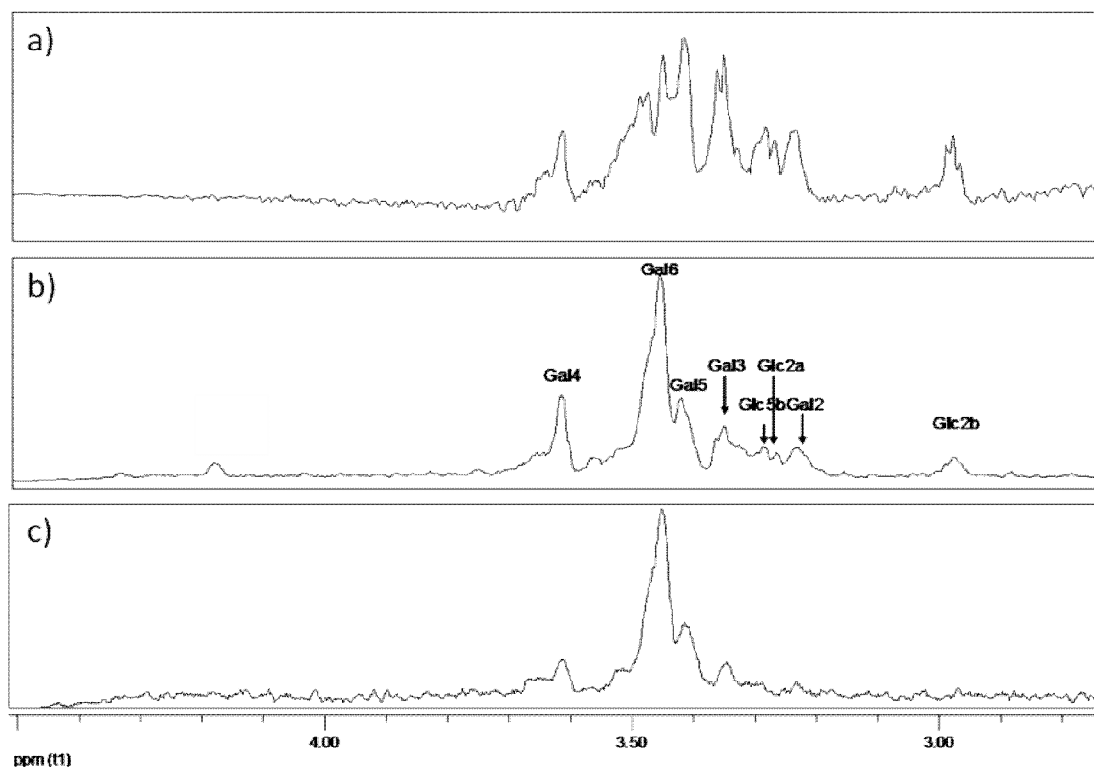


Figura 4.18: Comparación de espectros STD de la muestra compuesta por galectina CG-14 y lactosa; a) Espectro obtenido con un experimento 1D-STD convencional (off-on con tiempo de saturación 2 s), b) 1D-STD-Homodesacoplado (off-on con tiempo de saturación 2 s en región aromática de la galectina) y c) 1D-STD-Homodesacoplado (tiempo de saturación 1 s en región aromática de la galectina). Las etiquetas a y b hacen referencia a los anómeros α y β de la glucosa.

Comparando las señales de los espectros de Figura 4.18a y 18b la diferencia principal ocurre en la señal a ~ 3.45 ppm. Mientras que esta señal aparece como un complejo multiplete en el espectro de la Figura 4.18a, ésta se simplifica en el espectro 1D-STD-homodesacoplado de la Figura 4.18b en dos señales singlete parcialmente

solapadas que fueron asignadas como Gal-6 y Gal-5 con ayuda del espectro 2D-HSQC-wg.

El efecto de incrementar el tiempo de saturación en la intensidad de las señales STD puede comprobarse en los espectros 1D STD-Homodesacoplados de las Figuras 4.18b y 4.18c, correspondientes a tiempo de saturación de 2 y 1 s, respectivamente. En el primer caso, el mayor tiempo de saturación aplicado genera STDs con protones tanto del residuo de galactosa como del de glucosa (ambos componentes del disacárido lactosa). Los STDs con la galactosa son claramente más intensos que los que aparecen con la glucosa. Por otro lado, en el espectro de la Figura 4.18c, adquirido con menor tiempo de saturación, existen exclusivamente STDs con el residuo de galactosa, y las intensidades son similares a las que presentan las señales equivalentes en el espectro de la Figura 4.18b. Estos resultados indican que los STDs con la glucosa se deben a efectos de difusión de spin en el complejo, mientras que los STDs relativamente fuertes de Gal-3, Gal-4, Gal-5 y Gal-6, observables en ambos espectros (Figuras 4.18b y 4.18c), deben corresponder a STDs directos entre la lactosa y el lugar de interacción de CG-14, es decir, al epítipo de unión. Evidentemente, el extremo no reductor del azúcar se sitúa en las proximidades del centro activo de la lectina, como se ha observado por diversos medios, incluido análisis cristalográfico (Solís1996, Varela1999, Alonso-Plaza2001, Lopez-Lucedo2004). Esta confirmación apoya la validez del resultado del análisis de interacción realizado.

4.5 Conclusiones

Las secuencias de filtrado desarrolladas, *Drev* y *CPMG-Drev* consiguen, en muestras mezcla de macromoléculas con pequeños compuestos orgánicos, un alto grado de supresión de la característica señal de fondo de las macromoléculas, permitiendo visualizar las señales debidas a las moléculas pequeñas. Comparando los resultados que se obtienen con la de los filtros conocidos de resultado equivalente, T_2 y $T_{1\rho}$, experimentalmente se llega a conseguir un factor de supresión un orden de magnitud mayor. La fácil implementación de estos filtros en experimentos de STD los hace convenientes para su uso en el *screening* de pequeños ligandos por RMN. A pesar de que por su naturaleza estos filtros implican una menor sensibilidad, son una alternativa para la detección de señales STD de ligandos en regiones del espectro que se encuentren ampliamente pobladas por señales de la macromolécula receptora. Por otro lado, los nuevos filtros propuestos son también muy recomendables para suprimir señales de macromoléculas en casos en que las señales de la molécula de pequeño tamaño tengan una relajación transversal muy eficaz y con ello se limiten las posibilidades de utilizar los mencionados filtros de resultado equivalente T_2 y $T_{1\rho}$.

La transformación del espectro 2D-STD- $T_{1\rho}$ -J-resolved-watergate en un espectro 1D-STD-homodesacoplado da como resultado un espectro monodimensional en el que, al colapsar la multiplicidad sobre el centro del desplazamiento químico del pico, se obtiene una menor cantidad de picos y por lo tanto proporciona una herramienta útil a la hora de abordar problemas de solapamiento de señales en experimentos STD.

4.6 Bibliografía

- Alonso-Plaza, J.M.; Canales, M.A.; Jiménez, M.; Roldán, J.L.; García-Herrero, A.; Hurrino, L.; Asensio, J.L.; Cañada, F.J.; Romero, A.; Siebert, H.-C.; André, S.; Solís, D.; Gabius, H.-J.; Jiménez-Barbero, J., “NMR investigations of protein–carbohydrate interactions: insights into the topology of the bound conformation of a lactose isomer and β -galactosyl xyloses to mistletoe lectin and galectin-1”, *Biochim. Biophys. Acta*, **2001**, *1568*, 225-236.
- André, S.; Kaltner, H.; Lensch, M.; Russwurm, R.; Siebert, H.-C.; Fallsehr, C.; Tajkhorshid, E.; Heck, A.J.R.; von Knebel Doeberitz, M.; Gabius, H.-J.; Kopitz, J., “Determination of structural and functional overlap/divergence of five proto-type galectins by analysis of the growth-regulatory interaction with ganglioside GM1 in silico and in vitro on human neuroblastoma cells”, *Int. J. Cancer*, **2005**, *114*, 46–57.
- Aue, W.P.; Karhan, J.; Ernst, R.R., “Homonuclear broad band decoupling and two-dimensional J- resolved NMR spectroscopy”, *J. Chem. Phys.*, **1976**, *64*, 4226-4227.
- Breg, J.; Romijn, D.; van Halbeek, H.; Vliegthart, J.F.G.; Visser, R.A.; Haasnoot, C.A.G., “Characterisation of four lactose monophosphates by application of ^{31}P -, ^{13}C -, and ^1H -NMR spectroscopy”, *Carbohydr. Res.*, **1988**, *174*, 23-36.
- Borlototti, V.; Fantazzini, P.; Gombia, M.; Greco, D.; Rinaldin, G.; Sykora, S., “PERFIDI filters to suppress and/or quantify relaxation time components in multi-component systems: An example for fat–water systems”, *J. Magn. Reson.*, **2010**, *206*, 219-226.
- Chen, H.-C.; Chen, S.-H., “Diffusion of Crown Ethers in Alcohols”, *J. Phys. Chem.*, **1984**, *88*, 5118-5121.
- Chen, A.; Shapiro, M.J., “NOE Pumping: A Novel NMR Technique for Identification of Compounds with Binding Affinity to Macromolecules”, *J. Am. Chem. Soc.*, **1998**, *120*, 10258-10259.
- Chen, A.; Shapiro, M.J., “NOE Pumping: 2. A High-Throughput Method To Determine Compounds with Binding Affinity to Macromolecules by NMR”, *J. Am. Chem. Soc.*, **2000**, *122*, 414.
- Claasen, B.; Axmann, M.; Meinecke, R.; Meyer, B., “Direct Observation of Ligand Binding to Membrane Proteins in Living Cells by a Saturation Transfer Double Difference (STDD) NMR Spectroscopy Method Shows a Significantly Higher Affinity of Integrin $\alpha_{\text{IIb}}\beta_3$ in Native Platelets than in Liposomes”, *J. Am. Chem. Soc.*, **2005**, *127*, 916-919.

- Cobas, J.C.; Martín-Pastor, M., “A homodecoupled diffusion experiment for the analysis of complex mixtures by NMR”, *J. Magn. Reson.*, **2004**, *171*, 20-24.
- Dalvit, C.; Pevarello, P.; Tatò, M.; Veronesi, M.; Vulpetti, A.; Sundström, M., “Identification of compounds with binding affinity to proteins via magnetization transfer from bulk water”, *J. Biomol. NMR*, **2000**, *18*, 65-68.
- Dalvit, C.; Fogliatto, G.P.; Stewart, A.; Veronesi, M.; Stockman, B., “WaterLOGSY as a method for primary NMR screening: Practical aspects and range of applicability”, *J. Biomol. NMR*, **2001**, *21*, 349-359.
- Dam, T.K.; Gabius, H.-J.; André, S.; Kaltner, H.; Lensch, M.; Brewer, C.F., “Galectins Bind to the Multivalent Glycoprotein Asialofetuin with Enhanced Affinities and a Gradient of Decreasing Binding Constants”, *Biochemistry*, **2005**, *44*, 12564-12571.
- Diercks, T.; Ribeiro, J.P.; Cañada, F.J.; André, S.; Jiménez-Barbero, J.; Gabius, H.-J.; “Fluorinated Carbohydrates as Lectin Ligands: Versatile Sensors in ¹⁹F-Detected Saturation Transfer Difference NMR Spectroscopy”, *Chem. Eur. J.*, **2009**, *15*, 5666-5668.
- Espinosa, P.J.; Garcia de la Torre, J., “Theoretical prediction of translational diffusion coefficients of small rigid molecules from their molecular geometry”, *J. Phys. Chem.*, **1987**, *91*, 3612-3616.
- Esturau, N.; Espinosa, J.F., “Optimization of diffusion-filtered NMR experiments for selective suppression of residual nondeuterated solvent and water signals from ¹H NMR spectra of organic compounds”, *J. Org. Chem.*, **2006**, *71*, 4103-4110.
- Fielding, L., “Determination of Association Constants (K_a) from Solution NMR Data”, *Tetrahedron*, **2000**, *56*, 6151-6170.
- Guenneau, F.; Mutzenhardt, P.; Grandclaude, D.; Canet, D., “Measurement of Longitudinal and Rotating Frame Relaxation Times through Fully *J*-Decoupled Homonuclear Spectra”, *J. Magn. Reson.*, **1999**, *140*, 250-258.
- Hajduk, P.J.; Olejniczak, E.T.; Fesik, S.W., “One-Dimensional Relaxation- and Diffusion-Edited NMR Methods for Screening Compounds That Bind to Macromolecules”, *J. Am. Chem. Soc.*, **1997**, *119*, 12257-12261.
- Johnson, C.S., “Diffusion ordered nuclear magnetic resonance spectroscopy: principles and applications”, *Progress in NMR Spectroscopy*, **1999**, *34*, 203-256.
- Johnson, M.A.; Pinto, B.M., “Saturation Transfer Difference 1D-TOCSY Experiments to Map the Topography of Oligosaccharides Recognized by a Monoclonal Antibody

Directed Against the Cell-Wall Polysaccharide of Group A Streptococcus”, *J. Am. Chem. Soc.*, **2002**, *124*, 15368-15374.

- Khera, S.; Grillo, M.; Schnier, P.; Hollis, S., “Application of diffusion-edited NMR spectroscopy for the structural characterization of drug metabolites in mixtures”, *J. Pharm. Biomed. Anal.*, **2010**, *51*, 164-169.

- Kopitz, J.; Russwurm, R.; Kaltner, H.; André, S.; Dotti, C.G.; Gabius, H.-J.; Abad-Rodriguez, J., “Hippocampal neurons and recombinant galectins as tools for systematic carbohydrate structure–function studies in neuronal differentiation”, *Dev. Brain. Res.*, **2004**, *153*, 189-196.

- Kowalewski, J.; Maler, L., “Nuclear Spin Relaxation in liquids: Theory, Experiments, and Applications”, Moore, J.H., Spencer, N.D., Eds.; CRC Press, Taylor & Francis Group: Boca Ratón, Florida **2006**.

- Leonidas, D.D.; Vatzaki, E.H.; Vorum, H.; Celis, J.E.; Madsen, P.; Acharya, K.R., “Structural basis for the recognition of carbohydrates by human galectin-7” *Biochemistry*, **1998**, *37*, 13930-13940.

- López-Lucendo, M.F.; Solís, D.; André, S.; Hirabayashi, J.; Kasai, K.-I.; Kaltner, H.; Gabius, H.-J.; Romero, A., “Growth-regulatory Human Galectin-1: Crystallographic Characterisation of the Structural Changes Induced by Single-site Mutations and their Impact on the Thermodynamics of Ligand Binding”, *J. Mol. Biol.*, **2004**, *343*, 957-970.

- Ma, L.; Philipp, E.; Led, J.J., “Determination of the electron self-exchange rates of blue copper proteins by super-WEFT NMR spectroscopy”, *J. Biomol. NMR*, **2001**, *19*, 199-208.

- Mari, S.; Serrano-Gomez, D.; Cañada, F.J.; Gorbi, A.L.; Jiménez-Barbero, J., “1D Saturation Transfer Difference NMR Experiments on Living Cells: The DC-SIGN/Oligomannose Interaction”, *Angew. Chem. Int. Ed.*, **2005**, *44*, 296-298.

- Matsushima, N.; G. Danno, et al. “Three-dimensional structure of maize alpha-zein proteins studied by small-angle X-ray scattering”, *Biochim. Biophys. Acta*, **1997**, *1339*(1), 14-22.

- Mayer, M.; Meyer, B., “Characterization of ligand binding by saturation transfer difference NMR spectroscopy”, *Angew. Chem. Int. Ed. Eng.*, **1999**, *38*, 1784-1788.

- Mayer, M.; Meyer, B., “Group Epitope Mapping by Saturation Transfer Difference NMR To Identify Segments of a Ligand in Direct Contact with a Protein Receptor”, *J. Am. Chem. Soc.*, **2001**, *123*, 6108-6117.

- Meiboom, S.; Gill, D., “Modified Spinecho method for measuring nuclear relaxation times”, *Rev. Sci. Instr.*, **1958**, *29*, 688-691.

- Meyer, B.; Peters, T., “NMR spectroscopy Techniques for Screening and identifying ligand binding to protein receptors”, *Angew. Chem., Int. Ed.*, **2003**, *42*, 864-890.
- Möller, H.; Serttas, N.; Paulsen, H.; Burchell, J.M.; Taylor-Papadimitriou, J.; Meyer, B., “NMR-based determination of the binding epitope and conformational analysis of MUC-1 glycopeptides and peptides bound to the breast cancer-selective monoclonal antibody SM3”, *Eur. J. Biochem.*, **2002**, *269*, 1444-1455.
- Morris, G.A., “Two-Dimensional J-Resolved Spectroscopy” in *Multidimensional NMR Methods for the Solution State*, G.A. Morris and J.W. Emsley Eds; EMR handbooks; Willey: Chichester (U.K.), **2010**; pp 145-160.
- Nie, L.; Wu, S.; Lin, X.; Zheng, L.; Rui, L., “Approximate Derivative Calculated by Using Continuous Wavelet Transform”, *J. Chem. Inf. Comput. Sci.*, **2002**, *42*, 274-283.
- Piotto, M.; Saudek, V.; Sklenář, V., “Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions”, *J. Biomol. NMR*, **1992**, *2*, 661-665.
- Platzer, N.; Davoust, D.; Lhermitte, M.; Bauvy, D.; Meyer, D.M.; Derappe, C., “Structural analysis of five lactose-containing oligosaccharides by improved, high-resolution, two-dimensional ¹H-NMR spectroscopy”, *Carbohydr. Res.*, **1989**, *191*, 191-207.
- Price, W.S.; “Pulsed-field gradient nuclear magnetic resonance as a tool for studying translational diffusion: Part II. Experimental aspects”, *Concepts in Nucl. Magn. Reson.*, **1998**, *10*, 197-237.
- Rastrelli, F.; Jha, S.; Mancin, F., “Seeing through Macromolecules: T₂-Filtered NMR for the Purity Assay of Functionalized Nanosystems and the Screening of Biofluids”, *J. Am. Chem. Soc.*, **2009**, *131*, 14222-14224.
- Solís, D.; Romero, A.; Kaltner, H.; Gabius, H.-J.; Diaz-Mauriño, T., “Different Architecture of the Combining Site of the Two Chicken Galectins Revealed by Chemical Mapping Studies with Synthetic Ligand Derivatives”, *J. Biol. Chem.*, **1996**, *271*, 12744-12748.
- Sousa, F.F.O.; Luzardo-Álvarez, A.; Blanco-Méndez, J.; Martín-Pastor, M., “NMR techniques in drug delivery: application to Zein protein complexes”, **2011**, en preparación.
- Strecker, G.; Wieruszkeski, J.M.; Michalski, J.C.; Montreuil, J., “Assignment of the ¹H- and ¹³C-NMR spectra of eight oligosaccharides of the lacto-N-tetraose and neotetraose series”, *Glycoconj. J.*, **1989**, *6*, 67-83.

-
- Subramanian, S.; Sampath, S., “Adsorption of zein on surfaces with controlled wettability and thermal stability of adsorbed zein films”, *Biomacromolecules*, **2007**, 8(7), 2120-2128.
 - Vogtherr, M.; Peters, T., “Application of NMR Based Binding Assays to Identify Key Hydroxy Groups for Intermolecular Recognition”, *J. Am. Chem. Soc.*, **2000**, 122, 6093-6099.
 - Wang, Y., Rakotonirainy, A.M.; Padua, G.W., “Thermal behavior of zein-based biodegradable films”, *Starch*, **2003**, 55, 25-29.
 - Wang, Y.-S.; Liu, D.; Wyss, F.D., “Competition STD NMR for the detection of high-affinity ligands and NMR-based screening”, *Magn. Reson. Chem.*, **2004**, 42, 485–489.
 - Wu, D.; Chen, A.; Johnson, C.S., “An improved diffusion-ordered spectroscopy experiment incorporating bipolar gradient pulses”, *J. Magn. Reson. ser A*, **1995**, 115, 260-264.
 - Wu, A.M.; Singh, T.; Liu, J.H.; Krzeminski, M.; Russwurm, R.; Siebert, H.C.; Bonvin, A.M.; André, S.; Gabius, H.J., “Activity-structure correlations in divergent lectin evolution: fine specificity of chicken galectin CG-14 and computational analysis of flexible ligand docking for CG-14 and the closely related CG-16.”, *Glycobiology*, **2007**, 17, 165-184.

Capítulo 5:
Métodos de clasificación automatizada para
espectros de RMN. Clasificación automatizada de
muestras de crudos basada en espectros 2D-HMBC

5.1 Introducción

Petróleo (también llamado petróleo crudo o sencillamente crudo) es el término usado para describir multitud de compuestos de origen natural, generalmente aunque no exclusivamente fluidos, formados por una mezcla heterogénea de compuestos orgánicos, principalmente hidrocarburos insolubles en agua, de origen fósil, que se acumula normalmente en reservorios subterráneos. Los crudos acumulados en diferentes depósitos presentan una gran variabilidad en todas sus propiedades fisicoquímicas, debido a las diferencias tanto en los compuestos que pueden presentar como de concentración de estos mismos compuestos que los conforman. De hecho, en un mismo reservorio la conformación de la mezcla suele variar en función de la profundidad. Además de hidrocarburos, en el petróleo pueden aparecer compuestos orgánicos con azufre, nitrógeno, oxígeno, así como trazas de elementos metálicos y otros elementos (Speight2002).

Las primeras clasificaciones de crudos basadas en los métodos fisicoquímicos tradicionales surgen del interés comercial del mismo, y su utilidad es la de servir como guía para los operarios de refinería, aportando unas nociones básicas de las condiciones necesarias para el procesado de cada crudo particular. Este sistema es válido a la hora de hacer una generalización de la calidad comercial de un crudo, tomando como referencia la inspección de alguna/s de estas propiedades. Sin embargo, plantea dificultades para hacer una clasificación rigurosa debido a la propia complejidad y variabilidad intrínseca a los crudos. Aunque existen muchas clasificaciones que emplean estos métodos, éstas son fundamentalmente válidas como descripciones de los diferentes tipos, y no resultan lo suficientemente precisas como para poder conformar un sistema de clasificación global. La química analítica ha intentado establecer clasificaciones atendiendo a una

única propiedad, como la gravedad API (ASTM D-287, ASTM D-1298, ASTM D-3142) o la viscosidad cinemática (ASTM D-445)*, con un éxito muy limitado. Se han llegado a establecer ciertas correlaciones entre diferentes propiedades físicas interrelacionadas para diversos crudos, pero mientras que propiedades como la viscosidad, la densidad, el punto de ebullición y el color pueden variar considerablemente de unos a otros, el análisis elemental de cada crudo muestra unas diferencias de composición mucho más pequeñas de lo esperable. Por lo tanto, parece que el problema de clasificar crudos estriba sobretodo en decidir qué propiedades son relevantes y cómo correlacionarlas con la composición del crudo, sumado a la necesidad de un método analítico que pueda detectar de forma cuantitativa las pequeñas variaciones en la composición (Speight2002). En las últimas décadas los métodos espectroscópicos han jugado un papel importante en la evaluación y clasificación de los crudos, y muchos de estos métodos se utilizan a día de hoy como estándares para el análisis de las materias primas y los productos del refinado.

La resonancia magnética nuclear ha demostrado ser una herramienta muy potente para el estudio y clasificación de mezclas complejas que pueden presentarse en disolución, estado sólido o semi-sólido. Algunas de sus características son especialmente apropiadas: cuantitatividad, alta reproducibilidad y robustez, además de que normalmente las muestras requieren poca manipulación para su preparación y las medidas pueden realizarse bajo condiciones experimentales relevantes (disolvente, pH, temperatura, o incluso dentro del propio tejido o medio celular en el caso de muestras biológicas). En los últimos años está en expansión el uso de la RMN en quimiometría (Winning2008). El análisis de mezclas de origen biológico por RMN y el interés por la

* Los códigos ASTM se refieren a estándares desarrollados por la American Society for Testing and Materials. Están disponibles online en la página web www.astm.org

metabolómica y metabonómica (Nicholson2008) son posiblemente las áreas que más han contribuido a asentar las bases de este tipo de RMN (Pelczer2003). Se han estudiado por RMN muestras de origen alimentario (Belton2003) o médico (Lee2006, Mayr2009), e incluso muestras complejas de la más diversa índole (Lindon2007). Los datos espectrales se analizan normalmente mediante métodos quimiométricos (Van der Greef2005, Trygg2007).

Los estudios de RMN de crudos se basan frecuentemente en espectros 1D-RMN de ^1H y ^{13}C . El área de las señales contenidas en determinadas regiones del espectro de RMN proporciona una caracterización general y cuantitativa de la composición media de una muestra (Saab2005, Molina2007) relativa al contenido del crudo en hidrocarburos, asfaltenos, resinas y otros compuestos presentes en los crudos. Se han establecido correlaciones entre la composición media determinada por RMN (espectro 1D de protón) con sus propiedades fisicoquímicas (Molina2007), demostrando así ser una poderosa herramienta para el problema de la clasificación y caracterización de los crudos. La aplicación de espectros 2D-RMN de correlación ^1H - ^{13}C como 2D-HMBC para este tipo de muestras no está muy extendida. Dado que el experimento 2D-HMBC establece correlaciones entre protones y carbonos conectados a través de dos o tres enlaces químicos en una molécula, y a pesar de su menor sensibilidad y mayor duración total de experimento frente a los espectros monodimensionales de protón y carbono, la mayor dispersión de señales que presenta puede ser de gran ayuda a la hora de revelar detalles estructurales relativos a los componentes mayoritarios en un crudo.

En este capítulo se presentan los resultados de clasificar 58 crudos extraídos de pozos de distintas características y procedencia geográfica a partir de sus correspondientes espectros 2D-HMBC ^1H - ^{13}C . Los espectros se clasificaron de forma

automática por métodos quimiométricos, asumiendo que espectros similares se corresponden con estructuras químicas similares.

Una parte de los cálculos realizados para la clasificación de las muestras de los crudos se realizó con dos herramientas de software que se han desarrollado en nuestro laboratorio (programas *similarity.pas* y *kmeans.pas*). Estas herramientas están basadas en métodos conocidos de clasificación aunque están especialmente adaptadas para analizar rápidamente datos provenientes de espectros de RMN mono y multi-dimensional. Estas herramientas pueden utilizarse para tratar problemas de similares características.

5.2 Teoría

5.2.1 Quimiometría y análisis de datos:

La quimiometría (Winning2008, Alam2004, Eads2004, Filzmoser2009, Siebert2001) es una ciencia multidisciplinar consistente en aplicar los principios de las matemáticas y la estadística multivariantes para extraer de manera eficaz el máximo de información relevante a partir de los datos obtenidos de un sistema químico. La quimiometría es útil tanto en la solución de problemas descriptivos como predictivos, siendo sobretodo de gran utilidad en química analítica y en metabolómica.

Las técnicas espectroscópicas (UV/visible, IR/Raman, RMN) y espectrométricas (espectrometría de masas) son capaces de aportar esta información química. Su análisis mediante métodos multivariantes ha sido históricamente una de las fuerzas impulsoras de la quimiometría, beneficiada además por el gran avance experimentado por la informática desde los años 70.

El análisis exploratorio de los datos multivariantes requiere muchas veces de técnicas y procedimientos que permitan reconocer patrones en los propios datos, a través de algún tipo de medida de similitud, que facilite la extracción de las “variables latentes” en los datos, de forma que toda la casuística de éstos pueda reducirse a un sistema de coordenadas más compacto que dependa de un conjunto pequeño de variables. Existen varios tipos de métodos de agrupamiento de datos (en inglés *clustering*) usados en quimiometría [refs. anteriores] que, *grosso modo*, pueden diferenciarse entre supervisados y no-supervisados. Por su popularidad en quimiometría, destacan entre los primeros el análisis de mínimos cuadrados parciales (PLS), y entre los segundos el análisis de componentes principales (PCA). La reducción

de las variables en el conjunto de los datos que puede conseguirse con estos métodos facilita su análisis posterior con técnicas numéricas como la regresión, la clusterización o el reconocimiento de patrones.

La validación cruzada es, dentro de las técnicas de regresión, una forma efectiva para decidir, por ejemplo, sobre el número de variables relevantes y la fiabilidad (error asociado) a un modelo de clasificación. El principio básico de la validación cruzada consiste en dejar fuera parte de los datos (conjunto de entrenamiento), construir un modelo con los datos restantes, y a continuación predecir con este modelo las muestras que faltan. La capacidad de predicción es el porcentaje de datos del conjunto de entrenamiento que son clasificados correctamente por el modelo. La validación cruzada procede repitiendo esta operación con diferentes conjuntos de entrenamiento. La forma idónea de elegir el conjunto de entrenamiento puede variarse en función de cuán grande sea el conjunto de muestras a validar, por ejemplo se pueden elegir subconjuntos al azar cuando el número de muestras es elevado, o bien, por ejemplo elegir sistemáticamente una a una cada muestra cuando el número de muestras es pequeño (método del objeto excluido). Existen diferentes métodos para realizar la validación cruzada que han sido revisados recientemente en su aplicación a resultados de PCA (Bro2008, Filzmoser2009).

A continuación se describe la teoría de los métodos quimiométricos que fueron utilizados en el desarrollo de este capítulo.

5.2.2 Etapa previa de normalización o calibración de los datos experimentales:

Se trata en general de operaciones horizontales o verticales en la tabla de datos, con objeto de permitir la comparación de las características relevantes de las muestras,

eliminando variaciones y/o errores sistemáticos en las medidas que son conocidas *a priori*. Es habitual de esta etapa la normalización o calibración de los datos experimentales respecto a algún patrón de referencia y/o para eliminar el efecto del diferente nivel de ruido que pueda ocurrir en las medidas.

5.2.3 Análisis de componentes principales (PCA):

El análisis de las componentes principales es una de las más antiguas y más utilizadas de las técnicas de análisis multivariante. El concepto fue ideado por primera vez por Karl Pearson en 1901 (Pearson1901) y desarrollado posteriormente por Harold Hotelling (Hotelling1933). La idea central del análisis de componentes principales es la reducción de la dimensionalidad inicial del conjunto de datos, en los que existen un alto número de variables interrelacionadas, en un número menor de variables no correlacionadas, las variables latentes o componentes principales, que son combinaciones lineales de las variables originales. El efecto del PCA sobre las variables puede entenderse como una rotación en el espacio de las variables que las lleve a maximizar su varianza total cuando éstas se representan respecto a una serie de ejes principales que apuntan en la dirección de la máxima varianza (Figura 5.1).

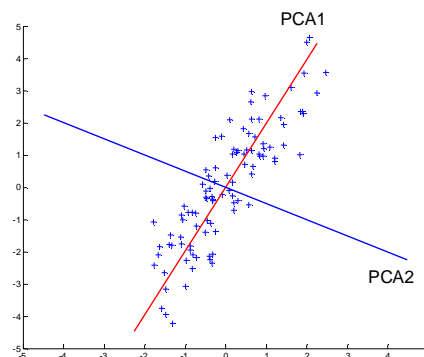


Figura 5.1: Determinación de los dos ejes principales PCA1 y PCA2 en un conjunto de datos. Estos ejes son ortogonales entre sí y apuntan en la dirección de máxima varianza en los datos.

El número total de componentes principales que se obtienen son tantos como variables tiene el conjunto de datos, y éstos describen el total de la variabilidad presente en los datos. Los componentes principales pueden ordenarse por relevancia de mayor a menor según la variabilidad que representan de los datos originales, para así poder elegir finalmente un número reducido de componentes que expliquen en mayor medida la parte determinista de la tabla de datos original, de forma que la información irrelevante más el ruido aleatorio queden eliminados.

La técnica extrae las causas de la variabilidad en los datos obtenidos de las muestras y los ordena, en forma de componentes principales, por importancia. La primera componente principal se define por el perfil espectral en el cual los datos describen la mayor parte de la varianza. La segunda componente es la segunda que describe una mayor cantidad de varianza, y es ortogonal a la primera, y así sucesivamente. Las sucesivas componentes describen cada vez menos varianza. Dado que la varianza sistemática es descrita principalmente por las primeras componentes, en quimiometría no suele ser demasiado crítica la elección del número de componentes, con tal que éstos den cuenta de un buen porcentaje de la varianza total (Winning2008). Por otro lado, el PCA es extremadamente sensible a las diferencias sutiles en los espectros, en concreto a la contribución del ruido espectral en la clusterización del PCA (Halouska2006).

Se han establecido algunos criterios objetivos para establecer el número ideal de componentes (Gonzalez1995, Geladi1986). Este número dependerá del objetivo buscado. Si únicamente se desea una caracterización global de la similitud o correlación entre muestras, bastará con representar gráficamente las primeras componentes y estudiar cómo se distribuyen los datos. Si lo que se desea es modelar los datos

originales eliminando contribuciones irrelevantes y el ruido, se requiere un mayor número de componentes principales y supervisar con un criterio estadístico y/o químico la decisión de dónde establecer el corte.

El PCA se reduce a la solución de un problema de autovalores-autovectores de una matriz simétrica semidefinida positiva. Supongamos que x es un vector de m variables aleatorias, y que las varianzas de las m variables aleatorias y la estructura de las covarianzas o de las correlaciones entre las m variables es de interés. Salvo en el caso de que el valor de m sea pequeño, o de que la estructura sea muy simple, lo más probable es que no sea de ninguna ayuda la observación directa de las varianzas de m y de todas las $\frac{1}{2}m(m - 1)$ correlaciones o covarianzas. Una alternativa a este problema es el intentar buscar unas pocas variables derivadas que contengan la mayoría de la información contenida en las variables y en sus correlaciones originales.

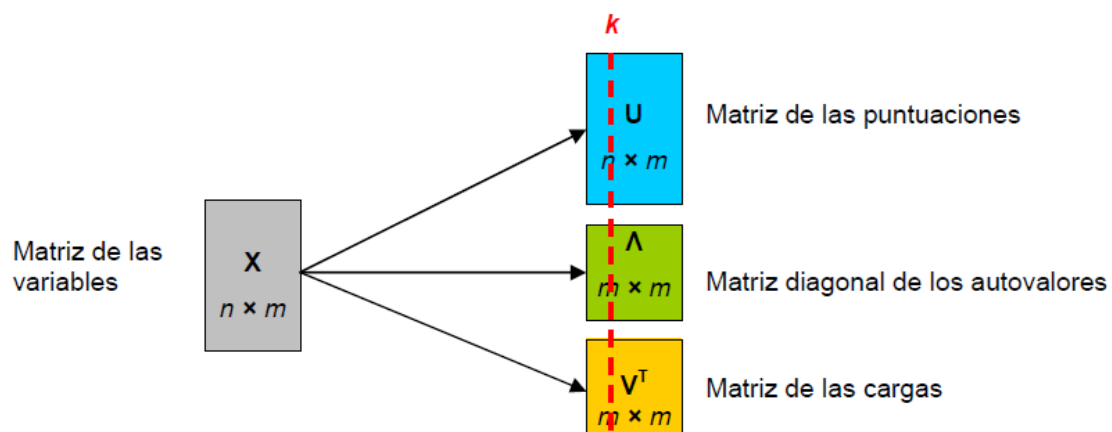


Figura 5.2: Esquema matricial de PCA (Referencia: José Ignacio Santos González, Tesis Doctoral, Figura 1.25 pag. 100)

Si creamos una matriz, denominada X , de dimensiones $n \times m$, donde cada fila de n representa un elemento del campo de estudio y cada columna m representa una variable medida para cada elemento, el PCA será una transformación lineal ortogonal

que trasladará los datos de la matriz \mathbf{X} a un nuevo sistema de coordenadas, con la mayor varianza de los datos situada sobre la primera coordenada, la segunda mayor varianza sobre la segunda y así sucesivamente. La transformación pasa por la descomposición de la matriz \mathbf{X} original en tres matrices: Una matriz \mathbf{U} , compuesta por los autovectores de \mathbf{X} , de dimensiones $n \times m$, llamada comúnmente matriz de puntuaciones y que será la que contenga las componentes principales, una matriz Λ diagonal, de dimensiones $m \times m$ compuesta en su diagonal por los autovalores de la matriz \mathbf{X} y cuya suma total será la varianza total de los datos iniciales y una matriz \mathbf{V}^T , llamada matriz de cargas, que representa los coeficientes de partición de las variables originales en los componentes principales. El interés del PCA es conseguir una reducción de la dimensionalidad de los datos, por lo que se tomarán las k componentes principales que expliquen la mayor cantidad de varianza de los datos.

Una explicación extensa del procedimiento matemático para calcular las diferentes componentes principales se puede encontrar en la bibliografía (Jolliffe2002)

5.2.4 Análisis de Clusters

El análisis del agrupamiento, en adelante análisis de clusters, (del inglés *cluster analysis*) (Jain1999) es un método no supervisado de clasificación de elementos en grupos, de forma que los elementos que se encuadran en el mismo cluster son similares entre ellos. El análisis de clusters es de amplia aplicación en diversas disciplinas, lo que lo coloca como uno de los pasos de gran utilidad cuando se realiza un análisis exploratorio de datos.

Se define como una técnica no supervisada pues se trata de establecer categorías o grupos (*clusters*) a partir de un conjunto de objetos no categorizados sin tener una

referencia *a priori* de cuáles son esas categorías. Esto es, se trata de encontrar una estructura oculta en los datos.

Normalmente, tras la recogida y normalización de datos, y un proceso opcional de reducción de variables por ejemplo a componentes principales mediante técnicas como PCA, el proceso de clusterización sigue los siguientes pasos (Jain1988):

1 Representación de los patrones

En este paso se trabaja sobre las variables disponibles para introducir en el algoritmo de clusterización. Se divide en dos partes, la selección de características, que será el proceso de identificar el mejor subgrupo de variables de todas las disponibles y la extracción de características, que consiste en el uso de las transformaciones necesarias de los datos de las variables que se usarán para adaptarlas al algoritmo de clusterización.

2 Definición de la medida de proximidad (similitud) de los objetos, apropiada a las características de los datos de las variables medidas en los objetos

Normalmente la similitud será la medida de la distancia definida entre parejas de datos análogos de una determinada variable. Existe una amplia gama de distancias útiles para la clusterización; por nombrar algunas distancia euclídea, distancia Manhattan, distancia Mahalanobis y distancia de Hamming.

Para facilitar la comparación de distancias cuando se considera más de una variable se puede utilizar alguno de los llamados índices de similitud, que reducen el conjunto de distancias a un solo número a través de algún tipo de normalización de

éstas. En esta Tesis se ha utilizado para el análisis de muestras de crudos el índice de similitud MCHORD (*modified CHORD*) (Boyce2003, Varmuza2003), que se describe más adelante.

3 Clusterizado o agrupado

El clusterizado de los objetos puede llevarse a cabo de diferentes maneras. Una división “dura” donde se reparten los objetos en diferentes clusters definidos, o “difusa” donde cada objeto tiene un porcentaje de pertenencia a cada cluster.

En función de la manera a la que se llega a la clusterización final, tenemos dos grandes aproximaciones, los algoritmos jerárquicos de clusterizado, que trabajan partiendo del conjunto de los objetos dispersos y los va agrupando en función de su proximidad, y los algoritmos de clusterizado particional, que hacen una búsqueda de la partición que optimiza el clusterizado, partiendo de un numero de clusters preasignado.

El algoritmo k-means (McQueen1967, Steinhaus1957, Lloyd1982) se encuadra dentro de los algoritmos de clusterizado particionales. Dada una clusterización inicial arbitraria de los datos, el algoritmo k-means actúa en una serie de iteraciones reordenando los clusters bajo las condiciones siguientes:

- i) minimizar la desviación intra-cluster con su centroide correspondiente
- ii) maximizar la desviación inter-clusters con los centroides de los demás clusters.

El algoritmo k-means es uno de los más sencillos entre los que utilizan el criterio del error cuadrático mínimo. Este criterio se puede definir como

$$e^2(\mathcal{X}, \mathcal{L}) = \sum_{j=1}^K \sum_{i=1}^{n_j} \|x_i^{(j)} - c_j\|^2 \quad \text{Ecuación 5.1}$$

para un clusterizado \mathcal{L} de un conjunto de objetos \mathcal{X} , con K clusters, donde $x_i^{(j)}$ es el elemento i que pertenece al cluster j y c_j es el centroide del cluster j .

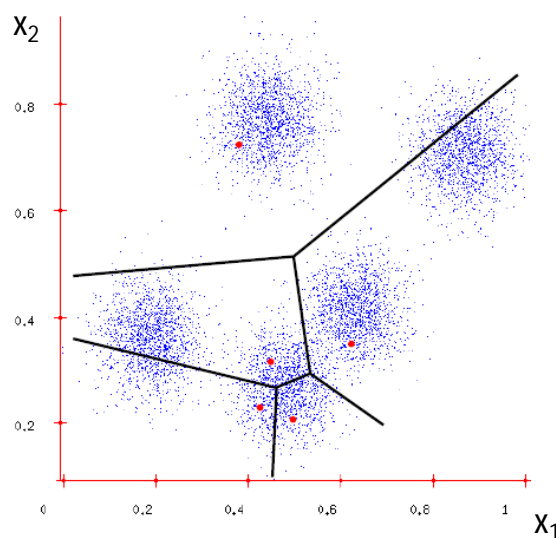


Figura 5.3: Ejemplo del funcionamiento de k-means para agrupar un número de muestras (en azul) en cinco clusters ($k = 5$). La pertenencia de cada muestra a un cluster se selecciona inicialmente al azar, o mejor, en función de su proximidad a unos centroides posicionados inicialmente al azar (en rojo). En cada iteración de k-means los objetos pueden cambiar de cluster para minimizar la función de error (Ecuación 5.1) y se recalculan los puntos centroides.

En cada iteración de k-means los objetos se asignan al cluster de cuyo centroide está más próximo. El algoritmo k-means comienza con una partición en clusters aleatoria, y realiza reparticiones de los objetos entre clusters en base a la similitud (menor distancia) entre los objetos y los centroides de los clusters, buscando minimizar el error cuadrático. En cada nueva iteración, se crea un nuevo conjunto de clusters (con nuevas posiciones de sus centroides) en base a la distribución de las muestras del paso

anterior, y se repite la repartición de todos los objetos en función de su distancia a los nuevos centroides creados (Figura 5.3).

Normalmente el algoritmo se detiene al encontrarse con un criterio de convergencia, como puede ser que no se produzca movimiento de objetos entre clusters en sucesivas iteraciones, o que no se produzcan cambios en el valor del error cuadrático tras un número determinado de iteraciones.

5.2.5 Descripción del software específico realizado:

Se detallan a continuación las funcionalidades incorporadas de los principales programas realizados. Se dan algunos detalles técnicos de su manejo así como de otros programas adicionales realizados en el apéndice.

-Programa similarity.pas: Implementación de un algoritmo para calcular índices de similitud entre muestras

Este programa permite realizar rápidamente el cálculo de índices de similitud entre un conjunto de regiones idénticas (Regiones de interés o ROI's) seleccionadas en espectros de RMN (mono o multidimensional). Incorpora los siguientes elementos:

1) Se pueden elegir entre diferentes esquemas para la normalización de los valores de integral de los buckets obtenidos de las regiones seleccionadas. Se pueden normalizar respecto a los siguientes criterios: i) División por el correspondiente nivel de ruido promedio del espectro para convertir las integrales en valores de relación SN (opción de normalización *average_noise*), ii) Normalización de la suma total del conjunto de integrales a un valor constante (*sum_of_intensities*), y iii) Normalización

mediante división por la integral con la máxima intensidad del espectro (*maximum_intensity*).

2) Se pueden elegir entre diversos criterios para el filtrado de los buckets. Uno de estos criterios es establecer un requerimiento mínimo de relación SN de la integral del bucket para ser incluido en el cálculo del índice de similitud. Esta opción permite eliminar la contribución al índice de similitud de buckets que sólo contienen ruido en todos los espectros. Un efecto adicional de descontar estos buckets sin señal para el índice de similitud es que éste se desplaza hacia valores más bajos en términos absolutos haciéndose más evidente las diferencias entre muestras. Un segundo criterio de filtrado que puede incluirse es establecer un límite superior en la integral, de forma que aquellos buckets que superen un cierto valor de SN se igualen a ese valor límite. Esta última opción puede ser útil para compensar un peso excesivo de estos buckets en el índice de similitud, especialmente para aquellos índices de similitud que no estén normalizados (por ej. EUC y RMSD).

3) El índice de similitud que se calcula para cada pareja de muestras puede elegirse entre alguno de los siguientes descritos en las referencias (Boyce2003, Varmuza2003, Li2006):

3.1 Índice MCHORD (*modified CHORD index*)

$$MCHORD^{AB} = 1 - \sqrt{1 - \frac{\sum_{j=1}^n SN_j^A \cdot SN_j^B}{\sqrt{\sum_{j=1}^n (SN_j^A)^2 \cdot \sum_{j=1}^n (SN_j^B)^2}}} \quad \text{Ecuación 5.2}$$

Los sumatorios se extienden para todos los valores de integral de los n buckets de las gradillas de integración aplicadas sobre las ROI (2194 buckets en este caso). SN_j^A y SN_j^B se refiere a los valores de integral normalizados de los buckets en los espectros

A y B, respectivamente. El índice de similitud MCHORD es un valor en el rango comprendido entre 1 y 0. Un valor de 1 implica la completa similitud, mientras que el 0 implica completa diferencia.

3.2 Índice EUC (*Squared Euclidean Cosine index*)

$$EUC = \frac{(\sum_j SN_j^A SN_j^B)^2}{\sum_i (SN_j^A)^2 \sum_i (SN_j^B)^2} \quad \text{Ecuación 5.3}$$

3.3 Índice TAN (*Tanimoto Index (binary data)*)

$$TAN = \sum_j \frac{Both SN_j^A SN_j^B}{only SN_j^A + only SN_j^B + both SN_j^A SN_j^B + neither SN_j^A SN_j^B} \quad \text{Ecuación 5.4}$$

El índice Tanimoto tiene en cuenta la presencia o no de valor de integral en un punto determinado en cada par de espectros considerado. No tiene en cuenta el valor de la integral, solamente la existencia o no de valor. El índice TAN es mayor al aumentar la coincidencia de integrales con valor diferente de cero entre ambos espectros.

3.4 Coeficiente PEA (*Pearson Correlation coef.*)

$$PEA = \frac{\sum_{j=1}^n (SN_j^A - \overline{SN_j^A})(SN_j^B - \overline{SN_j^B})}{\sqrt{\sum_{j=1}^n (SN_j^A - \overline{SN_j^A})^2} \sqrt{\sum_{j=1}^n (SN_j^B - \overline{SN_j^B})^2}} \quad \text{Ecuación 5.5}$$

El coeficiente de Pearson entre dos variables se define como la covarianza de ambas variables dividido por el producto de sus desviaciones estándar. El valor absoluto del coeficiente toma valores entre 0 y 1, de forma que +1 y -1 corresponden a dos variables totalmente correlacionadas y que se pueden situar linealmente. Cuando el valor del coeficiente es positivo indica que las variables aumentan o disminuyen

conjuntamente, mientras que el valor negativo indica que el comportamiento de una variable es el opuesto al de la otra. El valor de 0 indica la inexistencia de correlación entre ambas variables. Una propiedad fundamental del índice PEA es que es invariante frente a cambios de localización o de escala de las variables.

3.5 Coeficiente COR (Square Correlation coef.)

$$\mathbf{r} = \frac{\mathbf{z}_A^T \mathbf{z}_B}{\|\mathbf{z}_A\| \cdot \|\mathbf{z}_B\|} \quad \text{Ecuación 5.6}$$

El índice de similitud COR se basa en el coeficiente de correlación entre dos espectros calculado a partir de las medias centradas de las intensidades de ambos espectros a comparar, que se definen como

$$\mathbf{z}_A = \mathbf{SN}_j^A - \mathbf{1} \cdot \overline{\mathbf{SN}_j^A}$$

$$\mathbf{z}_B = \mathbf{SN}_j^B - \mathbf{1} \cdot \overline{\mathbf{SN}_j^B}$$

donde \mathbf{z}_A y \mathbf{z}_B son las medias centradas, x_A y x_B la intensidad de cada bucket y \bar{x}_A y \bar{x}_B la media aritmética de todos los valores de integrales, para los espectros A y B respectivamente. El índice COR toma valores entre 0 y 999, al hacer

$$\mathbf{COR} = \frac{999(\mathbf{r} + 1)}{2}$$

Implicando un mayor valor de COR mayor similitud entre espectros.

3.7 Índice RMSD (*Root Mean Square Index*)

$$\mathbf{RMSD} = \sqrt{\frac{\sum_{j=1}^n (SN_j^A - SN_j^B)^2}{n}} \quad \text{Ecuación 5.7}$$

El índice RMSD se usa comúnmente para la medida de diferencias entre valores predichos para una variable por un modelo o por una estimación, y los valores de esa misma variable tomados experimentalmente. El índice RMSD es un buen medidor de la precisión del ajuste entre dos conjuntos de datos o variables.

3.6 Índice NRMSD (*Normalized-Root Mean Square deviation index*)

$$\mathbf{NRMSD} = \frac{\sqrt{\frac{\sum_{j=1}^n (SN_j^A - SN_j^B)^2}{n}}}{SN_{max} - SN_{min}} \quad \text{Ecuación 5.8}$$

El índice NRMSD es igual que el anterior, pero normalizado. Suele expresarse como un percentil.

3.8 Índice BUB (*Baroni-Urbani & Buser Index*)

$$\mathbf{BUB} = \frac{a + \sqrt{a \cdot d}}{a + b + c + \sqrt{a \cdot d}} \quad \text{Ecuación 5.9}$$

El índice BUB se creó originalmente (Baroni-Urbani1976) para la comprobación de similitud entre datos binarios en clasificaciones taxonómicas Toma en cuenta 4 parámetros, a, b, c, y d que son, respectivamente, el número de atributos en común entre dos elementos (a), el numero de atributos presentes en el primer elemento pero no en el segundo (b), los presentes en el segundo elemento pero no en el primero (c) y los atributos no presentes en ninguno de los dos elementos comparados (pero si en el resto de elementos de la serie a comparar) (d). Originalmente se trata de un índice de

comparación binaria (Boyce2001), pero puede transformarse en un índice que tenga en cuenta cuantitativamente el atributo, y no sólo la presencia o ausencia del mismo. (Boyce2003, Tamás2001), quedando de la siguiente forma:

$$\frac{\sum_{i=1}^n \min(SN_i^A, SN_i^B) + \sqrt{(\sum_{i=1}^n \min(SN_i^A, SN_i^B))(\sum_{i=1}^n (\max_A(SN_i^A) - \max(SN_i^A, SN_i^B)))}}{\sum_{i=1}^n \max(SN_i^A, SN_i^B) + \sqrt{(\sum_{i=1}^n \min(SN_i^A, SN_i^B))(\sum_{i=1}^n (\max_A(SN_i^A) - \max(SN_i^A, SN_i^B)))}} \quad \text{Ecuación 5.10}$$

Un ejemplo del resultado del cálculo de índice de similitud se da en la Figura 5.4.

4) Permite estudiar lo robusto que es el valor de índice de similitud considerando la posible variación de éste por efecto de la relación señal ruido (SN) del espectro. El programa realiza una exploración sistemática de SN dentro de unos límites que pueden ser definidos por el usuario y proporciona tablas de índices de similitud correspondientes a la mayor y menor similitud posible.

5) Si previamente se introduce información sobre la clasificación de las muestras en clusters (conocida *a priori* o bien obtenida ésta por métodos como PCA o k-means). El programa calcula un parámetro denominado índice de discriminación global, GDI (*Global Discrimination Index*) que define el grado de relevancia de una determinada variable para lograr la clasificación deseada. Su funcionamiento se describe a continuación.

Dada una determinada agrupación de muestras en clusters, A, B, ..., Z, el cálculo de GDI procede a partir de la correspondiente matriz de confusión, CM (*Confussion Matrix*) (Ecuación 5.11)

$$CM = \begin{bmatrix} rmsd_{AA} & \cdots & rmsd_{AZ} \\ \vdots & \ddots & \vdots \\ rmsd_{ZA} & \cdots & rmsd_{ZZ} \end{bmatrix} \quad \text{Ecuación 5.11}$$

CM es matriz simétrica, sus términos son los índices RMSD (Ecuación 5.7) calculados con los datos de la variable entre la pareja de clusters elegida. Los términos diagonales de CM definen la variabilidad intra-cluster, mientras que los de fuera de la diagonal definen la variabilidad inter-cluster. El GDI de la variable en cuestión se calcula como un cociente entre la suma de términos fuera de la diagonal y suma de términos diagonales de la matriz CM por la siguiente ecuación (Ecuación 5.12).

$$GDI = \frac{\sum_{i=1}^n \sum_{j=1}^n \frac{CM_{ij}}{CM_{i,i}}}{\sum_{i=1}^n 1} \quad \text{Ecuación 5.12}$$

donde n es la dimensión de la matriz CM. El valor de GDI es útil para identificar las variables más relevantes para la clusterización propuesta. Un valor de GDI elevado significa que la variable es más relevante para esa clasificación, ya sea porque los datos intra-cluster son muy similares y/o los datos inter-cluster son muy diferentes. Lo contrario sucede para una variable con un valor GDI bajo. En la implementación realizada, el programa proporciona el GDI y la contribución a éste desglosada para cada cluster en particular.

En lo que se refiere a espectros de RMN, las variables consideradas para el GDI se refieren a los buckets en los que se ha integrado la/s señal/es del espectro. La Figura 5.4 muestra un ejemplo del gran impacto del GDI que consigue determinar los buckets del espectro de RMN que mejor distinguen dos tipos de muestra de vino, albariño y godello. Cuando no se establece ningún agrupamiento los índices de similitud son solamente un poco mayores para las muestras de un mismo tipo de vino (Figura 5.4a).

Cuando se introduce información sobre el agrupamiento se encuentran por GDI aquellos buckets del espectro que son más relevantes para diferenciarlos (Figura 5.4b).

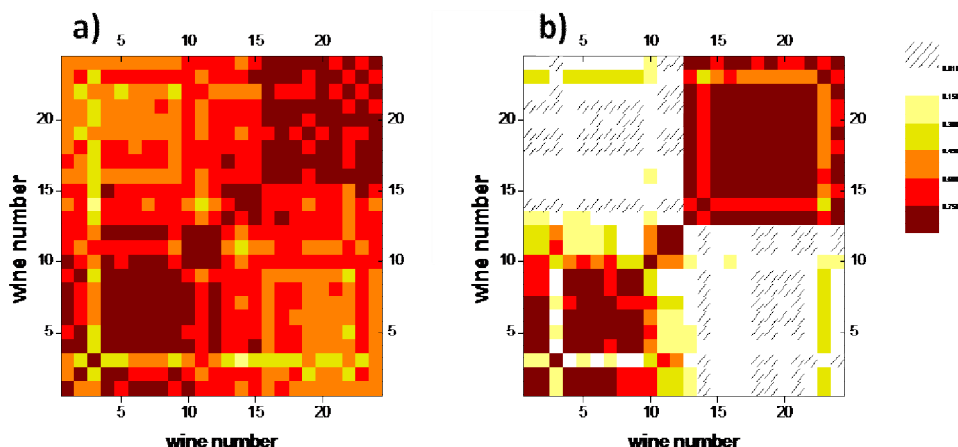


Figura 5.4: Representación de índices de similitud MCHORD calculados para espectros 1D ^1H watergate de 24 muestras de vino de dos tipos diferentes; tipo godello (muestras 1 a 12) y tipo albariño (muestras 13 a 24). a) Índices MCHORD con integración de 160 buckets que cubren toda la región espectral. No se ha definido ningún tipo de agrupamiento entre las muestras. b) Índices MCHORD para una selección de los 10 buckets (del total de 160) con mayor índice GDI cuando las muestras se clasifican previamente en dos clusters correspondientes a cada tipo de vino. En el código de colores utilizado más oscuro indica mayor índice de similitud y viceversa.

-Programa kmeans.pas: Implementación del algoritmo k-means para agrupamiento (clustering) automático de muestras definidas por coordenadas tridimensionales.

En la implementación del método k-means se ha utilizado el algoritmo de Lloyd (Lloyd1982), también llamado algoritmo k-means, debido a lo extendido de su uso, Existen variaciones, muchas de ellas enfocadas a la obtención de una buena clusterización inicial, que lleve hacia el mínimo global (Maitra, no publicado). Las modificaciones introducidas en la implementación realizada permitieron hacer frente a los siguientes problemas: i) dependencia de la partición inicial, ii) varias soluciones

compatibles, iii) eliminación de datos extremos (*outliers*). Estas modificaciones se describen a continuación.

i) *Dependencia de la partición inicial*: Uno de los problemas que pueden ocurrir con el algoritmo k-means convencional es que el resultado final puede depender mucho de la partición inicial (aleatoria) de los clusters, ya que ésta puede conducir hacia un mínimo local y arrojar un resultado no satisfactorio. Una forma habitual de evitar este problema consiste en repetir el cálculo k-means completo con particiones iniciales diferentes para encontrar la mejor solución. La bondad de una solución se puede definir a través de los valores normalizados promedio de la variabilidad intra- e inter-cluster (respectivamente parámetros *aveintrdev* y *aveinterdev* en Figura 5.5). En el programa realizado la partición inicial se generan de forma pseudo-aleatoria por medio de un algoritmo genético que realiza entrecruzamiento de soluciones anteriores del algoritmo k-means. El algoritmo genético controlado por algunos parámetros ajustables (por ej. hay una función de memoria del acervo genético de soluciones, un parámetro que controla el máximo número de entrecruzamientos fallidos entre éstas antes de introducir mutaciones, etc.) (Figura 5.6).

ii) *Varias soluciones compatibles*. En el algoritmo k-means convencional puede suceder que después del cálculo se llegue no a una solución única, sino a un número de soluciones diferentes pero estables y con un criterio de error muy similar. Una forma de discriminar entre las soluciones obtenidas consiste en medir la frecuencia de aparición de cada solución durante las diversas repeticiones aleatorias del cálculo para proporcionar finalmente como idóneas aquella/s partición/es que aparecen con mayor frecuencia. En el programa *kmeans* implementado además de la solución óptima, se obtiene como resultado todas las soluciones diferentes y su frecuencia de aparición.

iii) *Eliminación de datos extremos (outliers)*. Cuando en el algoritmo k-means se fuerza a todos los datos a pertenecer a alguno de los clusters entre los k seleccionados, existe la posibilidad de que existan datos extremos o anómalos que producen un gran aumento en la variabilidad intra-cluster e introducen confusión en cuanto a sus límites. En el algoritmo implementado se puede eliminar de la clusterización este tipo de datos anómalos, que son detectados por tener un bajo nivel de estabilidad en cuanto a su pertenencia a un determinado cluster. El nivel de estabilidad que se calcula para cada muestra es un valor entre 1 y 0 que corresponde respectivamente a una estabilidad del 100% y 0% entre el conjunto de soluciones obtenidas, respectivamente.

La siguiente figura muestra una salida típica del programa implementado *kmeans.pas*

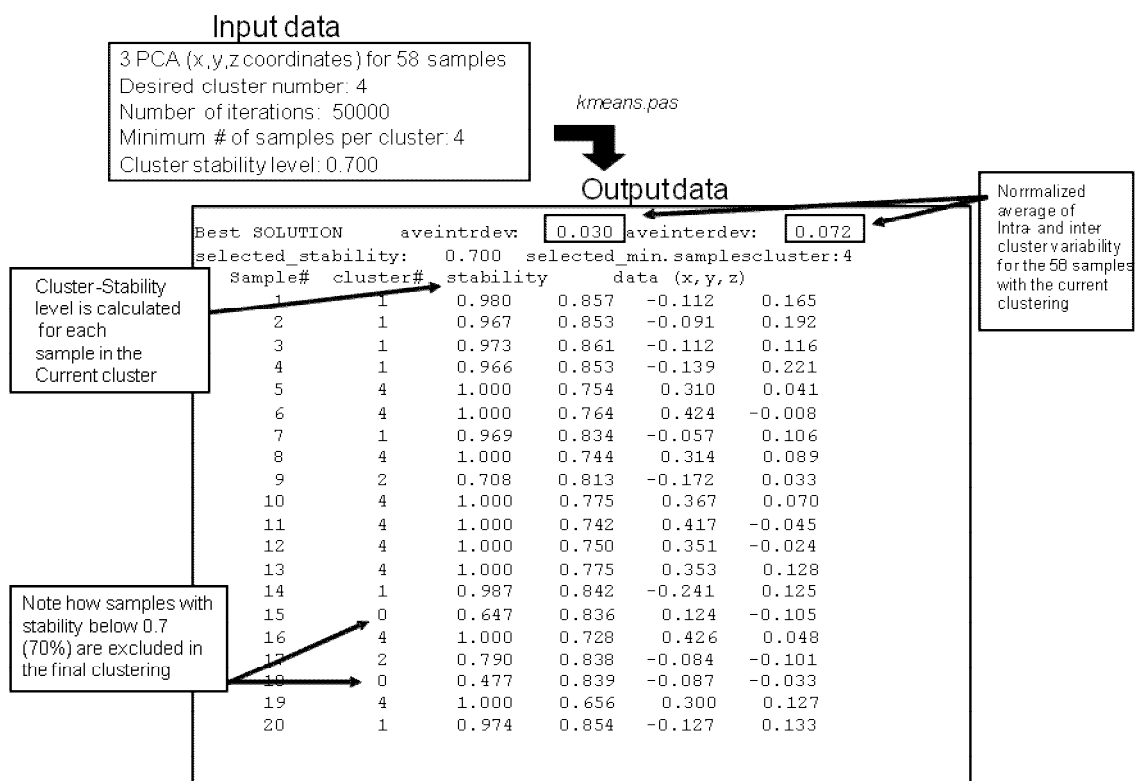


Figura 5.5: Extracto de una salida del programa *kmeans.pas* mostrando la clasificación obtenida para las 20 primeras muestras de crudo repartidos en 4 clusters. Los datos introducidos (x, y, z) fueron las coordenadas obtenidas en el cálculo PCA.

El funcionamiento del programa k-means implementado consta de 3 etapas

Etapa 1: El cálculo con el algoritmo k-means se repite durante un número de ciclos (por ej. $2 \cdot 10^5$ ciclos en el caso de las muestras de crudo). En cada ciclo se parte de una clusterización inicial generada por el algoritmo genético y se llega a una solución que es una clusterización optimizada de las muestras (mínimo local o global).

El algoritmo genético crea una clusterización inicial mediante el entrecruzamiento de alguna pareja de soluciones diferentes conseguidas en anteriores ciclos del algoritmo. Estas dos soluciones se toman de forma aleatoria de una reserva genética que contiene las mejores soluciones diferentes (por ej. 50 en el caso de las muestras de crudo) ya calculadas en anteriores iteraciones del algoritmo k-means. Inicialmente esta reserva de soluciones se genera de forma aleatoria y se va actualizando con las soluciones mejores y diferentes obtenidas después de cada ciclo del cálculo (Figura 5.6).

```
• begin
•   repeat
•     Inicializa partición de puntos aleatoria;
•     kmeans;
•     Evaluar partición y si procede su memorización;
•   until (criterio de terminación1)
•
•   repeat
•     Reproducción (memoria_particiones);
•     Entrecruzamiento(memoria_particiones);
•     Mutación (memoria_particiones);
•     kmeans;
•     Evaluar partición y si procede su memorización;
•   until (criterio de terminación2);
• end.
```

Figura 5.6: Pseudo-código del algoritmo genético implementado para el algoritmo k-means. El primer bucle repeat-until particiona al azar, aplica k-means, y memoriza las mejores particiones. El segundo bucle aplica el algoritmo genético con ellas.

Etapa 2: Se utiliza el método Monte-Carlo para calcular la estabilidad de cada muestra de crudo dentro del cluster al que se le ha asignado, para una clusterización de partida obtenida en el anterior paso. La estabilidad de cada muestra en su cluster se calcula como el porcentaje de veces que esa muestra es clasificada en ese cluster respecto del total de iteraciones utilizadas en el cálculo del algoritmo k-means.

Etapa 3: Se realiza una clusterización k-means con múltiples ciclos de algoritmo similar a la hecha en el paso 1, pero se desechan previamente las muestras que no obtengan un nivel mínimo de estabilidad en su cluster de destino, según lo calculado en el paso 2.

Al completarse las tres etapas se obtiene una o varias soluciones de clusterización y el nivel de estabilidad de cada muestra dentro del cluster que le asigne la solución.

5.3 Procedimiento experimental

Para este estudio se han utilizado 58 muestras de crudos de diferente procedencia. A lo largo del texto se hace referencia a cada crudo por su nombre propio o bien por el número identificativo que se da en Tabla 5.1. Las muestras fueron proporcionadas por M.M.C. Fernández-Feal, provenientes del Laboratorio de Combustibles, Centro Invest. Tecnológicas, Universidade da Coruña (campus de Esteiro).

5.3.1 Preparación de muestras:

Cada muestra de crudo se preparó disolviendo 0.1 ml de crudo extraído en 0.4 ml de CDCl_3 . La muestra fue traspasada a un tubo de RMN convencional convenientemente filtrada, para evitar la presencia de partículas sólidas.

5.3.2 Medidas de RMN:

Todos los espectros RMN fueron adquiridos a 300 K en un espectrómetro Bruker DRX-500 equipado con una sonda de detección inversa TXI $^1\text{H}/^{13}\text{C}/^{15}\text{N}$ equipada con gradientes apantallados en eje Z. Los espectros se procesaron con software MestRe-C (Cobas2003) o MestreNova (MestreLab Research).

Los espectros 2D-HMBC se adquirieron con la secuencia *inv4gplplrnd* de la librería de experimentos Bruker. El filtro J de paso bajo para la supresión de los acoplamientos a un enlace $^1\text{H}-^{13}\text{C}$ se especificó a un valor nominal de 145 Hz y se colocó un tiempo de espera de 85 ms para permitir la evolución de los acoplamientos de largo alcance de $^1\text{H}-^{13}\text{C}$. El número de puntos adquiridos en las dimensiones F2 y F1 fueron 1024 y 256 respectivamente y el número de adquisiciones por incremento de t_1

fue de 16. En el procesado de los espectros, el dominio del tiempo fue apodizado con una función seno y se aplicó un llenado de ceros en ambas dimensiones. A continuación se aplicó la transformada de Fourier en ambas dimensiones, obteniendo una matriz final de 2048x512 puntos reales en F2xF1. Cada espectro HMBC fue representado en magnitud y cuidadosamente referenciado en ambas dimensiones, a partir de la señal del CHCl₃ residual del CDCl₃ que aparece desdoblada en F2 por el acoplamiento ¹JCH. El conjunto de espectros obtenidos fue alineado visualmente para una mejor superposición.

5.3.3 Cálculos con espectros de RMN:

Integración en buckets de los espectros 2D HMBC

Cada uno de los 58 espectros 2D HMBC de los crudos se integraron de forma semi-automática con software MestRe-C en zonas del espectro que contenían señales. Se integraron un total de 2194 buckets con un tamaño de 64 y 295 Hz por buckets en F2 y F1 respectivamente.

Cálculo del índice de similitud entre espectros

Los cálculos de similitud entre los espectros 2D HMBC de los 58 crudos se realizaron con el software *similarity.pas* desarrollado para esta Tesis.

Previamente al cálculo del índice de similitud se utilizaron las siguientes opciones en el programa para la normalización y filtrado de los 2194 buckets obtenidos para cada muestra: i) se normalizaron los buckets para convertirlos desde unidades arbitrarias a su valor de señal sobre ruido (SN) (Craig2006). Esta normalización consiste en dividir los buckets de un espectro por el ruido medio del correspondiente

espectro. El ruido medio se calculó tomando el 10% de los valores de integrales de bucket con menor valor de integral (que se corresponden con buckets que no contienen ninguna señal) y ii) Se utilizó en el cálculo la opción de filtrado para eliminar automáticamente los buckets que sólo contienen ruido en todos los espectros; éstos se seleccionan como aquellos que no superan un valor mínimo requerido de SN en ninguno de los 58 espectros de crudos. Los tres valores de SN utilizados fueron 3, 5 y 10.

Tras las etapas de normalización y filtrado el programa calculó para cada pareja de espectros 2D-HMBC de los crudos el índice de similitud MCHORD (*modified* CHORD) (Boyce2003, Varmuza2003) (Ecuación 5.2) generándose con el conjunto de espectros una matriz de índices de similitud MCHORD de dimensiones 58 x 58.

Análisis de componentes principales (PCA) sobre la matriz de similitud

Los cálculos de PCA se llevaron a cabo con el software XLSTAT 5.1 (Addinsoft, <http://www.xlstat.com>). Los datos de entrada para el análisis PCA provienen de cada una de las matrices de similitud MCHORD obtenidas en el apartado anterior. Los factores de carga (*factor loadings*) obtenidos en el cálculo PCA se representaron con software Origin v8.0 (Originlab inc) a lo largo de los tres ejes principales de mayor variabilidad (PCA1, PCA2 y PCA3) que en todos los casos describen más de un 70% de la variabilidad total.

Análisis de clusters k-means sobre resultados PCA

El resultado de cada cálculo PCA del apartado anterior fue analizado con el programa *k-means* desarrollado para esta Tesis. Se realizó una clusterización automática

con la distribución de los factores de carga de las 58 muestras de crudo representados a lo largo de los tres ejes principales PCA1, PCA2 y PCA3.

Se realizaron una serie de cálculos k-means explorando el número de clusters desde 2 a 6 en pasos de 1. En cada cálculo el número mínimo de crudos para formar un cluster se estableció en 4 y la estabilidad mínima requerida para que una muestra forme parte de un cluster se estableció en $\geq 70\%$. El número de ciclos del algoritmo para cada una de sus 3 etapas fue de $2 \cdot 10^5$ donde cada ciclo lleva a una clusterización optimizada.

5.3.4 Cálculo de los buckets característicos de cluster:

Las muestras que se encuentran encuadradas dentro del mismo cluster en una solución válida obtenida presentan unos espectros 2D-HMBC similares. Es importante encontrar exactamente qué zonas/picos del espectro son las que presentan similar intensidad en todos los espectros, esto es, las que son características de ese cluster. Para su localización se procedió de la siguiente forma:

El índice de similitud MCHORD (Boyce2003) (Ecuación 5.2) se calculó para todas las parejas de buckets con $SN \geq 3$ entre los espectros de las muestras encuadradas en el mismo cluster, según están clasificadas en la Tabla 5.1. Un índice MCHORD medio se calculó para cada bucket por promediado simple de todos los índices obtenidos para cada bucket en todas las diferentes parejas de espectros consideradas.

Para la localización y representación gráfica de los buckets representativos de cada cluster, se clasificaron, de acuerdo con el índice MCHORD promedio calculado, como similitud muy alta (>0.9), alta (>0.7) o media (>0.5).

5.4 Resultados y discusión

Para una serie de 58 muestras de crudo, se adquirieron los espectros 2D-HMBC ^1H - ^{13}C con el fin de establecer una clasificación de esos crudos, basándose en la composición química mayoritaria de los mismos, y asumiendo que los crudos de similar composición presentarán espectros similares. Los crudos proceden de lugares de extracción diferentes y/o diferentes fechas de extracción. Además, algunas de las muestras fueron analizadas de nuevo un mes después de la primera preparación, para detectar posibles degradaciones en las muestras de crudos (Tabla 5.1).

Crude #	Crude name	Clustering			
		All SN	SN 3	SN 5	SN 10
1	IranPesado101671	1	1	1	1
2	UralLigero124214	1	1	-	-
3	SantaBarbara139570	1	1	-	1
4	SaharaBlend139658	-	-	-	-
5	Sharara143988	-	-	-	-
6	Maya149590	3	3	3	3
7	Maya200768200	3	3	3	3
8	Maya200773140	3	3	3	3
9	Sahara200736816*	-	3	2	3
10	Sahara	3	3	3	3
11	SaharaBL	3	3	3	3
12	Schiehallion	3	3	3	3
13	Sirtica200761202	3	3	3	3
14	SantaBarbara	-	-	-	-
15	Statford	3	3	3	3
16	Vasconia	3	3	3	3
17	Gulfaks	2	-	2	-
18	Gulfaks*	3	3	-	3
19	Norma	-	3	2	3
20	Sirtica200776501	3	3	-	3
21	Gulfaks200758984	-	-	-	-
22	Maya200773140*	3	3	3	3
23	Maya200768200	2	-	-	2
24	Maya200768200*	2	2	-	2
25	Maya200773140	2	2	-	2
26	Norma200780250	-	-	-	-
27	Sahara200736816	1	1	1	1

Crude #	Crude name	Clustering			
		All SN	SN 3	SN 5	SN 10
28	SaharaBL200721555				
29	SantaBarbara200739	1	1	1	1
30	Sirtica20076120				
31	Sirtica200776501*	1	1	1	1
32	Statford200775241	1	1	1	1
33	Vasconia200772945				
34	Vasconia200772945*				
35	Foroozan20070224	1	1	1	1
36	Gullfaks20070204	1	1	1	1
37	Maya20060801	2	2		2
38	Maya20060824	1		1	
39	Maya20060920	1		1	
40	Maya20061001	2		2	
41	Maya20061106	1		1	
42	Maya20061218	1		1	
43	Maya20070116	2	2	2	2
44	Maya20070127	2	2		2
45	Maya20070221	1		1	
46	Norne20060913	1	1		1
47	Norne20070314	1	1	1	1
48	Saharablend2006072	1	1	1	1
49	Saharablend2006120	1	1	1	1
50	SaharaBlend2007012				
51	SaharaBlend2008053	1	1	1	1
52	SantaBarbara200611	1	1	1	1
53	SantaBarbara200702	1	1	1	1
54	Sharara20061115	2		2	
55	Sirtica2007026				
56	Sirtica20060921	2		2	
57	Sirtica20061026				
58	Sirtica20070105				

Tabla 5.1: Clasificación de las 58 muestras de crudo basada en el análisis de los espectros 2D-HMBC con el método propuesto. Las muestras etiquetadas con un asterisco se repitieron aproximadamente un mes después de la primera medición. SN se refiere al nivel señal-ruido mínimo requerido para considerar en el análisis un bucket del espectro. Se elimina el bucket si sólo contiene ruido en todos los espectros. Las muestras clasificadas en cualquiera de los tres clusters tienen una estabilidad de cluster $\geq 70\%$. La columna etiquetada como All SN es una clusterización de consenso basada en la combinación de resultados de SN 3,5 y 10.

La sensibilidad del espectro HMBC depende del hardware del espectrómetro y de las condiciones experimentales utilizadas en la adquisición del espectro (número de

scans, calibración de pulsos, tiempo de espera entre *scans*, etc.). Se estimó que bajo las condiciones experimentales utilizadas, en los espectros HMBC de los crudos estudiados se puede detectar la presencia de sustancias a partir de una concentración mínima de aprox. 10 mM.

5.4.1 Integración en buckets:

Los espectros 2D HMBC de los crudos estudiados presentan una distribución de picos mayoritariamente concentrada en cinco regiones etiquetadas con números romanos I a V en la Figura 5.7. Estas regiones tienen las siguientes características:

- Región I: (^1H 6 a 9 ppm, ^{13}C 3 a 48 ppm) contiene los picos de correlación entre los protones aromáticos y los carbonos alifáticos u olefínicos.

- Región II (^1H 6.7 a 7.7 ppm, ^{13}C 69 a 106 ppm) contiene los picos de correlación entre los protones aromáticos y carbonos en grupos funcionales como alquinos o C-OR.

- Región III (^1H 5.2 a 9.1 ppm, ^{13}C 109 a 167 ppm) contiene picos de correlación entre protones aromáticos a carbonos aromáticos, heteroaromáticos o de dobles enlaces conjugados.

- Región IV (^1H 0.7 a 3.7 ppm, ^{13}C 1 a 48 ppm) contiene los picos de correlación entre los protones alifáticos con los carbonos alifáticos u olefínicos.

- Región V (^1H 0.9 a 3.4 ppm, ^{13}C 116 a 163 ppm) contiene picos de correlación entre protones alifáticos a carbonos heteroaromáticos o de dobles enlaces conjugados.

Estas regiones se dividieron en 2194 buckets con un tamaño de buckets homogéneo y más grande que la anchura de línea de los picos en ambas dimensiones. Esto último permite en buena medida contener dentro del mismo buckets los pequeños movimientos de los picos que podrían causar errores sistemáticos para la comparación posterior entre espectros (Powers2009 y sus referencias). Estos buckets del espectro HMBC constituyen la base para la clasificación de muestras mediante el protocolo que sigue a continuación.

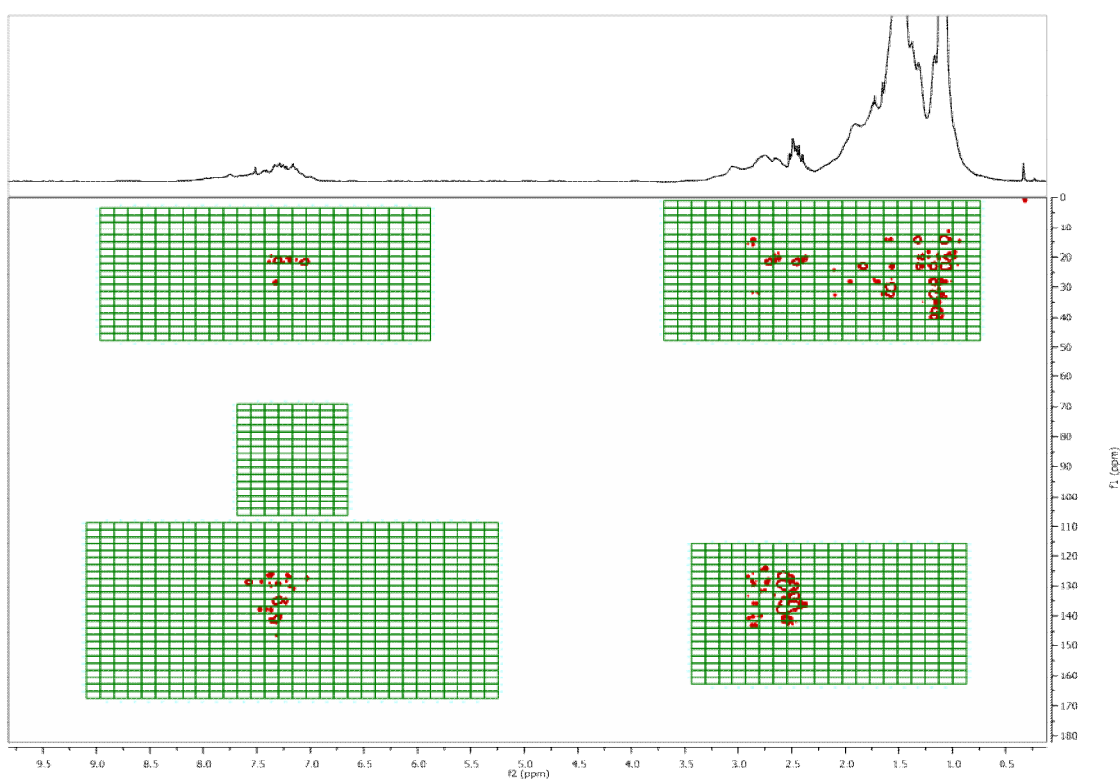


Figura 5.7: Espectro 2D-HMBC del crudo maya2000768200 (crudo #7 en el cluster 3, Tabla 5.1). Se muestran en verde las 5 regiones seleccionadas (de I a V) para la integración en gradillas compuestas por buckets, aplicadas sobre cada una de las 58 muestras de crudo. La parte superior muestra el espectro 1D-1H correspondiente a la muestra.

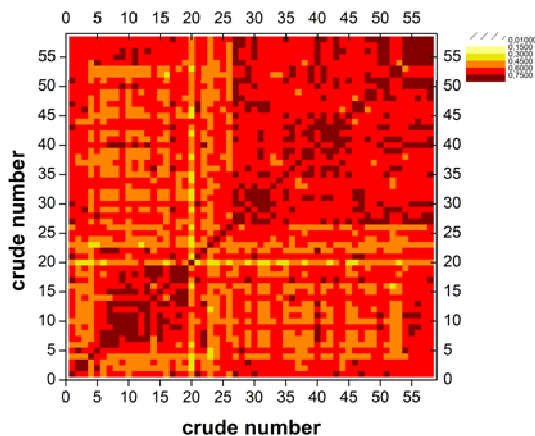
5.4.2 Cálculo del índice de similitud entre espectros:

La similitud entre pares de espectros 2D-HMBC de los crudos se puede determinar cuantitativamente en términos del índice de similitud (Boyce2003, Bodis2007 y sus referencias) que se obtiene de la comparación entre pares de los respectivos valores de integral normalizada correspondientes a cada cuadrícula de las 2196 obtenidas para cada espectro.

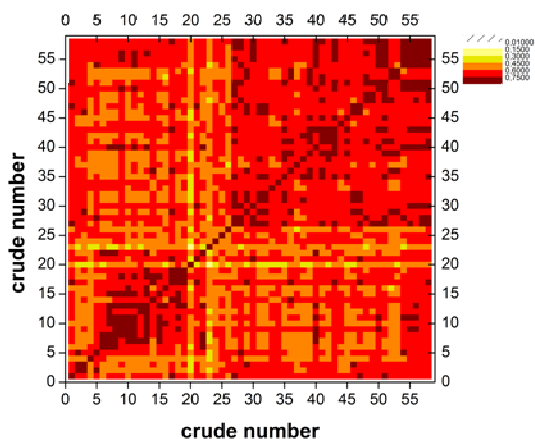
Entre los distintos índices de similitud el índice de similitud CHORD modificado (MCHORD) (Boyce2003) cuya forma funcional está expresada en la Ecuación 5.2 fue el que se utilizó para ese análisis debido a que proporciona un valor normalizado de la similitud entre 0 y 1, donde 0 significa diferencia total y 1 identidad completa, respectivamente. Se probó también con otros índices de similitud normalizados que también están descritos en la sección de teoría, como NRMSD (Ecuación 4.8) o el BUB (Ecuación 4.10) (Boyce2003) y se obtuvieron resultados muy comparables.

El índice de similitud MCHORD se calculó para todas las combinaciones por pares de los 58 espectros HMBC de las muestras de crudos. Un aspecto importante a considerar antes de proceder al cálculo de un índice de similitud es determinar si es necesario realizar una normalización previa y/o un filtrado del ruido los datos espectrales para que el resultado tenga relevancia. En el caso de datos de RMN hay que tener en cuenta que aún utilizando condiciones de preparación de muestra similares, hay aspectos que pueden influir en el nivel de ruido del espectro y/o alterar el valor de la integral en términos absolutos por ej. *shimming*, sintonía de la sonda, ganancia del receptor, número de *scans*, conductividad de la muestra, presencia de agentes paramagnéticos, etc. (Wider2006).

a) SN = 3



b) SN = 5



c) SN = 10

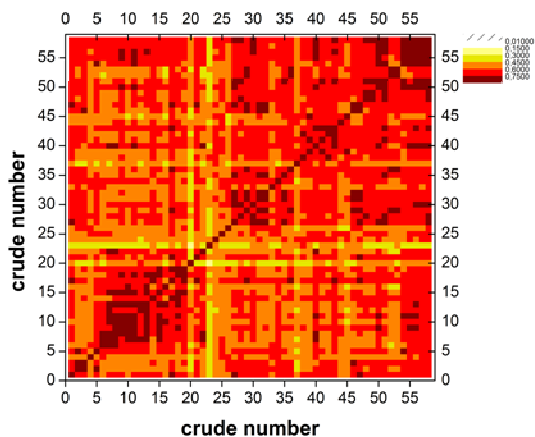


Figura 5.8: Matrices de similitud MCHORD obtenidas mediante comparación por parejas de buckets de los espectros 2D-HMBC de las 58 muestras de crudo del estudio. a) MCHORD-3 es la matriz calculada a un valor de corte mínimo de SN de 3, b) MCHORD-5 es el correspondiente a un valor de corte SN = 5 y c) MCHORD-10 es el correspondiente para SN = 10. En la escala de color usada, los puntos más oscuros corresponden a la alta similitud, mientras que los más claros indican las mayores diferencias. Las escalas de los ejes corresponden al número de referencia de los crudos de la Tabla 5.1.

Un método que se utilizó para reducir este efecto fue la normalizaron los valores de integral de los buckets de cada espectro dividiéndolos por su nivel de ruido, para convertirlos así desde unidades arbitrarias de integral a unidades comparables de relación señal sobre ruido (SN). Por otra parte, dado que los buckets se habían seleccionado de forma semi-automática como una gradilla (Figura 5.7) y muchos de ellos no contienen señales (Figura 5.7), para hacer que el índice de similitud MCHORD sea más sensible a las diferencias entre la pareja de espectros comparados, se utilizó la opción de filtrado que consiste en eliminar automáticamente del cálculo los buckets que sólo contienen ruido en todos los espectros; éstos se seleccionan como aquellos que no superan un valor mínimo requerido de SN en ninguno de los 58 espectros de crudos.

Con este tipo de normalización y filtrado se calculó el índice de similitud MCHORD para todas las combinaciones posibles de dos espectros entre los 58 espectros 2D-HMBC de los crudos. Esta información se puede expresar de forma más compacta como una matriz de índices de similitud MCHORD como las representadas en la Figura 5.8. Se trata de una matriz cuadrada, simétrica y donde cada elemento genérico de la forma (x, y) contiene el índice de similitud MCHORD calculado entre el par de espectros x e y . El tamaño de la matriz es en ambas dimensiones el número total de muestras presentes en el estudio (58 en nuestro caso) y los elementos de la diagonal son siempre 1, dado que este valor implica la comparación de los buckets de una muestra consigo mismo.

Las matrices de índices MCHORD de la Figura 5.8 permiten comparar la similitud de los 58 espectros HMBC de muestras de crudo para tres valores de corte de SN fijados en 3, 5 y 10 (MCHORD-3, MCHORD-5 y MCHORD-10 en Figura 5.8). En esta figura se pueden identificar visualmente grupos de muestras que presentan una alta

similitud. Puede verse que las muestras 54, 55, 56, 57 y 58 por un lado, y las muestras 8, 9, 10, 11 y 12 por otro lado, forman grupos compactos en cuanto a su similitud. En la Figura 5.8 también puede observarse el efecto de aumentar el valor SN de corte. La tendencia, como es de esperar, es que el índice de similitud disminuye en todas las muestras al aumentar el valor mínimo SN ya que se evita la introducción de muchos buckets sin señal en el cálculo del índice. Al pasar de un valor SN = 3 (Figura 5.8a) a SN = 5 (Figura 5.8b) las diferencias son mínimas, mientras que el efecto es más observable al situar el corte en SN = 10 (Figura 5.8c). En el caso de los dos grupos compactos mencionados antes, no parece que exista una disminución apreciable en su similitud.

5.4.3 Análisis de componentes principales (PCA) sobre la matriz de similitud:

Aunque visualmente se pudieron encontrar en la matriz dos grupos (clusters) de crudos con alta similitud, no es trivial encontrar qué muestras deben componer los grupos de similitud, excepto en que se tratase de pocas muestras y/o de índices de similitud muy extremos. Un método automático y eficaz para encontrar patrones es mediante el análisis de componentes principales (PCA) (Hotelling 1933, ver sección de teoría). PCA es muy usado en quimiometría para reducir o comprimir datos multidimensionales a un conjunto de datos con dimensiones más manejables para su análisis.

La matriz de índices de similitud calculada en el paso anterior (Figura 5.8) representa una reducción de variables importante de los datos originales de los espectros, concentra todas las variables en un solo número que refleja una característica de interés, la similitud entre muestras en este caso. La aplicación de PCA sobre la matriz de similitud es una alternativa que aporta estabilidad para observar diferencias

globales entre muestras, debido al filtrado de datos inherente al cálculo de este tipo de índice, y computacionalmente menos costosa frente al más habitual análisis PCA sobre el conjunto de variables originales (en el caso de los crudos la matriz de similitud tiene dimensiones 58 x 58 mientras que la matriz de variables es 58 x 2194). Una ventaja adicional es que la matriz de índices MCHORD es una matriz cuadrada simétrica, por lo que su diagonalización para el cálculo del PCA está asegurada.

Los cálculos de PCA se realizaron independientemente para cada una de las matrices de similitud MCHORD-3, MCHORD-5 y MCHORD-10 de la Figura 5.8. Los factores de carga (factor loadings) obtenidos en cada cálculo se representaron a lo largo de los tres ejes principales (PCA1, PCA2 y PC3) de mayor variabilidad (Figura 5.9). Cada punto en estas gráficas representa una muestra de crudo y la relación de proximidad entre los diferentes puntos refleja a su vez el grado de similitud entre sus respectivos espectros 2D-HMBC. Como observación, en las tres distribuciones de puntos de la Figura 5.9 no se observan diferencias remarcables, no parece que tenga mucho efecto el nivel SN utilizado en los cálculos.

En la siguiente sección se describe como establecer de forma automática una división óptima de las distribuciones de puntos de la Figura 5.9, para determinar el número de clusters que se debe considerar.

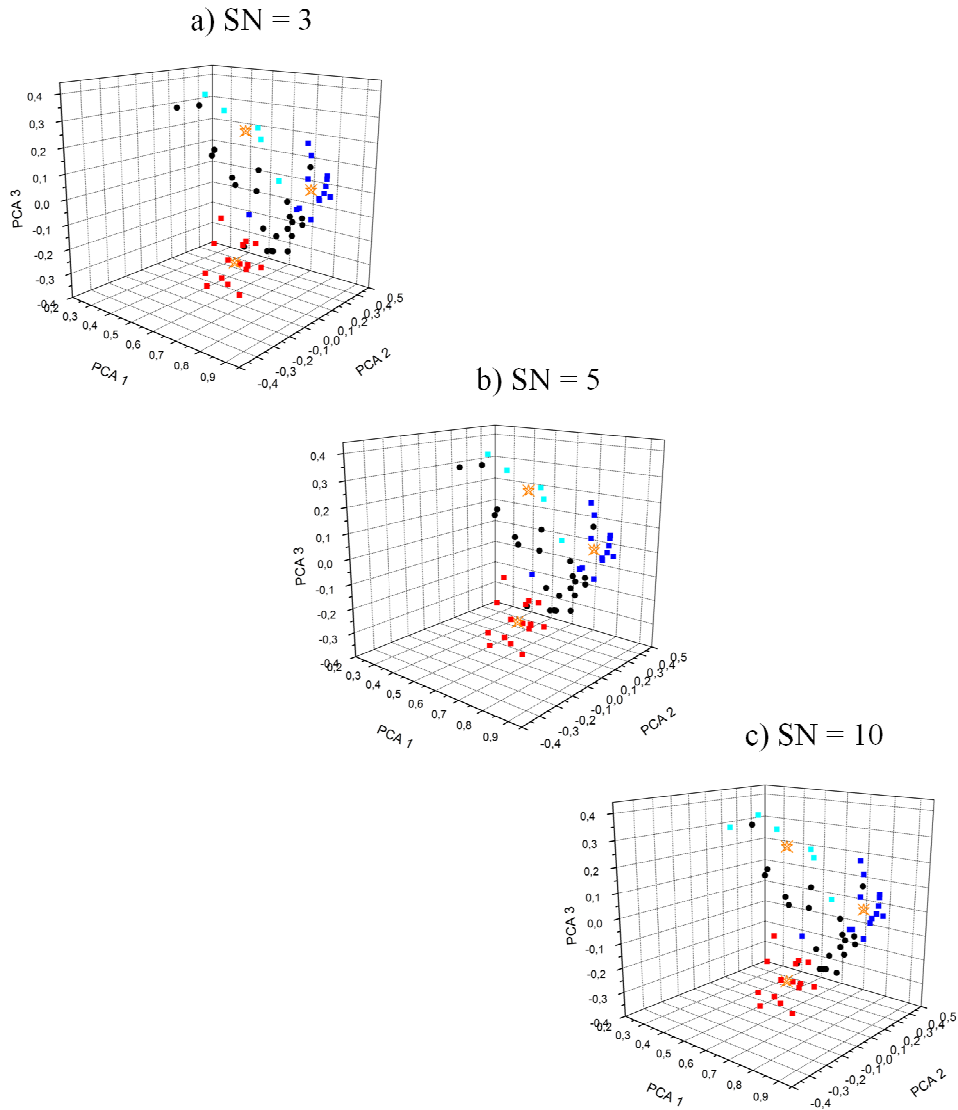


Figura 5.9: Cálculo de PCA sobre las matrices de índices MCHORD. Cada muestra se ha representado con un punto cuyas coordenadas son los 3 ejes principales, PC1, PC2 y PC3. a) PCA de la matriz MCHORD-3, b) PCA de la matriz MCHORD-5 y c) PCA de la matriz MCHORD-10. El color y los símbolos usados para los puntos representan el resultado posterior de la clusterización k-means. los puntos de color rojo, azul claro y azul oscuro representan crudos que pertenecen a los clusters 1, 2 y 3 respectivamente. Los círculos negros corresponden a crudos no clusterizados debido a que no cumplen la condición de estabilidad mínima impuesta. (70%). La X marca el centroide de cada cluster.

5.4.4 Análisis de clusters *k-means* guiado por PCA:

El algoritmo *k-means* (se describe en la sección de teoría) es una manera conveniente de realizar la clusterización de las muestras en el sub-espacio del PCA. Esta estrategia se conoce como *k-means* guiado por PCA (Ding2004). Se realizó una clusterización automática con el software realizado *kmeans* para cada distribución de los crudos de la Figura 5.9 que habían sido obtenidas en el anterior apartado mediante PCA.

Aunque *k-means* garantiza que se encuentre siempre una solución para la clusterización de datos arbitrarios. El software realizado permite una clusterización racional de las muestras, excluyendo aquellas que presentan una clusterización inestable, como puede ocurrir con datos que estén en las zonas limítrofes entre diferentes clusters. En los cálculos *k-means* se estableció un requerimiento de estabilidad mínima del 70% para cada muestra en un cluster.

Otro de los aspectos del algoritmo *k-means* es que se requiere elegir el número deseado de clusters para la clasificación. Esta información es inicialmente desconocida para el problema planteado del análisis de los crudos y por tanto debe ser computada a partir de los propios datos.

Una forma de determinar el número óptimo de clusters consiste en medir como varía la calidad de la clusterización a medida que se aumenta el número de clusters. La calidad de una clusterización viene dada por dos parámetros, la desviación media intra-cluster (*avintra*) y la desviación media inter-cluster (*avinter*), con ambos parámetros calculados respecto de los centroides de cada cluster. Estos parámetros se pueden combinar en un solo parámetro, denominado desviación global de cluster (*globdev*), que se obtiene de la siguiente ecuación (Ecuación 5.13):

$$globdev = |ave_{intra} - Min(ave_{intra})| + |ave_{inter} - Max(ave_{inter})| \text{ Ecuación 5.13}$$

Cuanto menor sea el valor de *globdev* mejor será la clusterización, ya que explicará mejor la varianza de los datos.

Para la distribución de PCA de las muestras de crudos de la Figura 5.9a se realizaron una serie de cálculos de k-means en los que se varió el número de clusters *k* desde 2 hasta 6 (con las condiciones de estabilidad impuestas a *kmeans* no se encuentra solución válida para valores de $k \geq 7$). Para cada cálculo el programa *kmeans* determinó el valor de *globdev* de la clusterización óptima y éste se representó frente al número de clusters (Figura 5.10).

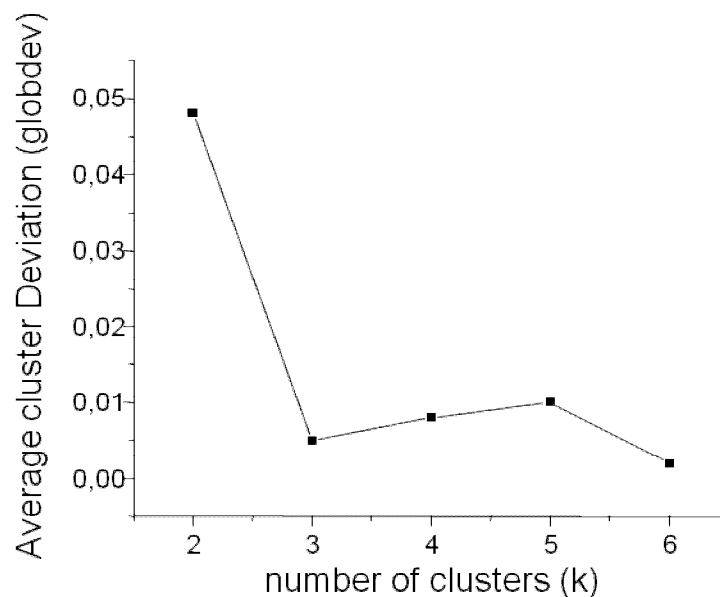


Figura 5.10: Gráfica de la desviación global media de cluster (*globdev*) frente al número de clusters para la distribución PCA de las 58 muestras de crudos de la Figura 5.9a, (SN = 3).

Se puede observar en la Figura 5.10 que *globdev* disminuye drásticamente cuando aumentamos el número de clusters *k* de 2 a 3. Para valores de *k* superiores a 3 *globdev* se ve afectado proporcionalmente en mucha menor medida, lo que implica que la clusterización mejora poco con el aumento de clusters por encima de 3. Por lo tanto,

un criterio de codo (Chen2007) para estos resultados sugiere que $k = 3$ es el número óptimo de clusters que se debe usar para la clasificación de los crudos.

Los resultados de la clusterización de las 58 muestras de crudo con la aplicación del k-means con $k = 3$ se recogen en la Tabla 5.1. Estos resultados se representaron además en las distribuciones de PCA de la Figura 5.9 con el uso de diferentes símbolos y colores.

5.4.5 Efecto del nivel SN escogido en la clasificación:

El resultado de la Tabla 5.1 indica que para un porcentaje importante de los crudos, la clasificación es suficientemente estable ($\geq 70\%$) y básicamente no se ve afectado por el nivel de SN mínimo utilizado en el análisis de las integrales de los espectros, en el rango 3 a 10. Hay en cambio alguna excepción que pertenecen a dos categorías: a) Muestras que pueden o no estar clusterizadas en función del límite SN establecido (*SantaBarbara139570*, *gulfaks*, *gulfaks** y otras) y b) muestras con suficiente estabilidad para estar siempre clusterizadas, pero no siempre en el mismo, en función el límite SN marcado. Solo hay dos muestras que presenten este comportamiento, *Norma* y *Sahara200736816** y ambas se encuentran muy cerca de la intersección de los clusters 2 y 3 en la distribución espacial que se puede observar en la Figura 5.9.

Para el primer caso, la razón de que algunas muestras no se puedan situar en clusters es por su baja estabilidad en cualquiera de ellos, por debajo del límite del 70% impuesto. Por lo tanto, la muestras en esta categoría pueden ser incluidas como parte de la clasificación en clusters con relativa seguridad, mientras que a las muestras de la categoría b) es más seguro tratarlas como no clasificadas. La aplicación de esta regla a

los clusters obtenidos en el estudio, y expuestos en la Tabla 5.1 a los niveles explorados de SN 3, 5 y 10, dan el clusterizado de consenso presentado en la columna "All SN", en esa misma Tabla.

La clasificación final resultante, reflejada en la Tabla 5.1, sugiere firmemente que hay tres tipos básicos de espectros 2D-HMBC en el conjunto de 58 muestras de crudo estudiadas, definidos estos tipos básicos por los hallados con la clusterización k-means. El espectro 2D-HMBC representativo de cada clase se muestra en la Figura 5.11.

La Figura 5.11 permite poner de manifiesto algunas diferencias cualitativas entre los espectros 2D-HMBC de las muestras de crudo clasificadas en los diferentes clusters. El Cluster 1, representado por el espectro 2D-HMBC de la Figura 5.11a, contiene una amplia distribución de picos en las regiones I y III (especialmente a partir de los 7.3 ppm en la dimensión de ^1H) y también en las regiones IV y V. Por otro lado, el Cluster 3 representado en la Figura 5.11c muestra menor cantidad de señales en esas mismas regiones.

El Cluster 2, representado por el espectro de la Figura 5.11b muestra una distribución de picos en las regiones comentadas que parecen un punto intermedio entre la Figura 5.11a y la Figura 5.11c.

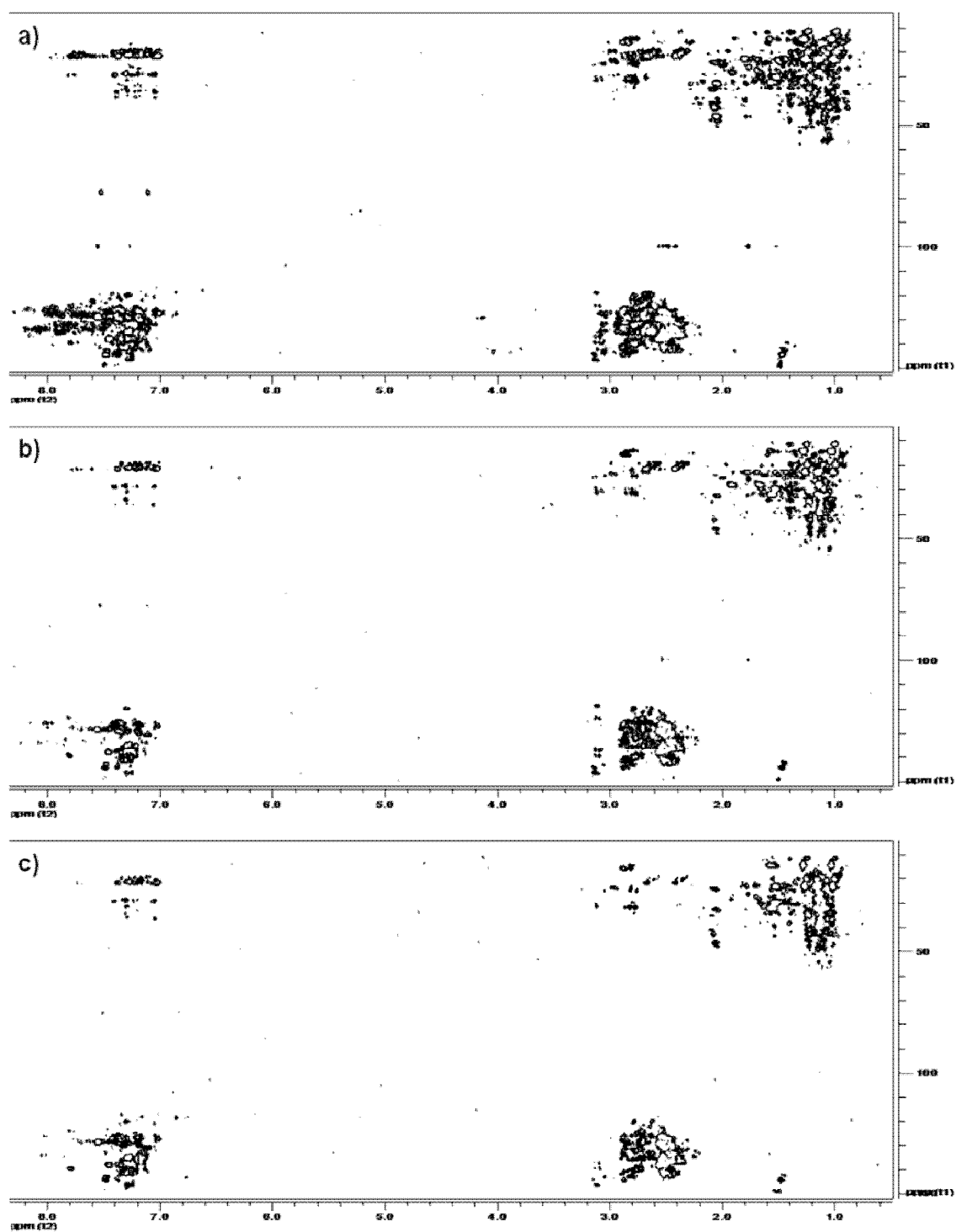


Figura 5.11: Los tres tipos básicos de espectros 2D-HMBC de las muestras de crudos de acuerdo al resultado obtenido por el análisis de clusters. a) Cluster 1 (IranPesado101671) b) Cluster 2 (Maya20070116) c) Cluster 3 (Maya149590). Los espectros están representados con el mismo nivel de contornos, ligeramente por encima del nivel del ruido.

5.4.6 Estudio de la evolución en el tiempo de las muestras de crudos:

Las muestras de crudo pueden contener componentes volátiles o ser propensas a la degradación, lo que puede suponer que exista una dependencia con el tiempo a sufrir cambios en su composición, y que pueden por tanto reflejarse en cambios en sus espectros HMBC. En ese caso, estos cambios espectrales podrían detectarse con la metodología aquí propuesta de clasificación de crudos en función de su espectro HMBC

Para comprobar la reproducibilidad a largo plazo de los espectros de las muestras de crudos, se analizaron de nuevo seis de las muestras de crudos, que habían sido almacenadas a temperatura ambiente, un mes después de haber adquirido el resto de los espectros HMBC. Estas seis muestras se escogieron para representar todos los casos diferentes de tipos de muestra que habían sido encontrados (ver Tabla 5.1). Los resultados del re-análisis a través de estos nuevos espectros se incluyen también en la Tabla 5.1 con el nombre de la muestra original, seguido de un asterisco (*).

Los resultados expuestos en la Tabla 5.1 muestran que cuatro de las seis muestras re-analizadas han sufrido un cambio por motivo de su degradación y visualmente ya apreciable en el espectro, que les lleva a situarse en otro cluster diferente al original en la clasificación.

5.4.7 Buckets exclusivos representativos de los clusters en los espectros 2D-HMBC de los crudos:

Una vez que establecidas las diferencias globales que presentan los espectros característicos de cada una de las clases de los crudos obtenidas mediante su clasificación en clusters, se puede estudiar la existencia (o la inexistencia) de ciertos buckets (que podrían identificarse con picos) en los espectros que puedan aparecer con

similar intensidad en todas las muestras pertenecientes a un mismo cluster. De existir, se catalogarán como buckets representativos.

Para los tres tipos básicos de espectros HMBC de las muestras de crudos obtenidos en el análisis de crudos que se dan en la Tabla 5.1, los buckets representativos se calcularon como aquéllos que tienen un índice de similitud MCHORD alto y con una desviación estándar representativa. Este cálculo se llevó a cabo para cada uno de los 2194 buckets analizados en cada espectro, y se realizó teniendo en cuenta sólo los espectros de los crudos encuadrados dentro del mismo cluster.

Los buckets representativos obtenidos para los tres clusters se representan en la Figura 5.12. La distribución de estos buckets representativos recuerda a la distribución de picos vista en la Figura 5.11. Para el Cluster 1 (Figura 5.12a), los buckets representativos se encuentran predominantemente en la Región IV. Para el Cluster 3 (Figura 5.12c), la zona con mayor presencia de buckets representativos es la Región III. En el caso del Cluster 2 la cantidad de buckets representativos (Figura 5.12b) encontrada es mucho menor que en los casos anteriores, y no se encuentran con predominancia en ninguna Región concreta.

La presencia de buckets representativos en una ROI determinada en mayor medida que en cualquier otra implica que esa ROI presenta mayor similitud entre los espectros de un mismo Cluster. En el Cluster 2, por lo tanto, no parece que ninguna zona destaque como representativa del grupo (en términos de similitud)

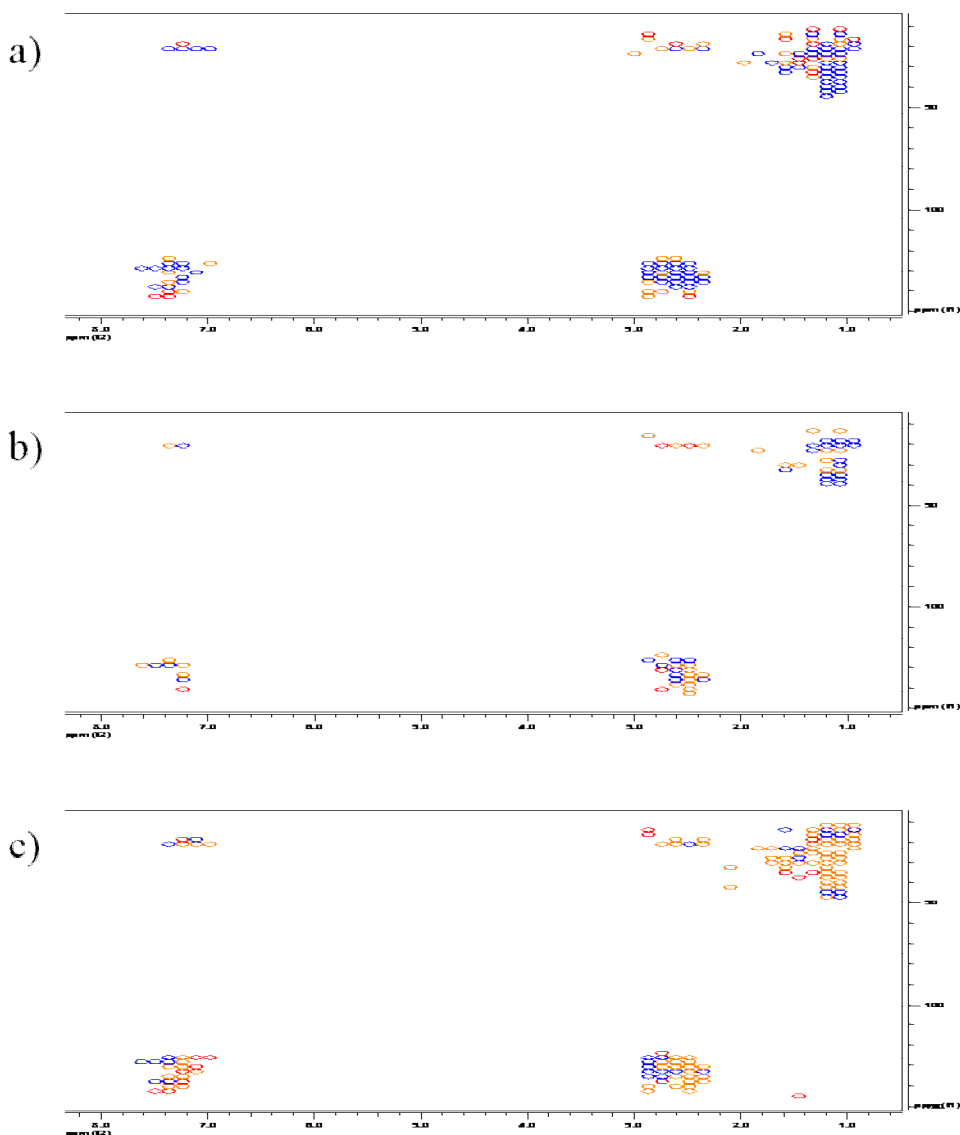


Figura 5.12: Buckets representivos de cluster en los espectros HMBC de las muestras de crudo de cada cluster. a) Cluster 1, b) Cluster 2, c) Cluster 3. Los buckets representados usan el siguiente código de colores: rojo ($1 \geq$ índice MCHORD medio ≥ 0.9), naranja ($0.9 >$ índice MCHORD medio ≥ 0.7) y azul ($0.7 >$ índice MCHORD medio ≥ 0.5).

En estudios de metabonomía por RMN en los que se conoce *a priori* el número de grupos a clasificar (por ej. grupo experimental y grupo control, o también 3 grupos de crudos en el caso que nos ocupa) y se dispone de datos espectrales de cada uno de ellos, es habitual querer reducir el problema de la clasificación a un pequeño número de buckets que sean los más similares dentro de su grupo y a la vez lo más diferentes posibles de los demás grupos. Este cálculo se aplicó para las muestras de crudo. Los

resultados se muestran en las Figuras 5.13 y 5.14. Los buckets representados en la Figura 5.14 son los que mejor permiten clasificar una muestra de crudo que pertenezca a uno de los 3 clusters.

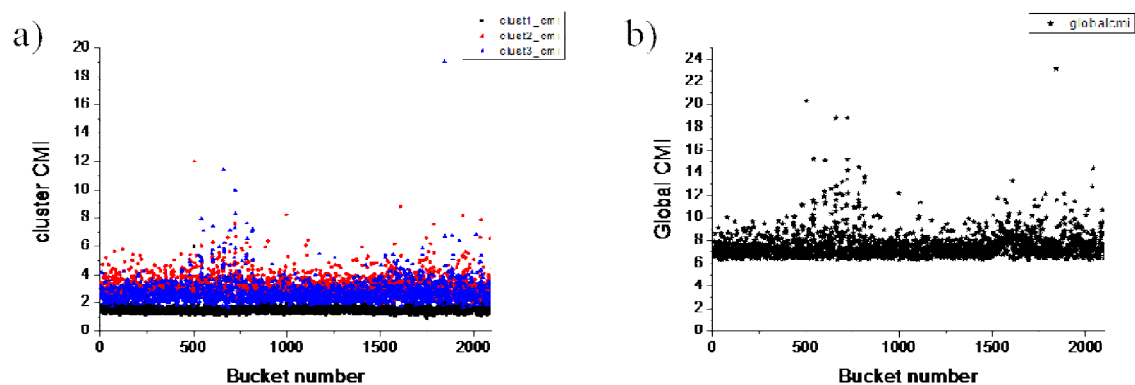


Figura 5.13: Representación de los buckets respecto a su valor de CMI para los 3 clusters encontrados en las muestras de crudo. a) Valor de CMI independiente para clusters 1, 2 y 3. b) Valor global de cmi de los 3 clusters.

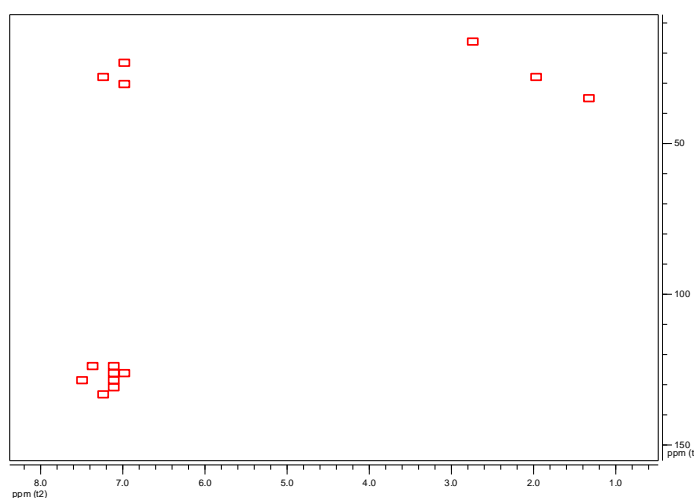


Figura 5.14: Buckets de integración del espectro HMBC más relevantes para diferenciar entre los 3 tipos de muestra de crudo encontradas. Los 14 buckets representados seleccionados entre los 2194 buckets considerados corresponden a valores de $\text{global_CMI} \geq 13$ (ver Figura 5.13b).

5.5 Conclusiones

Se ha desarrollado una metodología para la clasificación no supervisada de conjuntos de espectros de RMN (1D ó nD) basado en el análisis cuantitativo de los espectros. El método se desarrolla a lo largo de las siguientes etapas: i) procesado de espectros e integración en cuadrícula, ii) cálculo de la similitud espectral, mediante el índice de similitud MCHORD, iii) análisis de componentes principales (PCA), y iv) análisis de clusters por medio del algoritmo k-means en el subespacio definido por el resultado del PCA. En principio, la metodología aquí expuesta es aplicable al estudio de cualquier tipo de mezcla, simple o compleja, que esté sujeta a variabilidad, del tipo de muestras que pueden encontrarse en metabonomía, proteómica, industria alimentaria o en el control de calidad de medicamentos.

Se ha aplicado la metodología para el estudio y clasificación de un conjunto de muestras de crudo en base a su espectro 2D-HMBC. Se ha escogido este experimento por su sensibilidad frente a detalles estructurales de los componentes más abundantes de la mezcla. El resultado de la clusterización establece la existencia, en base al criterio de similitud, de tres tipos generales de crudos, e identifica un grupo de muestras de crudos que no pueden ser clasificados en ninguno de estos grupos y que no conforman un grupo por sí mismas tampoco. Es posible identificar, dentro de cada cluster, una serie de picos característicos, que definen detalles de la composición, que presentan alto grado de similitud en todos los espectros de las muestras de crudo encuadradas en ese grupo. Es posible igualmente encontrar un número reducido de picos que mejor permiten clasificar una muestra que pertenezca a uno de los clusters encontrados.

La metodología de clasificación propuesta en este capítulo no se restringe sólo a datos de RMN sino que puede aplicarse igualmente a datos de otras técnicas de análisis

químico. Por ejemplo recientemente se ha probado con éxito en nuestro laboratorio para la clasificación de muestras mediante datos de ICP-MS (espectrometría de masas con plasma acoplado inductivamente). Esta metodología ofrece una fuente de conocimiento extra, que en combinación con otras propiedades fisicoquímicas conocidas de las muestras puede ser de valor en la optimización del procesado industrial del crudo. Por ejemplo, se puede maximizar el rendimiento de productos de alto valor en el refinado de un crudo mediante la elección acertada del crudo o mezcla de crudos a procesar. Otra posibilidad de la metodología propuesta es la fácil obtención de información acerca de la biodegradación de un crudo por microorganismos u otros agentes en el tiempo, a partir de una muestra inicial conocida.

5.6 Referencias

- Alam, T.M.; Alam, M.K., “Chemometric analysis of NMR spectroscopy data: a review”, *Ann. Rep. NMR Spectrosc.*, **2004**, *54*, 41-80.
- Belton, P.S., “Magnetic resonance in food science: latest developments”, Gil, A.M.; Webb, G.A.; Rutledge D., Eds.; The Royal Society of Chemistry: Cambridge, UK, **2003**.
- Bodis, L.; Ross, A.; Pretsch, E., “A novel spectra similarity measure”, *Chemometr Intell Lab Syst*, **2007**, *85*, 1-8.
- Boyce, R.L.; Ellison, P.C., “Choosing the best similarity index when performing fuzzy set ordination on binary data”, *J. Veg. Sci.*, **2001**, *12*, 711-720.
- Boyce, R.L., “Choosing the Best Similarity Index when Performing Fuzzy Set Ordination on Abundance Data”, Poster #12 en el ESA 2003 Annual Meeting.
- Bro, R.; Kjeldahl, K.; Smilde, A.K.; Kiers, H.A.L., “Cross-validation of component models: A critical look at current methods”, *Anal. Bioanal. Chem.*, **2008**, *390*, 1241–1251.
- Chen, F.; El-Dawlatly, S.; Jin, R.; Oweiss, K., “Identifying and Tracking the number of independent clusters of functionally interdependent neurons from biophysical models of population activity”, En *Neural Engineering, 2007. CNE '07. 3rd International IEEE/EMBS Conference on; USA*, **2007**. pp. 542-545.
- Cobas, J.C.; Sardina, F.J., “Nuclear magnetic resonance data processing. MestRe-C: A software package for desktop computers”, *Concepts in Magn. Reson.*, **2003**, *19A*, 80-96.
- Craig, A.; Cloarec, O.; Holmes, E.; Nicholson, J.K.; Lindon, J.C., “Scaling and Normalization Effects in NMR Spectroscopic Metabonomic Data Sets”, *Anal. Chem.*, **2006**, *78*, 2262-2267.
- Ding, C.; He, X., “K-means Clustering via Principal Component Analysis”, En *Proceedings of the 21st International Conference on Machine Learning; ACM New York: New York*, **2004**.
- Eads, C.D.; Furnish, C.M.; Noda, I.; Juhlin, K.D.; Cooper, D.A.; Morrall, S.W., “Molecular factor analysis applied to collections of NMR spectra”, *Anal. Chem.*, **2004**, *76*, 1982-1990.
- Filzmoser, P.; Serneels, S.; Maronna, R.; Van Espen, P.J., “Robust multivariate methods in chemometrics”, en *Comprehensive Chemometrics: Chemical and*

Biochemical Data Analysis, Walczak, B.; Tauler, R.; Brown, S., Eds; Elsevier Science: Oxford, UK, **2009**; vol 3 pp. 681-722.

- Geladi P.; Kowalski, B., "Partial least-squares regression: a tutorial", *Anal. Chim. Acta*, **1986**, *185*, 1-17.

- González, A.G.; González-Arjona, D., "Statistical assessment of a new criterion for selecting the number of factors in factor analysis", *Anal. Chim. Acta*, **1995**, *314*, 251-252.

- Halouska, S.; Powers, R., "Negative impact of noise on the principal component analysis of NMR data", *J. Magn. Reson.*, **2006**, *178*, 88-95.

- Hotelling, H., "Analysis of a complex of statistical variables into principal components", *J. Educ. Psychol.*, **1933**, *24*, 417-441.

- Jain, A.K.; Dubes, R.C., "Algorithms for Clustering Data" Prentice-Hall advanced reference series: Computer Science. Prentice Hall College Division: Upper Saddle River, New Jersey, **1988**.

- Jain, A.K.; Murty, M.N.; Flynn P.J., "Data Clustering, A Review", *Computing Surveys*, **1999**, *31*, 264-323.

- Jolliffe I.T, "Principal Component Analysis, 2ed.", Springer, Springer-Verlag: New York, **2002**.

- Lee, M.B.; Storer, M.K.; Blunt, J.W.; Lever, M., "Validation of ¹H NMR spectroscopy as an analytical tool for methylamine metabolites in urine", *Clinica Chimica Acta*, **2006**, *365*, 264-269.

- Li, J.; Hibbert, D.B.; Fuller, S.; Vaughn, G., "A comparative study of point-to-point algorithms for matching spectra", *Chemometr Intell Lab Syst*, **2006**, *82*, 50-58.

- Lindon, J.C., "The Handbook of Metabonomics and Metabolomics", Nicholson, J.K.; Holmes, E., Eds.; Elsevier: Oxford, UK, 2007.

- Lloyd, S.P., "Least Squares Quantization in PCM", *IEEE Trans Inf Theory*, **1982**, *28*, 129-137.

- MacQueen, J.B., "Some Methods for classification and Analysis of Multivariate Observations", En *Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability, Vol 1*; Le Cam, L.M.; Neyman, J., Ed.; University of California Press: Berkeley, CA, USA, **1967**, pp. 281-297.

- Maitra, R.; Peterson, A.D.; Ghosh A.P., “A systematic evaluation of different methods for initializing the K-means clustering algorithm” *in preparation*, available via web: <http://www.public.iastate.edu/~apghosh/files/IEEEclust2.pdf>.
- Mayr, M.; Liemd, D; Zhangd, J.; Lid, X.; Avliyakuovd, N.K.; Yangd, J.I.; Youngd, G.; Vondriskad, T.M.; Ladrouec, C.; Madhub, B.; Griffithsb, J.R.; Gomed, Aldrin.; Xua, Q.; Ping, P., “Proteomic and metabolomic analysis of cardioprotection: Interplay between protein kinase C epsilon and delta in regulating glucose metabolism of murine hearts”, *J. Mol. and Cell. Cardiology*, **2009**, *46*, 268-277.
- Molina, D.V.; Navarro-Uribe, U.; Murgich, J., “Partial Least-Squares (PLS) Correlation between Refined Product Yields and Physicochemical Properties with the ¹H Nuclear Magnetic Resonance (NMR) Spectra of Colombian Crude Oils”, *Energy & Fuels*, **2007**, *21*, 1674-1680.
- Nicholson, J.K.; Lindon J.C., “Metabonomics”, *Nature*, **2008**, *455*, 1054-156.
- Pearson, K., “On Lines And Planes Of Closest Fit To Systems Of Points In Space”, *Philosophical Magazine*, **1901**, *2*, 559-572.
- Pelczer, I., “Structural biology, ligand binding, metabomics, the changing face of high-field, high-resolution NMR spectroscopy”, *J Mol Struct (Theochem)*, **2003**, *666–667*, 499-505.
- Powers, R., “NMR: metabolomics and drug discovery”, *Magn. Reson. Chem.*, **2009**, *47*, S2–S11.
- Saab, J.; Mokbel, I.; Razzouk, A.C.; Ainous, N.; Zydowicz, N.; Jose, J., “Quantitative Extraction Procedure of Naphthenic Acids Contained in Crude Oils. Characterization with Different Spectroscopic Methods”, *Energy & Fuels*, **2005**, *19*, 525-531.
- Siebert, K.J., “Chemometrics in Brewing - A Review.”, *J. Am. Soc. Brew. Chem.*, **2001**, *59(4)*, 147-156.
- Speight, J.G., “Handbook Of Petroleum Product Analysis”, John Wiley & Sons: West Sussex, UK, **2002**, pp. 29-33.
- Steinhaus, H., “Sur la division des corps matériels en parties”, *Bull. Acad. Polon. Sci. Cl. III.*, **1956**, *4*, 801-804.
- Tamás, J.; Podani, J; Csontos, P., “An extension of presence/absence coefficients to abundance data: a new look at absence”, *J. Veg. Sci.*, **2001**, *12*, 401-410.

- Trygg, J.; Holmes, E.; Lundstedt, T., “Chemometrics in Metabonomics”, *J. Proteom. Res.*, **2007**, *6*, 469-479.

- Van der Greef, J.; Smilde, A.K.; “Symbiosis of chemometrics and metabolomics: past, present, and future”, *J. Chemom.*, **2005**, *19*, 376-386.

- Varmuza, K.; Karlovits, M.; Demuth, W., “Spectral similarity versus structural similarity: infrared spectroscopy”, *Analytica Chimica Acta*, **2003**, *490*, 313-324.

- Wider, G.; Dreier, L., “Measuring Protein Concentrations by NMR Spectroscopy”, *J. Am. Chem. Soc.*, **2006**, *128*, 2571-2576.

- Winning, H.; Larsen, F.H.; Bro, R.; Engelsen, S.B.; “Quantitative analysis of NMR spectra with chemometrics”, *J. Magn. Reson.*, **2008**, *190*, 26-32.

5.7 Apéndice

5.7.1 Detalles sobre el funcionamiento de los programas creados para este estudio:

Para la realización de los diferentes pasos se desarrollaron una serie de programas. Para algunos de ellos existen alternativas comerciales, que nos valieron para comprobar la bondad de los resultados conseguidos por los programas creados.

Peakgeneratorclst

Para la generación de los buckets utilizados en los espectros 2D-HMBC, se creó el programa “peakgeneratorclst”. Este programa funciona creando un archivo de salida en formato “.xml”, que puede ser importado por el programa MestRe-C, generando las gradillas de integración sobre el espectro deseado. El archivo de entrada que hemos de aportar al programa “peakgeneratorclst” debe tener un nombre de la forma “.dat” y debe contener el número de puntos en cada dimensión, la anchura espectral de las dos dimensiones y la frecuencia más baja para ambas dimensiones. Además, podemos definir individualmente cada una de las ROI's que definamos en nuestro problema, que se deben introducir en el fichero de entrada en un formato de 5 columnas, indicando el número de ROI, la posición (en ppm, coordenadas en ambas dimensiones) del centro del ROI y la amplitud en hertzios en ambas dimensiones. Finalmente, se puede introducir en el archivo de entrada una tercera parte, en la que asignamos a los diferentes clusters una serie de números de picos concretos. Introduciendo esto en el archivo de entrada, la salida serán n archivos “.xml”, donde n será el número de clusters introducidos en el archivo de entrada, dentro de los cuales se definen los buckets de integración asignados

a cada cluster. El archivo con las dos primeras partes se utiliza al comienzo del estudio, para situar las gradillas de integración en las ROI's, y así poder obtener la lista de integrales que se utilizarán en el cálculo del índice de similitud. La introducción de la tercera parte del archivo será útil para el último proceso del estudio, para poder representar en el espectro procesado de cada muestra los picos característicos que definen el cluster al que se asigna esa muestra.

Similarityindex

Para el cálculo de la matriz de similitudes, compuesta de los índices de similitud entre pares de espectros de las muestras del estudio se creó el programa “similarityindex”. El fichero de entrada para este programa debe tener la extensión “.dat” y contiene una tabla en que los datos de cada muestra ordenadas por columnas. En una muestra la primera fila de su columna contiene un nombre identificativo y las que vienen a continuación contienen los valores de las intensidad de los buckets obtenidas en el paso anterior, cuando se importan sobre cada espectro. Por lo tanto, la tabla tendrá un tamaño de n columnas, donde n es el número de muestras, y m filas, donde m es el número de buckets definidos mas la fila de nombres. Para generar el fichero “.dat” de forma automática la tabla completa con el conjunto de muestras debe ser exportada en formato ASCII y con separación entre columnas definida por espacios. El programa utiliza otro archivo de entrada con extensión “.cfg” que contiene una serie de parámetros de configuración para el cálculo; el número de iteraciones para la búsqueda de similitud explorando la relación SN, porcentaje de valores con intensidad mínima que definen el valor de SN de un espectro entre otros. Además, cabe la opción de incluir un archivo de entrada “.clst” en el que podemos preasignar muestras a clusters concretos para que el programa lo tenga en cuenta.

En este trabajo se ha utilizado únicamente el índice de similitud obtenido a partir del MCHORD, pero el programa nos ofrece la opción de escoger entre diferentes maneras de calcular el índice de similitud.

La salida del programa consiste en 5 archivos diferentes. Uno con los datos de entrada, pero normalizados según los parámetros dados (que serán los datos que utilice el programa para calcular el índice de similitud). Este archivo tendrá por nombre el nombre del input original, con extensión “.normal”. Dos archivos, con extensión “.rmsdlike” y “.rmsdunlike”, que contienen la matriz de similitud calculada. Además, en el caso de haber introducido clusters y muestras asignadas a esos clusters en el archivo “.clst”, en estos archivos tendremos reflejados la selección de picos que resultan ser característicos para cada cluster. Esta lista es la que podemos añadir como tercera entrada en el archivo “.dat” con el que trabaja el programa “peakgeneratorclst” para conseguir la representación en el espectro de estos picos característicos. La salida también proporciona un cuarto archivo, de extensión “.peaksallowed”, en el que obtenemos una serie de matrices de dimensiones $n \times n$ (donde n es el número de muestras) donde se nos indican los picos tomados en cuenta para cada comparación pico-pico, el factor de normalización, el ruido medio y el factor de concentración. Cada punto de cada una de las matrices se corresponde con el dato relativo a la comparación del espectro n con el espectro m en el cálculo del índice de similitud. Por último, si además del fichero de entrada “.dat” se incluye un fichero con una clusterización deseada de las muestras, archivo de extensión “.clst”, el programa proporciona como salida un archivo con extensión “.cmi” que contiene el valor global y particular de CMI de cada bucket, esto es, una medida de su concordancia con la clusterización introducida.

Program_PCA

Este programa realiza el análisis PCA a partir de los datos de la matriz de similitud. El fichero de entrada para este programa será el archivo “.rmsdlike” obtenido como fichero de salida del programa “similarityindex”, sin necesidad de modificaciones. En el programa tendremos que introducir manualmente el número de filas y columnas de la matriz. El fichero de salida es un fichero con nombre igual al fichero de entrada, pero con la extensión “.out”. En este fichero obtendremos una lista de autovalores, que nos indican el porcentaje de varianza que explica cada nueva componente tomada en cuenta en el espacio del PCA.

K-means8

Este programa se encarga de la clusterización k-means de las muestras. Puede utilizarse como entrada la distribución (coordenadas) de las muestras obtenida en un análisis PCA. El programa trabaja directamente con el fichero de salida obtenido del programa “Program_PCA” sin modificaciones. Antes del cálculo, el programa requiere introducir el número deseado de clusters y el número mínimo requerido de muestras por cluster. Además, se puede incluir información sobre el número de iteraciones con las que trabajará el algoritmo, la estabilidad mínima deseada para cada muestra en el cluster de destino y también asignar una ordenación *ad-hoc* de muestras en los clusters con la que empiece a trabajar el programa. Finalmente, podemos escoger entre que el programa busque la solución mejor o la solución que más se haya repetido. El programa trabajará a lo largo de las tres etapas nombradas anteriormente; una primera clusterización k-means, la aplicación del método Monte-Carlo para calcular la estabilidad de las muestras en los clusters y por último vuelve a hacer la clusterización,

pero teniendo en cuenta los datos de estabilidad obtenidos mediante el Monte-Carlo. El archivo de salida más importante es el de extensión “.kmeans”, en el que se detalla la solución óptima de clusterización k-means encontrada, en función de la opción elegida ésta es la de menor valor de desviación o bien la más popular. En este archivo se muestra el cluster al que se asigna cada muestra, su estabilidad dentro de ese cluster, y las posiciones de los centroides de todos los clusters. Además se facilita también en la solución las coordenadas de las muestras (en el cálculo realizado son las del subespacio del PCA). Este archivo se puede importar directamente a una hoja de cálculo para representar gráficamente el resultado. El segundo archivo que el programa proporciona como salida es un archivo de extensión “.more”, en el que se detallan las soluciones obtenidas y almacenadas en la reserva genética, sus valores correspondientes de *avintra* y *avinter*, y el número de veces que se ha obtenido cada solución en el conjunto de las iteraciones (popularidad).

OBTENCIÓN RÁPIDA DE ESPECTROS DE RMN MULTIDIMENSIONALES MEDIANTE INCREMENTO DE RELAJACIÓN POR ADICIÓN DE SUSTANCIAS PARAMAGNÉTICAS

RAPID ACQUISITION OF MULTIDIMENSIONAL NMR SPECTRA BY RELAXATION ENHANCEMENT ADDING PARAMAGNETIC SUBSTANCES

Marino Vega-Vázquez y Manuel Martín-Pastor*

Universidad de Santiago de Compostela, Unidad de Resonancia Magnética, RIAIDT, Edif. CACTUS, campus Sur, Santiago de Compostela, A Coruña 15706 (España)

*autor de contacto (e-mail: manuel.martin@usc.es)

Recibido: 12/02/2010 - Evaluado: 07/03/2010 - Aceptado: 19/03/2010

RESUMEN

Se propone un método para acelerar la obtención de experimentos de RMN multidimensionales (nD-RMN) mediante la adición de sustancias paramagnéticas que inducen un incremento de la relajación en el soluto que se quiere estudiar. Escogiendo adecuadamente el tipo de sustancia paramagnética y ajustando su concentración en la muestra, el efecto llamado PRE (Paramagnetic Relaxation Enhancement) puede modularse adecuadamente permitiendo pérdidas mínimas de sensibilidad y una reducción sustancial del tiempo en los experimentos nD-RMN. Este tipo de metodología denominada Fast-PRE-NMR se ha probado para los experimentos 2D-COSY, 2D-TOCSY, 2D-HMQC y 2D-HMBC con moléculas de tamaño pequeño y mediano, utilizando sustancias paramagnéticas de bajo coste, como Mn^{2+} o Fe^{3+} . El método permitió una reducción del tiempo de obtención en un factor 3 proporcionando espectros de buena calidad y sin artefactos. Basándose en estos resultados, se propone un protocolo sencillo para la optimización de la concentración de paramagnético para unas condiciones experimentales particulares.

ABSTRACT

A method is proposed to speed-up the acquisition of multidimensional NMR experiments (nD-NMR) by addition of paramagnetic substances, which increases relaxation in the studied solute. By properly choosing the type of paramagnetic and adjusting its concentration in the sample, the PRE effect (Paramagnetic Relaxation Enhancement) is conveniently modulated, allowing a substantial reduction of experiment time of many nD-NMR, with minimum losses in sensitivity. This type of Fast-PRE-NMR methodology was optimized for 2D COSY, 2D TOCSY, 2D HMQC and 2D HMBC experiments for small and medium size- molecules and using low cost paramagnetic substances, such as Mn^{2+} . This method allowed a reduction of the experimental time by a factor of three, providing good quality spectra and no use of artifacts. Based on these results, it is proposed a general protocol for rapid set-up of optimum paramagnetic concentration for particular solute and sample experimental conditions.

Palabras clave: RMN rápida, saturación progresiva, incremento de relajación paramagnética
Keywords: Fast-NMR, progressive saturation, fast-NMR, paramagnetic relaxation enhancement

INTRODUCCIÓN

En años recientes, se han desarrollado varios métodos para acelerar la adquisición de experimentos de Resonancia Magnética Nuclear (NMR) (Freeman y Kupce, 2003). Además del interés de esos métodos para acortar el tiempo de medida de los experimentos multidimensionales de RMN (nD-RMN), pueden ser la única alternativa para estudios de RMN en escalas de tiempo cortas o intermedias, o para realizar medidas con muestras relativamente inestables.

Hay un tipo de métodos rápidos de RMN que tienen como objetivo reducir el tiempo de espera necesario para que se produzca la relajación completa de las señales entre acumulaciones (*scans*) sucesivas. Algunos ejemplos recientes son los experimentos SOFAST/BEST (Schanda y Brutscher, 2005; Schanda et al., 2005. Schanda y Brutscher, 2006; Gal et al., 2007) aplicables a péptidos y proteínas enriquecidas isotópicamente en ^{15}N . El tiempo de espera entre *scans* sucesivos normalmente supone una parte considerable del tiempo que dura un experimento de RMN, éste debe ser al menos 5 veces el máximo valor de relajación longitudinal T_1 de los núcleos a estudiar, para así garantizar que todas las resonancias afectadas por el anterior *scan* han vuelto al equilibrio (> 99 %). Si no se cumple esta condición, se produce saturación que afecta a la integral de la señal (pérdida de cuantitatividad) y provocar aparición de ruido t_1 y otros artefactos en los espectros nD (Kupce y Freeman, 2007).

La interacción intermolecular entre el soluto y los iones paramagnéticos añadidos a la muestra pueden producirle al soluto una disminución de los tiempos de relajación longitudinal (T_1) y transversal (T_2) de sus espines nucleares. Este efecto, denominado incremento de velocidad de relajación mediante paramagnéticos (PRE), ha sido muy usado para el estudio de estructura y propiedades dinámicas de moléculas (Bertini y Luchinat, 1996), y también históricamente para incrementar la velocidad de adquisición de espectros 1D de ^{13}C así como para acelerar la relajación de la señal correspondiente al H_2O . Recientemente la estrategia PRE se ha usado con éxito en la aceleración de la adquisición de experimentos nD en moléculas grandes (Cai et al., 2006). El efecto PRE en el caso de pequeñas y medianas moléculas es teóricamente menos favorable que para macromoléculas (Cai et al., 2006) porque teóricamente además de acortarse los tiempos T_1 también pueden acortarse significativamente los T_2 , lo que conduce a un ensanchamiento de señal y la correspondiente pérdida de sensibilidad en los espectros. En este punto, uno de los objetivos de este artículo es poner de manifiesto que si se eligen cuidadosamente las condiciones experimentales, el efecto PRE también permite acelerar los experimentos nD-RMN de moléculas medianas-pequeñas con mínimas pérdidas de sensibilidad y resolución.

En este estudio se ha considerado el uso de PRE para acelerar experimentos de rutina 2D RMN que proporcionan correlaciones a través del acoplamiento escalar (^1H - ^1H , ^1H - ^{13}C ,..., etc.). En concreto se ha utilizado PRE para los experimentos 2D g-COSY, 2D TOCSY, 2D g-HMQC, y 2D g-HMBC en moléculas de tamaño mediano-pequeño, un método al que se denominará fast-PRE-NMR. Mediante la optimización en cada uno de estos experimentos del efecto PRE óptimo, en concordancia con una duración optimizada de los eventos de la propia secuencia de pulsos, se puede obtener la mejor relación de sensibilidad y velocidad del experimento posibles. Finalmente, se presenta un protocolo sencillo para la rápida optimización de las condiciones de fast-PRE-NMR en el conjunto de los experimentos considerados para cualquier tipo de soluto.

MATERIALES Y MÉTODOS

Preparación de muestras

Se prepararon varias muestras de sacarosa a concentración 9.75 mM, disolviendo 20 mg de sacarosa en 0.6 ml de D_2O a concentración de Mn^{2+} 0, 0.5, 1, 1.5 y 2 mM. Para la adquisición de experimentos de medida de tiempos de relajación T_1 y T_2 se prepararon muestras adicionales con concentraciones de Mn^{2+} de 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8 y 0.9 mM.

Espectroscopía RMN

Los espectros se adquirieron a 300 K en un espectrómetro BRUKER DRX-500. Los datos fueron procesados y analizados con MestRe-C v 4.9 (*Mestrelab Research*). El ajuste de los datos de intensidad/integral a funciones matemáticas y su representación gráfica fueron realizados con Origin v 6.1 (*Originlab Corporation*). La asignación de las señales de ^1H y ^{13}C de sacarosa se llevó a cabo con los espectros adquiridos, completando la asignación de la literatura (Bock et al., 1984). Para nombrar las señales de RMN de la sacarosa se usan las iniciales g y f para el monosacárido de glucosa y fructosa, respectivamente. A no ser que se indique otra cosa, todos los espectros fueron adquiridos con 32 *dummy-scans* para asegurar el estado estacionario de espín antes de la adquisición de las FIDs. En los experimentos 2D se insertó un bloque de purga $90^\circ(^1\text{H})$ -gradiente- G_z justo antes del periodo de relajación d_1 , para purgar magnetización residual del scan anterior y así reducir la presencia de artefactos (Kupce y Freeman, 2007).

Medidas de relajación longitudinal (T_1) y relajación transversal (T_2)

Los tiempos de relajación longitudinal T_1 y transversal T_2 de protón se determinaron a partir de 14 muestras de sacarosa con concentraciones de metal paramagnético entre 0 y 2 mM.

Tiempos de relajación longitudinal T_1 : Se determinaron con el experimento de saturación progresiva (Mitrovic et al., 2001) (d_1 - 90° -*acq.*). Para cada muestra se adquirieron 16 espectros a los siguientes tiempos de d_1 $4 \cdot 10^{-6}$, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1, 1.1, 1.2, 1.3, 1.4 y 1.5 s. Cada espectro fue adquirido con 128 *dummy-scans* seguidos por 32 *scans*. Las FIDs fueron detectadas con 1024 puntos complejos y el tiempo de adquisición fue de 0.225 en cada experimento. El tiempo de relajación longitudinal para cada señal no solapada de protón de sacarosa fue determinada por un ajuste no lineal de la integral de las señales siguiendo la siguiente ecuación (Mitrovic et al., 2001):

$$\frac{I(\tau)}{I^{\max}} = 1 - \exp(-\tau / T_1) \quad (1)$$

Donde $\tau = d_1 + aq$, $I(\tau)$ es la integral de la señal medida en el espectro al valor correspondiente de τ e I^{\max} es la integral de la señal al mayor valor de τ

Tiempos de relajación transversal T_2 : Las medidas se hicieron usando la secuencia clásica de CPMG (Meiboom y Gill, 1958) (d_1 - 90° - $(\delta - 180^\circ - \delta)_n$ -*acq.*). El número de ciclos n se varió linealmente en 50 pasos para obtener duraciones variables del bloque CPMG entre 3 y 700 ms. Los espectros fueron adquiridos con 16 *dummy scans* seguidos por 16 *scans*. Las FIDs fueron detectadas con 2048 puntos. El tiempo de relajación T_2 para cada señal de protón no solapada de (1) fue determinada por el ajuste no lineal de la integral a la siguiente ecuación:

$$\frac{I(\tau)}{I_0} = \exp(-\tau / T_2) \quad (2)$$

Donde $\tau = 2n\delta$ es la duración del bloque CPMG. $I(\tau)$ es la integral de la señal al valor correspondiente de τ , e I_0 es la integral de la señal al mínimo valor de τ .

Tiempos de relajación longitudinal efectivos en experimentos 2D (T_1^{eff}): Los tiempos de relajación efectivos de protón (T_1^{eff}) se determinaron para los picos de cruce no solapados de sacarosa en los experimentos 2D g-COSY, 2D TOCSY, 2D g-HMQC y 2D g-HMBC. Se utilizó el método de saturación progresiva (Mitrovic et al.,

2001) aplicado al experimento en cuestión. En cada experimento, el T_1^{eff} de un pico de cruce se determina por ajuste no lineal a la Ec. 1 de la intensidad del pico de cruce a los valores de d_1 que se utilizaron en cada caso (ver abajo).

Espectros 2D g-COSY

Se adquirieron en modo magnitud con el experimento COSY que emplea gradientes para la selección de coherencias (secuencia *cosygp* de Bruker). Para cada relación de sacarosa: Mn^{2+} se determinó el T_1^{eff} explorando el valor de d_1 en un rango desde 4.10^{-6} a 1.4 s obteniéndose 15 espectros. La única excepción es la muestra con Mn^{2+} 2mM en la que se exploró un rango de d_1 desde 4.10^{-6} a 0.7 s. Los espectros se adquirieron con 8 *scans* y el tiempo de adquisición fue de 0.255 s. El número de puntos detectados fue de 128 y 1024 para las dimensiones F1 y F2, respectivamente.

Espectros 2D TOCSY

Se adquirieron en el modo sensible a la fase (secuencia *mlevst* de Bruker). Para cada relación de sacarosa: Mn^{2+} se determinó el T_1^{eff} explorando el valor de d_1 en un rango desde 4.10^{-6} a 1.4 s obteniéndose 15 espectros. La única excepción es la muestra con Mn^{2+} 2mM en la que se usó un rango de d_1 desde 4.10^{-6} a 0.7 s. Los espectros se adquirieron con 8 *scans* y el tiempo de adquisición fue de 0.255 s. Se usó el esquema MLEV-17 para el periodo de mezcla isotrópica con un tiempo de mezcla de 80 ms. El número de puntos detectados fue de 128 y 1024 en las dimensiones F1 y F2, respectivamente.

Espectros 2D g-HMQC

Se adquirieron en modo magnitud usando la secuencia HMQC que emplea los gradientes para la selección de coherencias (*inv4gp* en la biblioteca de Bruker). Para cada muestra de sacarosa: Mn^{2+} se determinó el T_1^{eff} adquiriendo 15 experimentos con valores de d_1 en un rango desde 4.10^{-6} a 1.4 s. Las únicas excepciones fueron las muestras con Mn^{2+} 1.5 y 2 mM en las que se usó un rango de d_1 desde 4.10^{-6} a 0.72 s. Los espectros se adquirieron con 16 *scans* y los tiempos de espera en la secuencia se ajustaron a un valor nominal de ${}^1J_{\text{CH}}$ de 145 Hz. Se utilizó desacoplamiento de banda ancha de carbono durante el tiempo de adquisición (0.12 s) con el esquema GARP. El número de puntos detectados fue de 128 y 384 en las dimensiones F1 y F2, respectivamente.

2D g-HMBC

Estos experimentos se adquirieron en modo magnitud usando la secuencia HMBC que emplea los gradientes para la selección de coherencias (secuencia *inv4gplprnd* de Bruker). Para cada muestra de sacarosa: Mn^{2+} se determinó el T_1^{eff} adquiriendo 15 experimentos con valores de d_1 en un rango desde 4.10^{-6} a 1.4 s para las muestras con Mn^{2+} 0 y 0.5 mM, mientras que para las muestras con Mn^{2+} 1, 1.5 y 2 mM se usó un rango de d_1 desde 4.10^{-6} a 0.72 s. Los espectros se adquirieron con 16 *scans* y un tiempo de adquisición de 0.255 s. Los tiempos de espera en la secuencia se ajustaron a unos valores nominales de ${}^1J_{\text{CH}}$ de 145 Hz y ${}^nJ_{\text{CH}}$ of 7.7 Hz. Se utilizó desacoplamiento de banda ancha de carbono durante el tiempo de adquisición (0.12 s) con ese esquema GARP. El número de puntos detectados fue de 200 y 1024 en f1 y f2 respectivamente.

TEORÍA

Tiempo de medida del experimento de RMN, sensibilidad y relajación longitudinal T_1

En general, en los experimentos de RMN la relación señal-ruido por unidad de tiempo de medida (SN_t), referido como la sensibilidad del experimento por *scan*, viene dada por Ec. 3 (Ernst et al., 1987; Pervushin et al., 2002).

$$SN_i = c(1 - \exp(-\frac{1}{T_1} d_1)) / (\sqrt{d_0 + d_1}) \quad (3)$$

Donde T_1 es el tiempo de relajación longitudinal del espín a estudiar, d_0 es la duración de la propia secuencia desde el primer pulso de excitación hasta el final de la adquisición de la FID, d_1 es el tiempo de espera de relajación antes del primer pulso del scan y c es un coeficiente de proporcionalidad.

Para obtener espectros cuantitativos se debe cumplir la condición $d_1 + d_0 \geq 5 T_1$ con objeto de asegurar que se recupera >99% de la señal en cada scan. Cuando no se desea obtener espectros cuantitativos sino la máxima sensibilidad por unidad de tiempo de medida (Pervushin et al., 2002), ésta viene dada por $SN_i^{\max} = \frac{c \cdot 0.71}{\sqrt{d_1^{\text{opt}}}} \sim$

$c \cdot 0.64 \sqrt{1/T_1}$ lo que resulta en un tiempo de espera óptimo d_1^{opt} dado por Ec. 4 (Ernst et al., 1987; Pervushin et al., 2002):

$$d_1^{\text{opt}} = 1.25 \cdot T_1 - d_0 \quad (4)$$

De las relaciones anteriores se deduce que la relajación longitudinal T_1 de los espines define la sensibilidad y duración de los experimentos de RMN cuando se realizan muchas repeticiones de ciclos de excitación/detección (scans).

Incremento de relajación mediante sustancias paramagnéticas (PRE)

Los tiempos de relajación T_1 y T_2 que pueden medirse en un soluto al añadir sustancias paramagnéticas son la suma de las contribuciones diamagnética y paramagnética a la relajación. Si la concentración no es muy alta, el PRE es proporcional a la concentración del paramagnético (Ec. 5) (Cai et al., 2006)

$$\frac{1}{T_i^{\text{obs}}} = \frac{1}{T_i^{\text{dia}}} + c \frac{1}{T_i^{\text{para}}} \quad i = 1, 2 \quad (5)$$

Donde T_i se refiere a las relajaciones longitudinal y transversal, T_1 y T_2 respectivamente, c es un parámetro proporcional a la concentración de la especie paramagnética que también tiene en cuenta el intercambio químico (Cai et al., 2006). Cuando el soluto es una molécula mediana-pequeña, puede asumirse que todos los protones son igualmente accesibles al disolvente y que por tanto, los cambios esperados en T_1^{obs} e independientemente en T_2^{obs} serán esencialmente iguales para todos ellos, con una dependencia lineal con la concentración dada por Ec. 5. Desviaciones de esa linealidad podrán ser útiles para refutar la presunción inicial de la ausencia de interacciones específicas con el soluto (ej. formación de aductos). Las condiciones más favorables para el método fast-PRE-NMR propuesto son aquellas que producen la menor relación $T_1^{\text{obs}}/T_2^{\text{obs}}$.

Dado que no existen contacto directo entre los espines que interactúan, solo la interacción intermolecular dipolo-dipolo entre un núcleo y electrones desapareados pueden dar lugar a PRE significativo (relajación "outer-sphere") (Freed, 1978; Ayant et al., 1975). A campos magnéticos altos (frecuencia de $^1\text{H} \geq 400$ MHz) las velocidades de relajación paramagnéticas dipolares $R_1^{\text{para,dd}}$ y $R_2^{\text{para,dd}}$ (son las inversas de los tiempos de relajación $1/T_1^{\text{para,dd}}$ y $1/T_2^{\text{para,dd}}$, respectivamente) vienen dadas por la teoría de Solomon-Bloembergen (Ecs. 6 y 7) (Solomon, 1955; Bloembergen y Morgan, 1961):

$$R_1^{para,dd} = \frac{2}{15} \left(\frac{\mu_0}{4\pi} \right)^2 \gamma_I^2 g^2 \mu_B^2 S(S+1) r^{-6} \left\{ \frac{3\tau_c}{1+(\omega_H \tau_c)^2} + \frac{7\tau_c}{1+(\omega_S \tau_c)^2} \right\} \quad (6)$$

$$R_2^{para,dd} = \frac{1}{15} \left(\frac{\mu_0}{4\pi} \right)^2 \gamma_I^2 g^2 \mu_B^2 S(S+1) r^{-6} \left\{ 4\tau_c + \frac{3\tau_c}{1+(\omega_H \tau_c)^2} \right\} \quad (7)$$

Donde r es la distancia entre el centro paramagnético y el núcleo observado; μ_0 la permeabilidad del vacío; γ_I la constante giromagnética; g el factor g electrónico; μ_B el magnetón de Bohr; S el número cuántico de espín electrónico ($S=5/2$ para Mn^{2+} , $S=1/2$ para Cu^{2+}), ω_H y ω_S las frecuencias de Larmor de 1H y del electrón, respectivamente, y $\omega_S = 658 \omega_H$. El tiempo de correlación τ_c en las Ecs. 6 y 7 viene dado por:

$$\frac{1}{\tau_c} = \frac{1}{T_{1e}} + \frac{1}{\tau_M} + \frac{1}{\tau_R} \quad (8)$$

Donde T_{1e} es el tiempo de relajación electrónico, τ_R es el tiempo de correlación rotacional del soluto y τ_M es el tiempo de residencia del paramagnético en la superficie del soluto. De estos tres, el proceso más rápido será el que tenga un mayor peso en τ_c .

T_{1e} depende del ión metálico paramagnético, con unos valores situados en el rango de 10^{-12} a 10^{-9} s. τ_R es aproximadamente 10^{-10} a 10^{-8} en moléculas medianas-pequeñas ($MW < 2000$). τ_M ha sido calculado (Pintacuda y Otting, 2002) usando $\tau_M = x^2/6D$, donde D es la distancia de difusión x requerida por el paramagnético para salir de la interacción del complejo, para Mn^{2+} la distancia x es 34 \AA (Iwahara y Clore, 2006), para distancias mayores que este valor el efecto del paramagnético se vuelve indetectable.

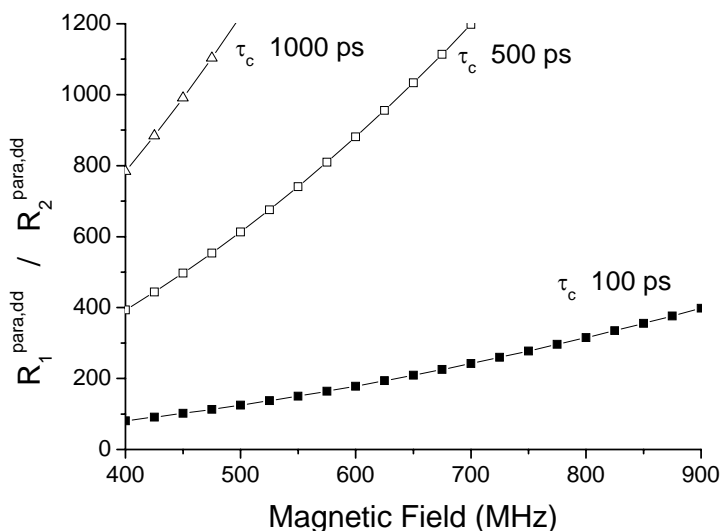


Fig. S1: Dependencia teórica de la relación $R_1^{para,dd} / R_2^{para,dd}$ con el campo magnético. Las curvas representadas corresponden a tiempos de correlación τ_c 1000 ps (triángulos), 500 (cuadrados sin relleno) y 100 ps (cuadrados con relleno).

A campos magnéticos altos, $\omega_{S, H} \cdot \tau_c \gg 1$ en Ecs. 6 y 7 y la función de densidad espectral $J(\omega)$ disminuye, así pues, para la misma concentración de paramagnético el PRE disminuye con el aumento de la fuerza del campo. Por otro lado, a mayor fuerza de campo magnético se incrementa la relación $R_1^{\text{para,dd}} / R_2^{\text{para,dd}}$ especialmente en casos de mayor τ_c y por lo tanto son favorables para el método fast-PRE-NMR aquí propuesto (Fig. S1). Para moléculas de tamaño mediano-pequeño en solventes no viscosos, τ_M es mucho mayor que τ_R y por lo tanto no contribuye al valor de τ_c en la Ec. 8. Dado que el más rápido de ambos procesos T_{1e} o τ_R será el que más influya en τ_c , la única condición que se le requiere al paramagnético para obtener la mayor relación $R_1^{\text{para,dd}} / R_2^{\text{para,dd}}$ es que $T_{1e} \geq \tau_R$. Existen iones paramagnéticos bastante asequibles que satisfacen esta condición tales como, Mn^{2+} and Fe^{3+} que tienen un T_{1e} en el rango de 100 a 1000 ps, lo cual está cerca del τ_R típico de moléculas pequeñas y medianas

RESULTADOS Y DISCUSIÓN

Para moléculas de pequeño y mediano tamaño a frecuencias de resonancia entre 400 y 500 MHz, el efecto PRE depende principalmente de las propiedades de la sustancia paramagnética (espín electrónico, tipo de coordinación del metal y tiempo de relajación electrónica T_{1e}) y de su concentración en la muestra. El efecto PRE introduce una disminución en los tiempos de relajación T_1 's y T_2 's del núcleo de interés (ver sección de teoría). Mientras que un efecto PRE en T_1 es útil para la reducción del tiempo de espera entre scans (d_1^{opt} en Ec. 4) de los experimentos de RMN, el efecto PRE en T_2 causa ensanchamiento de señales y pérdida de sensibilidad en los espectros. Las condiciones más favorables para fast-PRE-NMR pueden encontrarse experimentalmente de forma que se produzca la menor relación PRE- T_1 / PRE- T_2 en moléculas de tamaño mediano-pequeño. Esto requiere el ajuste de la concentración del paramagnético a su valor idóneo en la muestra y para el experimento de RMN en cuestión.

La forma práctica de optimizar fast-PRE-NMR para una sustancia paramagnética que se utiliza por primera vez, es determinar la concentración óptima con una muestra de referencia de un soluto de tamaño medio-pequeño (datos de referencia PRE) bajo las condiciones experimentales a utilizar (campo magnético, temperatura, disolvente, etc.). Es esperable que esta concentración óptima sea fácilmente extrapolable a otros solutos de similares características y que presumiblemente vayan a interactuar de forma similar con el paramagnético.

En lo que sigue se presenta un análisis completo de efecto PRE con una muestra de referencia de sacarosa en D_2O utilizando como paramagnético el ión Mn^{2+} . Basándonos en estos resultados de referencia, finalmente se propone un protocolo simplificado para la determinación rápida de la concentración óptima de esta especie paramagnética en solutos de características similares.

Optimización de PRE- T_1 y PRE- T_2

El efecto PRE- T_1 y PRE- T_2 de un paramagnético se determina en una muestra de referencia mediante una valoración con el paramagnético elegido. En este estudio se obtuvieron datos de referencia PRE con una muestra de sacarosa en D_2O y el paramagnético Mn^{2+} . A cada concentración de paramagnético se midieron los tiempos de relajación T_1 y T_2 de los protones no solapados de la sacarosa. Estos resultados se dan en la Tabla 1 y se representan en las Figs. 1a y 1b. Para cada protón analizado en la Fig. 1a ó 1b, la correspondiente curva claramente se separa de la linealidad predicha por la Ec. 5, sugiriendo la formación de aductos (Slappendel et al., 1982). Cada curva puede descomponerse en dos zonas con cierto carácter lineal, una entre 0 a 0.3 mM que sugiere la formación de aducto 1:1, y otra en el rango 1 a 2 mM que sugiere un aducto 1:2.

En las Figs. 1a y 1b puede verse que la mayor caída de los tiempos T_1 y T_2 se produce en el rango de $[Mn^{2+}]$ entre 0 a 1.0 mM. A concentraciones superiores, el decaimiento es mucho más moderado y asintótico. Es interesante apuntar que la magnitud del PRE es mayor cuanto mayor es el T_1 (o T_2) original en ausencia de

paramagnético. Como consecuencia de esto, el incremento de $[Mn^{2+}]$ tiende a reducir la dispersión en los T_1 and T_2 observados de los diferentes protones de la sacarosa considerados.

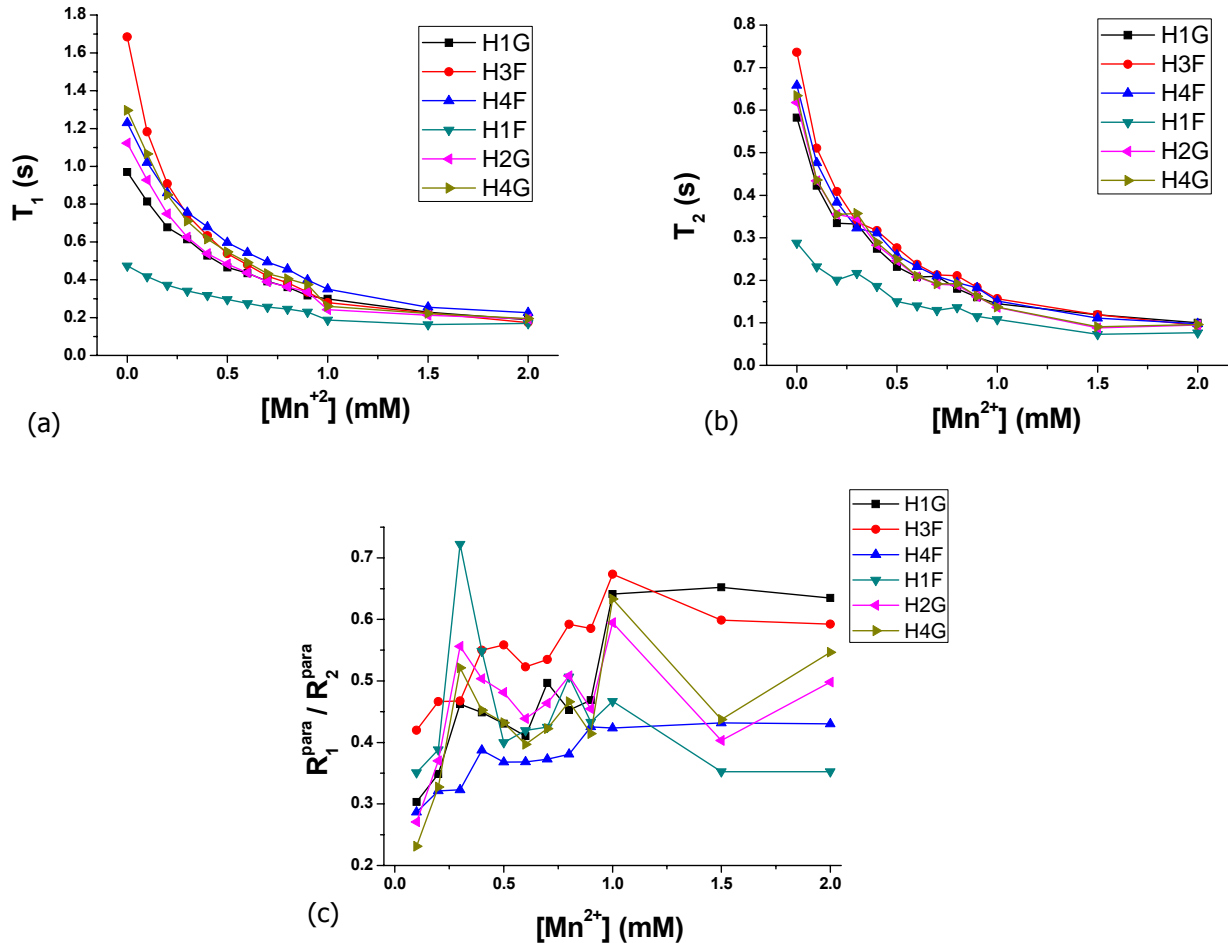


Fig. 1: Curvas de valoración para los protones no solapados de la sacarosa 1 en función de la concentración de Mn^{2+} (mM). a) Tiempos de relajación longitudinal T_1 (s), b) Tiempos de relajación transversal T_2 (s) y c) relación experimental R_{1para}/R_{2para} .

La contribución paramagnética a la relajación R_1^{para} y R_2^{para} se calculó para los protones de la sacarosa a cada concentración de Mn^{2+} usando los datos de la Tabla 1 con las Ecs. 3 y 4, respectivamente y la relación experimental R_1^{para} / R_2^{para} está representada en la Fig. 1c. Para la mayor parte de los protones de la Fig. 1c hay una tendencia significativa en el cociente R_1^{para} / R_2^{para} a aumentar con la concentración del Mn^{2+} , indicando que el paramagnético causa una mayor caída en el T_1 que en el T_2 de los protones. La única excepción es la señal H1f que tiene un comportamiento errático; aumenta rápidamente en las muestras con concentración de Mn^{2+} 0 a 0.25 mM y entonces disminuye para concentraciones más altas. La interacción paramagnética dipolar entre Mn^{2+} y la sacarosa por sí sola no explica las curvas representadas en la Fig. 1c, dado que la teoría de Solomon-Bloembergen en las Ecs. 5 y 6 predice un cociente R_1^{para} / R_2^{para} constante a concentraciones diferentes de Mn^{2+} . La conclusión es que otros mecanismos deben tener una contribución relevante en la relajación paramagnética en este sistema (por ej. contacto y/o mecanismos de Curie). Estos resultados señalan la influencia del

paramagnético y del tipo de muestra a medir cuando se quiere optimizar la relación PRE- T_1 / PRE- T_2 de cara a utilizar el método fast-PRE-NMR.

Tabla 1: Tiempos de relajación de protón T_1 (s) y T_2 (s) determinados a 500 MHz para las señales seleccionadas de la sacarosa a diferentes concentraciones de Mn^{2+} .

Mn²⁺ (mM)	T₁(s) H-1g	T₁(s) H-2g	T₁(s) H-4g	T₁(s) H-1f	T₁(s) H-3f	T₁(s) H-4f
0.0	0.970	1.122	1.296	0.474	1.684	1.230
0.1	0.814	0.928	1.066	0.417	1.182	1.021
0.2	0.677	0.748	0.849	0.372	0.908	0.860
0.3	0.614	0.626	0.710	0.341	0.743	0.756
0.4	0.527	0.538	0.615	0.318	0.633	0.680
0.5	0.465	0.483	0.548	0.296	0.538	0.597
0.6	0.435	0.437	0.490	0.275	0.480	0.544
0.7	0.392	0.389	0.433	0.256	0.419	0.494
0.8	0.361	0.362	0.405	0.246	0.385	0.456
0.9	0.317	0.336	0.375	0.229	0.335	0.399
1.0	0.299	0.242	0.260	0.187	0.279	0.351
1.5	0.229	0.213	0.222	0.163	0.224	0.255
2.0	0.189	0.193	0.195	0.170	0.174	0.226
Mn²⁺ (mM)	T₂(s) H1g	T₂(s) H2g	T₂(s) H4g	T₂(s) H1f	T₂(s) H3f	T₂(s) H4f
0.0	0.583	0.618	0.634	0.288	0.736	0.658
0.1	0.422	0.434	0.436	0.232	0.510	0.476
0.2	0.334	0.355	0.355	0.201	0.409	0.383
0.3	0.332	0.346	0.357	0.217	0.337	0.323
0.4	0.274	0.283	0.289	0.186	0.317	0.311
0.5	0.232	0.246	0.249	0.150	0.276	0.259
0.6	0.208	0.208	0.210	0.140	0.237	0.232
0.7	0.209	0.191	0.192	0.130	0.212	0.210
0.8	0.180	0.189	0.192	0.136	0.211	0.195
0.9	0.160	0.161	0.163	0.115	0.183	0.182
1.0	0.145	0.136	0.137	0.108	0.157	0.151
1.5	0.119	0.088	0.091	0.073	0.119	0.111
2.0	0.100	0.095	0.096	0.077	0.094	0.098

Fast-PRE-NMR con moléculas de tamaño mediano-pequeño

Para el desarrollo de un protocolo general de fast-PRE-NMR para acelerar los experimentos 2D g-COSY, 2D TOCSY, 2D g-HMQC y 2D g-HMBC, en paralelo con el estudio de PRE- T_1 y PRE- T_2 descrito arriba, se hizo una valoración con el paramagnético Mn^{2+} para determinar el efecto PRE en los tiempos de relajación longitudinal efectivos, PRE- T_1^{eff} de aquellos picos no solapados de cada experimento 2D. Esto se llevó a cabo utilizando el método de saturación progresiva (Mitrovic et al., 2001) con el propio experimento 2D (parte experimental).

El PRE- T_1^{eff} se determina para un pico del espectro mediante el ajuste a la Ec. 1 de su intensidad normalizada (I/I_{max}) para una serie de valores de d_1 . El PRE- T_1^{eff} determina cómo de rápido un pico de cruce recupera su

máxima intensidad de equilibrio en función del tiempo de espera entre *scans* d_1 para una concentración dada de paramagnético. El $\text{PRE-T}_1^{\text{eff}}$ depende del tiempo de relajación de todas las coherencias generadas durante la secuencia de pulsos, no solo aquellas detectadas, y por lo tanto es más apropiado que el T_1 para definir el tiempo óptimo de relajación entre *scans* en un experimento 2D (Pervushin et al., 2002).

La Fig. 2 muestra dos curvas de saturación progresiva de 2D g-COSY para el pico de cruce H1g-H2g de sacarosa a concentraciones de Mn^{2+} 0 y 1.0 mM. Para $d_1 = 0$ s, la cantidad de señal recuperada en la primera muestra es solo de $\sim 30\%$ de la máxima intensidad, mientras que es $\sim 75\%$ en el segundo caso. Los valores correspondientes de $\text{PRE-T}_1^{\text{eff}}$ 1.84s, y 0.25 s para Mn^{2+} 0 y 1.0 mM, respectivamente, son consistentes con estas observaciones.

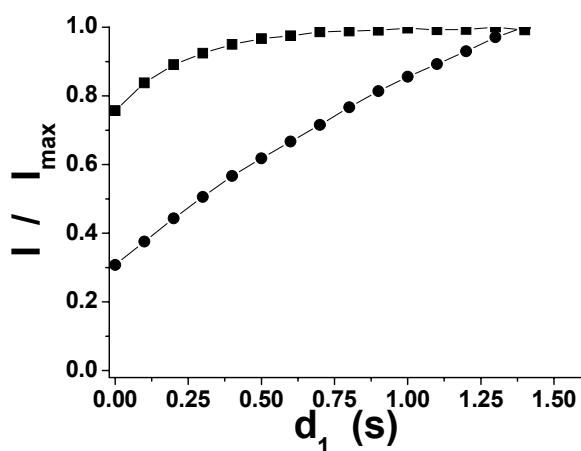


Fig. 2: Grafica de la intensidad de señal normalizada (%) vs. tiempo entre scans d_1 (s) obtenido para el pico de cruce H1G-H2G de 1 con el experimento 2D g-COSY. Los círculos y las cajas corresponden a Mn^{2+} 0 mM y 1.0 mM, respectivamente.

Los valores de $\text{PRE-T}_1^{\text{eff}}$ determinados para diversos picos de cruce de la sacarosa en los espectros 2D g-COSY, 2D TOCSY, 2D g-HMQC y 2D g-HMBC están representados en las Figs. 3a, 3b, 3c, and 3d, respectivamente. En estas figuras se puede observar que el mayor cambio en $\text{PRE-T}_1^{\text{eff}}$ ocurre a concentraciones de Mn^{2+} situadas en el rango de 0 a 1 mM en todos los espectros. Para concentraciones por encima de 1.0 mM los $\text{PRE-T}_1^{\text{eff}}$ siguen disminuyendo pero su caída es menos acusada y por lo tanto no causan un gran impacto en la cantidad de señal recuperada en esos experimentos, por el contrario, puede amplificar efectos adversos para la sensibilidad como incrementar PRE-T_2 e incluso disminuir la sensibilidad de la sonda de RMN en el caso de sondas de detección inversa con un elevado factor-Q y poca tolerancia salina.

Los resultados de $\text{PRE-T}_1^{\text{eff}}$ de la Fig. 3 indican que si la duración se hace similar en los diferentes tipos de espectros 2D considerados ($d_0 =$ tiempo de pulsos + adquisición FID ~ 0.3 s), una concentración de Mn^{2+} entre 1.0 y 1,5 mM resulta la mejor elección para maximizar la cantidad de señal recuperada en los diferentes picos de cruce analizados al mínimo valor de d_1 .

Cuando se introducen en la Ec. 4 los datos de $\text{PRE-T}_1^{\text{eff}}$ obtenidos para sacarosa a concentración Mn^{2+} 1.5 mM, se calcula un d_1^{opt} de 0.253 s, 0.196 s, 0.205 s, y 0.181 s para 2D g-COSY, 2D TOCSY, 2D g-HMQC y 2D g-HMBC, respectivamente. Estos valores de d_1^{opt} son mucho más cortos que los generalmente asumidos de d_1 1.2 a 1.6 s para este tipo de experimentos y suponen un considerable ahorro de tiempo de espectrómetro.

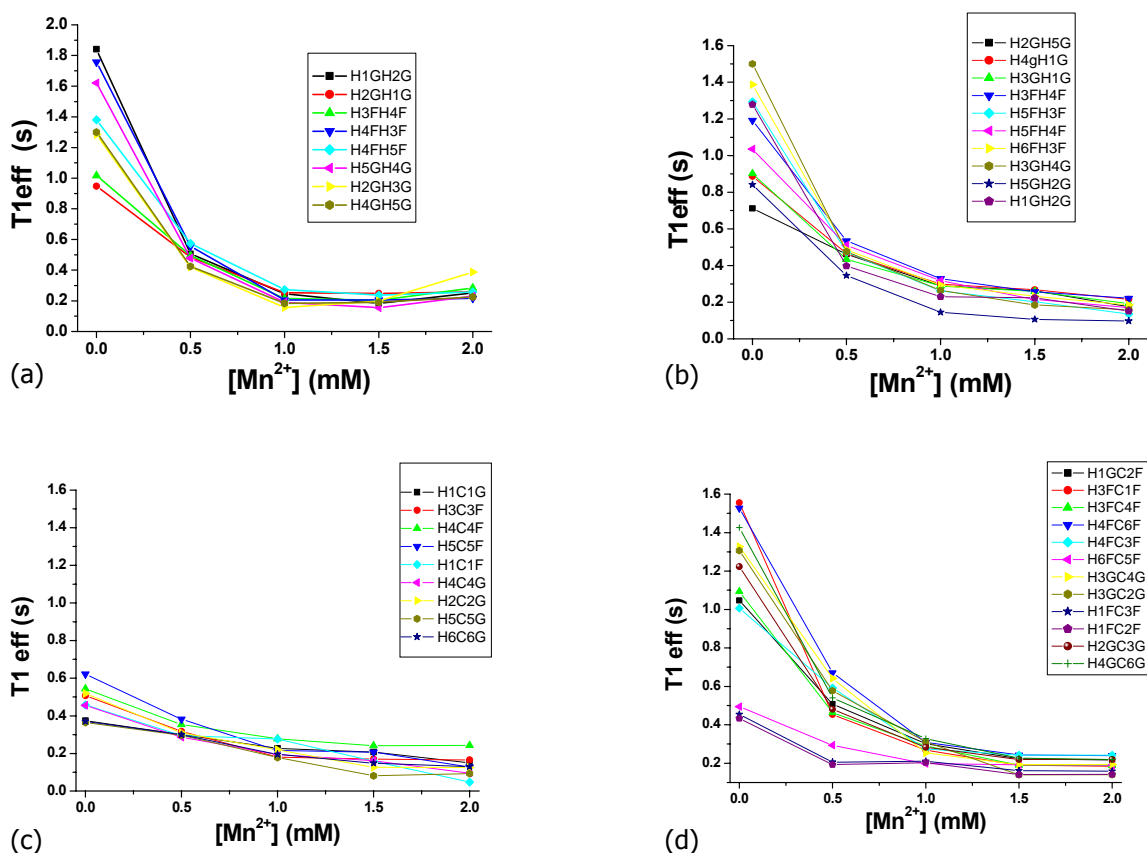


Fig. 3: Tiempos de relajación longitudinal efectivos (T_{1eff}) para diferentes picos de correlación de los espectros de 1 a diferentes concentraciones de Mn^{2+} . a) T_{1eff} obtenidos de 2D g-COSY con saturación progresiva, b) T_{1eff} obtenidos de de 2D TOCSY con saturación progresiva, c) T_{1eff} obtenidos de de 2D g-HMQC con saturación progresiva, d) T_{1eff} obtenidos de de 2D g-HMBC con saturación progresiva.

Para el caso del experimento 2D-g-HMQC, dado que se utiliza desacoplamiento heteronuclear, en vez del d_1^{opt} calculado de 0.205 s, se alargo a 0.3 s para prevenir daños en la sonda como resultado de los ciclos de trabajo excesivamente rápidos. Existen algunas versiones recientes del experimento HMQC como el 2D IPAP-HMQC (Kern et al., 2008) que no emplean desacoplamiento heteronuclear y por lo tanto no sería necesario aumentar el d_1^{opt} .

Los espectros de las Figs. 4 ejemplifican los resultados obtenidos con el método fast-PRE-NMR para la muestra de sacarosa. El espectro 2D-gHMBC de la Fig. 4a se adquirió en 17 minutos sin la adición de Mn^{2+} y sufre los efectos perniciosos de utilizar un d_1 muy corto (d_1 0.181 s). La baja calidad de este espectro se refleja en las muchas bandas de ruido t_1 y la pérdida evidente de sensibilidad cuando se le compara con el espectro convencional (Fig. 4b) adquirido en 60 min. (d_1 1.4 s). Sin embargo, cuando a la muestra se le añade Mn^{2+} 1.5 mM, el espectro puede adquirirse en 17 min. (Fig. 4c) obteniéndose picos perfectamente resueltos y con una calidad similar a la del espectro convencional y sin ningún tipo de artefactos.

La Tabla 2 resume los resultados experimentales del método fast-PRE-NMR aplicado a 2D-g-COSY, 2D-TOCSY, 2D-g-HMQC y 2D-g-HMBC con la muestra de sacarosa a una concentración de Mn^{2+} 1.5 mM. A esa

concentración de paramagnético, el tiempo de experimento se reduce en un factor de 3 sin la aparición de artefactos y con una resolución de pico de cruce igual a la de los experimentos convencionales.

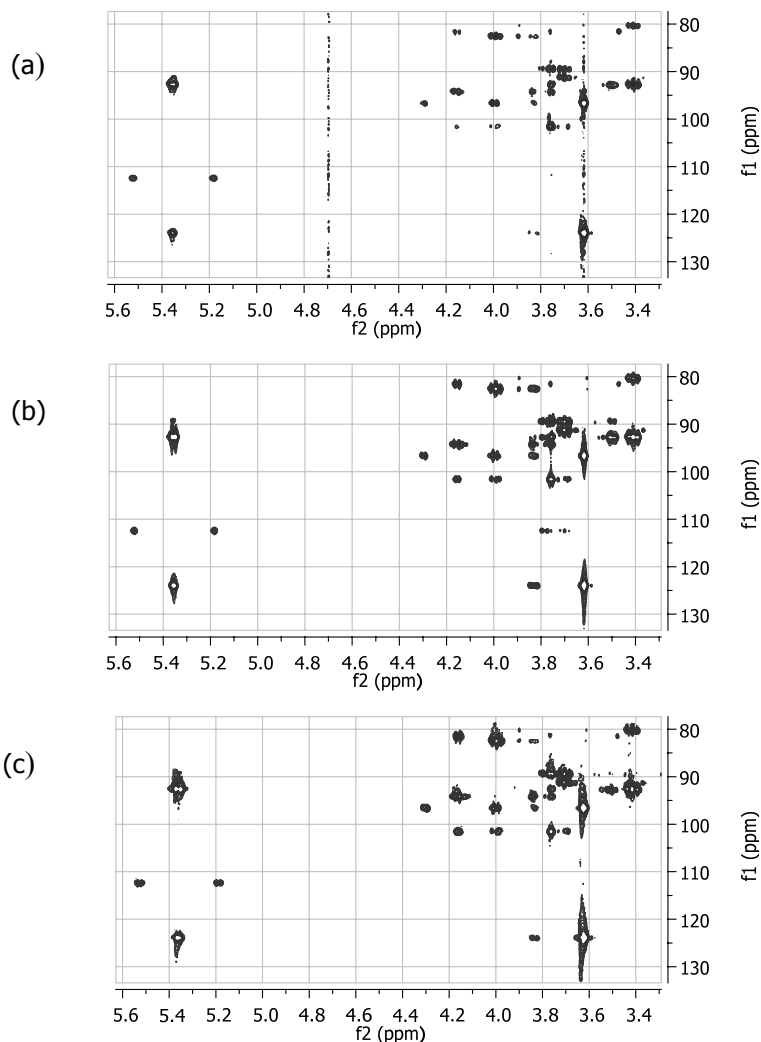


Fig. 4: Comparación de espectro 2D g-HMBC de 1. a) Experimento convencional con d1 0.181 s en una muestra sin Mn2+ adquirido en 17 minutos b) Experimento convencional con d1 1.4 s en una muestra sin Mn2+ adquirido en 60 minutos y c) Experimento Fast-PRE-NMR con d1 0.181 s en una muestra con Mn2+ 1.5 mM adquirido en 17 minutos. El resto de condiciones experimentales fue igual en los 3 espectros.

Un aspecto importante del experimento fast-PRE-NMR se refiere a la sensibilidad absoluta respecto al experimento convencional. El efecto PRE- T_2 asociado al uso de paramagnéticos inevitablemente introduce pérdidas de sensibilidad en los espectros. Los resultados de la Tabla 2 muestran que los experimentos fast-PRE-NMR 2D g-COSY, 2D-TOCSY y 2D-g HMQC tienen una sensibilidad siempre por encima del 45% en las muestras de oligosacáridos estudiados. La situación más desfavorable ocurre para el fast-PRE 2D g-HMBC en el cual la sensibilidad de ciertos picos se reduce hasta un 20% respecto del experimento convencional (Tabla 2). La razón de esa baja sensibilidad está relacionada con que las coherencias heteronucleares ^1H - ^{13}C de múltiple cuanto gastan un tiempo relativamente mayor en este experimento (65 ms) expuesto a los procesos de relajación PRE- T_2 .

Tabla 2: Resultados experimentales para Fast-PRE-NMR de 2D-COSY, 2D-TOCSY, 2D-HMBC y 2D-HMQC para la muestra de sacarosa y Mn²⁺ 1.5 mM. Los resultados se comparan con el experimento análogo de la muestra sin paramagnético bajo las mismas condiciones exceptuando que el tiempo de espera entre scans d1 es 1.4 s.

Experimento	Fast-PRE-NMR d ₁ ^{opt(a)} (s)	Fast-PRE-NMR Exp. time ^(b) (min.)	Conventional NMR Exp. time ^(c) (min.)	Time saving factor	Fast-PRE-NMR Sensitivity ^(f) (%)
2D g-COSY	0.253	9	28	3.1	56 to 80%
2D TOCSY ^(d)	0.196	19	60	3.1	45 to 64%
2D g-HMQC	0.300 ^(e)	10	30	3.0	56 to 87%
2D g-HMBC	0.181	17	60	3.5	20 to 40%

(a) d_{1opt} está calculado con Eq. 4 para un tiempo T₁ nominal de protón de 0.225 s.

(b) 8 scans para 2D g-COSY, 2D TOCSY y 2D g-HMQC, 16 scans para 2D g-HMBC. Numero de incrementos en t₁ 128. Tiempo de adquisición 0.255 en todos los experimentos menos 2D g-HMQC en que fue 0.12 s.

(c) Los experimentos convencionales utilizaron las mismas condiciones que los fast-PRE-NMR excepto el d₁, que fue de 1.4 s.

(d) Experimento sensible a la fase.

(e) El d_{1opt} teórico fue 0.205 s pero el valor final utilizado fue 0.300 s debido a los requisitos del ciclo de trabajo del desacoplamiento heteronuclear.

(f) Calculado como una media de la relación señal-ruido de los picos de en fast-PRE-NMR respecto de los experimentos convencionales.

En conjunto, la reducción de la sensibilidad obtenida en Tabla 2 para los experimentos fast-PRE-NMR indica que el método es más apropiado para estudiar muestras que no estén limitadas por la sensibilidad del espectrómetro.

Método rápido de ajuste de condiciones de fast-PRE-NMR

Los resultados obtenidos de PRE-T₁, PRE-T₂ y PRE-T₁^{eff} con la muestra de sacarosa y el paramagnético Mn²⁺ (Figs. 1 y 3) proporcionan la base para conseguir resultados óptimos con el método fast-PRE-NMR en otras moléculas de tamaño mediano-pequeño utilizando este paramagnético. Básicamente, lo que se requiere es ajustar la concentración de paramagnético hasta que el efecto PRE-T₁ esté aproximadamente en el valor óptimo PRE-T₁^{opt} ~ 0.225 s (valor medio de PRE-T₁ en Fig. 1a para Mn²⁺ 1.5 mM) que se ha comprobado que es adecuado para el conjunto de experimentos 2D considerados aquí.

Una receta para poder ajustar rápidamente la concentración necesaria de paramagnético en un soluto cualquiera para conseguir PRE-T₁^{opt} se da en la Fig. 5. La concentración de paramagnético se incrementa en pequeños pasos y su efecto PRE-T₁ se determina con el experimento de inversión-recuperación (180°-τ-90°-Acq). Este ajuste puede hacerse en un solo paso mediante el método conocido como anulación de señal T₁ (T₁-null). Para conseguir PRE-T₁^{opt} ≤ 0.225 s, la condición es que se anule o cambie de negativa a positiva la señal del experimento de inversión-recuperación al utilizar un valor τ_{null} = T₁^{opt} * ln(2) = 0.156 s.

Dado que PRE-T₂ no se ha considerado para este protocolo de ajuste rápido de la condición de fast-PRE-NMR, es aconsejable comprobar que la calidad del propio espectro de protón (anchuras de línea y sensibilidad) es aceptable para estas condiciones.

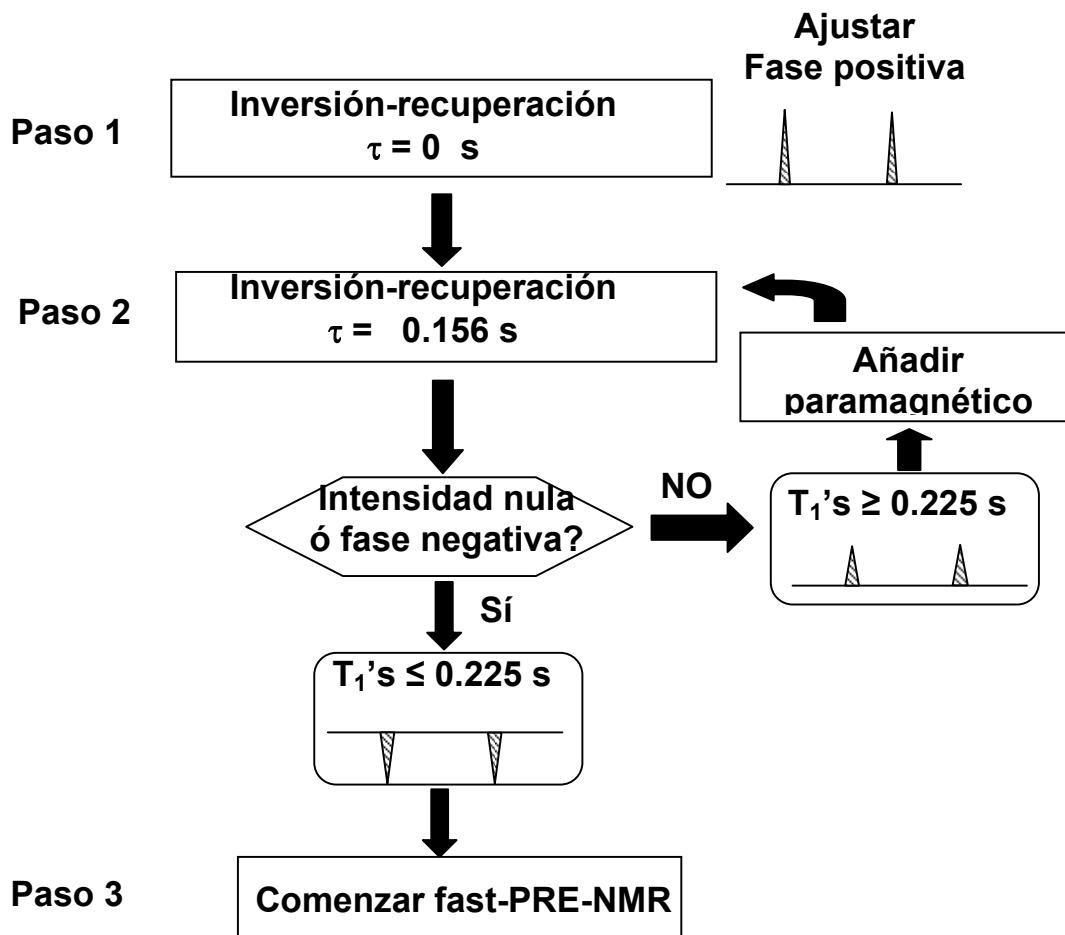


Fig. 5: Protocolo propuesto para el ajuste rápido de la concentración de paramagnético. Se utiliza el experimento de inversión-recuperación con el método de T1-null. La concentración de paramagnético se aumenta en pequeños pasos hasta que el T1max del soluto es igual o inferior a 0.225 s.

CONCLUSIONES

La validez de esta estrategia de ajuste rápido de la condición de fast-PRE-NMR se ha probado con excelentes resultados en el laboratorio de la Unidad de Resonancia Magnética de la Universidad de Santiago de Compostela, para varios oligosacáridos (lactosa y un trisacárido formado por manosas). En definitiva, puede concluirse que el método fast-PRE-NMR permite adquirir experimentos nD-RMN con un ahorro considerable de tiempo de espectrómetro evaluado en un factor 3 para los experimentos fast-PRE-NMR 2D g-COSY, 2D TOCSY, 2D g-HMQC and 2D g-HMBC obteniéndose espectros de una calidad similar a los convencionales.

La simplicidad del método Fast-PRE-NMR, que no requiere secuencias de pulsos o procesamiento de datos especiales, y el hecho de que se puedan utilizar especies paramagnéticas de bajo coste (ej. Mn^{2+} o Fe^{3+}) hace que pueda tener aplicaciones para la aceleración de análisis rutinarios de nD-RMN de otros solutos de tamaño mediano-pequeño por ejemplo para acelerar determinados protocolos de control de calidad basados en RMN.

ABREVIATURAS

nD	RMN multidimensional
PRE	(Paramagnetic Relaxation Enhancement) Incremento de Relajación mediante adición de sustancias Paramagnéticas
SNR	Relación Señal-Ruido
FID	(Free Induction Decay) Caída libre de la Inducción, señal de RMN
d_1	Tiempo de espera para relajación antes de pulsar un scan
d_0	Tiempo desde el primer pulso de excitación hasta el final de la adquisición de la FID.
T_1^{eff}	Tiempo de relajación longitudinal o espín-red efectivo
T_1	Tiempo de relajación longitudinal o espín-red
T_2	Tiempo de relajación transversal o espín-espín
$Scan$	Acumulación con adquisición de FID
$Dummy-Scan$	Acumulación sin adquisición de FID

AGRADECIMIENTOS

M. V-V agradece la financiación de una beca de la Comisión Europea (proyecto artizymes FP6-2003-NEST-B3 15471).

REFERENCIAS

- Ayant, Y.; Belorizky, E.; Alizon, J.; Gallice, J. (1975): "*Calcul des densités spectrales résultant d'un mouvement aléatoire de translation en relaxation par interaction dipolaire magnétique dans les liquides*", J. Phys.: 36, 991-1004.
- Bertini, I.; Luchinat, C. (1996); "*NMR of paramagnetic substances*"; Coord. Chem. Rev.: 150, Elsevier. Amsterdam, pp 1-300.
- Bloembergen, N.; Morgan, L.O. (1961); "*Proton Relaxation Times in Paramagnetic Solutions. Effects of Electron Spin Relaxation*", J. Chem. Phys.: 34, 842-850.
- Bock, K.; Pedersen, C.; Pedersen, H. (1984); "*Carbon-13 Nuclear Magnetic Resonance Data for Oligosaccharides*", Adv. Carbohydr. Chem. Biochem.: 42, 193-225.
- Cai, S.; Seu, C.; Kovacs, Z.; Sherry, A.D.; Chen, Y. (2006); "*Sensitivity Enhancement of Multidimensional NMR Experiments by Paramagnetic Relaxation Effects*", J. Am. Chem. Soc.: 128, 13474-13478.
- Ernst, R.R.; Bodenhausen, G.; Wokaum, A. (1987); "*The Principles of Nuclear Magnetic Resonance in One and Two Dimensions*" Clarendon, Oxford.
- Freeman, R.; Kupce, E. (2003); "*New methods for fast multidimensional NMR*", J. Biomol. NMR: 27, 101-113.
- Freed, J.H. (1978); "*Dynamic effects of pair correlation functions on spin relaxation by translational diffusion in liquids. II. Finite jumps and independent T1 processes*", J. Chem. Phys.: 68, 4034-4037.
- Gal, M.; Schanda, P.; Brutscher, B.; Frydman, L. (2007); "*UltraSOFAST HMQC NMR and the Repetitive Acquisition of 2D Protein Spectra at Hz Rates*", J. Am. Chem. Soc.: 129, 1372-1377.
- Iwahara, J.; Clore, G.M. (2006); "*Detecting transient intermediates in macromolecular binding by paramagnetic NMR*", Nature: 440, 1227-1230.

- Kern, T.; Schanda, P.; Brutscher, B. (2008); "*Sensitivity-enhanced IPAP-SOFAST-HMQC for fast-pulsing 2D NMR with reduced radiofrequency load*", J. Magn. Reson.: 190, 333-338.
- Kupce, E.; Freeman, R. (2007); "*Fast multidimensional NMR by polarization sharing*", Magn. Reson. Chem.: 45, 2-4.
- Meiboom, S.; Gill, D. (1958); "*Modified Spin-Echo Method for Measuring Nuclear Relaxation Times*". Rev. Sci. Instr.: 29, 688-691.
- Mitrovic, V.F.; Sigmund, E.E.; Halperin, W.P. (2001); "*Progressive saturation NMR relaxation*", Phys. Review B.: 64, 24520-24527.
- Pervushin, K.; Vögeli, B.; Eletsky, A. (2002); "*Longitudinal 1H Relaxation Optimization in TROSY NMR Spectroscopy*", J. Am. Chem. Soc.: 124, 12898-12902.
- Pintacuda, G.; Otting, G. (2002); "*Identification of Protein Surfaces by NMR Measurements with a Paramagnetic Gd(III) Chelate*", J. Am. Chem. Soc.: 124, 372-373.
- Schanda, P.; Brutscher, B. (2005); "*Very Fast Two-Dimensional NMR Spectroscopy for Real-Time Investigation of Dynamic Events in Proteins on the Time Scale of Seconds*", J. Am. Chem. Soc.: 127, 8014-8015.
- Schanda, P.; Kupče, E.; Brutscher, B. (2005); "*SOFAST-HMQC Experiments for Recording Two-dimensional Deteronuclear Correlation Spectra of Proteins within a Few Seconds*", J. Biomol. NMR: 33, 199-211.
- Schanda, P.; Brutscher, B. (2006); "*Hadamard frequency-encoded SOFAST-HMQC for ultrafast two-dimensional protein NMR*", J. Magn. Reson.: 178, 334-339.
- Slappendel, S.; Aasa, R.; Falk, K.E.; Malmstrom, B.G.; Vliegthart, T.V. (1982); "*¹H-NMR spectroscopic study on the binding of alcohols to soybean lipoxygenase-1*", Biochim. et Biophys. Acta: 708, 266-271.
- Solomon, I. (1955); "*Relaxation Processes in a System of Two Spins*", Phys. Rev.: 99, 559-565.

Fast multidimensional localized parallel NMR spectroscopy for the analysis of samples

Marino Vega-Vazquez,^a Juan Carlos Cobas,^b and Manuel Martin-Pastor^{a*}



A parallel localized spectroscopy (PALS) method is presented to speed up the acquisition of multidimensional NMR (nD) spectra. The sample is virtually divided into a discrete number of nonoverlapping slices that relax independently during consecutive scans of the experiment, affording a substantial reduction in the interscan relaxation delay and the total experiment time. PALS was tested for the acquisition of three experiments 2D COSY, 2D DQF-COSY and 2D TQF-COSY in parallel, affording a time-saving factor of 3–4. Some unique advantages are that the achievable resolution in any dimension is not compromised in any way: it uses conventional NMR data processing, it is not prone to generate spectral artifacts, and once calibrated, it can be used routinely with these and other combinations of NMR spectra. Copyright © 2010 John Wiley & Sons, Ltd.

Supporting information may be found in the online version of this article.

Keywords: NMR; ¹H; fast-NMR; localized spectroscopy; localized COSY

Multidimensional NMR (nD) spectroscopy has expanded dramatically in recent years and is a powerful tool for the study of structure, dynamics, and interactions of both low molecular weight compounds and relative large biomolecules. Recording a series of nD spectra can be time-demanding and there is the interplay between achieving a given signal resolution and/or sensitivity within a given time frame. In recent years, an increasing interest for developing fast-NMR techniques has arisen.^[1–13] Apart from the interest of those methods to save NMR spectrometer time, they can be the only alternative for NMR studies at short-to-intermediate timescales^[2] when it is necessary to obtain data with low concentrated samples or to avoid sample degradation.

Among the fast-NMR strategies, there is a class aimed to reduce the time spent waiting for interscan signal relaxation, usually known as d_1 delay. Strategies to diminish the d_1 delay that has been applied to nD NMR includes the so-called SOFAST and BEST methods^[9–12] and paramagnetic relaxation enhancement of the solute by mixing the sample with paramagnetic substances that reduce the longitudinal T_1 s of the solute under study.^[14]

For quantitative NMR spectroscopy, the interscan relaxation delay d_1 must be set at least 5 times the longest T_1 of the molecule. For nonquantitative studies, the maximum signal-to-noise ratio per measurement time, referred to as the *sensitivity* of the experiment, can be achieved at the optimal interscan delay d_1^{opt} given by Eqn (1)^[15]

$$d_1^{\text{opt}} = 1.25 T_1 - d_0 \quad (1)$$

where d_0 is the duration of pulse sequence plus the FID detection and T_1 is the longitudinal relaxation time of the excited spins. Arbitrarily reducing the d_1 delay in nD spectra may cause lack of quantitativity and/or sensitivity and the appearance of artifacts such as t_1 noise among others. Therefore, especially for nD experiments, where many repetitions of excitation/detection cycles are performed, the time waiting for signal relaxation usually represents a considerable part of NMR acquisition time.

The fast-NMR method PALS (parallel localized spectroscopy) that is proposed herein is aimed to reduce the interscan delay

by employing spatially localized spectroscopy.^[16] By combining semiselective radiofrequency (RF) pulses simultaneously with a pulse field gradient, spatially selective pulses can be introduced in an NMR pulse sequence. The PALS method virtually divides the sample volume in a discrete number of nonoverlapping slices and then acquires a different nD spectrum in each slice. Thus, if the slice used is changed sequentially in successive scans, the time required to acquire a scan in a given slice is employed by the spins in the other parts of the sample to relax toward equilibrium. Thus, considering the case of the same scan duration for the slices, the optimum interscan delay d_1^{opt} of Eqn (1) is divided by the number of slices selected.

Despite the losses in sensitivity associated with the slice selection employed in PALS, the method affords a considerable saving of NMR time whenever a number of nD spectra want to be acquired for a sample which is not very limited by concentration. For nD of small- and medium-sized molecules typically requiring d_1 delay of ~ 1 – 1.5 s and FID acquisition time of ~ 0.3 s, the best compromise in sensitivity *versus* timesaving is obtained when the sample is divided into approximately three or four separated slices, with each slice thickness covering as much as possible of the effective sample volume.

In principle, PALS permits the parallelization of the acquisition of spectra that involves either the same or a number of different types of nuclides in the indirect dimensions, without the requirement of special hardware such as multiple receptors^[17,18]

* Correspondence to: Manuel Martin-Pastor, Unidad de Resonancia Magnética, RIAIDT, Edif. CACTUS, Campus Sur, Universidad de Santiago de Compostela, Santiago de Compostela, A Coruña 15706, Spain.
E-mail: manuel.martin@usc.es

^a Unidad de Resonancia Magnética, RIAIDT, Edif. CACTUS, Campus Sur, Universidad de Santiago de Compostela, Santiago de Compostela, A Coruña 15706, Spain

^b Mestrelab Research S. L., Feliciano Barrera 9B - Bajo, 15706 Santiago de Compostela, A Coruña 15706, Spain

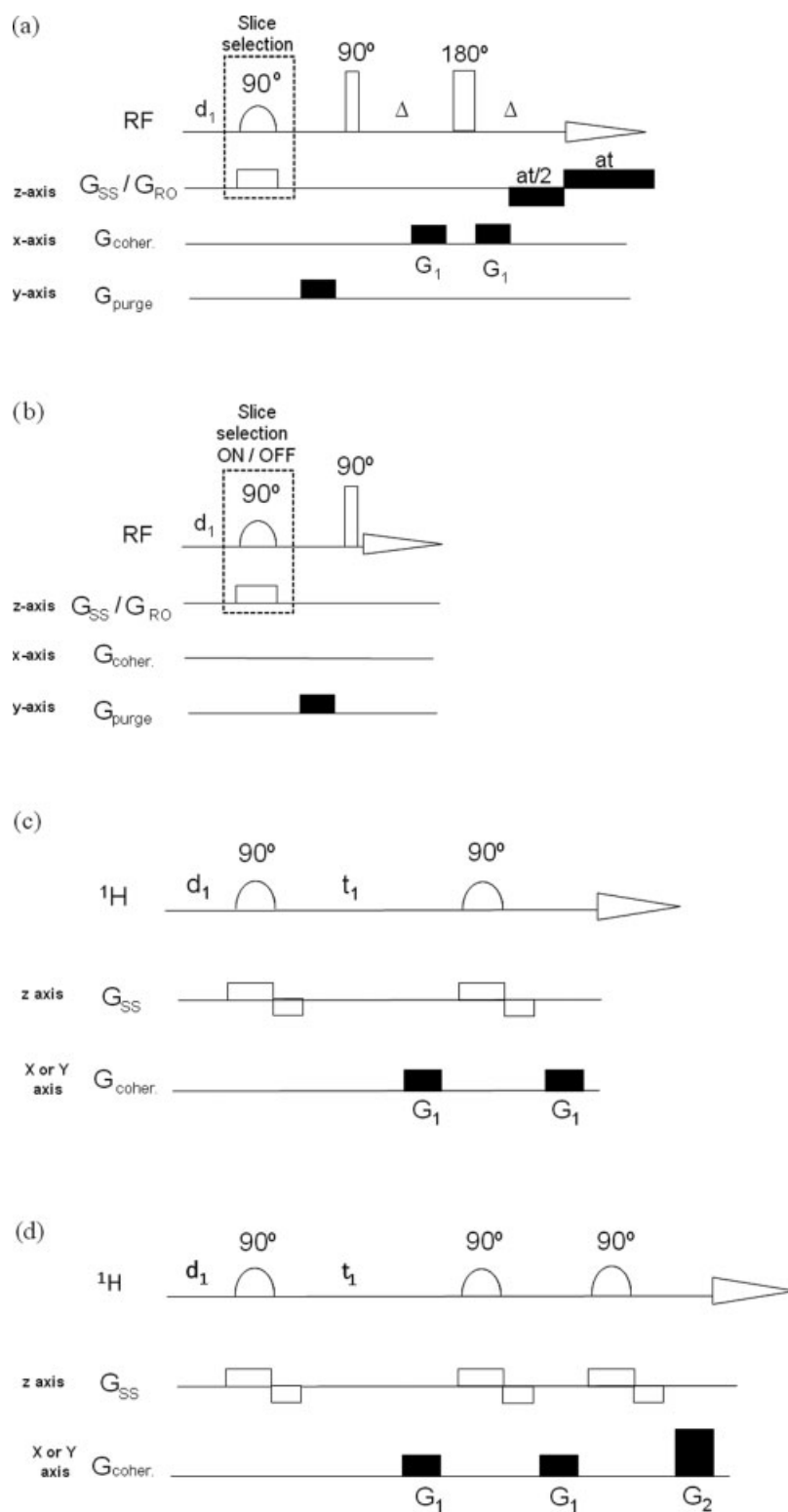


Figure 1. (a) One-dimensional slice-profile mapping sequence. The slice selection and readout gradients (G_{SS} and G_{RO} , respectively) are both applied in the z-direction. G_{purge} is a 1-ms gradient of 10 G/cm. The echo delay Δ was set to 15 ms. (b) Pulse sequence to obtain a quantitative slice-selective 1D spectrum with pure absorptive phase. A 1D reference spectrum of the whole sample is obtained when the slice-selection element of this sequence is switched OFF. The slice-selective 1D quantitative spectrum is obtained upon subtraction of the OFF and ON spectra under the same experimental conditions. (c) 2D slice-selective magnitude COSY and (d) 2D slice-selective magnitude MQF-COSY. In all sequences, square pulses correspond to nonselective hard pulses, the shape of the gradients is rectangular and the slice-selective pulse is of the *sinc* type with a width of 500 μ s. It is applied simultaneously with a slice-selection gradient G_{SS} of 5 G/cm along the z-axis. The duration of gradients G_1 and G_2 is 1 ms and the strength of G_1 is 10 G/cm. The same phase is applied for all the pulses and the receiver. In sequences (b) and (c), after the slice-selective 90° pulse, a gradient of opposite sign and half area than G_{SS} is applied during 1 ms for spatial refocusing. In sequence (c), the gradient ratio $G_1 : G_2$ is 1 : 2 and 1 : 3 for DQF-COSY and TQF-COSY, respectively.

or multiple-coil probes.^[19] The PALS method simply requires spectrometer equipped with pulse-field shielded gradient capability which are standard today. NMR pulse sequences that require pulse field gradients would benefit from using triple axis gradients hardware, since it is easier to prevent interscan refocusing of undesired coherences. Nevertheless, the PALS method should also be feasible using single z-shielded gradient probes using appropriate values of the gradients to prevent undesired refocusing of spatially selected coherences.

The PALS method was tested with a sample of sucrose 500 mM in D₂O in a Varian INOVA-750 (Varian inc.) spectrometer equipped with an inverse detection probe and triple axis gradients to obtain in parallel the three localized spectra 2D COSY, 2D DQF-COSY and 2D TQF-COSY.^[20–22] The starting point for the PALS method is the selection of the required number of nonoverlapping slices along the main axis of the sample tube (z-axis). The 1D slice-mapping experiment of Fig. 1a was used to experimentally set the position and thickness of three nonoverlapping slices along the sample tube. In the sequence, the magnetization excited by the first slice-selective pulse is immediately destroyed by the purge gradient applied on the y-axis. The next part of the sequence is a gradient spin echo image^[16] which employs hard pulses to excite the whole sample volume with spatial detection of the proton density along the readout direction (z-axis). Thus, the part of the magnetization destroyed by the gradient after the first slice-selective pulse generates a hole in the sample profile image (spin tagging). Using the sequence of Fig. 1a, three nonoverlapping slices were calibrated directly with the test sample (Figure S1). The calibration was performed interactively with the 1D slice-mapping spectrum. The following parameters were optimized: (i) the G_{ss} gradient strength controls the slice thickness, (ii) the frequency offset of the shaped pulse controls the slice position, and (iii) the power of the shaped pulse was optimized to maximize the depth of hole that is created in the slice-profile spectrum.

The NMR signal sensitivity obtained upon selection of each one of these three slices in the sample was measured with the experiment of Fig. 1b that provides a 1D quantitative slice-selective spectrum with pure phase (Figure S2). The sensitivity achieved in each slice referred to that of the whole active sample volume was 27.9%, 29.2%, and 30.9%. These results are conveniently close to the theoretical maximum sensitivity per slice of the PALS method of 33.3%, corresponding to 100% divided by the number of slices, and represent 88% of the whole sample signal intensity.

The PALS method was applied to speed up the acquisition of 2D COSY, 2D DQF-COSY and 2D TQF-COSY with the test sample. The pulse sequences of these experiments are given in Fig. 1b and c and were simply adapted for localized spectroscopy.^[22–25] The same slice is chosen consistently to pulse at a given PALS spectrum in order to achieve perfect coherence of phase and intensity along the t_1 dimension of each spectrum. The three localized spectra were acquired with the previous calibration of the slices of Figures S1 and S2 and with the total scan time exclusively dedicated to pulsing and acquisition of each FID data points. Using a single scan and 128 increments in the t_1 dimension, the acquisition of the three spectra took only 3 min with ~ 0.3 s per scan.

The quality of the spectra was comparable to those acquired conventionally with good sensitivity and complete lack of artifacts (Figure S3). Figure 2 compares the standard TQF-COSY using 90° hard pulses with the whole active sample volume and the PALS localized TQF-COSY, each of them acquired in 1 min. It is evident in the spectrum of Fig. 2a, the presence of a number of artifacts

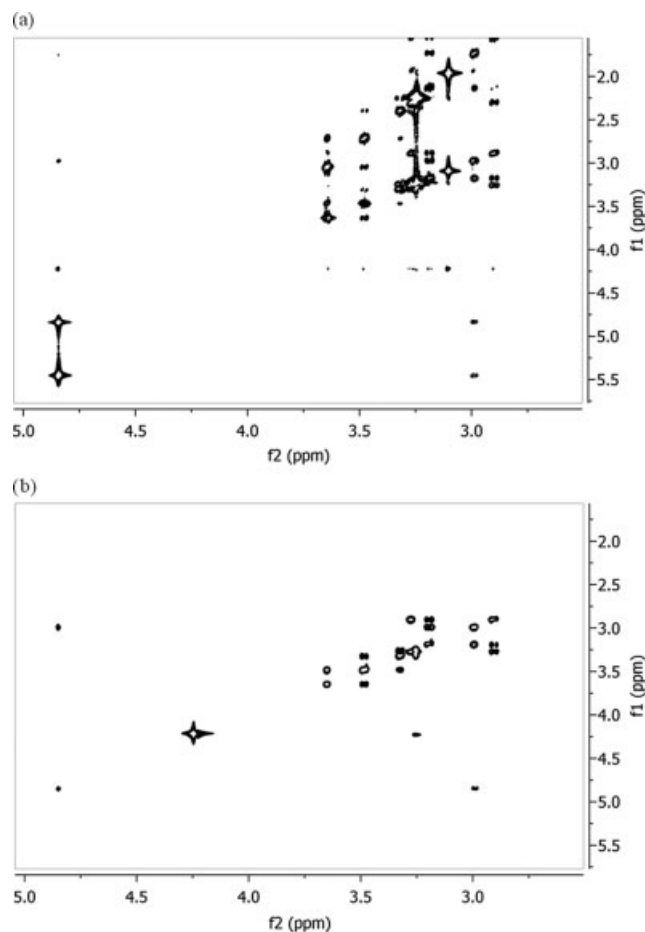


Figure 2. (a) Conventional 2D TQF-COSY using hard 90° pulses (full-sample volume) and standard acquisition and (b) PALS 2D TQF-COSY (slice selective). Each spectrum was acquired in 1 min, with ~ 0.3 s per scan, 128 t_1 points and a single scan per t_1 increment. They are processed identically and represented at the same contour level.

that include a double-quantum diagonal generated by interscan refocusing of undesired coherences due to the fast acquisition. The PALS DQF-COSY spectrum of Fig. 2b does not have any of these artifacts.

The processing of the three PALS spectra obtained requires a simple data shuffling operation to produce three spectral data matrices formed by extraction of the correspondingly t_1 row of each 2D spectrum. Once these matrices are created, they can be transformed using traditional 2D-NMR data processing.

In conclusion, the parallel PALS method experiment affords a considerable timesaving in the acquisition of a number of nD NMR spectra. The method is recommended for samples not very limited by concentration and for nD experiments with enough sensitivity per scan. The penalty in sensitivity depends on the sample thickness dedicated to each spectrum. With proper calibration, a penalty factor of ~ 3 was obtained in each of the three PALS spectra presented herein. Some of the unique advantages of this fast-NMR method are that it does not impose limits on the number of points to be detected in any dimensions, it is not prone to generate artifacts and the data can be processed with well-established conventional methods. The method is also compatible with solvent-suppression schemes that rely in conventional selective pulses (i.e. nonspatially selective) such as WET.^[26] An extension of this fast-NMR strategy to other nD NMR

experiments should be relatively straightforward, for instance, to speed up the measurements of relaxation times or diffusion coefficients among other experiments. Furthermore, the method could also be used in combination with some of the fast-NMR methods already published.

Acknowledgements

M.V.V. thanks the funding of the European Commission (project artizymes FP6-2003-NEST-B3 15471). C. C thanks the financial support of the Xunta de Galicia (Grant PGIDT05SIN012E).

Supporting information

Supporting information may be found in the online version of this article.

References

- [1] R. Freeman, E. Kupce, *J. Biomol. NMR* **2003**, *27*, 101.
- [2] C. Bougault, L. M. Feng, J. Glushka, E. Kupce, J. H. Prestegard, *J. Biomol. NMR* **2004**, *28*, 385.
- [3] J. H. Chen, V. A. Mandelshtam, A. J. Shaka, *J. Magn. Reson.* **2000**, *146*, 363.
- [4] S. Kim, T. Szyperski, *J. Am. Chem. Soc.* **2003**, *125*, 1385.
- [5] H. R. Eghbalnia, A. Bahrami, M. Tonelli, K. Hallenga, J. L. Markley, *J. Am. Chem. Soc.* **2005**, *127*, 12528.
- [6] H. R. Eghbalnia, A. Bahrami, M. Tonelli, K. Hallenga, J. L. Markley, *J. Am. Chem. Soc.* **2005**, *127*, 12528.
- [7] L. Frydman, T. Scherf, A. Lupulescu, *Proc. Natl. Acad. Sci. USA* **2002**, *29*, 15858.
- [8] M. Mishkovsky, L. Frydman, *Annu. Rev. Phys. Chem.* **2009**, *60*, 429.
- [9] P. Schanda, B. Brutscher, *J. Am. Chem. Soc.* **2005**, *127*, 8014.
- [10] P. Schanda, E. Kupče, B. Brutscher, *J. Biomol. NMR* **2005**, *33*, 199.
- [11] P. Schanda, B. Brutscher, *J. Magn. Reson.* **2006**, *178*, 334.
- [12] E. Lescop, P. Schanda, B. Brutscher, *J. Magn. Reson.* **2007**, *187*, 163.
- [13] D. M. Parish, T. Szyperski, *J. Am. Chem. Soc.* **2008**, *130*, 4925.
- [14] S. Cai, S., C. Seu, Z. Kovacs, A. D. Sherry, Y. Chen, *J. Am. Chem. Soc.* **2006**, *128*, 13474.
- [15] K. Pervushin, B. Vögeli, A. Eletsky, *J. Am. Chem. Soc.* **2002**, *124*, 12898.
- [16] P. T. Callaghan, *Principles of Nuclear Magnetic Resonance Microscopy*, Clarendon Press: Oxford, **1991**.
- [17] E. Kupce, R. Freeman, B. K. John, *J. Am. Chem. Soc.* **2006**, *128*, 9606.
- [18] E. Kupče, S. Cheatham, R. Freeman, *Magn. Reson. Chem.* **2007**, *45*, 103.
- [19] H. Wang, L. Ciobanu, A. Webb, *J. Magn. Reson.* **2005**, *173*, 134.
- [20] Y. Cohen, L. H. Chan, L. Litt, T. L. James, *J. Magn. Reson.* **1989**, *85*, 203.
- [21] A. A. Shaw, C. Salaun, J. F. Dauphin, B. Ancian, *J. Magn. Reson.* **1996**, *120*, 110.
- [22] A. P. Prescott, A. S. K. Dzik-Jurasz, M. O. Leach, B. Sirohi, R. Powles, D. J. Collins, *J. Magn. Reson. Imag.* **2005**, *22*, 541.
- [23] N. M. Loening, M. J. Thrippleton, J. Keeler, R. G. Griffin, *J. Magn. Reson.* **2003**, *164*, 321.
- [24] M. A. Thomas, K. Yue, N. Binesh, P. Davanzo, A. Kumar, B. Siegel, M. Frye, J. Curran, R. Lufkin, P. Martin, B. Guze, *Magn. Reson. Med.* **2001**, *46*, 58.
- [25] F. Delmas, J.-C. Beloeil, B. P. J. van der Sanden, K. Nicola, B. Gillet, *J. Magn. Reson.* **2001**, *149*, 119.
- [26] S. H. Smallcombe, S. L. Patt, P. A. Keifer, *J. Magn. Reson. Ser. A* **1995**, *117*, 295.

A NMR reverse diffusion filter for the simplification of spectra of complex mixtures and the study of drug receptor interactions

M. Vega-Vázquez,^a J. C. Cobas,^b F. F. Oliveira de Sousa^c
and M. Martin-Pastor^{a*}



A reverse diffusion filter NMR experiment (Drev) is proposed for the study of small molecules in binding with macromolecules. The filtering efficiency of Drev to eliminate the signals of the macromolecule is shown to be superior to conventional transverse relaxation filters at least for macromolecules containing a significant fraction of flexible residues. The Drev filter was also a useful complement for ligand-based NMR screening in combination with saturation transfer difference experiments. Copyright © 2011 John Wiley & Sons, Ltd.

Supporting information may be found in the online version of this article.

Keywords: STD; diffusion filter; relaxation filter; $T_{1\rho}$ filter; NMR screening

Introduction

One of the principal NMR characteristics which turn it into such a powerful tool is its versatility.^[1] Nowadays, NMR allows not only the study of the structures and dynamics of small or macromolecules in their pure state but also the analysis of complex raw mixtures without a need to perform physical separation of the compounds,^[2] or to study ligand–receptor complex formation.^[3–6]

In this context, the relentless advances in hardware and the development of new pulse sequences have opened up the possibility to choose, with great selectivity, which compounds may be observed in the presence of other compounds. Thus, for example, in binding studies of small ligands to macromolecular receptors in which the signals from the macromolecule may overlap with those of the ligand and obscure the analysis, it is possible to resort to experiments which include period/s of signal filtering according to a given NMR property so as to introduce a higher degree of signal attenuation for certain signals in the mixture, typically the macromolecule, thus contributing to simplifying the spectrum. The most common strategies to achieve this purpose are based on relaxation filters, normally transversal relaxation, although longitudinal relaxation filters have also been described,^[7–9] and diffusion filters.^[7,10] In this work we present the reverse diffusion filter (Drev filter), a tool for signal filtering which is based in the difference spectroscopy method published by Hajduk *et al.*,^[7] and which represents a complement to currently used methods for the identification of hits in the early stage of drug discovery, including high-throughput screening (HTS), Fragment-based drug discovery (FBDD)^[11] and detection of interactions of small drugs with receptors expressed on the surface of cells by NMR.^[12,13] In this context, the main advantage of the Drev filter is that it achieves an efficient suppression of the signals of the macromolecule/s in a way that is basically independent of their relaxation. The method is especially relevant for the suppression of signals of the most flexible residues of the macromolecule/s (e.g. certain aromatic

or aliphatic protons in pendant chains) whose local dynamics and therefore relaxation properties differ significantly from those of the macromolecular core.^[14] Thus, the enhanced mobility of those flexible residues could result in those signals presenting considerably longer transversal relaxation times and therefore being inefficiently suppressed by transversal relaxation filters.

In what follows, the Drev experiment is described and some of its applications are tested, both as a standalone tool and in combination with relaxation filters and STD experiments.

Results and Discussion

The pulse sequence of the reverse diffusion filter (1D Drev) is shown in Fig. 1. 1D Drev is based on the stimulated echo sequence with bipolar gradients for the measurement of diffusion coefficients. Its most unique and distinct feature is the use of difference spectroscopy in the experiment: FIDs corresponding to two consecutive scans, one of them with the diffusion gradient codification (G_z) at maximum power and the other with the diffusion gradient G_z switched off or at minimum power (G_z on/off in Fig. 1), are subtracted.

The result of applying the Drev filter is a spectrum in which signals that are less attenuated by the diffusion encoded gradients

* Correspondence to: M. Martin-Pastor, Unidad de Resonancia Magnética, Edif. CACTUS, RIAIDT, Campus Vida, Univ. Santiago de Compostela, A Coruña 15706, Spain. E-mail: manuel.martin@usc.es

a Unidad de Resonancia Magnética, Edif. CACTUS, RIAIDT, Campus Vida, University of Santiago de Compostela, A Coruña 15706, Spain

b Mestrelab Research S. L., Feliciano Barrera 9B –Bajo, 15706 Santiago de Compostela, A Coruña 15706, Spain

c Facultade de Farmacia, Campus Vida, University of Santiago de Compostela, A Coruña 15706, Spain

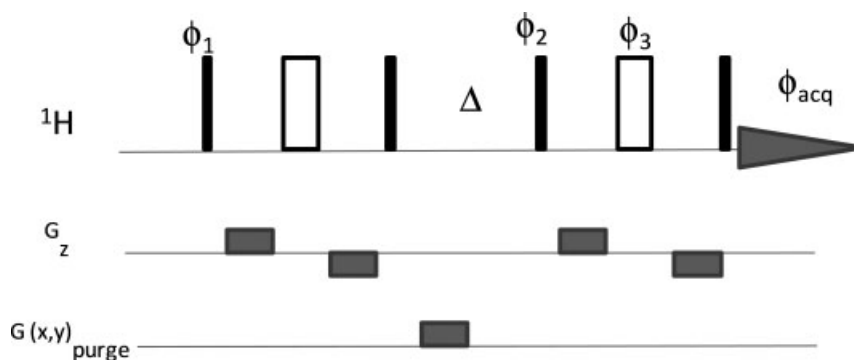


Figure 1. Pulse sequence for the 1D experiment. On the ^1H channel narrow and wide lines correspond to hard pulses of 90° and 180° , respectively. The phase cycle is $\phi_1: 0^{\wedge}4, 1^{\wedge}4; \phi_2: 0, 0, 2, 2, 0, 0, 2, 2, 1, 1, 3, 3, 1, 1, 3, 3, 2, 2, 0, 0, 2, 2, 0, 0, 3, 3, 1, 1, 3, 3, 1, 1; \phi_3: 0^{\wedge}16, 2^{\wedge}16; \phi_{\text{acq}}: 2, 0, 0, 2, 3, 1, 1, 3, 1, 3, 3, 1, 2, 0, 0, 2, 2, 0, 1, 3, 3, 1, 3, 1, 1, 3, 0, 2, 2, 0, 0, 2, 3, 1, 1, 3, 3, 1, 2, 0, 0, 2; G_z$ on/off: 1, 0 (power of gradient G_z has its maximum and minimum value on alternating scans). Non-explicit phases correspond to those on the x-axis. Δ is a constant period for diffusion evolution.

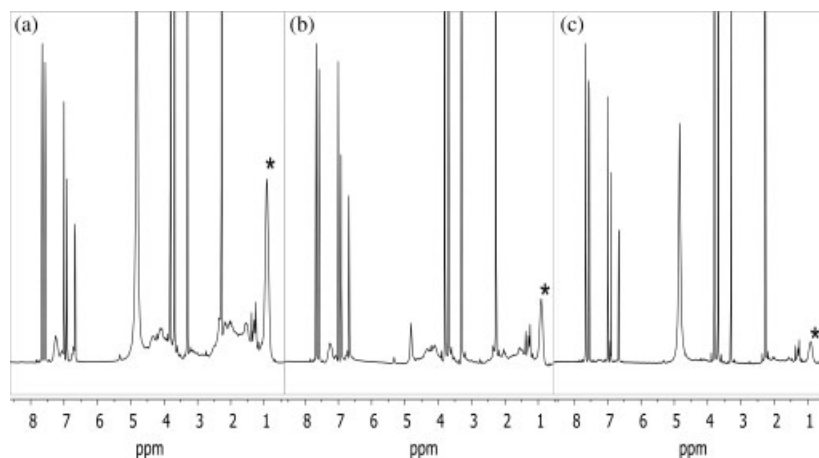


Figure 2. Spectra from an IND/Zein 40: 1 mixture (a) 1D ^1H , (b) 1D CPMG 50 ms and (c) 1D Drev 50 ms. The spectra were acquired with 128 scans and are displayed with the same vertical scale for the IND aromatic signal at 2',6'. The Zein aliphatic signal used for attenuation quantification is labeled with an asterisk.

of each scan, i.e. those of macromolecules or high molecular weight aggregates, are more significantly attenuated in such difference spectrum. The enhancement factor achieved for the signals of the small molecule over those belonging to the macromolecule is proportional to their difference in diffusion coefficient. The most critical parameter requiring optimization in the Drev experiment is the diffusion time Δ . Using the Stejskal–Tanner equation and neglecting relaxation effects, the optimized diffusion time, Δ_{opt} , that affords the maximum enhancement factor, is given by Eqns (1) and (2) for experiments using rectangular and half-sinusoidal diffusion gradients, respectively^[15]:

$$\Delta_{\text{opt}} = \frac{\ln(D_s/D_f)}{-\gamma^2 G_{\text{max}}^2 \delta^2 (D_f - D_s)} + \frac{\delta}{3} \quad (1)$$

$$\Delta_{\text{opt}} = \frac{4 \ln(D_s/D_f)}{-\gamma^2 G_{\text{max}}^2 \delta^2 (D_f - D_s)} + \frac{4\delta}{\pi^2} \quad (2)$$

where D_s and D_f are the diffusion coefficients of the slow and fast species, respectively, G_{max} is the maximum gradient power (the minimum gradient power is 0), δ is the length of the pulsed field gradients and γ is the magnetogyric constant of the nucleus. When gradients of common duration around 1–3 ms are used and their power is alternated in different scans of the experiment from

maximum to minimum available, typically small values of Δ_{opt} are required (e.g. 50–100 ms).

We have tested the performance of the 1D Drev experiment with a mixture of the ligand indomethacin (IND) at a concentration of 33 mM in CD_3OD (Supporting Information Fig. S1) and Zein protein^[16–18] prepared at a molar ratio IND/Zein of 40:1. A previous study carried out by our research group^[19] had shown that IND has two association constants with Zein of K_a 1859 ± 955 and $14.35 \pm 10 \text{ M}^{-1}$. Their range is in a convenient range for STD-based experiments that are discussed later in the text.

The spectra of Fig. 2 show the result of applying the Drev filter to the IND/Zein 40:1 sample. In the 1D ^1H spectrum (Fig. 2(a)), the Zein signals appear as a signal background clearly identifiable by its greater width as compared with the IND signals. Taking as a reference the intensity of the aromatic signal 2',6' of IND, the aliphatic Zein signals at ~ 0.92 ppm are attenuated in the 1D Drev spectrum (Fig. 2(c)) by a factor of 9.0 with respect to the original 1D ^1H spectrum. It can also be shown that the IND signals in the 1D Drev spectrum (Fig. 2(c)) still retain in good measure the original ratio between integrals. As a comparison, in the experiment with transversal relaxation filter 1D CPMG (Fig. 2(b)) with a filter time of 50 ms, the same aliphatic signals of Zein are attenuated only by a factor of 2.9.

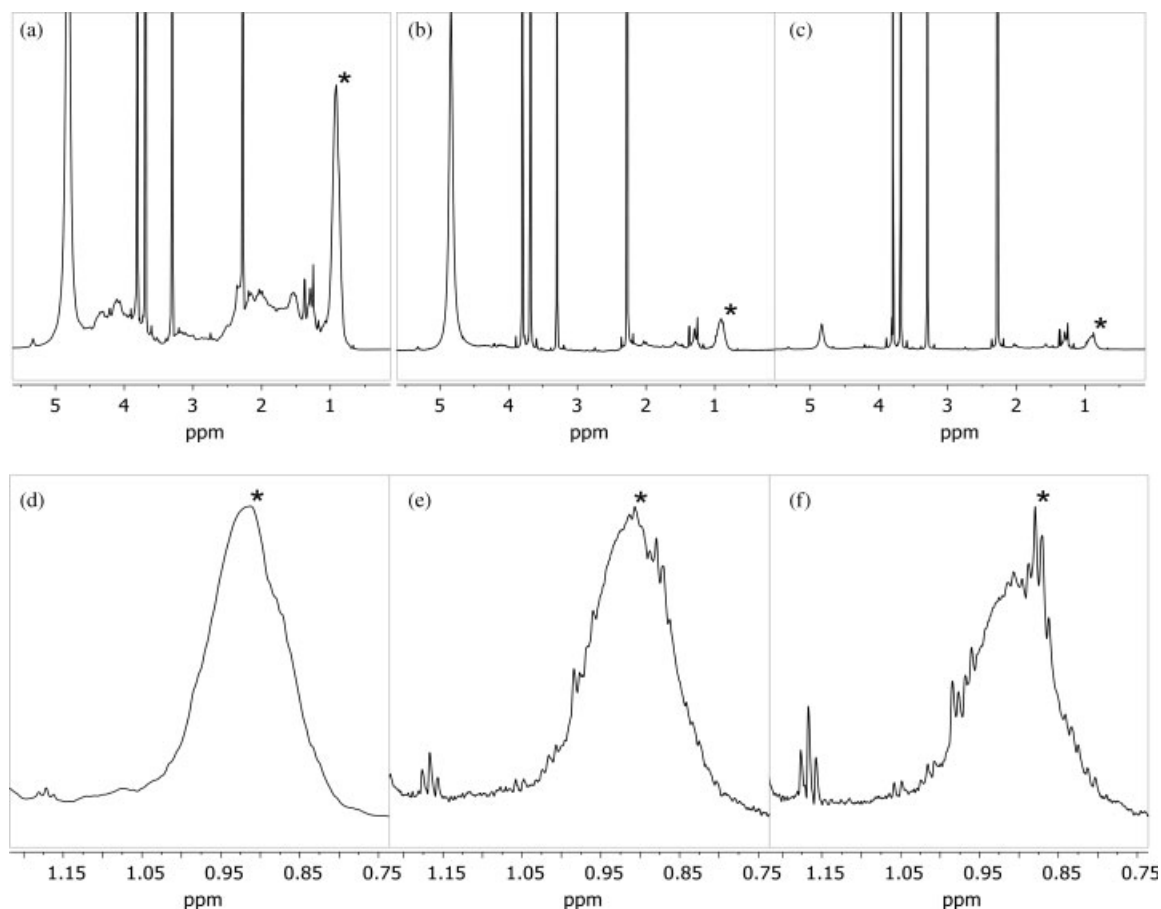


Figure 3. Spectra of the IND/Zein 40 : 1 mixture. (a) 1D ^1H , (b) 1D Drev and (c) 1D CPMG-Drev. The duration of the CPMG and Drev filters in these spectra was 50 ms each. Spectra a, b and c were acquired with 128 scans and are represented with the same vertical scale for the IND aromatic signal at 2',6'. Spectra d, e and f show an expansion of the aliphatic region from spectra a, b and c, respectively. The Zein aliphatic signal used for attenuation quantification is labeled with an asterisk.

Combined transversal relaxation filter and reverse diffusion filter: 1D CPMG-Drev

The filtering effect for signals of macromolecules or large size aggregates can be further enhanced by the combination in the same experiment of a transversal relaxation filter such as CPMG and the Drev filter. The combined filter CPMG-Drev was implemented in the pulse sequence of Supporting Information Fig. S2. In this sequence, both filters are applied sequentially and thus the net filtering effect to be expected is given by the product of the two filters.

The spectra in Fig. 3 and Supporting Information Fig. S3 show the performance of the 1D CPMG-Drev filter for the IND/Zein 40 : 1 sample with the total filter length of 100 ms. The attenuation factor in peak signal intensity achieved is 19.9 for the Zein aliphatic signal at 0.92 ppm, as compared with attenuation factors of 2.9 and 8.4 for the 1D CPMG spectra with 50 and 100 ms filters, respectively. The expansion of this spectral region shown in Fig. 3(d–f) shows how the utilization of the 1D CPMG-Drev sequence reveals, below the wide Zein aliphatic signal, some narrow signals undetectable in the 1D ^1H spectrum, which are, in all likelihood, impurities from compounds with low molecular weight.

A systematic exploration of the influence of the filter duration on the efficacy of 1D CPMG, 1D Drev and 1D CPMG-Drev filters also confirmed the superior capacity of the later two methods in the range of filter lengths up to 100 ms to suppress the signals of the

macromolecule in the sample IND/Zein (Supporting Information Fig. S4). The remarkable filtering capability and spectral quality obtained by 1D Drev or 1D CPMG-Drev has no paragon with other methods. These filters afford the capability to gain precise information regarding chemical shifts and couplings constants of small molecules together with quantitative or semi-quantitative signal integrals for molecules of similar size. They could find applications in many situations, in particular in the determination of binding affinities by titration studies,^[20] or to detect ligand signals in target proteins containing paramagnetics in the binding site.

Saturation transfer difference experiments with reverse diffusion filter: 1D STD-Drev and 1D-STD CPMG-Drev

In the last decade, a number of ligand-based NMR screening methods have been proposed for the detection of ligands with sufficient affinity to form complexes with a target macromolecular receptor. Some of these experiments are: saturation transfer difference (STD),^[3–5] WaterLOGSY (water–ligand observed via gradient spectroscopy),^[21,22] NOE pumping^[23] and reverse NOE pumping.^[24] These methods not only can show the evidence of the binding of a certain drug in a mixture of compounds but also identify the signals involved in the binding epitope, i.e. the areas of the ligand molecule which are interacting directly with the target receptor. Moreover, under controlled competitive conditions they

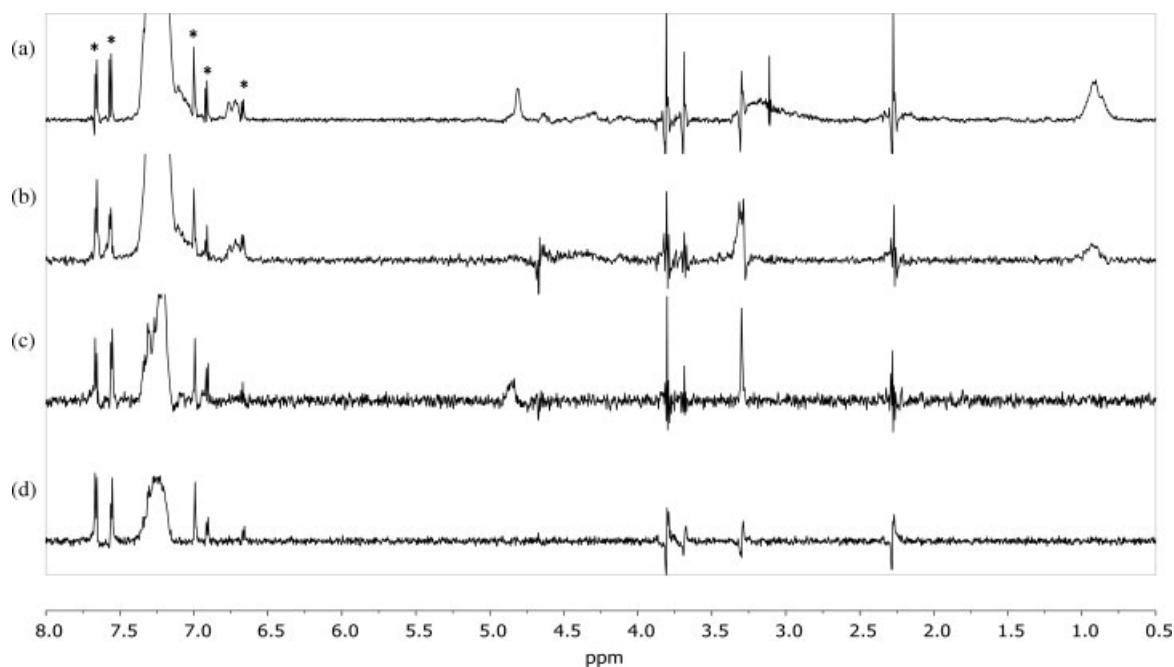


Figure 4. Spectra for the IND/Zein 40:1 mixture. (a) 1D STD (128 scans), (b) 1D STD-CPMG with a 50-ms filter (128 scans), (c) 1D STD-Drev with a 50-ms filter (512 scans) and (d) 1D-STD-CPMG-Drev with 50 ms CPMG filter and 50 ms Drev filter (512 scans). In spectrum, (a) the IND STD signals are marked with an asterisk; all other signals are Zein protein signals or subtraction artifacts. Additional experimental details are given in the Supporting Information.

allow the exploration of the specificity of the binding of a drug to the target active site.^[6]

The STD experiment is nowadays the most popular and robust experiment used for ligand screening. To eliminate the signals of the macromolecular receptor, the experiment is commonly used together with a transversal relaxation filter such as $T_{1\rho}$ ^[25] or CPMG^[26] (Supporting Information Fig. S5(a)). This work has tested the introduction of Drev and CPMG-Drev filters in STD experiments as alternative filtering methods. The new pulse sequences developed (named 1D STD-Drev and 1D STD-CPMG-Drev experiments) are shown in Supporting Information Fig. S5(b) and (c), respectively.

The mixture of IND/Zein 40:1 was used to test the performance of the 1D STD-Drev and 1D-CPMG-Drev experiments, and their results were compared with those obtained with the conventional 1D STD and 1D STD-CPMG experiments (Fig. 4). All the spectra in Fig. 4 show that the five aromatic signals of IND are the only ones showing STD responses and therefore are the ones which belong to the binding epitope of Zein. In addition, the influence of the filtering method in the attenuation of the protein signal can also be observed. Filtering is clearly enhanced in the 1D STD-Drev experiment and it is further improved in the 1D STD-CPMG-Drev experiment.

It must be noted that due to the stimulated echo method used for the Drev filter, the sensitivity of 1D STD-Drev and 1D STD-CPMG-Drev experiments is theoretically inferior by a factor 1/2 with respect to the analogous $T_{1\rho}$ or CPMG filtered experiments of the same duration. Further losses in sensitivity could occur if the length of the Drev filter (Δ) is not properly calibrated to maximize the signals of the ligand. However, with proper calibration of the Drev filter, the experimental results obtained for the sample IND/Zein in the spectra STD-Drev and STD-CPMG-Drev of Fig. 4(c) and (d), respectively were very consistent with the mentioned theoretical factor.

As conclusion, the new filtered experiments proposed Drev and Drev-CPMG can achieve a high degree of suppression of the background signals of macromolecules in mixtures with small ligands. Experimentally, suppression factors of more than one order of magnitude were obtained compared with the analogue T_2 or $T_{1\rho}$ filters of the same duration. The introduction of these new filters in STD experiments can be a convenient method for NMR-based screening of small ligands that bind to macromolecules. Despite their lower sensitivity, they afford a clean detection of the signals of the ligand-binding epitope in regions of the spectrum that overlap with the macromolecule. In this context, the new filters proposed are also recommended when the fast transversal relaxation for the signal of the ligand imposes limits to the maximum duration of the filters required to suppress the signals of the macromolecules.

Acknowledgement

We thank Santiago Dominguez for helpful discussions.

Supporting information

Supporting information may be found in the online version of this article.

References

- [1] W. Jahnke, *J. Biomol. NMR* **2007**, *39*, 87.
- [2] K. F. Morris, C. S. Johnson Jr., *J. Am. Chem. Soc.* **1992**, *114*, 3139.
- [3] M. Mayer, B. Meyer, *Angew. Chem. Int. Ed.* **1999**, *38*, 1784.
- [4] J. Klein, R. Meinecke, M. Mayer, B. Meyer, *J. Am. Chem. Soc.* **1999**, *121*, 5336.
- [5] B. Meyer, T. Peters, *Angew. Chem. Int. Ed.* **2003**, *42*, 864.
- [6] Y.-S. Wang, D. Liu, D. F. Wyss, *Magn. Reson. Chem.* **2004**, *42*, 485.
- [7] P. J. Hajduk, E. T. Olejniczak, S. W. Fesik, *J. Am. Chem. Soc.* **1997**, *119*, 12257.

- [8] L. Ma, E. Philipp, J. J. Led, *J. Biomol. NMR* **2001**, *19*, 199.
- [9] V. Borlototti, P. Fantazzini, M. Gombia, D. Greco, G. Rinaldin, S. Sykora, *J. Magn. Reson.* **2010**, *206*, 219.
- [10] N. Esturau, J. F. Espinosa, *J. Org. Chem.* **2006**, *71*, 4103.
- [11] C. Dalvit, *Drug Discov. Today* **2009**, *14*, 1051.
- [12] B. Claasen, M. Axmann, R. Meinecke, B. Meyer, *J. Am. Chem. Soc.* **2005**, *127*, 916.
- [13] S. Mari, D. Serrano-Gómez, F. J. Cañada, A. L. Corbí, J. Jiménez-Barbero, *Angew. Chem. Int. Ed.* **2005**, *44*, 296.
- [14] J. Kowalewski, L. Maler, in: *Nuclear Spin Relaxation in Liquids: Theory, Experiments, and Applications* (Eds: J. H. Moore, N. D. Spencer), Taylor & Francis: Oxford, **2006**, pp 244.
- [15] W. S. Price, *Concepts Magn. Reson.* **1998**, *10*, 197.
- [16] Y. Wang, A. M. Rakotonirainy, G. W. Padua, *Starch/Stärke* **2003**, *55*, 25.
- [17] S. Subramanian, S. Sampath, *Biomacromolecules* **2007**, *8*, 2120.
- [18] N. Matsushima, G. Danno, H. Takezawa, Y. Izumi, *Biochim. Biophys. Acta Protein Struct. Mol. Enzymol.* **1997**, *1339*, 14.
- [19] F. F. Oliveira de Sousa, A. Luzardo-Álvarez, J. Blanco-Méndez, M. Martín-Pastor, *Macromol. Biosci.* **2011**, (in press).
- [20] L. Fielding, *Prog. Nucl. Magn. Reson. Spectrosc.* **2007**, *51*, 219.
- [21] C. Dalvit, P. Pevarello, M. Tatò, M. Veronesi, A. Vulpetti, M. Sundström, *J. Biomol. NMR* **2000**, *18*, 65.
- [22] C. Dalvit, G. P. Fogliatto, A. Stewart, M. Veronesi, B. Stockman, *J. Biomol. NMR* **2001**, *21*, 349.
- [23] A. Chen, M. J. Shapiro, *J. Am. Chem. Soc.* **1998**, *120*, 10258.
- [24] A. Chen, M. J. Shapiro, *J. Am. Chem. Soc.* **2000**, *122*, 414.
- [25] M. Mayer, B. Meyer, *J. Am. Chem. Soc.* **2001**, *123*, 6108.
- [26] T. Diercks, J. P. Ribeiro, F. J. Cañada, S. André, J. Jiménez-Barbero, H.-J. Gabius, *Chem. Eur. J.* **2009**, *15*, 5666.

Article

Enhanced signal dispersion in saturation transfer difference experiments by conversion to a 1D-STD-homodecoupled spectrum

Manuel Martín-Pastor^{a,*}, Marino Vega-Vázquez^a, Antonia De Capua^b, Angeles Canales^c, Sabine André^d, Hans-Joachim Gabius^d & Jesús Jiménez-Barbero^{c,*}

^aLaboratorio Integral de Dinámica e Estructura de Biomoléculas José R. Carracido, Unidade de Resonancia Magnética, Edificio CACTUS, RIAIDT, Universidade de Santiago de Compostela, 15706 Santiago de Compostela, Spain; ^bDipartimento di Scienze Ambientali, Seconda Università degli Studi di Napoli, via Vivaldi 43, 81100 Caserta, Italy; ^cDepartamento de Estructura y función de proteínas, Centro de Investigaciones Biológicas, CSIC, Ramiro de Maeztu 9, 28040 Madrid, Spain; ^dInstitut für Physiologische Chemie, Tierärztliche Fakultät, Ludwig-Maximilians-Universität, Veterinärstr. 13, 80539 München, Germany

Received 25 January 2006; Accepted 7 July 2006

Key words: J-resolved spectroscopy, lectin, saturation transfer difference, sugar code, $T_{1\rho}$ relaxation filter

Abstract

The saturation transfer difference (STD) experiment is a rich source of information on topological aspects of ligand binding to a receptor. The epitope mapping is based on a magnetization transfer after signal saturation from the receptor to the ligand, where interproton distances permit this process. Signal overlap in the STD spectrum can cause difficulties to correctly assign and/or quantitate the measured enhancements. To address this issue we report here a modified version of the routine experiment and a processing scheme that provides a 1D-STD homodecoupled spectrum (i.e. an experiment in which all STD signals appear as singlets) with line widths similar to those in original STD spectrum. These refinements contribute to alleviate problems of signal overlap. The experiment is based on 2D-J-resolved spectroscopy, one of the fastest 2D experiments under conventional data sampling in the indirect dimension, and provides excellent sensitivity, a key factor for the difference experiments.

Abbreviations: J-resolved – J-resolved spectroscopy; STD – saturation transfer difference; wg – watergate.

Introduction

Current developments in NMR techniques are aimed to provide and refine tools for drug discovery that allow to probe specificity, affinity and structural aspects of receptor–ligand interaction (Meyer and Peters, 2003; Peng et al., 2004). Among the NMR methods suitable for analyzing

ligand binding such as transferred NOE (trNOE) (Meyer et al., 1997; Mayer and Meyer, 2000), inter-ligand NOE (Li et al., 1999; London, 1999; Li et al. 2001), diffusion (Lin and Shapiro, 1996, Lin et al., 1997, Yan et al. 2002), relaxation (Hadjuk et al., 1997, LaPlante et al., 2000), NOE pumping (Chen and Shapiro, 1998, 2000), quantification of the binding affinity by ligand release in competition experiments (Jahnke et al., 2002; Sirwardena et al., 2002) and water-LOGSY (Dalvit et al., 2000,2002). The elegant Saturation Transfer

*To whom correspondence should be addressed. E-mail: JJbarbero@cib.csic.es

Difference (STD) experiment has attained notable popularity (Mayer and Meyer, 1999). This experimental approach is based on magnetization transfer from a receptor after signal saturation to ligand protons in spatial vicinity. Acquisition of on- and off-resonance spectra with subsequent subtraction generates difference spectra that contain only signals of the ligand which are indicative of the contact area. There are several advantages which make the STD experiment attractive (1) it can be run without any labelling or immobilization of either the ligand or the receptor, (2) the size of the receptor does not become a factor limiting applicability, (3) it can be performed in experimental designs close to the physiological conditions using living cells (Claasen et al., 2005; Mari et al., 2005), (4) it can be used as a fast NMR screening method to pinpoint reactive compounds in a mixture (Mayer and Meyer, 1999; Klein et al., 1999), (5) it can define the critical epitope, that is, the portions of the ligand in close contact with the receptor (Mayer and Meyer, 2001), and (6) it can even provide information of the conformation of the bound ligand in the complex if strategically combined with modeling (Siebert et al., 2003; Hajduk et al., 2004; Di Micco et al., 2005; Yuen et al., 2005). A serious limitation is posed by signal overlap. To address this issue, this study has been performed.

Signal overlap in conventional 1D-STD spectra can preclude assignment or quantification of the enhancements. Therefore, modifications of the basic protocol of STD experiments such as 2D-STD-TOCSY (Mayer and Meyer, 1999; Vogtherr and Peters, 2000; Möller et al., 2002) or the selective version 1D-STD-selective TOCSY (Johnson and Pinto, 2002) have already been proposed. We report here a new approach for this type of NMR experiment and its protocol. Based on the emerging significance of protein-carbohydrate interactions for drug design and the regulation of various cellular activities we test our technique with an animal lectin (Reuter and Gabius, 1999; Gabius et al., 2004; Gabius, 2006). The introduced methodological modification enabled us to obtain a 1D-STD Homodecoupled spectrum (i.e. an experiment in which all STD signals appear as singlets) with line widths similar to the original STD spectrum. Our results suggest that the reported protocol adds a new option to deal with signal overlap in STD analysis.

Materials and methods

All spectra were acquired at 15°C on a Varian INOVA 750 MHz spectrometer equipped with a triple gradient shielded probe and processed with MestRe-C v 3.x software (Cobas and Sardina, 2004) (*Mestrelab research*). The test samples containing 12 mg of lactose and 3 mg of chicken galectin CG-14 was prepared in 600 µl NaCl/P_i (buffer concentration 20 mM, H₂O/D₂O = 9:1, 0.05 % NaN₃) at pH 7.0. The ligand to protein ratio was about ~200:1. Furthermore, the pulse sequence was also successfully tested using a second ligand, a complex tetrasaccharide:methyl 4-O-(3-O-(2-acetamido-2-deoxy-4-O-β-D-galactopyranosyl)-β-D-glucopyranosyl)-β-D-galactopyranosyl)-β-D-glucopyranoside. The molar ratio tetrasaccharide:lectin was also ~200:1.

Two 2D-STD-T_{1ρ}-Jres-wg (2D-STD-T_{1ρ}-J-resolved-watergate) spectra were acquired with the protocol steps given in Figure 1. The total saturation time was 1s (and d₁ = 1s) and 2s (and d₁ = 0s), respectively, while all other parameters to record and process the spectra were the same. The selective on-resonance/off-resonance saturation was placed at ~7.5 ppm (signals CG-14)/20 ppm (empty region). Saturation was accomplished by a series of Gaussian pulses of 50 ms pulse length and a 32 Hz field separated by a delay of 1 ms. The T_{1ρ} filter used to suppress protein signals is applied as a continuous-wave spin-lock pulse of 60 ms with a 2 kHz field. The selective 90° pulses used for solvent suppression with the Watergate method (Piotto et al., 1992) were set to the centre of the on-resonance water signal, shaped as *sinc* pulses with a pulse width of about 1.1 ms, and the power was adjusted to obtain optimal solvent suppression. Gradients were of rectangular shape, and their duration of 1 ms was followed by a stabilization delay period of 0.2 ms. The strength of G_w was adjusted to 30 G/cm, a value that was found to provide a convenient degree of water suppression. The spectral width in the acquisition (proton) and J-resolved dimensions was 9500 and 65 Hz, respectively. The spectrum was recorded with 128 scans and 48 increments in the J-resolved dimension. The total completion time of the experiment was ~4 h.

The processing of the 2D-STD-T_{1ρ}-Jres-wg experiment to obtain the 1D-STD-homodecoupled experiment is based on the method proposed by Guenneau et al. (1999). It starts with remov-

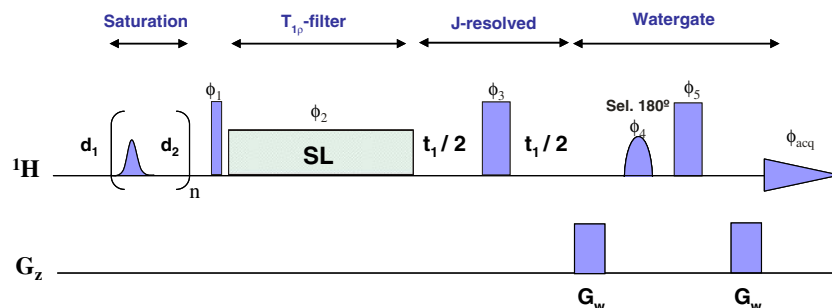


Figure 1. Scheme of the 2D-STD- $T_{1\rho}$ -Jres-wg (2D-STD- $T_{1\rho}$ -J-resolved-watergate) experiment. Filled and open rectangles correspond to 90° and 180° hard pulses, d_1 is an optional relaxation delay. The selective saturation of the protein is accomplished as a series of soft Gaussian pulses separated by a short delay (d_2). The $T_{1\rho}$ filter is operated as a continuous-wave spin-lock pulse. In the Watergate module, the gradient power level, G_w and the selective 180° pulse should be calibrated to suppress the strong solvent signals. On alternate control scans, saturation is tested outside the range of protein resonances, and subtraction is attained by phase cycling, i.e. ϕ_1 : $x, x, -x, -x, y, y, -y, -y, \phi_2$: $y, y, -y, -y, -x, -x, x, x, \phi_3$: $x, x, x, x, y, y, y, y, \phi_4$: $x, -x, -x, x, y, -y, -y, y, \phi_5$: $-x, x, x, -x, -y, y, y, -y, -y$, and ϕ_{acq} : $x, -x, -x, x, y, -y, -y, y$. Pulses whose phase is not indicated correspond to the x -axis.

ing the residual solvent signal using a high-pass filter function. The spectrum was then Fourier-transformed along the acquisition (proton) and J-resolved dimensions with a 45° and 90° shifted sinebell apodization in F2 and F1, respectively, and zero filling was applied in both dimensions to extend the data to 4096 and 128 complex points in F2 and F1, respectively. The spectrum was then tilted by 45° along the acquisition dimension, and a derivative algorithm was applied along F1 to enhance the resolution in this dimension. In this respect, instead of using the common discrete derivative method, a continuous wavelet derivative method (Wu et al., 2002) was applied with a factor of 10 for the dilation filter which provided an improved signal-to-noise ratio. This processing provided a 2D-STD-J-resolved spectrum that was represented in the power mode. The 1D STD-homodecoupled spectrum was generated by calculating the projection-sum along all the F2 rows (Guennenau et al., 1999; Cobas and Martín-Pastor, 2004) from the already processed 2D-STD- $T_{1\rho}$ -Jres-wg spectrum. A 1D-STD- $T_{1\rho}$ -wg (Wang et al., 2004) experiment was performed using conditions analogous to those described for the 2D-STD- $T_{1\rho}$ -Jres-wg experiment. For this purpose the same series as given in Figure 1 was used but omitting the t_1 evolution period from the protocol. The spectrum was recorded with 512 scans in 21 min. A 2D-HSQC-wg experiment was carried out under standard conditions. This spectrum was used as reference for the proton chemical shifts.

Results and discussion

Our aim was to design a new method to tackle the problem of severe signal overlap in STD experiments. In the 2D-STD- $T_{1\rho}$ -Jres-wg (2D-STD- $T_{1\rho}$ -J-resolved-watergate) sequence given in detail in Figure 1 the STD-NOE signals are spread along an indirect J-resolved dimension. Hereby, potential problems due to signal overlap which – when undetected – could cause errors in assignment or quantification of the STD-NOEs can be sorted out. The given procedure is suitable for binding studies with biomolecules. Subtraction errors in STD-NOEs caused by temperature fluctuations or sample aggregation during the acquisition are minimized by means of the phase cycling, attaining subtraction in consecutive scans during the on-resonance and off-resonance experiments. The duration of the $T_{1\rho}$ -filter element can be adjusted to destroy comparatively rapidly relaxing signals of a protein while preserving those of the small ligand under study (Hajduk et al., 1997). A strong degree of solvent suppression can be achieved by means of the Watergate element (grad, selective 180° , 180° , grad) (Piotto et al., 1992) which has been added after the t_1 incremented delay of the J-resolved element. Since our sequence extends the original 1D-STD protocol by a J-modulation period, which consists of a t_1 incremented delay and a 180° spin echo pulse by means of the Watergate element, its sensitivity per scan should not be compromised. In fact, we expect that the quality of this parameter is maintained. Because

the spectral width in Hz of the J-resolved dimension is relatively modest, the number of points required to digitize this dimension can even be relatively small. Typically, the J-resolved dimension can be acquired with ~ 64 points in the magnitude mode (i.e. not required to double the number of increments to obtain a pure phase-sensitive spectrum), which drastically reduces the total acquisition time in comparison to other 2D-STD spectroscopy (Mayer and Meyer, 1999; Vogtherr and Peters, 2000; Möller et al., 2002).

Following these methodological considerations, we document application of this type of 2D-STD- $T_{1\rho}$ -Jres-wg experiment to the study of the recognition of lactose by an animal lectin, belonging to the family of galectins (β -galactoside binding lectins) (Leonidas et al., 1998). This class of adhesion/growth-regulatory effectors influencing growth and invasion behavior of tumor cells by virtue of glycan binding is gaining increasing attention as target for drug design and as potential pharmaceuticals (Gabius, 1997; Gabius, 2001; Solís et al., 2001). As role model, we selected chicken galectin CG-14, one of the few monomeric galectins which had been tested previously in functional assays on tumor proliferation and neuronal differentiation (Kopitz et al., 2004; André et al., 2005). The binding affinity of galectins is typically in the micro- and millimolar range, satisfying the essential requirement for transferred-

NOE and STD experiments (Meyer and Peters, 2003; Dam et al., 2005). In fact, STD experiments are a very good alternative to the use of synthetically engineered carbohydrate ligands to deduce the binding epitope to lectin receptors (Rivera et al., 1992; Solís et al., 1993).

An essential step for the NOE study was to determine the exact positions of the proton resonances of lactose under the given experimental conditions. Thus, a $^1\text{H}/^{13}\text{C}$ 2D-HSQC-wg correlation experiment was instrumental to separate the proton resonances, which were then completely assigned based on earlier reports (Breg et al., 1988; Platzer et al., 1989).

For the interaction study, a ratio of $\sim 200:1$ with lactose and the lectin was prepared. The 2D-STD- $T_{1\rho}$ -Jres-wg difference spectrum obtained upon saturation at the aromatic region of CG-14 (sat. time 2 s) is shown in Figure 2. For comparison, the ^1H - $T_{1\rho}$ -filtered spectrum which shows all signals of lactose is added in the top part of this figure. The peaks observed in the 2D J-resolved spectrum are STD-NOEs generated through dynamic interaction of lactose with the lectin. At each proton chemical shift, the peaks are dispersed along the J-resolved dimension according to the respective scalar coupling multiplicity, aiding to identify any instance of signal overlap and to contribute to the reliability of signal assignment.

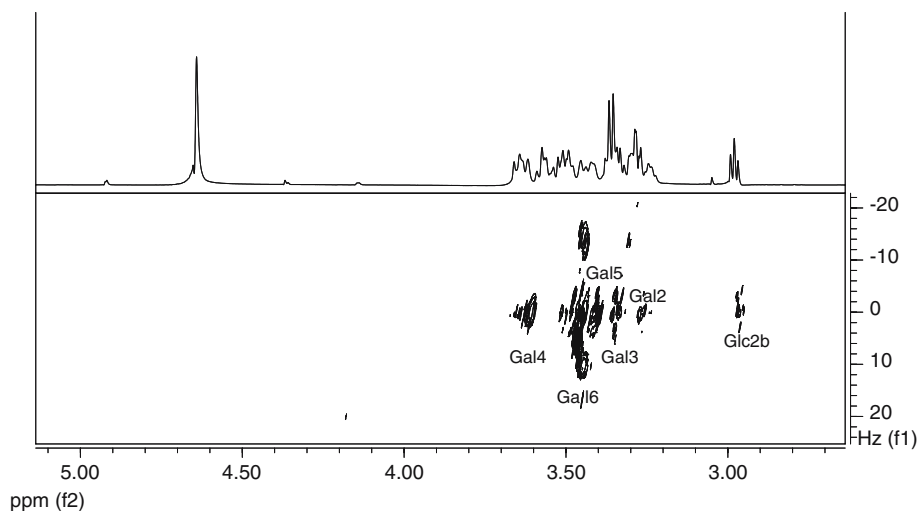


Figure 2. 2D-STD- $T_{1\rho}$ -Jres-wg difference spectrum of lactose and the chicken galectin CG-14 upon saturation of the aromatic signals of the protein (sat. time 2 s). The 1D ^1H - $T_{1\rho}$ -filtered spectrum of the sample is given for comparison (top panel). Protein signals were effectively filtered out in both spectra by means of the $T_{1\rho}$ filter. Proton labelled with b refer to the beta anomer of the glucose moiety of lactose.

Although it is certainly possible to analyze the 2D-STD-J-resolved spectrum in the common way to obtain the signal assignment as shown in Figure 2, an arising possibility described for J-resolved type spectra is to process the data to yield a 1D-Homodecoupled spectrum (i.e. a 1D spectrum in which all proton signals appears as singlets). Such a data presentation might improve the analysis.

From the different processing methods described in the literature to obtain a 1D-homodecoupled spectrum (Guenneau et al., 1999, and references therein), a particularly simple and robust method is given by Guenneau et al. (1999). In this method, a 2D J-resolved spectrum is apodized and processed with a Fourier transform along both dimensions, then the spectrum is tilted by 45° along the acquisition dimension and displayed in power mode to remove the phase-twist problem. The projection sum of all the F2 rows of the proton dimension directly provides the corresponding 1D-homodecoupled spectrum. The line widths achievable in this spectrum are similar to that in the original 1D proton spectrum and, of

note since the 1D-homodecoupled spectrum is built from a sum of 1D proton spectra, there is an important gain in terms of sensitivity. This parameter is important for experiments based on difference spectroscopy such as the STD. We therefore present the comparison regarding this factor next.

Figures 3a and b show the conventional 1D-STD- $T_{1\rho}$ -wg experiment (512 scans) and the 1D-STD-homodecoupled experiment (128 scans per increment), respectively, using a saturation time of 2 s. Apart from the obvious gain in terms of signal to noise ratio in the spectrum of Figure 3b, it can be seen that the complex multiplet signal at ~ 3.45 ppm in the 1D-STD spectrum of Figure 3a is separated into two partially overlapping singlets in the 1D-STD-homodecoupled spectrum of Figure 3b, while other the signals appear unaltered. Using a 2D-HSQC-wg spectrum as reference, it was possible to identify these two singlets as arising from Gal-6 and Gal-5 (Figure 3b). Other signals in the 1D-STD-homodecoupled spectrum do not show problems of overlap and were assigned without difficulties

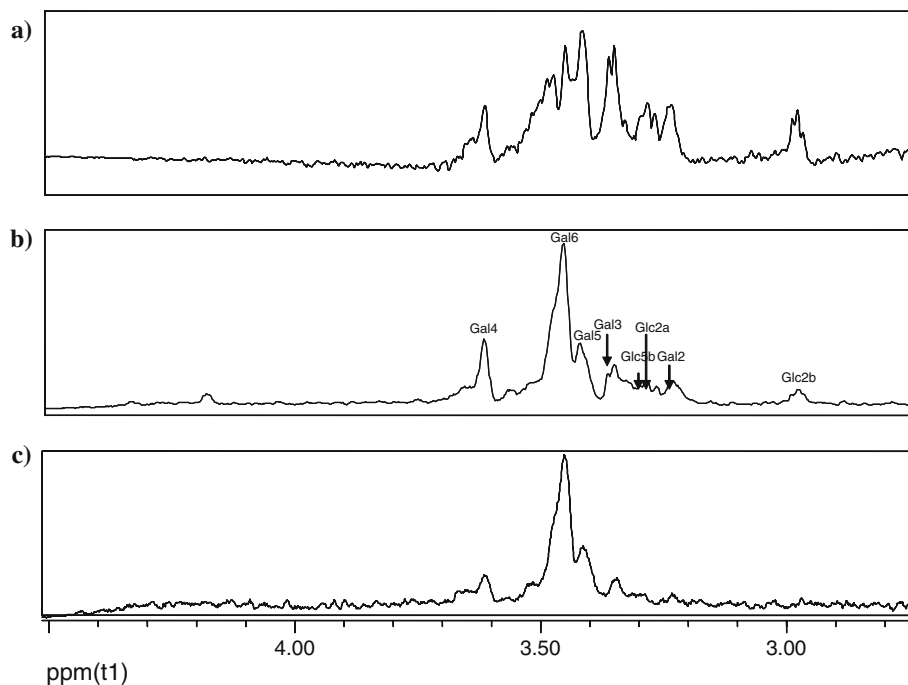


Figure 3. Comparison of STD spectra of the mixture of the chicken galectin CG-14 and its ligand lactose: (a) conventional 1D-STD experiment (sat time 2 s), (b) 1D-STD-Homodecoupled (sat. time 2 s), and (c) 1D-STD-Homodecoupled (sat. time 1 s). Protein signals were effectively filtered out in spectra by means of the $T_{1\rho}$ filter. Protons labelled a and b refer to the alpha and beta anomers of the glucose moiety of lactose, respectively.

(Figure 3b). One interesting aspect of homodecoupled spectra generated with the method of Guenneau et al. (1999) is that despite the fact that the intensities within the spectrum can not be compared because the information of the relative signal intensity is lost due to the use of the power mode during the processing stage of the spectrum, however it is still possible to determine quantitatively the modulation of the signal intensity with respect to a change in a certain acquisition parameter such as mixing time (Guenneau et al., 1999) or gradient strength (Cobas and Martín-Pastor, 2004), being equal the rest of the acquisition and processing parameters. Thus, for the present case of the 1D STD-homodecoupled spectrum, the saturation time can be changed along two or more spectra, and the modulation of any signal be calculated from the square-root of the intensity at the maximum peak height (Guenneau et al., 1999; Cobas and Martín-Pastor, 2004). This quantitative analysis of the 1D-STD-homodecoupled spectrum can be applied for the determination of STD build-up curves what can prove useful for the quantitative interpretation of the STD-NOE (Jayalakshmi and Krishna, 2004, 2005; Jayalakshmi et al., 2004) effect.

In analogy with the conventional 1D-STD experiment, there are situations in which the simple qualitative inspection of the intensities of 1D-homodecoupled experiment acquired with different saturation times allows to distinguish direct ligand-protein STD-NOEs involved in the binding epitope (Mayer and Meyer, 2001) from those NOEs mediated by spin-diffusion effects. This situation is exemplified in the two 1D-STD-homodecoupled spectra of Figures 3b and c acquired with sat. time of 2s and 1s, respectively. In Figure 3b, the extended saturation time generates STD-NOEs with the protons of both the galactose and glucose moieties of the disaccharide, the STD-NOEs with the galactose moiety being clearly more intense than those with the glucose moiety. On the other hand, in the spectrum of Figure 3c which was acquired with the shorter sat. time, there are STD-NOEs exclusively with the galactose moiety (Gal-3, Gal-4, Gal-5 and Gal-6), and their intensities are almost as strong as given in the spectrum of Figure 3b. These observations indicate that the relatively small STD-NOEs observed with the glucose moiety are likely caused

by spin diffusion effects, while the relative strong STD-NOEs of Gal-3, Gal-4, Gal-5 and Gal-6 observed in both spectra should correspond to direct STD-NOEs between lactose and the site of interaction with the lectin. Evidently, the non-reducing-end sugar unit comes into close vicinity to the lectin's binding site, as observed in studies using chemical mapping or crystallographic analysis. (Solís et al., 1996; Asensio et al., 1999; Varela et al., 1999; Alonso-Plaza et al., 2001; López-Lucedo et al., 2004) This confirmation attests the validity of the presented results of interaction analysis and thus the given protocol. Moreover, the pulse sequence was also applied to a more complex tetrasaccharide, methyl 4-O-(3-O-(2-acetamido-2-deoxy-4-O- β -D-galactopyranosyl)- β -D-glucopyranosyl)- β -D-galactopyranosyl)- β -D-glucopyranoside. The application of the pulse sequence of Figure 1 also indicated the recognition of the non reducing end residue, since only signals for this moiety were indeed observed in the 2D-STD-J-resolved spectrum (Fig. in the supporting information).

Acknowledgements

The generous support by an EC Marie Curie Research Training Network grant (contract No. HRTN-CT-2005-019561) is gratefully acknowledged. Dr. Juan Carlos Cobas (*Mestrelab research*) is acknowledged for helping discussions regarding the processing of the STD-homodecoupled spectra.

Electronic Supplementary Material is available to authorized users in the online version of this article at <http://www.dx.doi.org/10.1007/s10858-006-9055-4>.

References

- Alonso-Plaza, J.M., Canales, M.A., Jiménez, M., Roldán, J.L., García-Herrero, A., Hurrino, L., Asensio, J.L., Cañada, F.J., Romero, A., Siebert, H.-C., André, S., Solís, D., Gabius, H.-J. and Jiménez-Barbero, J. (2001) *Biochim. Biophys. Acta*, **1568**, 225–236.
- André, S., Kaltner, H., Lensch, M., Russwurm, R., Siebert, H.-C., Fallsehr, C., Tajkhorshid, E., Heck, A.J.R., von Knebel Doeberitz, M., Gabius, H.-J. and Kopitz, J. (2005) *Int. J. Cancer*, **114**, 46–57.
- Asensio, J.L., Espinosa, J.F., Dietrich, H., Cañada, F.J., Schmidt, R.R., Martín-Lomas, M., André, S., Gabius, H.-J. and Jiménez-Barbero, J. (1999) *J. Am. Chem. Soc.*, **121**, 8995–9000.

- Breg, J., Romijn, D., van Halbeek, H., Vliegthart, J.F.G., Visser, R.A. and Haasnoot, C.A.G. (1988) *Carbohydr. Res.*, **174**, 23–36.
- Chen, A. and Shapiro, M.J. (1998) *J. Am. Chem. Soc.*, **120**, 10258–10259.
- Chen, A. and Shapiro, M.J. (2000) *J. Am. Chem. Soc.*, **122**, 414
- Claasen, B., Axmann, M., Meinecke, R. and Meyer, B. (2005) *J. Am. Chem. Soc.*, **127**, 916–919.
- Cobas, J.C. and Martín-Pastor, M. (2004) *J. Magn. Reson.*, **171**, 20–24.
- Cobas, J.C. and Sardina, F.J. (2004) *Concepts Magn. Reson.*, **19A**, 80–96.
- Dalvit, C., Fasolini, M., Flocco, M., Knapp, S., Pevarello, P. and Veronesi, M. (2002) *J. Med. Chem.*, **45**, 2610–2614.
- Dalvit, C., Pevarello, P., Tato, M., Veronesi, M., Vulpetti, A. and Sundstrom, M. (2000) *J. Biomol. NMR*, **18**, 65–68.
- Dam, T.K., Gabius, H.-J., André, S., Kaltner, H., Lensch, M. and Brewer, C.F. (2005) *Biochemistry*, **44**, 12564–12571.
- Di Micco, S., Bassarello, C., Bifulco, G., Riccio, R. and Gomez-Paloma, L. (2005) *Angew. Chem. Int. Ed.*, **44**, 1–5.
- Gabius, H.-J. (1997) *Eur. J. Biochem.*, **243**, 543–576.
- Gabius, H.-J. (2001) *Anat. Histol. Embryol.*, **30**, 3–31.
- Gabius, H.-J. (2006) *Crit. Rev. Immunol.*, **26**, 43–79.
- Gabius, H.-J., Siebert, H.-C., André, S., Jiménez-Barbero, J. and Rüdiger, H. (2004) *ChemBiochem*, **5**, 740–764.
- Guenneau, F., Mutzenhardt, P., Grandclaude, D. and Canet, D. (1999) *J. Magn. Reson.*, **140**, 250–258.
- Hajduk, P.J., Mack, J.C., Olejniczak, E.T., Park, C., Dandliker, P.J. and Beutel, B.A. (2004) *J. Am. Chem. Soc.*, **126**, 2390–2398.
- Hajduk, P.J., Olejniczak, E.T. and Fesik, S.W. (1997) *J. Am. Chem. Soc.*, **119**, 12257–12261.
- Jahnke, W., Floersheim, P., Ostermeier, C., Zhang, X., Hemmig, R., Hurth, K. and Uzunov, D.P. (2002) *Angew. Chem. Int. Ed.*, **41**, 3420–3423.
- Jayalakshmi, V. and Krishna, N.R. (2005) *J. Am. Chem. Soc.*, **127**, 14080–14084.
- Jayalakshmi, V. and Krishna, N.R. (2004) *J. Magn. Reson.*, **168**, 36–45.
- Jayalakshmi, V., Biet, T., Peters, T. and Krishna, N.R. (2004) *J. Am. Chem. Soc.*, **126**, 8610–8611.
- Johnson, M.A. and Pinto, B.M. (2002) *J. Am. Chem. Soc.*, **124**, 15368–15374.
- Klein, J., Meinecke, R., Mayer, M. and Meyer, B. (1999) *J. Am. Chem. Soc.*, **121**, 5336–5337.
- Kopitz, J., Russwurm, R., Kaltner, H., André, S., Dotti, C.G., Gabius, H.-J. and Abad-Rodriguez, J. (2004) *Dev. Brain Res.*, **153**, 189–196.
- LaPlante, S.R., Aubry, N., Deziel, R., Ni, F. and Xu, P. (2000) *J. Am. Chem. Soc.*, **120**, 12530–12535.
- Leonidas, D.D., Vatzaki, E.H., Vorum, H., Celis, J.E., Madsen, P. and Acharya, K.R. (1998) *Biochemistry*, **37**, 13930–13940.
- Li, D., DeRose, E.F. and London, R.E. (1999) *J. Biomol. NMR*, **15**, 71–76.
- Li, D., Levy, L.A., Gabel, S.A., Lebetkin, M.S., DeRose, E.F., Wall, M.J., Howell, E.E. and London, R.E. (2001) *Biochemistry*, **40**, 4242–4252.
- Lin, M. and Shapiro, M.J. (1996) *J. Org. Chem.*, **61**, 7617–7619.
- Lin, M., Shapiro, M.J. and Wareing, J.R. (1997) *J. Am. Chem. Soc.*, **119**, 5249–5250.
- London, R.E. (1999) *J. Magn. Reson.*, **141**, 301–311.
- López-Lucendo, M.F., Solís, D., André, S., Hirabayashi, J., Kasai, K.-I., Kaltner, H., Gabius, H.-J. and Romero, A. (2004) *J. Mol. Biol.*, **343**, 957–970.
- Mari, S., Serrano-Gomez, D., Cañada, F.J., Corbi, A.L. and Jiménez-Barbero, J. (2005) *Angew. Chem. Int. Ed.*, **44**, 296–298.
- Mayer, M. and Meyer, B. (1999) *Angew. Chem. Int. Ed. Eng.*, **38**, 1784–1788.
- Mayer, M. and Meyer, B. (2000) *J. Med. Chem.*, **43**, 2093–2099.
- Mayer, M. and Meyer, B. (2001) *J. Am. Chem. Soc.*, **123**, 6108–6117.
- Meyer, B. and Peters, T. (2003) *Angew. Chem. Int. Ed.*, **42**, 864–890.
- Meyer, B., Weimar, T. and Peters, T. (1997) *Eur. J. Biochem.*, **246**, 705–709.
- Möller, H., Serittas, N., Paulsen, H., Burchell, J.M., Taylor-Papadimitriou, J. and Meyer, B. (2002) *Eur. J. Biochem.*, **269**, 1444–1455.
- Peng, J.W., Moore, J. and Abdul-Manan, N. (2004) *Prog. Nucl. Magn. Reson. Spectrosc.*, **44**, 225–256.
- Piotto, M., Saudek, V. and Sklena, V. (1992) *J. Biomol. NMR*, **2**, 661–665.
- Platzer, N., Davoust, D., Lhermitte, M., Bauvy, D., Meyer, D.M. and Derappe, C. (1989) *Carbohydr. Res.*, **191**, 191–207.
- Reuter, G. and Gabius, H.-J. (1999) *Cell. Mol. Life Sci.*, **55**, 368–422.
- Rivera, A., Cañada, J., Nieto, O., Jimenez-Barbero, J. and Martin-Lomas, M. (1992) *Eur. J. Biochem.*, **209**, 415–422.
- Siebert, H.-C., André, S., Lu, S.Y., Frank, M., Kaltner, H., van Kuik, J.A., Korchagina, E.Y., Bovin, N.V., Tojkorshid, E., Kaptein, R., Vliegthart, J.F.G., von der Lieth, C.-W., Jiménez-Barbero, J. and Gabius, H.-J. (2003) *Biochemistry*, **42**, 14762–14773.
- Siriwardena, A.H., Tian, F., Noble, S. and Prestegard, J.H. (2002) *Angew. Chem. Int. Ed.*, **41**, 3454–3457.
- Solís, D., Fernandez, P., Diaz-Maurino, T., Jiménez-Barbero, J. and Martin-Lomas, M. (1993) *Eur. J. Biochem.*, **214**, 677–683.
- Solís, D., Jiménez-Barbero, J., Kaltner, H., Romero, A., Siebert, H.-C., von der Lieth, C.-W. and Gabius, H.-J. (2001) *Cells Tissues Organs*, **168**, 5–23.
- Solís, D., Romero, A., Kaltner, H., Gabius, H.-J. and Diaz-Maurino, T. (1996) *J. Biol. Chem.*, **271**, 12744–12748.
- Varela, P.F., Solís, D., Diaz-Maurino, T., Kaltner, H., Gabius, H.-J. and Romero, A. (1999) *J. Mol. Biol.*, **294**, 537–549.
- Vogtherr, M. and Peters, T. (2000) *J. Am. Chem. Soc.*, **122**, 6093–6099.
- Wang, Y.-S., Liu, D. and Wyss, F.D. (2004) *Magn. Reson. Chem.*, **42**, 485–489.
- Wu, L.N.S., Lin, X., Zheng, L. and Rui, L. (2002) *J. Chem. Inf. Comput. Sci.*, **42**, 274–283.
- Yan, J., Kline, A.D., Mo, H., Zartler, E.R. and Shapiro, M.J. (2002) *J. Am. Chem. Soc.*, **124**, 9984–9985.
- Yuen, Y., Wen, X., Sanders, D.A.R. and Pinto, B.M. (2005) *Biochemistry*, **44**, 14080–14089.