Universidad de Navarra

Facultad de Farmacia

New quinoxaline and indole derivatives as MT$_1$/MT$_2$ receptor agonists

Saioa Ancizu Perez de Ciriza

Universidad de Navarra

Facultad de Farmacia

New quinoxaline and indole derivatives as MT$_1$/MT$_2$ receptor agonists

Memoria presentada por Dª Saioa Ancizu Perez de Ciriza para aspirar al grado de Doctor por la Universidad de Navarra

El presente trabajo ha sido realizado bajo nuestra dirección en la Unidad de I+D de Medicamentos (CIFA) y autorizamos su presentación ante el Tribual que lo ha de juzgar

Pamplona, 26 de Marzo de 2012.

Dra. Silvia Galiano Ruiz   Dr. Ignacio Aldana Moraza

ESTE ES UN DOCUMENTO CONFIDENCIAL

QUEDA PROHIBIDA TODA REPRODUCCIÓN, DIVULGACIÓN O DISCUSIÓN TOTAL O PARCIAL DE CUALQUIER DATO DE ESTA MEMORIA ANTES DEL DÍA 26 DE MARZO DEL 2022 SIN AUTORIZACIÓN ESCRITA DE LA PROFESORA Dra. SILVIA GALIANO RUIZ

LA CONSERVACIÓN Y DEPÓSITO DE ESTE DOCUMENTO DEBE ESTAR GARANTIZADA SIEMPRE POR LA CONFIDENCIALIDAD

"Comienza de nuevo. Cada vez que fracases empieza otra vez y te harás más fuerte hasta que finalmente logres tu propósito"

(Anne Sullivan)

"Resérvate el derecho de pensar: incluso equivocarse es mejor que no pensar nada"

(Hipatia)

"Me atrevería a aventurar que Anónimo, que tantos poemas escribió sin firmarlos, era a menudo una mujer"

(Adeline Virginia Woolf)

Acknowledgments/Agradecimientos

Quiero expresar mi agradecimiento a todas aquellas personas que han contribuido a la elaboración de esta tesis doctoral, tanto a las que han participado activamente en cualquiera de sus etapas como a las que sin entender muy bien qué es lo que estoy haciendo me han apoyado durante estos años dándome las fuerzas necesarias para seguir adelante. Por lo tanto, mi más sincero agradecimiento…

Al **Dr. Antonio Monge**, por haberme brindado la oportunidad de formar parte de su grupo de investigación, por su disponibilidad y por todas las enseñanzas transmitidas a lo largo de estos años.

A la **Dra. Silvia Galiano**, la **Gali**, por su implicación e interés constante, y sobre todo, por creer en mí y transmitirme tranquilidad para afrontar todos los problemas que han ido surgiendo. Gracias por ser una gran directora a la vez que una gran compañera.

Al **Dr. Ignacio Aldana,** por su disponibilidad, por todos sus consejos y por todo el apoyo recibido durante este tiempo.

A la **Dra. Silvia Perez-Silanes**, por su ayuda cada vez que la he necesitado y por aportar siempre el optimismo necesario para seguir adelante.

Au **Dr. Philippe Chavatte**, de m'avoir donné l'occasion de réaliser un séjour dans son équipe de recherche du Laboratoire de Chimie Thérapeutique 2 de l'Université Lille 2.

Au **Dr. Nicolas Renault**, **Nico**, pour son accueil, son pédagogie et surtout son altruisme qui m'a permis de ne pas me sentir seule.

À mes collègues du laboratoire de Lille, **Xavier** et **Amauri,** pour leurs précieuses aides et connaissances, sans oublier ces repas partagés qui ont concouru á rendre ce séjour très agréable. Je suis particulièrement reconnaissante envers **Xavier** pour sa disponibilité permanente ainsi que ses colloques de traditions françaises.

A **Laura Stokes**, por su inestimable ayuda con el inglés. Gracias por la buena disposición mostrada ante cada nueva petición y por su gran implicación.

A todas las personas que forman o han formado parte del laboratorio; a las antiguas quinoxalineras, **Raquel**, **Asun** y **Bea**, por enseñármelo todo en mis primeros pasos en el laboratorio y por los buenos ratos vividos. A los antiguos **"Organiqueros"**, (**Ceras**, **Mikel**, **Elena**, **Esther**, **Iranzu**, **Ylenia**, **Torres**, **Rosa**,...) por haber tenido la suerte de trabajar con vosotros y por crear un gran ambiente en el laboratorio. A todas las nuevas incorporaciones, **quinoxalineros**, **caspaseras**, **masterianos** e **ITLs** por todos los buenos ratos compartidos.

A la **Eli**, por estar siempre dispuesta a echar una mano y por tratarme como a una hija alegrándose siempre por mí e incluso poniéndome algún punto negro cuando ha sido necesario.

A **Elsa**, **Dani**, **Adeli** y **Ner**, por escucharme y apoyarme siempre. Por todos los debates tanto científicos como personales y por todos los ratos, risas, comidas y demás eventos compartidos. Gracias por ser aparte de mis compañeros, mis amigos.

A las de **"saski"** por hacer o haber hecho posible un lugar donde desconectar y cargar pilas para seguir adelante. Gracias tanto a las veteranas, **Zube**, **Su**, **Miren**, **Maider**, **Oihana**,

Nekane, **Maika** y **Silvi** como a las jovenzuelas, **Garazi**, **Sara**, **Marta**, **Lorea**, **Izaskun, Ana**, **Alba**, **Debo**, **Laura** y **Keka** y también a los coach **Piny** y **Dami**, porque lo mejor del deporte es las amistades que deja.

Bai **Oihane**ri eta bai kimikakoei, **Aitziber**, **Axi**, **Ixi**, **Karmele**, **Ander**, **Oihana**, **Xabi** eta **Adrian**, zuen laguntasunarengatik eta Donostian pasatako urteak paregabeak izan izana lortzeagatik.

A la **familia Nikolas-Barea**, por su disposición a ayudar en lo que sea, por su cariño y por hacer que me sienta una más.

A toda la **familia Ancizu** en general y a **Lourdes** y **Tasio**, **Jesús** y **Carmen**, **Andoni**, **Peio** e **Iñigo** en particular, por vuestro apoyo incondicional y por animarme siempre.

A mi **Amatxi,** por todos los cuidados brindados a lo largo de mi vida y por inculcarme que es cada uno quien se forma a sí mismo. Gracias además por ser mi modelo de constancia y tesón.

A toda la "**Family Perez de Ciriza"**, por ser uno de los pilares más importantes de mi vida. A mis tíos y tías, **Sor Juana**, **Luis** y **Tere**, **Feli**, **Txarito** y **Carlos**, **Florita** y **Antonio**, **Elena** y **Pitxitxi**, **Javito** y **Gloria**, **Sagra** y **Ramontxo**, **Titika** y **Alfredo**, por todo vuestro apoyo y cariño. A los "primates y allegaus", **Eva** y **Xabi**, **Javier** y **Soraya**, **Eleni** y **Unai**, **Itsa** y **Dani**, **Iru** e **Iker**, **Ana** y **Koldo**, **Asier** y **María**, **Carlos** y **María**, **Iñaki** y **Chris**, **Javi** y **Ele**, **Lota** y **Carlos**, **David** y **Rebeca**, **Iker** y **Marta**, **Mikel**, **Naia** y **Eoin**, **Aingeru**, **Izai** y **Urko**, por todos los debates, juegos y buenos ratos pasados y por todos los que vendrán. A todos mis sobris, **Jon**, **Iker**, **Ariadna**, **June**, **Aitor**, **Josu**, **Irati**, **Ane**, **Julen**, **Hugo** y **Maia**, y en especial a mi ahijada **Haizea,** por ser fuente inagotable de alegría.

A mi **abuelo Ángel**, por su buen humor sin fin, por ser mi modelo de superación personal y por enseñarme que todo puede hacerse ya sea de una forma u otra.

A mi **abuela Carmen**, por todo lo que nos ha dado sin pedir nada a cambio. Gracias por transmitirme tanto la importancia de la familia como la de ser mujeres independientes y animarme siempre a llegar lo más lejos posible.

Nire koadrila osoari, urteetan elkarrekin bizitako momentuengatik, nireaz arduratzeagatik eta beti nirekin gogoratzeagatik. A **"estos"**, **Ander**, **Tolo**, **Karlos**, **Álvaro**, **Anxel**, **Lalu**, y **Arkaitz** por hacer que la vida sea más fácil a golpe de carcajada. A **"estas"**, **Noelia**, **Lorea**, **Montse**, **Maialen**, **Naia**, **MaiteMu**, **Maitetxiki** y **Aitzi** por hacerme sentir bien y compartir conmigo tantas tonterías tan importantes.

A **Maitetxiki**, por estar ahí siempre siempre. Por escucharme y aconsejarme durante esos largos raticos. Txikia zara tamainaz baina haundia lagun lez.

Aitzi-ri, zure laguntasunarengatik, beti entzuteko prest egoteagatik, elkarrekin pasatako bizipenengatik eta zure ikuspegi baikorrarengatik.

A **Arkaitz**, por sacarme siempre una sonrisa, incluso en los momentos más difíciles. Gracias por ser como eres y por apoyarme en todo momento. Elkarrekin bizitako urteengatik eta etorriko direnengatik.

A **mis padres**, por su generosidad y apoyo incondicional, por todo el esfuerzo realizado para que yo pueda lograr todo lo que me proponga. A mi **aita** por ser mí modelo de trabajo, esfuerzo y valentía. A mi **ama** por ser el referente de la mujer que algún día me gustaría llegar a ser. Gracias a los dos porque sin vosotros no sería quien soy.

A mis "Aitas"

Arkaitzi

Abbreviations

0-9

$[^{35}S]GTP\gamma S$	$[^{35}S]$guanosine-5'-O-(3-thio)-triphosphate
^{125}I-MLT	2-$[^{125}I]$iodomelatonin
^1H NMR	Proton nuclear magnetic resonance
^{13}C NMR	Carbon nuclear magnetic resonance
2D	Two dimension
3D	Three dimension
5-HT	5-hydroxytryptamine

A

AA-NAT	Arylakylamine-*N*-acetyltransferase
AASM	American academy of sleep medicine
AcCN	Acetonitrile
AcOEt	Ethyl acetate
ACTH	Adrenocorticotropic hormone
AFMK	N^1-acetyl-N^2-formyl-5-methoxykynurenine
AMK	N^1-acetyl-5-methoxy-kynurenine
aMT6s	6-Sulfatoxymelatonin
AUC	Area under the curve

B

BFX	Benzofuroxan
B_{max}	Receptor density
BRAs	Benzodiazepine receptor agonists
Bs	Broad singlet
BZD	Benzodiazepine

C

cAMP	Cyclic adenosine monophosphate
Ch	Chapter
CHN	Elemental microanalyses of carbon, hydrogen and nitrogen
CHO	Chinese hamster ovary
CoMFA	Comparative molecular field analysis
CoMSIA	Comparative molecular similarity indices analysis
COSY	Correlation spectroscopy
CREB	cAMP response element-binding
CRSD	Circadian rhythm sleep disorders
CRY	Cryptochrome gene
CYP	Cytochrome P450

D

D	Doublet
DALY	Disability adjusted life years

DCM	Dichloromethane
Dd	Doublet of doublets
Ddd	Doublet of doublet of doublets
Dddd	Doublet of doublet of doublet of doublets
DDR	Drug data report
DMEM	Dulbecco's modified eagle medium
DMSO	Dimethyl sulphoxide
Dpm	Disintegrations per minute
Dq	Doublet of quartets
Dt	Doublet of triplets

E

EC_{50}	Effective concentration 50%
EDTA	Ethylenediaminetetraacetic acid
EEG	Electroencephalography
EL1-EL3	Extracellular loops (1-3)
E_{max}	Relative efficacy
EMEA	European medicines evaluation agency
EMG	Electromyography
EOG	Electroculography
Et_3N	Triethylamine

F

FDA	Food and drugs administration
FN	False negative
FP	False positive

G

GABA	γ-aminobutyric acid
GDP	Guanosine diphosphate
GPCR	G-protein coupled receptor
GS	Good sleepers
GTO	Gaussian-type orbitals
GTP	Guanosine triphosphate

H

HEK	Human embryonic kidney
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HIOMT	Hydroxyindole-O-methyltransferase
His	Histidine
HMBC	Heteronuclear Multiple Bond Coherence
HMPA	Hexamethyl phosphoramide
HMQC	Heteronuclear Multiple Quantum Coherence
HPA	Hypothalamic pituitary adrenal axis
HPLC	High performance liquid chromatography
Hz	Hertz

I

I	Inactive
IC_{50}	Inhibitory concentration 50
ICSD	International classification of sleep disorders
IL1-IL3	Intracellular loops (1-3)
IMC	Intermediolateral column
IP_3	Inositol-1,4,5-tris-phosphate
IR	Infrared

J

J	Coupling constant

K

K_D	Dissociation constant
Ki	Inhibition constant
Kir	Inwardly rectifying potassium
KNN	Kohonen Neural Network

L

LCAO	Linear combinations of atomic orbital

M

M	Medium
M	Multiplet
Mel_{1a}	See MT_1
Mel_{1b}	See MT_2
MeOH	Methanol
MEP	Molecular electrostatic potential
MF	Molecular formula
M-II	2-hydroxy-*N*-[2-(2,6,7,8- tetrahydro-1*H*-indeno[5,4-*b*]furan-8-yl)ethyl]propanamide
ML_1	High affinity MLT binding sites
ML_2	Low affinity MLT binding sites
MLT	Melatonin
MM	Molecular mechanics
MP	Melting point
MSLT	Multiple sleep latency test
MT_1	MLT binding receptor 1
MT_2	MLT binding receptor 2
MT_3	MLT binding receptor 3
MW	Molecular weight

N

N,N-DMF	*N,N*-dimethylformamide
N.D.	No data
N.T.	Not tested
NA	Noradrenaline
NaOMe	Sodium methoxide
NBZD	Non-benzodiazepine

n-Hex	Hexane
NMR	Nuclear magnetic resonance
NREM	Non-rapid eye movement

P

PACAP	Pituitary adenylate cyclase activating polypeptide
PBS	Phosphate buffered saline
Pd-C	Palladium on carbon
PER	Period gene
PET	Positron emission tomography
PGF_{2a}	Prostaglandin F_{2a}
PKA	Protein kinase A
PKC	Protein kinase C
PLC	Phospholipase C
PLMD	Periodic limb movement disorders
PLS	Partial least squares
ppm	Parts per million
PSG	Polysomnography
psi	Pounds per square inch
PVN	Paraventricular nuclei

Q

Q	Quadruplet
QM	Quantum mechanics
QR2	Quinone reductase 2
QSAR	Quantitative structure activity relationship

R

REM	Rapid eye movement
RHT	Retino-hypothalamic tract
RLS	Restless legs syndrome
ROC	Receiver operating characteristic
ROR	RAR (retinoic acid receptor) related orphan receptor
Rt	Room temperature
rms	Root mean square

S

S	Strong
S	Singlet
SAR	Structure activity relationship
SCF	Self-consistent field
SCG	Superior cervical ganglia
SCN	Suprachiasmatic nucleus
SEGA	Selective extrasynaptic GABA
Ser	Serine
SF-36	36-item short form health survey of the medical outcomes study
SI	Severe insomniacs
S_NAr	Aromatic nucleophilic substitution

SOM	Self-organizing map
STO	Slater-type orbitals
SWS	Slow-wave sleep

T

t	Triplet
TDA	Toluene/dioxane/acetic acid
td	Triplet of doublets
tdd	Triplet of doublets of doublets
TF	Transcription factors
THF	Tetrahydrofuran
TLC	Thin layer chromatography
TM1-TM7	Transmembrane helices (1-7)
TMS	Tetramethylsilane
TN	True negative
TP	True positive
t_R	Retention time
T_{reflux}	Refluxing temperature
Tris	Tris(hydroxymethyl)aminomethane
tt	Triplet of triplets
ttd	Triplet of triplets of doublets

V

vdW	Van der Waals
vs	Very strong

W

w	Weak

Symbol

δ	Chemical shift

Index

Chapter 1

Introduction

I. SLEEP AND SLEEP DISORDERS

1. SLEEP

Sleep definition, functions and historical brushstrokes

Contrary to what is popularly thought, sleep is not a total lack of activity, but rather it can be considered a specific, active and periodic biological state.[1] Sleep is defined as a recurrent state of lack of wakefulness, of consciousness suspension and from a neurological point of view, it is characterized by specific electroencephalographic criteria.[2]

This universal and reversible condition that occupies about one third of each day cannot be eliminated without causing deleterious consequences. Sleep is a state of immobility with decreased ability to react to stimuli, a fact that distinguishes sleep from quiet wakefulness. In contrast, the rapid reversibility that this state presents distinguishes sleep from coma or anesthesia. An additional defining characteristic of sleep is that when it is prevented, the body tries to recover the lost amount. The existence of sleep "rebound" after deprivation demonstrates that sleep is not simply a period of reduced activity or alertness. The amplitude of the changes in brain metabolism and neuronal activity that occurs during sleep exceeds those which occur during most waking periods. Therefore, the argument that sleep serves a vital function is compelling.[3, 4]

It is well known that sleep is important and that it is desirable to be well rested but there is still no consensus when attempting to answer the question regarding the functional role of sleep. However, there is no doubt that sleep affects the vast majority of body functions, from immunity to hormonal regulation to metabolism to thermoregulation. Thus, it has been proposed that it may be a single core function that requires sleep and adventitious functions that take advantage of it.[4] There is a vast amount of studies which state that sleep is necessary, direct or indirectly, for many functions, such as for encoding and consolidation of memory, for modulation, regulation and preparation of cognitive and emotional brain processes, for saving energy, for controlling the body and brain's temperature, for removing free radicals accumulated during wakefulness and for regulation of some neurotransmitters.[3, 5-7] All these functions are coexisting, complementary and essential for a correct adaptation of the human specie.[8]

Although more and more studies are currently being performed with regard to sleep and sleep-related pathologies, historically, this biological state has been one of the most forgotten human functions, at least from a medical point of view.[9] Even though it was already known that sleep disorders often appear as the result of other diseases, it wasn't until the middle of the 20th century when it was finally established that the primary disorders of sleep are frequent, serious and treatable. Later, in the 70s, specialized centers in sleep alterations came into being. The pathology of sleep became a medical sub-specialty, and its development led to the publication of a magazine, "Sleep", dedicated specifically to these disorders.[2]

1.1. Sleep physiology

1.1.1. Polysomnography

Before the discovery of electrical cerebral activity, sleep was considered to be a passive response of the brain resulting from a reduction of the stimulations received.[2] Although the first recording of brain electrical activity was made in the late 1800s, it was not until 1929 when Hans Berger published the changes in human brain electrical activity, recorded in a patient transitioning from wakefulness to sleep. In his work "*Über das Elektrenkephalogramm des Menschen*", the term electroencephalography (EEG) was first established for recording cortical brain activity. In the 1930s, Loomis and colleagues further explored the information contained in an EEG, distinguishing periods of wakefulness, sleep and dreaming. They also identified five sleep stages (A-E) and set up the terms "K-complex" and "spindles". Subsequently, in the 1950s, the "REM" stage of sleep was described in humans (figure 1).[10, 11] Since then, EEG analyses have been enormously refined.
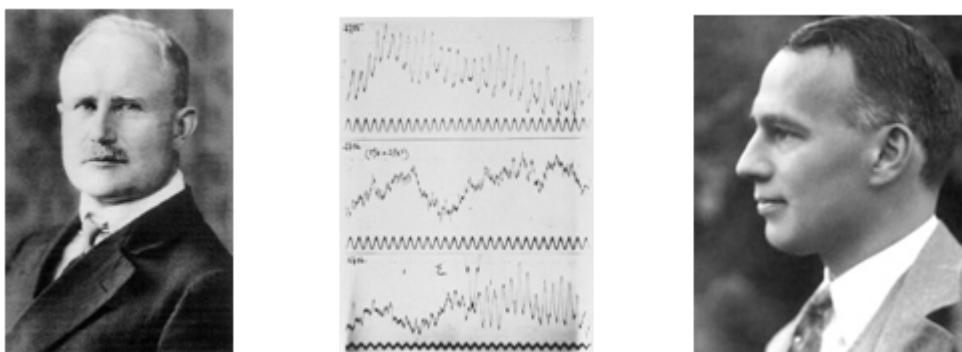
Figure 1. Hans Berger, some EEGs registered by Berger, and Alfred Lee Loomis [Figures adapted from *Deak, M. et all, Woo, J. and California-University*].[12-14]

It is now known that a typical EEG contains waves at multiple frequencies that depend on the level of consciousness. These waves are grouped in frequency bands defined as follows (figure 2):[10]
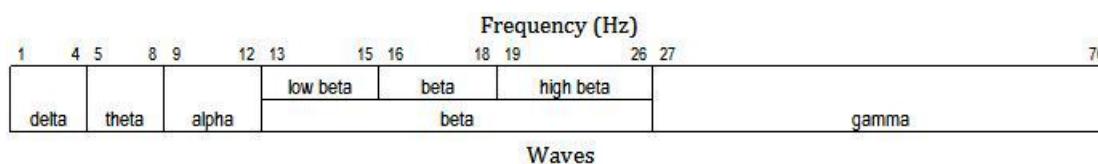
Figure 2. Frequency bands of EEG's wave groups.

Sleep is organized in different degrees of depth, presenting concrete physiological modifications in each stage. In order to further our knowledge regarding the functional changes that take place during sleep, three sleep indicators are studied: cortical brain activity or electroencephalography (EEG), eye movement or electrooculography (EOG) and muscular tone or electromyography (EMG). Polysomnography (PSG) is the simultaneous recording of these three neurophysiological and cardio-respiratory variables that allow us to evaluate the quantity and quality of sleep.[1, 10, 15]

1.1.2. Wakefulness-sleep stages

In 1968, the PSG was used to make a pioneer systematic classification of sleep-wakefulness stages which was then updated in 2007, by the American Academy of Sleep Medicine (AASM).[10]

1.1.2.1. WAKEFULNESS

In contrast to active wakefulness which is associated with fast high frequency beta waves, eye movement and muscle tone, the criteria for calm wakefulness (recorded at rest with closed eyes) are characterized by alpha waves on EEG, the presence of muscle tone, and lack of eye movements (figure 3).[16, 17] The alpha waves, discovered by Berger, are the background activity of a normal human adult EEG. He also observed that when patients opened their eyes, the EEG oscillations in the alpha band decreased in amplitude or disappeared completely. It is important to understand that oscillations in the EEG indicate periodic activity of large populations of synchronized neurons. Therefore, the disappearance of alpha waves after eye opening means that the oscillators in the brain get out of phase, meaning that they get "de-synchronized" in the wakefulness period.[18] On the contrary, when the depth of sleep is increased, a progressive "synchronized" activity is observed through the appearance of high voltage, low frequency waves in the EEG.[15]

1.1.2.2. SLEEP

Two types of sleep have been described; NREM (Non Rapid Eye Movement) sleep and REM (Rapid Eye Movement) sleep. NREM sleep, also termed as low waves sleep, is subdivided into four stages, corresponding to increasing depth of sleep, indicated by progressive dominance of the EEG by high voltage, low frequency ("synchronized") wave activity.[15, 16]

Sleep is distributed into cycles of NREM/REM sleep that alternate during the night. NREM sleep constitutes about 75-80% of total sleep time, and REM sleep consumes the remaining 20-25% of the sleep period. This cycle is repeated 3 to 6 times during the night. Generally, sleep begins with about 80 minutes of NREM sleep and 10 minutes of REM sleep. In normal adults, REM sleep length increases as the night progresses. Thus, the second and later cycles are longer lasting. As the sleep cycle progresses through the night, stages 3 and 4 of NREM sleep decrease, stage 2 begins to account for the majority of NREM sleep, and there is also more REM sleep (figure 3).[17, 19]

1.1.2.2.1. NREM sleep

NREM sleep or paradoxical sleep is divided into four stages (1-4), corresponding to increasing depth of sleep and more synchronized waves. Each of the four stages of NREM sleep is associated with different brain activity and physiology (figure 3).

- <u>Stage 1 sleep</u>

NREM stage 1 occurs at sleep onset and serves a transitional role in sleep-stage cycling. Aside from newborns and those with narcolepsy and other specific neurological disorders, the sleep cycle begins in the first stage of NREM sleep. This stage usually lasts 1 to 7 minutes in the

initial cycle. It constitutes 2-5% of total sleep, and is easily interrupted by a disruptive noise. Brain activity on the EEG in stage 1 changes from calm wakefulness to low voltage, mixed frequency waves (2-7 Hz). Stage 1 is also characterized by slow eye movements and preserved muscle tone.[16, 19]
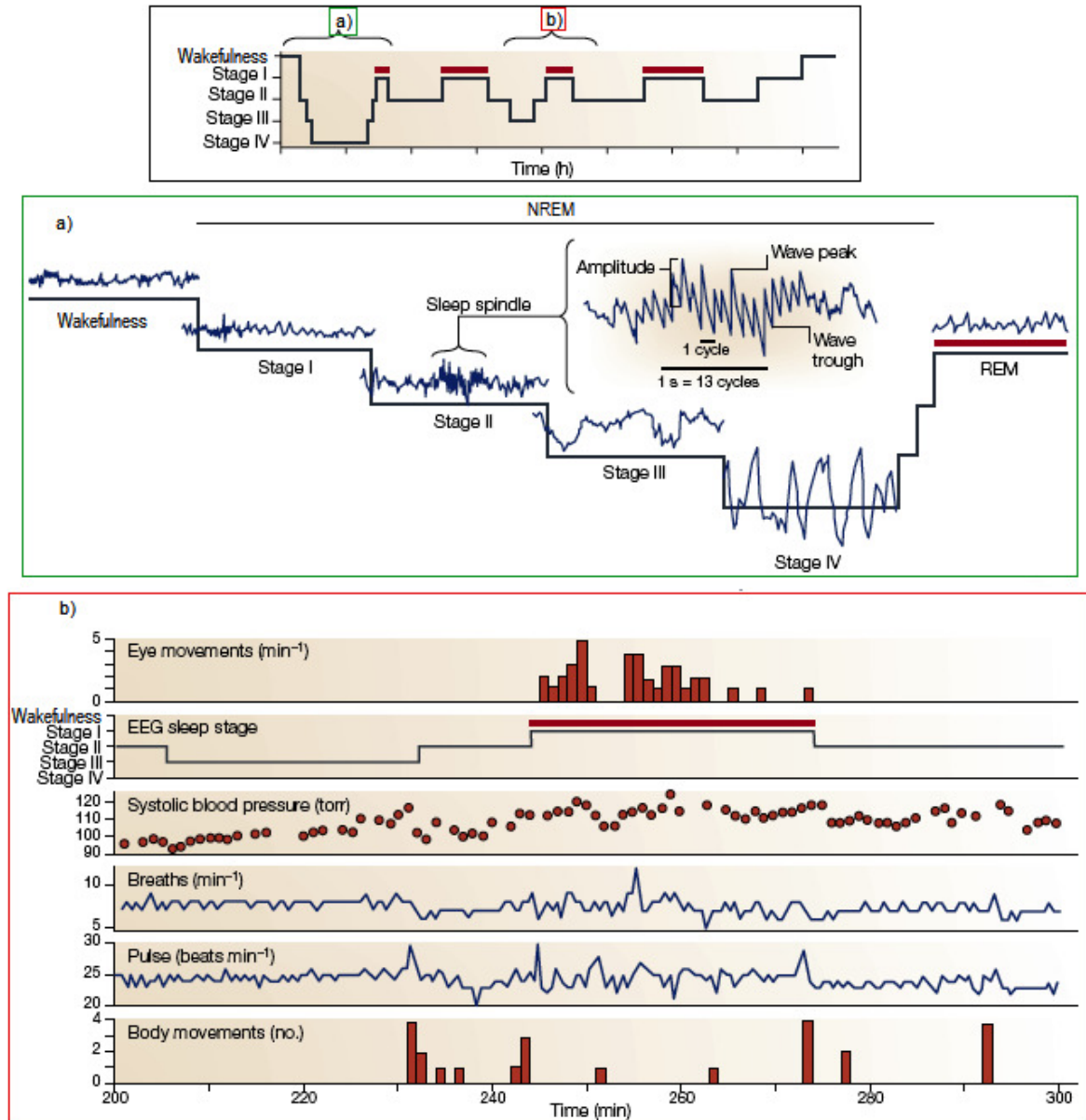
Figure 3. Stage changes over the course of a night's sleep describing the exact night period where panels a-b were recorded. **a)** Characteristic waveforms of the different sleep-wakefulness stages in an early-night sleep cycle. **b)** Changes in peripheral physiology associated with sleep stages in a late-night cycle [Figure adapted from *Pace-Schott, F.E. et al.].*[15]

- Stage 2 sleep

Stage 2 lasts approximately 10 to 25 minutes in the initial cycle but its length is increased in each successive cycle. This stage constitutes between 45 and 55% of the total sleep episode. A subject in stage 2 of its sleep cycle requires more intense stimuli to be awakened than when in stage 1. Brain activity in this stage is characterized by the presence of sleep spindles and K-

complexes that appear over a background activity of relatively low voltage mixed EEG frequencies. Sleep spindles are slow, sinusoidal waves with bursts of high amplitude peaks which present a frequency of 12-14 Hz and last at least 0.5".[17] K-complexes are negative sharp waves that are immediately followed by a slower positive component and last more than 0.5".[16, 19, 20] It is thought that sleep spindles are involved in memory consolidation. There are studies which demonstrate that individuals who have learned a new task have a significantly higher density of spindles at stage 2 than those in a control group. During stage 2, muscle tone is present and eye movements are absent.[16, 19]

- Stages 3 and 4, slow-wave sleep (SWS)

Stages 3 and 4 of NREM sleep are collectively referred to as slow-wave sleep (SWS). These stages generally take place during the first third of the night and during their occurrence, the EEG is synchronized. Although each stage has distinguishing characteristics, they are not very prominent.

Stage 3 is scored when a moderate amount (20-50%) of high amplitude slow wave activity (1-2 Hz) is observed in the EEG. It lasts only a few minutes and constitutes about 3-8% of sleep.

Stage 4 is the last NREM stage and it is characterized by increased amounts (>50%) of high voltage, slow-wave activity on the EEG. This stage is longer than stage 3; it lasts approximately 20 to 40 minutes in the first cycle and constitutes about 10-15% of sleep.[16, 17, 19]

1.1.2.2.2. REM sleep

REM sleep is defined by the presence of wake-like (low voltage, mixed frequency), "desynchronized" brain wave activity in the EEG, very low muscle tone or atonia and bursts of rapid eye movements (figure 3). Theta activity, slow alpha activity and "Sawtooth" characteristic wave forms, also defines REM sleep. The REM sleep during the first NREM/REM cycle may last only 1 to 5 minutes; however, it becomes progressively prolonged as the sleep cycles advances. Since dreaming is most often associated with REM sleep, it seems logical that the loss of muscle tone and reflexes are present in this stage because said loss protects an individual from "acting out" his or her dreams or nightmares while sleeping. Approximately 80% of vivid dream recall results after waking up from this stage of sleep. It is thought that REM sleep may be important for memory consolidation.[19]

1.1.3. Sleep patterns and age

The organization changes from one specie to another as well as within species. The length of each REM/NREM cycle increases proportionally to brain size growth across species and the depth and proportion of each NREM phase increases with brain maturation within subjects of the same species. NREM sleep reaches its maximum development in mature humans only to decline in post-mature age. Therefore, it can be concluded that sleep differentiation is a function of brain differentiation.[15]

Sleep architecture changes continuously and considerably with age. The initiation and maintenance of sleep changes from childhood to adulthood as well as the percentage of time

spent in each stage of sleep and the overall sleep efficiency. Although the consequences of decreased sleep efficiency among humans are relatively well reported, the reasons are complex and poorly understood.[21]
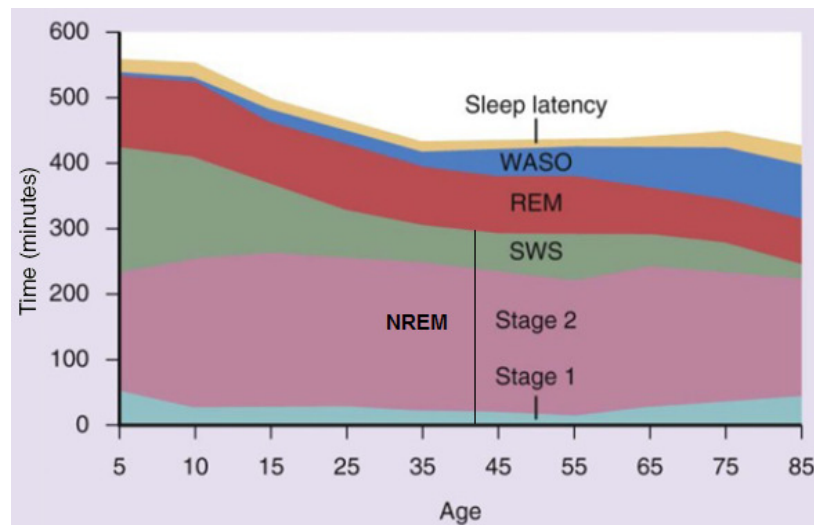
Figure 4. Changes in sleep with age [Figure adapted from *Carskadon, M. A. et al.*].[22]
WASO: amount of time spent awake after initially falling asleep / REM: rapid eye movement / NREM: non-rapid eye movement / SWS: slow-wave sleep

1.2. Sleep-wake regulation

In wakefulness-sleep cycle regulation, a very complex neuronal network takes place where different brain zones are activated or inhibited following a circadian rhythm that presents a length of 24 hours.

1.2.1. The two process model

"The two-process model" is one of the best established models of sleep regulation. In this model, it is stated that sleep-wake cycle is regulated by the interplay of two different processes; process C and process S. Process C is regulated by the circadian system and is responsible for promoting wakefulness as well as for the distribution of REM sleep. Process S is regulated by neurons that allow the brain to fall asleep, shutting down the wake system, and it acts as a homeostatic impulse that promotes sleep and controls SWS of NREM sleep.

The need for sleep accumulates during wakefulness; it peaks just before bedtime and dissipates throughout the night if adequate rest has been achieved. When the homeostatic drive for sleep is reduced, the circadian waking drive starts to increase and the sleep-wake cycle is restarted. If process C is eliminated, the total sleep time remains invariable, but it is randomly distributed over the day and night. Therefore, process C makes the sleep-wakefulness cycle coordinate with environmental light-dark cycles.[23-25]

1.2.2.   SCN and circadian rhythms synchronization

The suprachiasmatic nucleus (SCN) is responsible for synchronizing the circadian rhythms as well as for promoting sleep. Although it is still not very clear which mechanism is used, a great number of studies suggest that melatonin (MLT) is the principal neurochemical agent involved.[25]

There are several studies that demonstrate the importance of melatonin for the initiation, maintenance and quality of sleep. Although MLT's effect in sleep will be further discussed, it can be said that since MLT is mainly synthesized in the pineal gland. If pinealectomy is performed, sleep quality decreases. In addition, insomnia results if a total suppression of melatonin synthesis is induced by α_1/β-adrenoceptor agonists.[26]

SCN is responsible for regulating circadian rhythms in all organs. Thus, SCN coordinates many cyclic functions such as physical activity, food consumption, sleep-wake cycle, body temperature, heart rate, muscle tone, and hormone secretion.

In a much broader sense, it might be said that it is the environmental light that modulates the rhythm of melatonin secretion. There is some photopigment melanopsin in the retina which contains ganglions that act as a type of photodetector. When these ganglion cells perceive light variation, they send a signal to the SCN through the optic nerves in the monosynaptic retinohypothalamic tract (RHT).[25] These neurons vary the amount of released glutamate and pituitary adenylate cyclase activating polypeptide (PCACP) as a signal.[27] When the SCN receives these direct inputs from the retina, it transmits different signals to the rest of the brain and body in order to synchronize every daily cycle with external day/night cycle. The SCN especially controls the synthesis and secretion of melatonin by the pineal gland (figure 5). Melatonin, which is mainly secreted at night, acts to further consolidate the circadian rhythms as well as to provide a daily and a seasonal message to all tissues expressing melatonin receptors.[28, 29]
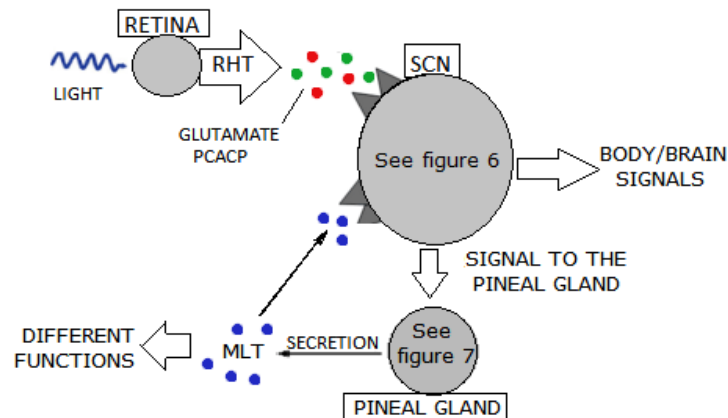
Figure 5. Circadian rhythms regulation by SCN and MLT.

In circadian rhythms regulation, a very complex network is involved where endogenous "clocks", synchronized inputs/outputs, and multiple oscillators are involved. This circadian network permits biological functions to be ready to react to the periodic changes of the environment before said changes actually occur. [24, 25] In mammals, the basis for these "clocks", are molecular pathways involving "clock" genes that are synchronously expressed in ~20,000 cells in the SCN. This "master clock" mechanism explains the nearly perfect 24-hour rhythmicity

of the circadian rhythms and consists of two main transcription factors (TF); CLOCK and BMAL1. When these TFs bind together, they form the CLOCK-BMAL1 heterodimer that moves into the nucleus of the cell in order to bind to a specific site in the DNA (E-box) of some genes. When CLOCK-BMAL1 heterodimer binds to the E-box of cryptochrome (CRY) and period (PER) genes, the transcription of these genes is activated, and the PER and CRY mRNAs are translocated to the cytoplasm where PER and CRY proteins are translated (figure 6). PER proteins interact with CRY proteins to form multimeric complexes that re-enter into the nucleus and inhibit the transcription mediated by CLOCK-BMAL1, stopping their own synthesis. But if CLOCK-BMAL1 heterodimer binds to the E-box of two other nuclear receptor transcription factors (Rev-Erbα and ROR) a second pathway is activated and as a result, ROR up-regulates and Rev-Erbα down regulates the expression of Bmal1 (figure 6). Although these TFs are not directly involved in rhythm generation, they are very important in controlling the phase and the amplitude of gene expression.[15, 28-32]
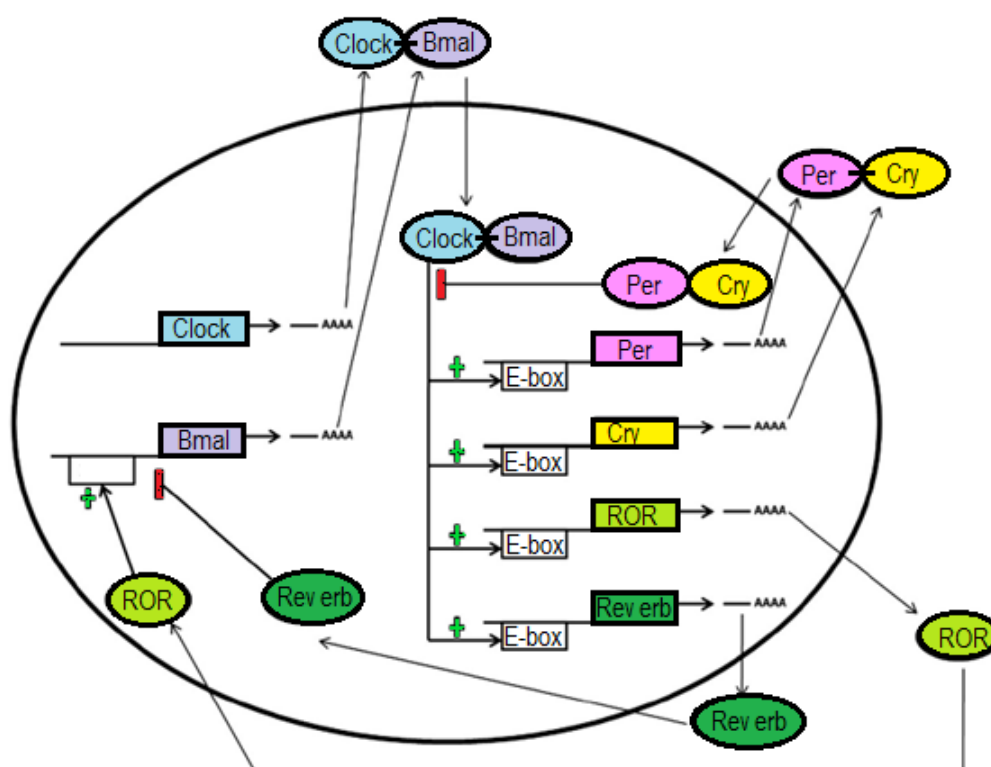
Figure 6. Molecular mechanisms involved in the regulation of circadian clock activity [Figure adapted from *Sukumaran, S. et al.*].[31]

Melatonin plays a feedback role in the SCN and circadian rhythm control through the melatonin receptors expressed in the cells of the SCN (figure 5). There are studies where the molecular analysis of the effects of melatonin on the SCN and peripheral tissues have shown that although MLT does not affect the expression of Per genes, it regulates the expression of RORs and Rev-Erbα genes (figures 6 and 8). Thus, MLT helps to regulate its own synthesis by the pineal gland following a circadian rhythmicity.[31]

1.2.3. Melatonin: "the darkness hormone"

Melatonin or *N*-acetyl-5-methoxytryptamine is a molecule mainly synthesized in the pineal gland, but not exclusively. There are many other areas where MLT is also synthesized such as the retina gastrointestinal track, skin, lymphocytes, platelets and bone narrow cells. MLT biosynthesis depends upon its precursor tryptophan, an essential amino acid coming from the diet. In the pineal glands, tryptophan is up taken from the blood by the 5-HT transporter sites on the surface of the pinealocytes. Once in the cytoplasm, tryptophan is first converted into 5-hydroxytryptophan by the enzyme tryptophan hydroxylase and then into serotonin or 5-hydroxytryptamine (5-HT) by the enzyme 5-HT-decarboxylase. Subsequently, 5-HT is acetylated to *N*-acetylserotonin by the enzyme arylalkylamine-*N*-acetyltransferase (AA-NAT) and then converted into MLT by the enzyme hydroxyindole-O-methyltransferase (HIOMT) (figure 7).[25]
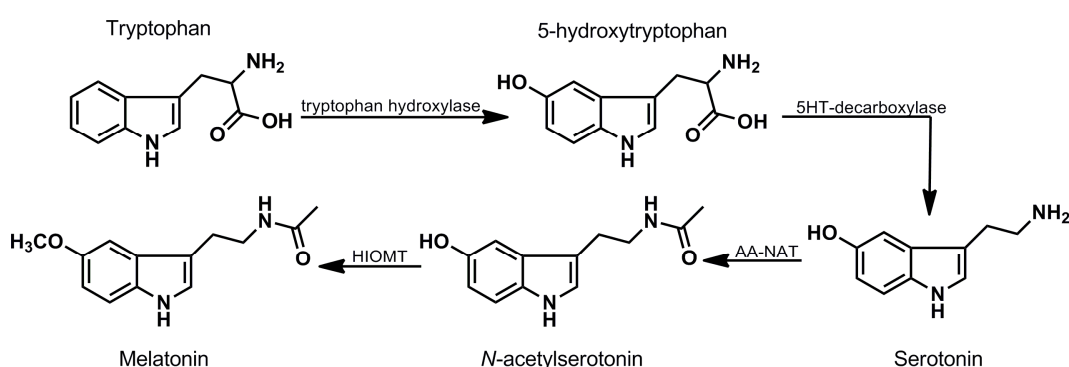
Figure 7. Biosynthesis of MLT in the pinealocytes.

The synthesis and secretion of MLT in the pineal gland is controlled by the SCN, being minimally secreted during the day and increasing progressively with the onset of night. When a light variation is detected by the ganglion cells in the retina, these cells send a signal (glutamate and PACAP) to the SCN via the optic nerves of the RHT. When an increase in light is detected, the amount of glutamate and PACAP released to the SCN rises and the neurons of the SCN release γ–aminobutyric acid (GABA) to the paraventricular nuclei (PVN) of the hypothalamus. The stimuli then pass from PVN through the intermediolateral column of the spinal cord (IMC) and superior cervical ganglia (SCG) to the pineal gland, where the release of noradrenaline (NA) from postganglionic sympathetic fibers is decreased. In contrast, at night an inhibition of GABA release is generated leading to an increase in NA release, which activates α_1-adrenoceptor and β-adrenoceptor receptors at pinealocyte membrane, leading to an increase in the intracellular level of cAMP and other second messengers such as protein kinase A (PKA) and cAMP response element binding (CREB) that stimulate the activity of AA-NAT and thus, MLT production (figure 8). AA-NAT is generally believed to be the first rate limiting enzyme in MLT production but a recent report has shown that HIOMT probably plays a more important role.[26, 33]

Once synthesized, MLT is not stored, but rather released into capillaries and mainly into the cerebrospinal fluid, to be rapidly distributed throughout all of the body tissues. As a highly lipophilic compound, MLT diffuses with ease through the biological membranes, reaching almost every cell where it provides a daily message to all tissues expressing melatonin receptors in order to synchronize circadian rhythms with external day/night cycle. Melatonin is secreted

mainly during the night, and is two hours after the peak of major secretion, when the highest increase in the trend to sleep is perceived. In addition, MLT is also able to communicate the change of seasons attending to the duration of the nocturnal melatonin peak; it increases with the lengthening of the dark phase of day/night cycle in winter and in contrast it decreases in summer when the dark phase is shortened. [25, 26, 32, 34]
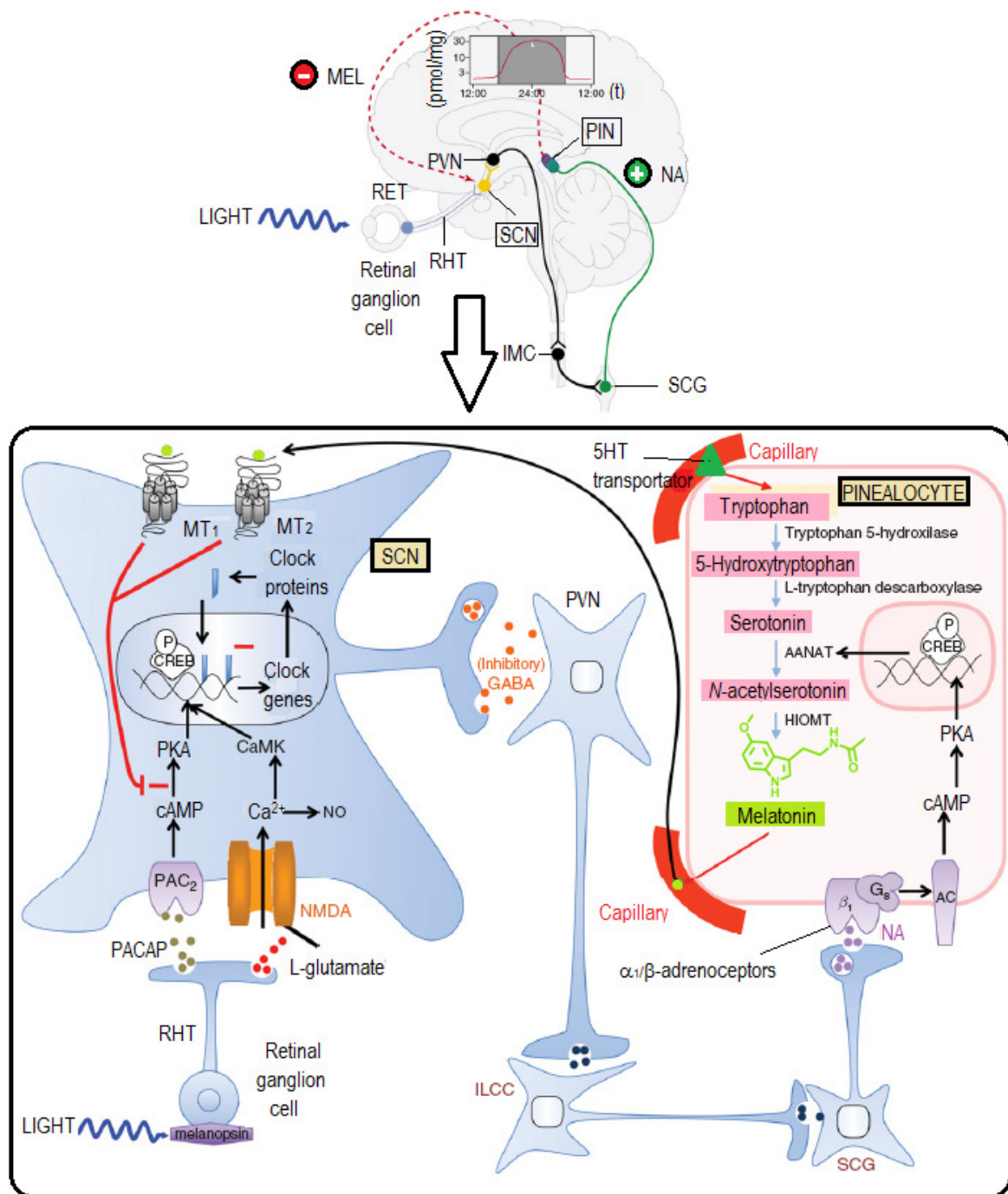
Figure 8. The graph shows the regulation of melatonin biosynthesis and secretion as well as the oscillation of its level in the human pineal gland with respect to clock-time [Figure adapted from *Maronde, E. et all* and *Reiter R.J. et al.].*[27, 33]

Circulating MLT is almost totally derived from the pineal gland and it is metabolized mainly in the liver. MLT is mainly hydroxylated at C6 position by cytochrome P450 (CYP) mono-oxygenase, resulting in 6-hydroxymelatonin formation. This product then undergoes further conjugations to give, among others, 6-sulfatoxymelatonin (aMT6s) which is the main urinary metabolite. Measurements of aMT6s in urine, saliva and plasma have been correlated with pineal melatonin biosynthetic activity in clinical studies. Although aMT6s is the main MLT metabolite, minor melatonin metabolites are also formed as a result of an alternative oxidative catabolic pathway which includes opening of indole core, leading first to the formation of an unstable intermediary compound N^1-acetyl-N^2-formyl-5-methoxykynurenine (AFMK), which is further metabolized to the more stable N^1-acetyl-5-methoxy-kynurenine (AMK). The importance of these metabolites is poorly understood, but they might contribute to some of the unexpected actions of MLT such as its antioxidant effect.[35] MLT can also be metabolized nonenzymatically in the cells or extracellularly by free radicals and other oxidants, and finally, a small amount of melatonin is excreted directly into urine.[25, 26, 36, 37]

Although some of the biological actions of MLT are receptor-independent, the majority of the effects produced by this hormone occur when MLT activates its membrane receptors; these effects include the initiation and maintenance of sleep.[26]

1.2.4. Melatonin receptors

The discovery of 2-[^{125}I]iodomelatonin (^{125}I-MLT) was a major breakthrough in the field of melatonin receptors research because it is a high affinity ligand (in the range of picomolar) and it allows the performance of autoradiographic studies. In fact, melatonin binding sites are expressed in very low density, even in the most sensitive tissues, probably due to the high affinity that melatonin presents for its endogenous receptors. Thus, the use of a radioligand with a lower specific activity would be prevented at least in tissues with low density melatonin binding sites.[38]

It is currently known that there are multiple receptor subtypes available to which melatonin can bind, expressed in a wide variety of tissues all throughout the body. To date, two kinds of MLT binding sites have been characterized, depending on their affinity for the radioligand ^{125}I-MLT; they are high affinity MLT binding sites (ML$_1$) and low affinity MLT binding sites (ML$_2$). The biological actions of MLT in humans are mainly exerted by the activation of two high affinity MLT binding sites denoted as MT$_1$ and MT$_2$ (initially termed as Mel$_{1a}$ and Mel$_{1b}$ receptors), as well as through a low affinity binding site, referred to as MT$_3$. The first two receptors, MT$_1$ and MT$_2$, belong to G-protein coupled receptor (GPCR) superfamily whereas the last one, MT$_3$, is a protein which can be classified in the family of the quinone reductase 2 (QR2). In addition, in the African clawed frog, the *Xenopus laevis*, another ML$_1$, the subtype Mel$_{1c}$, was characterized and cloned, but no mammalian homolog has been identified. Indeed, the fact that Mel$_{1c}$ was the first of the melatonin receptors to be cloned in 1994 is a curious detail.[38-41]
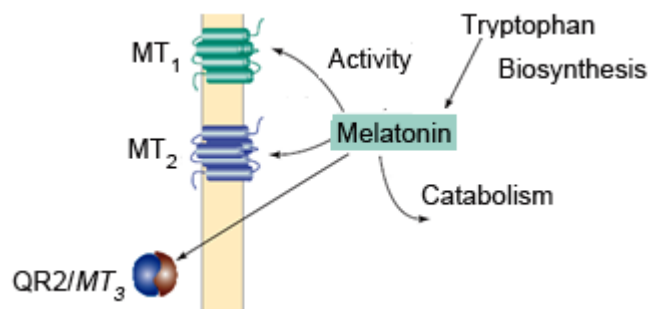
Figure 9. Diagram showing main MLT pathways in the body [Figure adapted from *Boutin, J.A. et al.*].[37]

In mammals, MT_1 receptor has been found in the kidney, retina and brain. The MT_2 receptor has been found in retina and brain also, but unlike MT_1, it has not yet been reported in the human SCN; this is most likely due to the limited amount of MT_2 receptors in this brain structure because it has been detected in the SCN of other species such as mouse and rat. Inside the human brain, MT_1 mRNA has been detected in SCN, cortex, hippocampus, thalamus and cerebellum. MT_2 mRNA appears to be more widely distributed but in smaller amounts; it has been detected in the retina, the hippocampus and in whole brain.[38, 40]

Apart from their distinct distribution, MT_1 and MT_2 receptors also differ in terms of their affinities with melatonin. In the case of humans, the respective Ki values of MT_1 and MT_2 receptors are 80.7 pM and 383 pM.[27] This fact is easily understood taking into account the different molecular structures that these receptors present, where only 60% of their amino acid sequence is identical and the fact that they posses different chromosomal localizations. Structurally, MT_1 is a 350 amino acid protein and MT_2 is a 362 amino acid protein, with calculated molecular weights of 39-40 kDa. As with all GPCRs, MT_1 and MT_2 are characterized by an extracellular *N*-terminus, followed by seven transmembrane helices (TM1-TM7) connected by three alternating intracellular loops (IL1-IL3) and three extracellular loops (EL1-EL3) and finally, an intracellular C-terminus. [40, 42]

Since these receptors belong to GPCRs superfamily, they are extremely difficult to isolate and crystallize because their inclusion in the cell membrane contributes to their structural stability. Thus, due to the impossibility of a direct determination of the structures of MT_1 and MT_2 receptors, the construction of theoretical models has been achieved based on the structure of the bovine rhodopsin (the only GPCR crystallographic structure available).[40, 42] According to the model published by *Farce et al.*, the binding space for MLT on receptor MT_1 appears to be relatively smaller than on MT_2. As can be observed in figure 10, this study states that for the binding of MLT to MT_1, the oxygen of the 5-methoxy group should accept a hydrogen from histidine 195 (His 195) placed in TM5 and the carbonyl oxygen and the hydrogen of the amide function should be engaged in two hydrogen bonds with serines 110 and 114 (Ser 110 and Ser 114) of TM3. In the binding of MLT to MT_2, the same functional groups take place but in this case, the oxygen of the 5-methoxy group binds to His 208 in TM5, the carbonyl of the amide function forms a hydrogen bond with Asn 175 in TM4, and the amine of the amide group forms a second hydrogen bond with the backbone carbonyl of Thr 191 belonging to the EL2.[37]

The orientation of MLT, as well as its conformation, changes upon binding to the two receptor subtypes. In both cases, the 5-methoxy group maintains almost the same orientation,

and the two-methylene linker is placed at a 90º angle from the plane of the indole. The major difference comes from the acetamide group, which in MT_1, appears to adopt a deployed conformation, with the amide bond directed away from the indole and parallel to it, whereas, in MT_2, this bond is directed toward the indole and forms an angle of approximately 30º with the plane of the aromatic ring. This critical conformational difference allows the more flexible compounds to fit in both receptors (figure 10).[40]
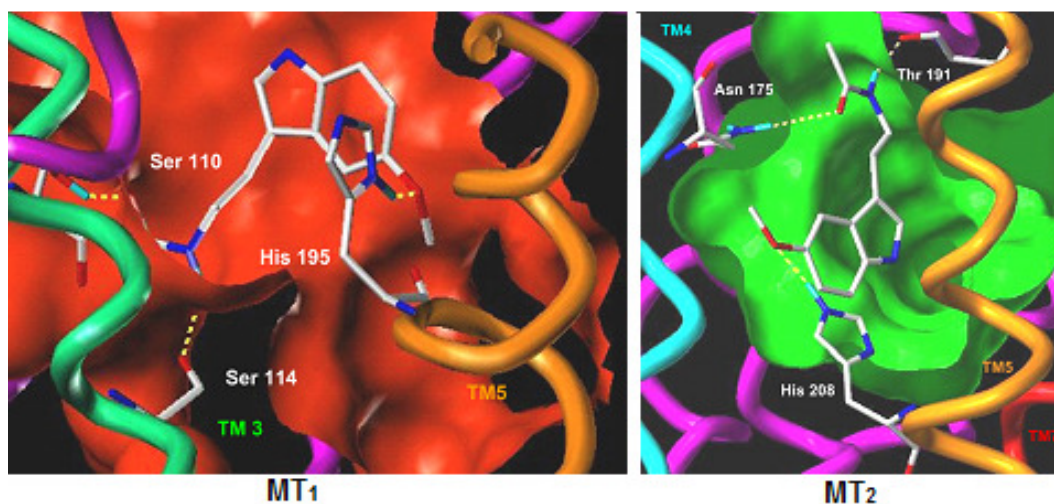
Figure 10. Melatonin binding to MT_1 and MT_2 receptors[Figure adapted from *Farce et al.*].[40]

The administration of MLT at dusk or dawn causes phase shifts in the circadian rhythms of electrical activity registered in the neurons of the SCN. MT_1 and MT_2 appear to be responsible for the effect of promoting sleep and modulating circadian rhythms produced by MLT. MT_2 receptors are responsible for inducing phase shifts and they are involved in the entrainment of circadian rhythmicity. MT_1 receptors are responsible for the suppression of firing in the neurons of the SCN or, in other words, they inhibit neuronal activity, meaning that they are involved in sleep promotion.[43] It is thought that MT_3 plays no role in the sleep-promoting effect or circadian rhythm modulating effect exerted by MLT.[25] Thus, this work will not focus on this receptor.

MLT exerts too many MT_1/MT_2-mediated effects to explain here in this dissertation. Therefore, only a few brief explanations will be given, the majority dealing with the effects on sleep and circadian rhythms regulation generated by MLT. Anyway, almost in every cell the classical signal transduction described for both receptors involves Gαi proteins which inhibit adenylate cyclase activity and as a result cause a decrease in cAMP formation, followed by decline in PKA activity and thus, a decrease in CREB phosphorylation, an early event in the signaling cascade that controls the circadian rhythmicity and MLT production (figures 8, 11 and 12).[42-45] MLT signaling via MT_1 and MT_2 is a very complex process which could not be explained uniquely as Gαi-mediated downregulation of cAMP-dependent processes. The existence of parallel or alternative signaling pathway through different G protein subforms and Gβγ heterodimers have been proved.[43, 44, 46]

Inhibition of neuronal activity and sleep-promoting effect of MLT appears to be mediated through the activation of MT_1 melatonin receptors. After activation of Kir channels (K^+ channels) through Gβγ heterodimers, external K^+ concentration increases and in turn, Ca^{2+} channels are

opened (figure 11). These findings on MT_1 signaling towards Kir and calcium channels can presumably explain the MT_1-mediated suppressive effects of melatonin on neuronal firing in the SCN. Moreover, modulation of ion channels by melatonin in the SCN is believed to contribute to the various anti-excitatory actions of MLT.[44] In addition, some authors have also proposed that the Gβγi may mediate the potentiation of phospholipase C (PLC) activation by prostaglandin $F_{2\alpha}$ ($PGF_{2\alpha}$), leading to an increase in IP_3 (inositol-1,4,5-tris-phosphate) which would also increase the intracellular calcium.[42]
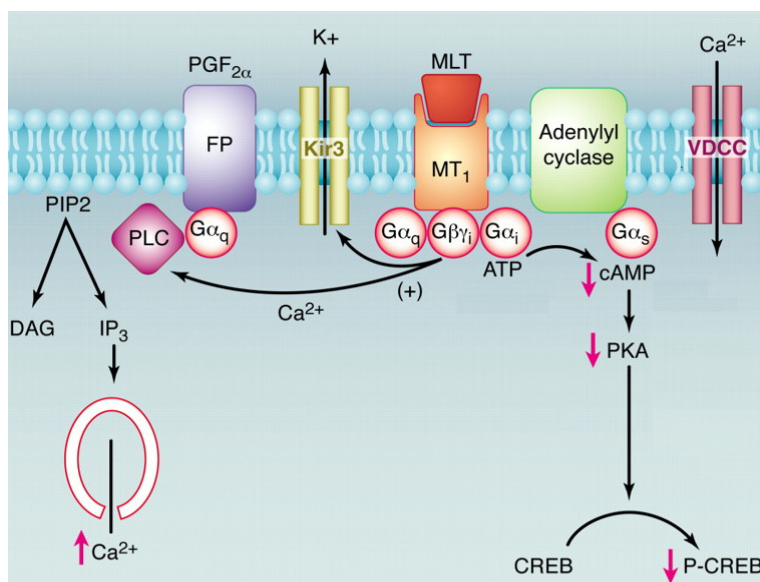
Figure 11: MLT signaling pathway at activating MT_1 receptor [Figure adapted from *Dubocovich et al.*].[47]

MT_2 activation appears to increase levels of PKC required for the phase shift of the circadian rhythm at dusk and dawn. Some G protein subunits are involved in the activation of PKC via activate phospholipase C (PLC) activation (figure 12).[42, 44]
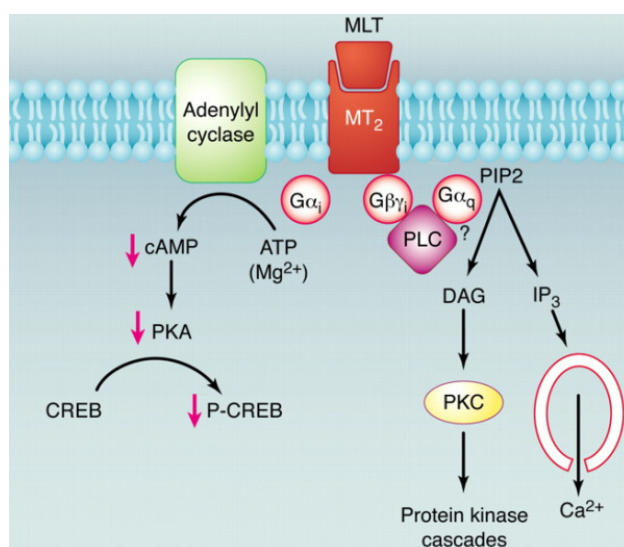
Figure 12. MLT signaling pathway at activating MT_2 receptor [Figure adapted from *Dubocovich et al.*].[47]

2. SLEEP DISORDERS

Sleep disorders include a very frequent pathology of diverse origins with clinical manifestations that vary from insomnia, to hypersomnia, to respiratory disorders, to complex motor disorders, etc. These disorders can appear alone or associated with each other, thereby making this pathology even more complex. In fact, most of the existing diseases have some type of sleep disorder among their symptoms and the majority of the diseases alter sleep in one way or another. From a medical point of view, sleep disorders have been a quite forgotten pathology; therefore, this is a relatively new medicinal area. In fact, the diagnostic and therapeutic advances that have been made in this field have all taken place during the last 40 years.[48]

Currently, sleep loss and sleep disorders are among the most frequently overlooked and readily treatable health problems. Failure to recognize sleep problems impedes diagnosis and treatment as well as the possibility of preventing their serious public health consequences that include from errors in judgment contributing to disastrous events, to impairment in normal functioning and quality of life, family well-being, and health care utilization. Some of these consequences, such as automobile crashes, can easily be linked to sleep problems but in the case of obesity or hypertension, this relationship is not so transparent at first sight. The implication and comorbidity of sleep with other health problems is an undeniable fact. The cumulative effects of sleep loss and sleep deprivation have been associated with a wide range of deleterious health consequences including an increased risk of hypertension, diabetes, obesity, depression, heart attack, and stroke.[49]

Taking into account the high prevalence that these disorders present, it could be said that at some given time throughout their professional life, almost all doctors, regardless of their specialty, will have to face at least one patient who describes some type of sleep alteration.[50] According to the "*National Institute of Neurological Disorders and Stroke*", every year, 40 million North Americans suffer from chronic, long-term sleep disorders and another 20 million suffer some type of occasional sleeping problem. As a result of these disorders, an estimated $16 billion dollars a year are spent in medical costs, and it is known that the indirect costs are probably much greater due to productivity loss and other factors related to sleep deprivation.[51]

2.1. Classification of sleep disorders

The classification of these diseases has evolved over time; first, classification was based on the clinical symptom and later, more emphasis was placed on the diseases themselves.[48] In 1990, the first International Classification of Sleep Disorders (ICSD) was established by the American Academy of Sleep Medicine (AASM) in association with the European Sleep Research Society, the Japanese Society of Sleep Research, and the Latin American Sleep Society. The primary aim of the classification system was to provide a unique code number for each sleep disorder so that the disorders could be efficiently tabulated for diagnostic, statistical, and research purposes by means of diagnostic, severity, and duration criteria present in the classification system.[52]

In 2005, a new classification called ICSD-2 was established, in which the symptoms once again gained more importance. In this new disposition, more than 90 sleep disorders are included (table 1). The aim of this new classification was to include all of the recognized sleep disorders and describe them from a scientific and clinical point of view in order to form a statistically valid rational structure, maintaining the unique code number described in the ICSD of 1990 as much as possible. As ICSD-2 is a more intuitive symptom-based approach, its use is preferred by the medical profession and presently, this classification continues to be the main reference of sleep disorders.[53] In this classification, sleep disorders are sorted into eight categories based on different factors such as a common complaints, a presumed basic etiology or the organ system from which the problems arise.[54] In the following lines, the eight main categories of the ICSD-2 are briefly explained:

- Insomnia: The ICSD-2's criteria for insomnia include complaints of individuals regarding the difficulty for initiating or maintaining sleep or because sleep is chronically non-restorative or poor in quality despite adequate conditions for a full night's sleep. In addition, these problems need to have lasted over a reasonable length of time and followed by daytime impairment.[55]

- Circadian rhythm sleep disorder (CRSD): These disorders are caused by a disruption between the endogenous circadian clock and the external 24 hours day/night cycle. As a result, complaints of insomnia and/or excessive sleepiness arise, in addition to impairment in normal functioning and quality of life.[56, 57]

- Sleep-related movement disorder: These disorders are characterized by relatively simple, usually stereotyped movements which cause sleep disruption. In the ICSD-2, restless legs syndrome (RLS), periodic limb movement disorders (PLMD) and bruxism have been included in this main category.[57, 58]

- Parasomnias: Parasomnias are involuntary and unpleasant experiences that occur during sleep but are not due to abnormalities of the wake-sleep processes. Parasomnias usually manifest themselves when brain activity is reorganizing during transition from one sleep state to another. They consist of abnormal sleep-related movements, behavior, emotions, perceptions, dreaming and autonomic nervous system functioning.[57, 59]

- Sleep-related breathing disorder: Sleep-related breathing disorders are a heterogeneous group of conditions that may be associated with alterations in the structure of sleep, in sleep quality, and in gas exchange or abnormal respiratory pattern during sleep.[60]

- Hypersomnia: This group involves sleep disorders such as obstructive sleep apnea and narcolepsy which cause excessive daytime sleepiness, not explained by volitional sleep deprivation such as disturbed nocturnal sleep or misaligned circadian rhythms. The diagnosis of the sleep disorders involved in this category is mainly based on the Multiple Sleep Latency Test (MSLT), which objectively quantifies daytime sleepiness.[57, 61]

- Isolated symptoms: This category includes sleep-related symptoms that are on the borderline between normal and abnormal sleep, such as sleep talking or snoring.[57]

- Other sleep disorders: This section involves disorders that are difficult to fit into any other group, such as environmental sleep disorders, in which sleep disturbance is caused by a disturbing environmental factor that prevents or interrupts sleep and leads to sleep-related complaints.[57]

Table 1. International classification of sleep disorders, second edition (ICSD-2).[62]

Insomnia	Adjustment insomnia (acute insomnia)
	Psychophysiological insomnia
	Paradoxical insomnia
	Idiopathic insomnia
	Insomnia due to mental disorder
	Inadequate sleep hygiene
	Behavioral insomnia of childhood
	Insomnia due to drug or substance
	Insomnia due to medical condition
	Insomnia not due to substance or known physiological condition, unspecified (nonorganic insomnia, NOS)
	Physiological (organic) insomnia, unspecified
Circadian rhythm sleep disorder	Circadian rhythm disorder, delayed sleep phase type (delayed sleep phase disorder)
	Circadian rhythm disorder, advanced sleep phase type (advanced sleep phase disorder)
	Circadian rhythm disorder, irregular sleep-wake type (irregular sleep-wake rhythm)
	Circadian rhythm disorder, free running type (nonentrained type)
	Circadian rhythm disorder, jet lag type (jet lag disorder)
	Circadian rhythm disorder, shift work type (shift work disorder)
	Circadian rhythm disorder, due to medical condition
	Other circadian rhythm disorder (circadian rhythm disorder, NOS)
	Other circadian rhythm disorder due to drug or substance
Sleep-related movement disorder	Restless legs syndrome
	Periodic limb movement disorder
	Sleep-related leg cramps
	Sleep-related bruxism
	Sleep-related rhythmic movement disorder
	Sleep-related movement disorder, unspecified
	Sleep-related movement disorder due to drug or substance
	Sleep-related movement disorder due to medical condition
Parasomnia	Disorders of arousal from NREM sleep
	-Confusional arousals
	-Sleepwalking
	-Sleep terrors
	Parasomnias usually associated with REM sleep
	-REM sleep behavior disorder
	-Recurrent isolated sleep paralysis
	-Nightmare disorder
	Other parasomnias
	-Sleep-related dissociative disorders
	-Sleep enuresis
	-Sleep-related groaning (catathrenia)
	-Exploding head syndrome
	-Sleep-related hallucinations
	-Sleep-related eating disorder
	-Parasomnia, unspecified
	-Parasomnias due to drug or substance
	-Parasomnias due to medical condition

Sleep-related breathing disorders	Central sleep apnea syndromes
	-Primary central sleep apnea
	-Central sleep apnea due to Cheyne Stokes breathing pattern
	-Central sleep apnea due to high-altitude periodic breathing
	-Central sleep apnea due to medical condition not Cheyne Stokes
	-Central sleep apnea due to drug or substance
	Primary sleep apnea of infancy
	Obstructive Sleep Apnea Syndromes
	-Obstructive sleep apnea, adult
	-Obstructive sleep apnea, pediatric
	Sleep-Related Hypoventilation/Hypoxemic Syndromes
	-Sleep-related nonobstructive alveolar hypoventilation, idiopathic
	-Congenital central alveolar hypoventilation syndrome
	Sleep-Related Hypoventilation/Hypoxemia Due to Medical Conditions
	-Sleep-related hypoventilation/hypoxemia due to pulmonary parenchymal or vascular pathology
	-Sleep-related hypoventilation/hypoxemia due to lower airways obstruction
	-Sleep-related hypoventilation/hypoxemia due to neuromuscular and chest wall disorders
	Other Sleep Apnea/Sleep-Related Breathing Disorder
	-Sleep apnea/sleep-related breathing disorder, unspecified
Hypersomnia of central origin not due to a circadian rhythm disorder, or other cause of disturbed nocturnal sleep	Narcolepsy with cataplexy
	Narcolepsy without cataplexy
	Narcolepsy due to medical condition
	Narcolepsy unspecified
	Recurrent hypersomnia
	-Kleine-Levin syndrome
	-Menstrual-related hypersomnia
	Idiopathic hypersomnia with long sleep time
	Idiopathic hypersomnia without long sleep time
	Behaviorally induced insufficient sleep syndrome
	Hypersomnia due to medical condition
	Hypersomnia due to drug or substance
	Hypersomnia not due to substance or known physiological condition (nonorganic hypersomnia, NOS)
	Physiological (organic) hypersomnia, unspecified (organic hypersomnia, NOS)
Isolated symptoms, apparently normal variants and unresolved issues	Long sleepers
	Short sleepers
	Snoring
	Sleep talking
	Sleep starts (hypnic jerks)
	Benign sleep myoclonus of infancy
	Hypnogogic foot tremor and alternating leg muscle activation during sleep
	Propriospinal myoclonus at sleep onset
	Excessive fragmentary myoclonus
Other sleep disorders	Other physiological (organic) sleep disorder
	Other sleep disorder not due to substance or known sleep disorder
	Environmental sleep disorder

In a full approach to sleep-related problems analysis, the term "sleep loss" should also be included. Sleep loss is a sleep of shorter duration than the average basal need of 7 to 8 hours per night and which is not caused by a formal syndrome or disorder. The main symptoms include excessive daytime sleepiness, depressed mood and poor memory or concentration and thus, have serious consequences affecting health, performance and safety. Sleep loss is a highly prevalent problem that continues to worsen in frequency as individuals grow.

Due to the great number of existing sleep disorders there is no reported cause, prevalence or treatment common to all of them. Each of the described disorders should be analyzed individually in order to be able to define its etiology, epidemiology, morbidity and other factors of interest. The majority of these sleep disorders deal with one or more of the following symptoms: excessive daytime sleepiness, difficulty initiating or maintaining sleep, or abnormal movements, behaviors, or sensations occurring during sleep.[49] The most common sleep disorders include insomnia, sleep apnea, restless legs syndrome, and narcolepsy, with insomnia being the most prevalent among them.[51]

2.2. Insomnia: The most prevalent sleep disorder

Although insomnia is the most commonly reported sleep problem, it often goes unrecognized and untreated despite its adverse impact on health and quality of life. The ICSD-2's diagnostic criteria for primary insomnia include complaints of individuals regarding the difficulty to initiate or maintain sleep or the fact that sleep is chronically non-restorative or poor in quality despite adequate conditions for a full night's sleep. In addition, these problems need to have occurred at least three times per week, and last for at least one month, and they have to be followed by daytime impairment such as fatigue, lack of energy, daytime sleepiness, attention or memory impairment, mood disturbance and/or irritability.

Two different categories of insomnia have been defined; primary insomnia which is not caused by/or associated with another medical, psychiatric or sleep disorder and comorbid insomnia, which is associated with another medical, psychiatric, or sleep disorder (e.g. depression, pain). Both insomnias will include the typical insomnia complaints described above.[63]

2.2.1. *Prevalence of insomnia*

There are different studies regarding the prevalence of insomnia that confirm its great frequency among the general population of many countries. However, the numbers vary from one study to another, often due to variations in the methodologies and the definitions used to assess insomnia.[64, 65] These different studies have been grouped according to the definition used by each author. The first group is made up of those studies in which the definition of insomnia includes the clinical manifestations of insomnia, such as the difficulty to initiate and/or maintain sleep or complaints regarding non-restful sleep, regardless of its duration or consequences. The studies of the second group cover different characteristics, including the manifestations relative to poor sleep, keeping in mind the dissatisfaction with the quantity or quality of sleep; these studies report the sleep problems to be of major concern. Finally, the studies in the third group

respond to the diagnostic criteria of insomnia according to ICSD-2 and relative classifications in which the clinical manifestations of insomnia are taken into account, as well as the diurnal consequences and the dissatisfaction with the quantity or quality of sleep and the duration of the problem.[65]

A compilation of the most recent studies confirmed that approximately one out of three adults in the general population presents at least one characteristic of insomnia. It has been reported that 16-21% of the population suffer these manifestations at least three times a week and 13-17% describe their trouble as being of major concern. When the clinical diurnal consequences of insomnia are added into the definition, the prevalence is placed between 9-15% and 8-18% of the population. The prevalence of the individuals with a diagnosis of insomnia decreases to 6% as shown in figure 13.[64, 65]
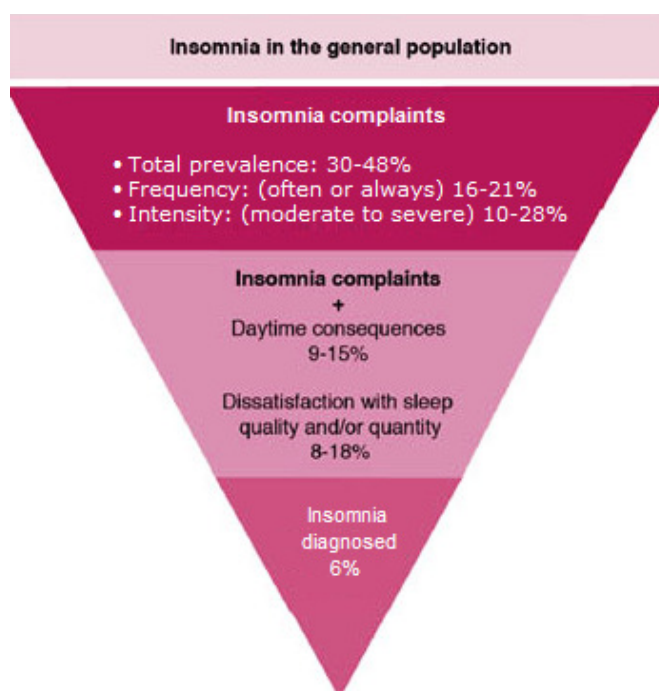
Figure 13. Prevalence of insomnia in the general population according to the different insomnia definitions [Figure adapted from *Grupo de Trabajo de la Guía de Práctica Clínica para el Manejo de Pacientes con Insomnio en Atención Primaria*].[65]

An analysis performed in USA shows that more than 50% of the total patients of primary care complain about insomnia only if they are specifically asked about sleep, 30% mention it on their own initiative, and only 5% come to the consultation for the main purpose of receiving treatment. A recent study also carried out in USA shows that only 15% of people with sleep problems are diagnosed. A total of 4% of these patients are diagnosed with insomnia and only half of them receive subsequent treatment.[62]

Although the studies regarding the prevalence of sleep disorders in Spain are very limited, they show that approximately 23% of the general population present some difficulty concerning sleep and that 11% claim to have insomnia.[63, 65] Disability adjusted life years (DALY) is a measure of overall disease burden, expressed as the number of years lost due to illness, disability, or premature death, thereby, combining mortality and morbidity into a single metric.

It is estimated that in 2006, the Spanish population bore a *burden of disease* equivalent to losing 37,628 years of DALY. Since there were no deaths registered as being caused by insomnia, the entire load is attributable to insomnia's morbidity and associated disabilities. This number of years lived with bad health is 2.6% of the load of total disease for neuropsychiatric diseases and 0.7% of the total DALY of Spain. These numbers are similar to other European countries.[65]

Other international studies where the geographical distribution of insomnia is observed, have demonstrated that the claim of insomnia is universal. These studies have reported that insomnia reaches considerably higher levels in USA (39%) than in Europe (28%) and Japan (21%). This is most likely due to factors which include variation in sleeping habits and climate as well as cultural differences which affect how questionnaires are answered rather than being due to real changes in the prevalence of insomnia. [64] In conclusion, it can be said that the socio-demographic characteristics of the sleepless patients involved in the studies influence the numbers of prevalence obtained to the same degree as the definition of insomnia being considered.[62]

2.2.2. Etiology of insomnia

Regardless of the definition being considered and of the socio-demographic characteristics of the subjects involved, almost all the studies show an increasing prevalence of insomnia with age. In addition, women appear to be at more risk of suffering insomnia than men (figure 14).[64, 65]
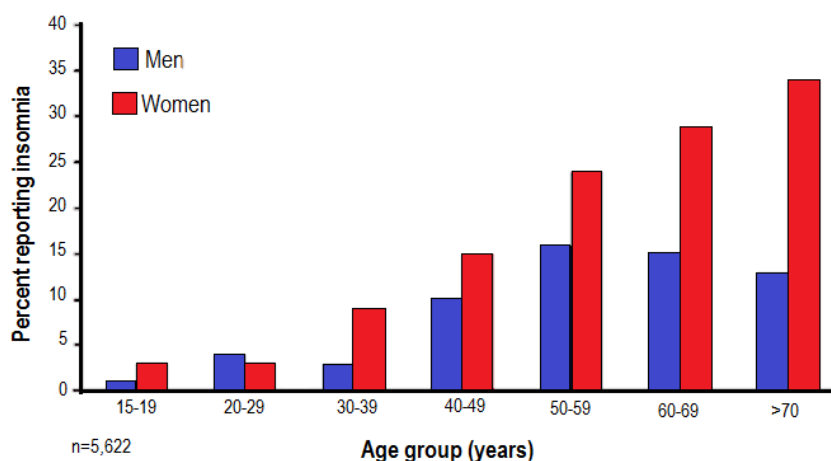
Figure 14. Progression of insomnia by age and gender. The increasing prevalence of insomnia with age and female gender is an undeniable fact [Figure adapted from *Neubauer, D. et al.*].[63]

Apart from female gender and age, there is a broad spectrum of risk factors such as divorce/separation/widowhood, psychiatric illnesses, medical conditions, cigarette smoking, alcohol consumption, coffee consumption, and certain prescription drugs. The analysis of the etiology of insomnia needs a multifactorial approach in which various causes and comorbidities of insomnia are included such as situational disturbances, psychological conditioning, poor sleep

habits, psychiatric disorders, medical disorders, medication effects, and even genetic predisposition.[63]

Insomnia is conceptualized as a disorder of increased activity of the wake-promoting mechanism also referred to as hyperarousal, which exhibits a state of hyper-vigilance during the day and difficulty initiating and maintaining sleep at night. This awakening can be explained cognitively as stressful episodes (e.g. worry and rumination about life events) that disrupt sleep, creating acute episodes of insomnia. These worries shift from life events to sleep itself and the daytime consequences of sleep deprivation, causing insomnia. This development of insomnia can be explained by the "3P model" which assumes that there are predisposing and precipitating factors that induce insomnia and perpetuating factors that prevent it from being resolved (figure 15). The predisposing factors include risk factors such as age, gender, genetics, medical conditions and even personality. The precipitating factors (any stressor) can lead to the development of insomnia. The perpetuating factors can be well-intentioned efforts to cope with insomnia that actually make it persist. For instance, repetitive difficulty falling asleep may result in anxiety at bedtime, thereby making it hard for the individual to fall asleep. Other perpetuating factors may include substance abuse and poor sleep hygiene. Even when the precipitating factors have been resolved, insomnia may persist as a result of these perpetuating factors.[63]
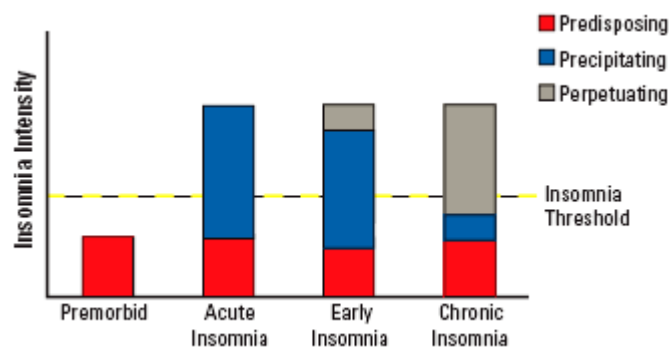
Figure 15. Progression of insomnia explained with 3P model. Predisposing factors themselves do not cause insomnia but when associated with precipitating factors, acute insomnia is reached. If perpetuating factors appear, insomnia becomes an illness. Finally, insomnia is chronic due to the perpetuating factors because without these factors, insomnia would not persist [Figure adapted from *Neubauer, D. et al.*].[63]

It is also thought that there may be a physiological explanation for insomnia, in which the hypothalamic pituitary adrenal axis (HPA) is activated and the stage for chronic insomnia is set up. Cortisol and adrenocorticotropic hormone (ACTH) which are released by HPA axis after stress exposure and are involved in the awakening process, present higher plasma and urinary levels in insomniacs than in normal sleepers. In addition, several lines of evidence from preclinical to sleep neuroimaging studies suggest that there are multiple interacting neural networks in the central nervous system which include a general arousal system, an emotion-regulating system and a cognitive system that contribute to insomnia complaints.[66, 67]

2.2.3. Consequences of insomnia

The influence of insomnia on health has been, epidemiologically speaking, scarcely studied. Nevertheless, the existence of an association between insomnia and a worse health state of individuals has been demonstrated. Insomnia always appears in association with diurnal fatigue and alterations of humor such as irritability, tension or a depressed frame of mind. [64, 67]

There are several studies where the "*36-item Short Form Health Survey of the Medical Outcomes Study*" (SF-36) has been used in order to measure the impairments in an individual's quality of life associated to insomnia. Insomniacs reported a detriment on virtually all dimensions of SF-36 which assess eight domains: physical functioning, role limitation due to physical health problems, bodily pain, general health perceptions, vitality, social functioning, role limitations due to emotional health problems and mental health.[63, 66]

Performance effects from sleep loss include involuntary micro-sleeps, slower response time, a decrease in short-term recall, impatience, deterioration in divergent thinking, increase in absenteeism, lower productivity, reduced acquisition of cognitive tasks, increase of work and traffic accidents, increment in ineffective solutions, etc. Therefore, the aforementioned effects worsen the quality of life and work. [65, 68]

The consequences of chronic insomnia affect not only the health of the patients but also their social and labor functioning. Insomnia is not a visible handicap or disability, so it is difficult for insomniacs to understand and explain that they are sick due to a poor night's sleep and that they need to rest more. Thus, insomniacs have to face a regular work load. As a result, insomniacs have been shown to be slower at work and have poorer career advancement than good sleepers. Insomniacs showed almost twice the rate of absenteeism of good sleepers and also a significantly higher rate of work accidents. Moreover, some researchers hypothesized that co-workers of the absent insomniacs are also more exposed to accidents due to their work overload. Overall, insomnia was found to be a significant predictor of sick leave. There is a study that explored the effects of insomnia on work and shows that during the month in which the study lasted, the number of severe insomniacs (SI) that have reported having made errors at work which could have resulted in serious consequences was greater than the number of good sleepers (GS). In addition, these errors were repeated more frequently in the SI group than in GS group. Reported difficulties completing complicated tasks at work are also greater among the SI group.[64]

2.2.4. Comorbidity of insomnia

Comorbid insomnia is the insomnia associated with another medical, psychiatric, or sleep disorder and these other disturbances are responsible for the development of insomnia.[63] When insomnia becomes chronic, it is associated with major comorbidity. Epidemiological studies have reported that typical manifestations of insomnia are related to respiratory diseases (pulmonary obstructive chronic disease, asthma, chronic bronchitis), rheumatic diseases, cardiovascular diseases (coronary disease, hypertension) and cerebro-vascular diseases (ictus), diabetes and diseases that involve pain, among others.[65] When the comorbid disorder is treated, the insomnia complaint often persists and may require further evaluation and treatment.[63]

As explained previously, insomnia is associated with the aging process. This fact could mainly be due to the impact of comorbid conditions. Among older populations, the amount of time in bed and the amount of time spent awake in bed, increases. This data suggests that although the need for sleep remains almost the same with age, the ability to fall asleep and remain asleep decreases. This fact appears to be a result of the high rate of comorbid conditions in the elderly leading to a disrupted sleep. A large study of the general population demonstrated that conditions increasing in prevalence as populations become older (e.g., congestive heart failure, angina, and arthritis) are associated with severe insomnia. The role that pain plays in disrupting sleep in disorders such as arthritis, post-trauma and/or other acute or chronic pain conditions is readily apparent; any movement may provoke pain and arouse the individual to move from a position that could damage the body. Sleep is also disrupted in conditions not conventionally linked to pain, such as diabetes, and these reasons have not yet been delineated.[63]

Population and clinically based studies have demonstrated that insomnia is more frequently associated with psychiatric disorders than any other medical illness. Traditionally, in depression and anxiety comorbidities, it has been assumed that insomnia is secondary to the psychiatric disorder. However, it is currently thought that given the chronicity of insomnia, in most cases, insomnia could precede the psychiatric disorder. Some studies showed that persistent insomnia over a number of years coincides with an increased probability of new onset psychiatric disorders, especially depression, anxiety or alcoholism (figure 16). Moreover, it has been observed that in insomniacs who have not resolved their problem after one year, the risk of developing a psychiatric disorder is significantly higher than in those who have resolved it. More specifically, the risk of developing depression is nearly 30-fold higher in the case that insomnia has not been resolved after one year (figure 16).[63, 65, 66]
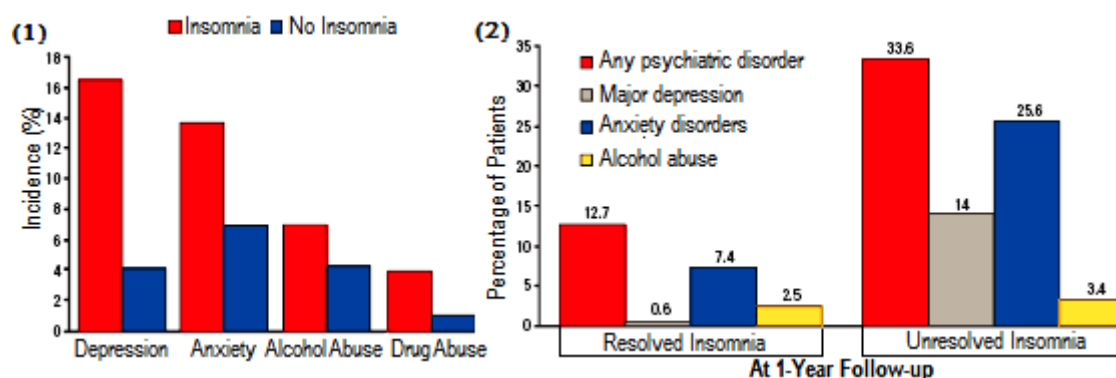
Figure 16. Comparative graphs regarding the relationship between the occurrence of psychiatric disorders and insomnia: (1) People who suffer from persistent insomnia have a higher probability of also suffering other psychiatric disorders than good sleepers do. (2) People with insomnia who have not resolved this problem after one year have a higher probability of suffering other psychiatric disorders than those who have resolved it [Figure adapted from *Neubauer, D. et al.*].[63]

2.2.5. *Insomnia and depression relationship*

Sleep disturbance is closely linked to major depression, as demonstrated in different clinical and epidemiological studies. To further understand the relationship between sleep and depression, several studies have examined the evolution of depression among insomnia patients and found that insomnia confers a substantial risk for developing a depressive disorder.[66, 69]

Some studies suggest that the link between insomnia and depression is bidirectional. For instance, approximately 20% of patients with insomnia exhibit some depressive symptoms whereas about three quarters of all depressed patients complain of difficulty either in initiating or in maintaining sleep and 41% of them report a number of insomnia symptoms sufficient enough to warrant an additional diagnosis of insomnia.[69]

Depressed patients who experience improvements in sleep will also experience a more rapid antidepressant response while those patients whose insomnia persists, present a shorter time to recurrence. At the same time, insomnia has been shown to become chronic despite a successful resolution of depressive symptoms.[66, 69]

In major depression, the patients present a defined EEG profile which includes a disruption of sleep continuity, deficit of slow wave sleep (SWS) of NREM sleep, especially during the first sleep cycle and prolonged REM sleep time. The EEG profile of the primary insomniacs is comparable to those of major depressive patients but with the difference being that in their EEG, no consistent REM sleep abnormalities have been shown. Positron Emission Tomography (PET) studies grossly corroborate sleep EEG findings in both, primary insomniacs and major depressed patients. PET studies show an increased activation of REM-related structures during REM sleep only in the case of depressive patients, and there are smaller indices of thalamo-cortical deactivation during NREM sleep in both cases. The changes in NREM sleep suggest that an increased activity of wake-promoting mechanisms is operating in both insomnia and depression.[66, 69]

As has been explained before, a key study showed that over a 24-hour period, adults with insomnia have higher levels of cortisol and adrenocorticotropic hormone (ACTH), than normal sleepers, which are hormones released by the hypothalamic-pituitary-adrenal axis after stress exposure.[67] A similar pathophysiological mechanism has been proposed in major depression in which, despite the removal of the stressful situation, an arousal reaction is maintained due to a maladaptive cognitive functioning. Indeed, stress-induced arousal responses that involve the HPA axis, have also been clearly demonstrated in depressed patients.[66] Thus, stress is thought to play a leading role in activating the hypothalamic-pituitary axis and setting the stage for chronic insomnia.[67]

2.2.6. *Treatment of insomnia*

The treatment of insomnia involves a great variety of treatments including psychological and behavioral therapies as well as pharmacological therapies that can be used jointly as part of a larger treatment plan, or independently. In many cases, combining the behavioral and pharmacological strategies results in the most efficacious treatment.[63, 66, 67]

2.2.6.1. NON-PHARMACOLOGICAL TREATMENT

- Sleep hygiene

The sleep hygiene interventions involve a series of conduct habits that facilitate the beginning or maintenance of sleep and that are the common denominator of any therapeutic intervention used for any sleep disorder. The interventions of sleep hygiene gather a series of recommendations that people with insomnia problems can adopt in order to minimize their incidence and favor a normal sleep. The aim is to acquire better habits that are convenient for everybody such as physical activity or changing those habits which may interfere with sleep, as in the case of caffeine intake. The fundamental interventions of sleep hygiene centered on insomnia are as follows:[65]

1. Go to bed only when feeling sleepy
2. Get up every day at the same hour, including the weekends
3. Avoid remaining awake in bed more time than is strictly necessary
4. Avoid taking a siesta during the day
5. Reduce or avoid the consumption of alcohol, caffeine and hypnotics
6. Avoid copious food ingestion before going to bed
7. Maintain environmental conditions adapted to sleep (temperature, ventilation, noises, light)
8. Avoid stressful activities in the hours previous to going to bed
9. Carry out moderate physical exercise at the end of the evening
10. Practice relaxation exercises before going to bed
11. Take water baths at corporal temperature due to its relaxing effect

The results published regarding the efficiency of the interventions of sleep hygiene coincide in that they all indicate that they are not capable of solving insomnia on their own and that it is when sleep hygiene is used in combination with other therapeutic interventions that the best results are obtained. [63, 65]

- Psychological intervention

The psychotherapeutic approach of insomnia arises mainly from the approach of behavioral and cognitive-behavioral interventions. The principal aim of the behavioral techniques applied to the treatment of insomnia is to modify the learned conducts and cognitions with regard to sleep which are factors that perpetuate sleep disturbances. These factors can be due to poor sleeping habits such as remaining too much time in bed, to irregular sleep-awake patterns such as waking up at different hours and to psychophysiological hyperactivation. The cognitive techniques are used to identify and analyze the dysfunctional thoughts and beliefs related to sleep in order to relate them to the symptoms of insomnia and thereby be able to construct more functional behavioral patterns. This technique is also used for the purpose of reducing the anticipatory anxiety produced before going to bed, related to the fear of not being able to fall sleep, and it is precisely this fear which prevents sleep. The cognitive-behavioral therapies more often used in the treatment of insomnia are listed below:[65]

1. Control of stimuli: Based on principles of the classic conditioning, this intervention tries to associate the use of bed with sleep. The waking up hour is fixed to stabilize the sleep-wake rhythm.

2. Restriction of time in bed: This intervention tries to consolidate sleep, shortening the time that is spent in bed, to increase the "efficiency" of sleep.

3. Relaxing and breathing: This technique is practiced with the aim of reaching a condition of hypoactivation that offsets the anticipatory anxiety.

4. Cognitive restructuring: This strategy is used to identify irrational or distorted thoughts regarding sleep and substitute them for more rational ones.

5. Paradoxical intention: In this approach it is assumed that a very important part of the sleeping problem is the anxiety produced on the part of the patient as result of the attempts made to try and solve the sleep-related problem. Thus, the patient is asked to do the opposite, to strain from sleeping.

2.2.6.2. PHARMACOLOGICAL TREATMENT

The pharmacological treatment is a helping measure within the integral treatment of insomnia. The selection of a given drug for the treatment of insomnia is based on diverse factors such as the symptoms, the aims of the treatment, the answers to past treatments, the patient's preference, the cost of the drug, availability of other treatments, the comorbidity conditions, contraindications and interactions with other medicines and side effects. In addition, it is necessary to consider the individual characteristics of each patient such as personality since certain types of personality predispose the individual to suffer dependence on drugs or substances. If this exists, it is necessary to consider it to be a contraindication to expire depressors of the central nervous system such as benzodiazepines (BZDs).[65]

In a brief overview of the drugs used in the pharmacological treatment of insomnia, currently approved and emerging pharmacotherapies are mentioned.

2.2.6.2.1. Hypnotics

The first pharmacological approaches were focused on GABA, the major inhibitory neurotransmitter in the central nervous system. Thus, the traditional hypnotics are benzodiazepine receptor agonists (BRAs), and exercise their action on the GABA-benzodiazepine-receptor complex, though the different types act on different subunits of the receptor. During approximately three decades, the pharmacological treatment of insomnia has been dominated by the use of benzodiazepines (BZDs). Later, the non-benzodiazepines (NBZDs) appeared with the aim of reducing the side effects promoted by the BZDs. Although these drugs improved binding selectivity and pharmacokinetic profiles, the enduring adverse effects highlighted the need for new treatment strategies. Among the emerging new treatments, the late-stage investigational compound gaboxadol, which influences GABAergic neurotransmission through a novel mechanism, appears on the scene.[65, 70]

- Benzodiazepines (BZDs)

BZDs (including flurazepam, triazolam, quazepam, loprazolam, midazolam, flunitrazepam, brotizolam, diazepam, lorazepam, nitrazepam, estazolam and temazepam) are central nervous system depressor agents that reduce the nervous activity in the brain, promoting physical and mental relaxation. BZDs allosterically enhances the ability of GABA; upon binding to the receptor, they provoke a change of configuration of the receptor in order to provide higher affinity for the GABA neurotransmitter. As a result, sleep latency and the number and duration of awakenings are significantly reduced, resulting in increased total sleep and improved sleep continuity.

These drugs can produce various secondary effects such as daytime sleepiness, headache, vertigo, nauseas, fatigue and loss of memory among others, mostly due to the relatively nonselective binding of the BZDs.[70] Moreover, clinically significant interactions with other drugs and alcohol have been described. In addition, if their use lasts longer than two weeks, the risk of developing tolerance as well as physical and psychical dependence increases enormously. It is also remarkable that when the treatment is suddenly interrupted, abstinence is reported. Thus, their use is recommended for only a short time when treating acute insomnia.[65]

With regard to their chemical structure, they can be catalogued into two different groups as can be observed in figure 17; the 1,4-benzodiazepines and the triazolobenzodiazepines. The two groups present different pharmacokinetic and pharmacodynamic properties, they do not present cross tolerance, and the adverse reactions profile is also different for each group. The triazolobenzodiazepines present better pharmacodynamic properties than the 1,4-benzodiazepines but this fact also potentiates their secondary effects. Since the triazolobenzodiazepines contain higher potency and a smaller elimination half-life, the probability of producing abstinence, tolerance, efficacy loss and dependence is even higher than in the 1,4-benzodiazepines. Therefore, they provide the worst risk/benefit profile.[65]

Flurazepam
(1,4-benzodiazepine)

Triazolam
(triazolobenzodiazepine)
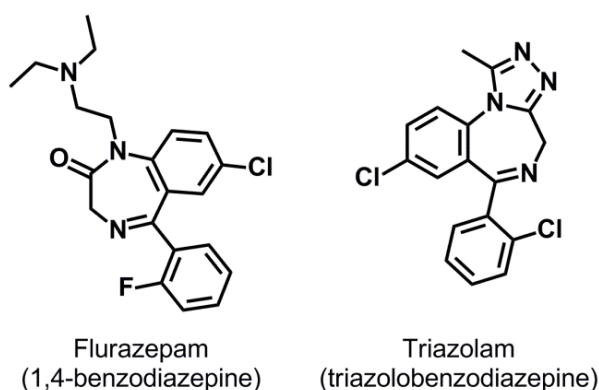
Figure 17. Some examples of BZDs and their chemical structures; flurazepam from the family of 1,4-benzodiazepines and triazolam from the family of triazolobenzodiazepines.

- Non-benzodiazepines (NBZDs)

During the past 25 years, several non-benzodiazepine hypnotics have been developed, belonging to three different classes depending on the chemical group they present (figure 18): cyclopyrrolones (zopiclone), imidazopyridines (zolpidem) and pyrazolopyrimidines (zaleplon).

The NBZDs have distinct pharmacological profiles from those of the classical BZDs and are regarded as an important advance in the treatment of insomnia. Receptor binding affinities and efficacy in modulating GABA response at different receptor subtypes vary among the different NBZDs but all of them have more selectivity than the BZDs and as result, their secondary effects are reduced. Almost all of the NBZDs have been approved for their use but only during a limited period of time (<35 days) because abuse of these compounds produces benzodiazepine-like effects such as the amnesic effects, next day residual sedation and physical dependence. Therefore, there is great interest in developing compounds that target different receptors that may have more specific effects on the regulation of sleep without these side effects. [65, 70]
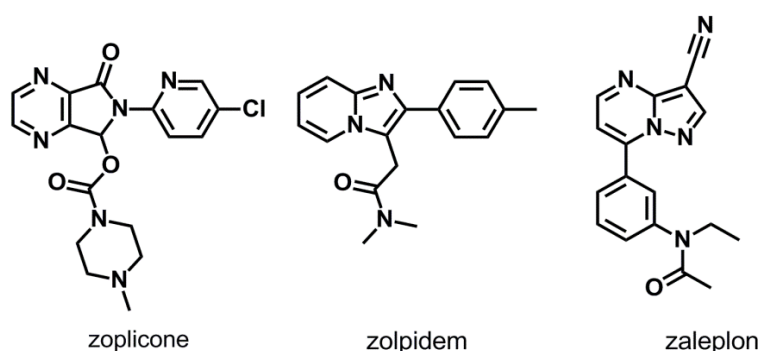
Figure 18. Some NBZD's and their chemical structures; zopiclone, zolpidem and zaleplon.

- Gaboxadol

Gaboxadol (figure 19), which is a selective extrasynaptic GABA (SEGA) agonist developed in the 70s as a possible anticonvulsive agent, has demonstrated benefits in the treatment of insomnia because, after administration, no evidence of adverse effects on next-day cognitive performance has been reported, although further investigation should be carried out in this area. This new class of GABAergic drugs selectively activates receptors located primarily at regions outside of the synapse, binding to the same site used by the endogenous ligand GABA. Therefore, gaboxadol is able to exert direct effects independent of GABA, in contrast to the allosteric enhancement mediated by BRAs.[65, 70]
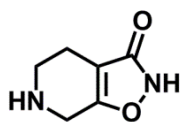
Figure 19. Chemical structure of gaboxadol

2.2.6.2.2. Antidepressants

Antidepressants have also been used as hypnotics for the treatment of insomnia but always in a smaller dose than when given to patients suffering depression. The efficiency and safety of antidepressants with sedative effect (trazodone, mirtazapine, doxepin, trimipramine and amitriptyline) have been demonstrated in the treatment of insomnia associated with depressive

disorders but there is not enough evidence regarding the efficiency and safety of these drugs in the treatment of primary insomnia. Nonetheless, their use for the treatment of primary insomnia has increased substantially in some countries over the last decade. There have been very few objective studies performed regarding the positive effects produced by these drugs in the treatment of primary insomnia and yet the severe side effects that they provoke are well known. In conclusion, due to the severe adverse effects that these drugs present and due to the lack of conclusive studies regarding their efficiency for the treatment of primary insomnia, the use of these antidepressants as sleep inducing agents is only indicated in the case of insomnia associated to depressive disorders.[65, 70]
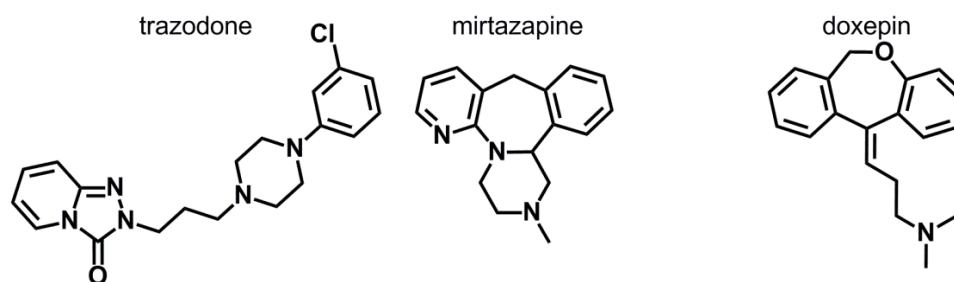
Figure 20. Some examples of antidepressant drugs and their chemical structures; trazodone, mirtazapine and doxepin.

2.2.6.2.3. Atypical antipsychotics and anticonvulsants

Apart from the potential side effects that they could present, the tests carried out regarding the efficiency of drugs such as the gabapentin, tiagabine, quetiapine, and olanzapine for the treatment of insomnia are insufficient, and therefore, their use is not recommended.
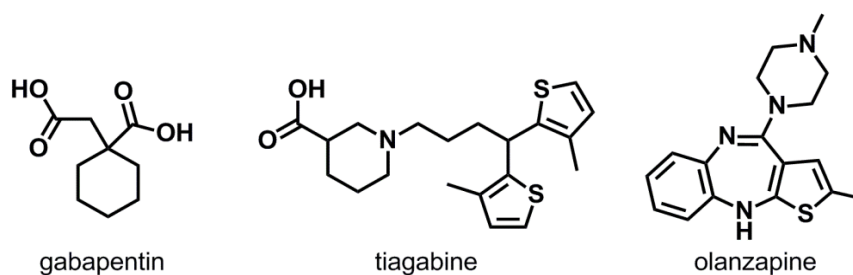
Figure 21. Some antipsychotic drugs and their chemical structures; gabapentin, tiagabine and olanzapine.

2.2.6.2.4. Non-recommended FDA approved drugs

Although drugs such as chloral hydrate, meprobamate and barbiturates (such as amobarbital) (figure 22) are approved by the FDA (*Food and Drugs Administration* of United States of America) their use is not recommended for the treatment of insomnia due to the lack of efficiency tests performed and the significant adverse effects that these drugs present.
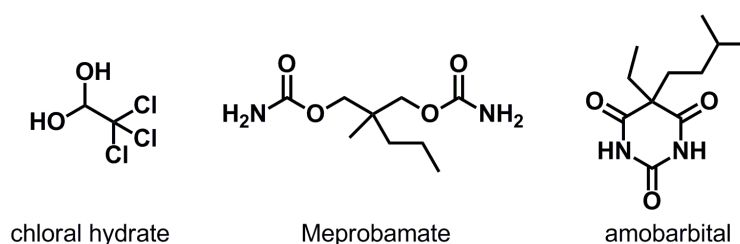
Figure 22. Chemical structures of some drugs approved by the FDA but whose use for treating insomnia is not recommended; chloral hydrate, meprobamate, amobarbital

2.2.6.2.5. MT_1 and MT_2 melatonin receptors agonists

In the past few years, melatonin receptors MT_1 and MT_2 have become one of the most interesting pharmacological targets for the treatment of many sleep disorders. Ever since the idea arose with regard to melatonin acting as a hormonal photoperiodic messenger controlling the seasonal physiological changes and the biological "clocks" of the individuals, the hypothesis that MLT administration might be a potential therapy to treat sleep disorders gained more acceptance. Following this hypothesis, the effect of resynchronizing circadian rhythms that are induced by MLT is of great interest in the treatment of circadian rhythm disorders produced as a result of jet lag, shift work and blindness as well as in physiopathological conditions of the advanced/delayed sleep phase disorder. On the other hand, the hypnotic effect that MLT provokes might go towards the prevention of insomnia.[71-74]

- Melatonin

Circadin® (Neurim Pharmaceuticals; Tel Aviv, Israel) is a prolonged-release 2 mg MLT (figure 23) formulation that was authorized for its commercialization in Europe by the European Medicines Evaluation Agency (EMEA) in June 2007.
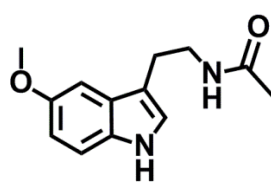
Figure 23. Chemical structure of melatonin (MLT).

When taken before bedtime, it maintains effective serum concentrations of the endogenous hormone throughout the night. Therefore, with regard to the low toxicity and absence of the severe side effects that this drug presents, it was thought to be the ideal hypnotic for short-term treatment of primary insomnia characterized by poor quality of sleep in patients aged 55 years and over. In clinical trials controlled with a placebo, it was demonstrated that, in addition to decreasing sleep latency, Circadin® also improves sleep quality and next day alertness. However, the effect provoked in the maintenance of sleep and the time of total sleep is quite poor. In contrast with traditional sedative hypnotics, Circadin® has not shown any proof of impairing cognitive and psychomotor skills, or of provoking effects of rebound, dependence or

abuse or of any other significant adverse effect compared to the placebo. Some studies demonstrated that six months of continuous administration of Circadin® is safe and efficacious for the treatment of insomnia, jet lag and shift work disturbances. Moreover, in exploratory studies Circadin® also improves sleep quality in patients with chronic schizophrenia as well as in patients with major depressive disorders.[72, 75]

MLT does not cause hardly any negative effects, but in turn, only presents short-term effects because it has a very short half-life in the organism. The discovery of new agonists of MLT receptors with a longer half-life and longer lasting effects can be of great interest in the treatment of sleep disorders in general, and more specifically, in the treatment of insomnia. Therefore, over the past few years, many research groups have centered their efforts on obtaining new MLT receptor agonists and this research has then led to the production of approved or experimental drugs, such as ramelteon, agomelatine, tasimelteon and one subtype selective MLT receptor agonist such as IIK7.[72, 73] It is important to point out that the obtainment of selective compounds is also of vital importance in order to broaden our knowledge regarding the exact biological functions displayed by each receptor and the physiological effects they activate as well as to secure new drugs that would allow a more direct and specific treatment against each type of insomnia and other sleep disorders.

- Ramelteon

Rozerem® (Takeda Pharmaceuticals Inc., Osaka, Japan) is the trademark of Ramelteon, (figure 24) a tricyclic synthetic analog of MLT, which contains an indenofuran ring replacing the indole ring found in MLT. It was approved by the FDA (July 2005) for the treatment of insomnia characterized by difficulty in sleep onset. In contrast, it has not been commercialized in the EU; since EMEA found the efficacy of ramelteon in improving sleep maintenance insufficient for marketing authorization. The EMEA stated that there is a lack of data regarding the pharmacological effect that repeated administration of ramelteon can cause and it can be expected that a repeated administration will result in an adaptation of the body to the effect. Thus, Takeda withdrew its European marketing authorization application.[72, 75]

Ramelteon binds to MT_1 and MT_2 receptors with higher affinity than MLT without any significant affinity for a large number of other CNS binding sites. Ramelteon is metabolized primarily via oxidation generating 2-hydroxy-*N*-[2-(2,6,7,8- tetrahydro-1*H*-indeno[5,4-*b*]furan-8-yl)ethyl]propanamide (M-II) as major metabolite. M-II retains good MLT receptor affinity (about 17 to 25 fold less potent than ramelteon *in vitro)* and there are studies that demonstrate that it can also present a potent sleep-promoting action. Therefore, it may also contribute to the pharmacological effects of ramelteon.[72]
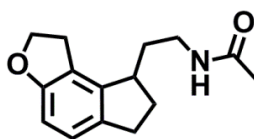
Figure 24. Chemical structure of ramelteon.

Beneficial effects of ramelteon on sleep as well as on training circadian rhythms have been confirmed in clinical studies. Considering the clinical information, ramelteon has demonstrated to have significant sleep-promoting effects, as evidenced by reductions in the sleep latency (approximately 10-13 minutes more than placebo) and increases in total sleep time (approximately 12 minutes) found between patients with chronic insomnia and subjects with transitory insomnia, during both short- and long-term treatments. Just like that which was reported for Circadin**®**, ramelteon has not been shown to have any severe next-day adverse effects such as impairing cognitive and psychomotor skills or provoking rebound, dependence or abuse effects.[72]

- Agomelatine

Valdoxan® (Servier, France and Novartis, USA) is a novel antidepressant with an innovative pharmacological profile, whose active principle is the naphthalenic MLT bioisostere agomelatine (figure 25), a potent melatonin receptor MT_1/MT_2 agonist and weak $5-HT_{2c}$ antagonist with no significant binding affinity to other CNS receptors. The efficacy, tolerability and safety of agomelatine have been assessed in several studies in adults of all ages, including the severely depressed and elderly depressed, and no dependence symptoms are observed after cessation of agomelatine treatment.

Valdoxan® was approved by the EMEA for the treatment of major depressive disorders in February 2009. In USA, Phase III clinical trials of agomelatine for depression have also been conducted but the results have not been revealed yet so, agomelatine is not yet approved by the FDA.[72]
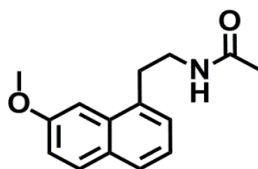
Figure 25. Chemical structure of agomelatine.

Apart from being an effective antidepressant, agomelatine diminishes the severity of depression associated anxiety, and, unlike other antidepressants, it has a notably positive impact on the synchronization of wake-sleep rhythms often disrupted in depressed patients, without affecting next-day vigilance. In addition, agomelatine effectively reduces sleep complaints in depressed patients. When treating depressed patients with agomelatine, the duration of NREM sleep increases without affecting REM sleep, thereby improving both sleep quality and continuity. In addition, it is reported that agomelatine normalizes the changes of NREM sleep found in depressed patients. Agomelatine is a dual action drug that can produce rapid antidepressive effects and also improve the quality of sleep. Clinically, this is a very important point because as explained previously, improvements in the sleep of depressed patients is associated with a reduced rate of recurrence of depressive symptoms and, inversely, the complaints of poor sleep made by depressed patients are associated with a poor response to subsequent antidepressant treatment. [72, 76]

- Tasimelteon

Tasimelteon (figure 26), a new experimental synthetic melatoninergic agonist developed by Vanda Pharmaceuticals under license from Bristol-Myers Squibb, reportes beneficial effects on sleep latency and maintenance and shows a good safety profile without significant side effects in comparison with a placebo. Moreover, the compound is effective in resetting the circadian rhythm, suggesting that it may be a good candidate for the treatment of circadian rhythm sleep disorders (CRSD), especially for jet lag, shift work and non-24-hour sleep-wake disorder (N24HSWD). In 2010, the FDA gave orphan drug designation status for tasimelteon in N24HSWD in blind individuals without light perception. Just like the previous melatoninergic drugs, tasimelteon does not present severe adverse effects. [72, 76, 77]

Figure 26. Chemical structure of Tasimelteon.

- IIK7

The tetracyclic indole derivative IIK7 (figure 27) was the first high affinity MT_2-selective agonist to be discovered. IIK7 produces a significant decrease in NREM sleep onset latency in rats, suggesting that the MT_2 receptor is the subtype involved in the acute sleep-promoting action of MLT. However, since rodents are active at night, when MLT concentrations are higher, these findings have to be further studied, for example, incorporating diurnal animals in the studies.[72]
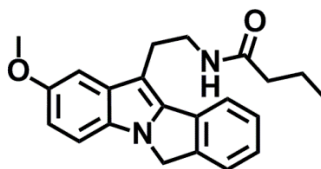
Figure 27. Chemical structure of IIK7.

2.2.7. *Importance of developing new MLT receptor agonists*

MLT receptor agonists represent a novel approach in the therapeutic management of some sleep disorders. Moreover, exogenous administration of MLT has been shown to decrease sleep latency and to help in the adaptation of the phase shifts of circadian rhythms with the day/night cycle. However, effects on sleep efficiency and total sleep have been inconsistent, most likely due to the short half-life of MLT.[70]

It is well known that MLT promotes sleep via its high affinity receptors MT_1 and MT_2 but in addition, it also displays a variety of receptor-independent responses. Thus, the multiple effects that MLT can promote in the body have motivated the design and synthesis of new agents capable of interacting selectively with melatonin MT_1 and MT_2 receptors, without promoting the receptor-independent responses of MLT. Moreover, obtaining compounds selective to one of the

two high affinity MLT receptor subtypes is of vital importance in order to broaden our knowledge regarding the exact biological functions displayed by each receptor and the physiological effects they activate as well as to secure new drugs that would allow a more direct and specific treatment against each type of existing disorders.

Over the past few decades, many MLT receptor agonists have been obtained but none of them fulfill all of the aforementioned requirements which are essential for these drugs to become ideal hypnotics. Therefore, at present, the discovery and development of more potent melatonin receptor agonists with a better pharmacokinetic profile, longer half-life and with complete absence of severe side effects is still needed.

II. PRECEDENTS AND JUSTIFICATION

3. MT₁/MT₂ RECEPTORS AGONISTS. STATE OF THE ART

In the last few decades, a multitude of research groups have centered their efforts on the obtainment of new agonists of melatonin receptors MT₁ and MT₂ in order to find a new drug for the treatment of diverse sleep disorders, as well as to recover a great number of structural variations that, in combination with activity information, would result in structure-activity relationship (SAR) studies.

An extensive bibliographic review has been carried out on the most prominent structures published in the years previous to the beginning of this project. As can be observed in table 2, among the MT₁/MT₂ receptor agonists published from 2000 to 2008, great structural diversity is found in both types because the rings used as central core and the chains replaced on said central core are different.

Table 2. Some of the MT₁/MT₂ agonists published in *Drug Data Report* (DDR) between 2000 and 2008.

DDR 2000.[78, 79]

CNRS (283685)
High affinity with melatonin receptors.
Full agonist. (20-fold higher affinity for
MT₁/MT₂ and 25-fold greater agonist
activity than melatonin)

ADIR (283669)
High affinity with melatonin receptors.

DDR 2001.[80, 81]

ADIR (294907)
High affinity with melatonin receptors.

BRISTOL-MYERS SQUIBB.
IC₅₀<10 nM with MT₁ human
receptor in NIH-3T3 cells

DDR 2003.[82, 83]

BRISTOL-MYERS SQUIBB.
High MT₂ affinity (IC₅₀=1.7 nM)
with >100 fold selectivity over MT₁.
Full agonist.

BRISTOL-MYERS SQUIBB.
High MT₁/MT₂ ffinity (IC₅₀<10 nM in
radioligand binding assays)

DDR 2004.[84]
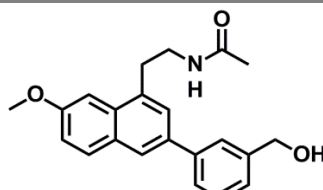
SERVIER (372483)
High affinity with MT_2 (Ki=0.36 nM
in radioligand binding assays)

DDR 2005.[85, 86]
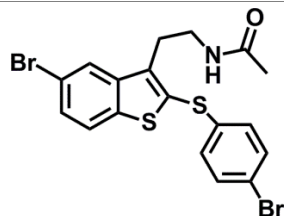
SERVIER (396644)
High affinity with human MT_1/MT_2 (IC$_{50}$=80
nM and IC$_{50}$=1 nM respectively with ^{125}I-MLT
displacement assay)
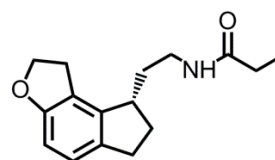
TAKEDA [ROZEREM ®]
(RAMELTEON) (255673)
Drug approved by the FDA for the
treatment of primary insomnia
characterized by difficulty with sleep onset.

DDR 2007.[87]
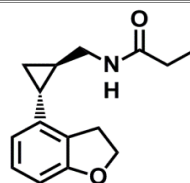
BRISTOL-MYERS SQUIBB.
VANDA PHARMACEUTICAL
MT_1/MT_2 dual agonist. Preclinic efficiency in
the treatment of sleep disorders resulting
from alterations on circadian rhythms

DDR 2008.[88-90]
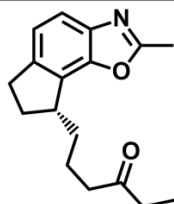
TAKEDA
High affinity with human MT_1/MT_2
(IC$_{50}$=0.030 nM and IC$_{50}$=0.049 nM
respectively in ^{125}I-MLT
displacement assays in CHO cells)
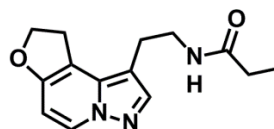
TAKEDA
High affinity with human
MT_1/MT_2 (IC$_{50}$<100 nM a in
^{125}I-MLT displacement
assays in CHO cells)
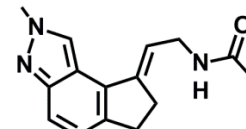
TAKEDA
High affinity with human
MT_1/MT_2 (IC$_{50}$<100 nM a in
^{125}I-MLT displacement
assays in CHO cells)

It is necessary to keep in mind that all the MT_1/MT_2 agonists appearing in table 2 contain, without exception, an aliphatic chain of variable length substituted on a central aromatic ring and linked to an alkylamide (or derivative). Moreover, in the majority of these cases, the structures also possess a methoxy group or an alkoxy group joined to the central core. In addition, the existing space between the oxygen atom of the alkoxy group and the nitrogen atom of the amide group is often a six-atom distance.

Therefore, some structural requirements can be highlighted in order to explain the binding of molecules to MT_1/MT_2 receptors.

- Methoxy group

As explained previously (*see Ch1 – 1.2.4*), the methoxy group of the MLT is a critical requirement for creating the hydrogen bonds needed for its binding to both receptors. Moreover, it has been observed that the removal of the methoxy group of MLT leading to *N*-acetyltryptamine results in a decrease of affinity for melatoninergic receptors of around 1000-fold (figure 28). In the same way, the replacement of the methoxy group of MLT by hydroxy, halogen or other alkoxy groups results in a decrease in receptor affinity.[73] As can be observed in table 2, any molecule capable of forming a hydrogen bond with the corresponding histidines of the melatoninergic receptors can have affinity for the MT_1 and MT_2 receptors, but only if other additional requirements are also met.

- *N*-alkylamide function attached to the central scaffold by an aliphatic linker
    - *Amide function*

Carbonyl oxygen and the hydrogen of the amide function found in MLT also appear to play a vital role in the formation of hydrogen bonds that allow the binding of this neurohormone to its receptors as previously explained (*see Ch1 – 1.2.4*). As can be observed in figure 28, 5-methoxytryptamine, a molecule resulting from the substitution of the amide function of the MLT by an amine, exhibited no significant affinity for the melatonin receptors.[73] Moreover, as can be observed in table 2, almost every agonist published in DDR over the last few decades contains an amide function or a derivative in their structure.

N-acetyltryptamine
Ki=730 nM

MLT
Ki=0.24 nM

5-methoxytryptamine
Ki=2528 nM
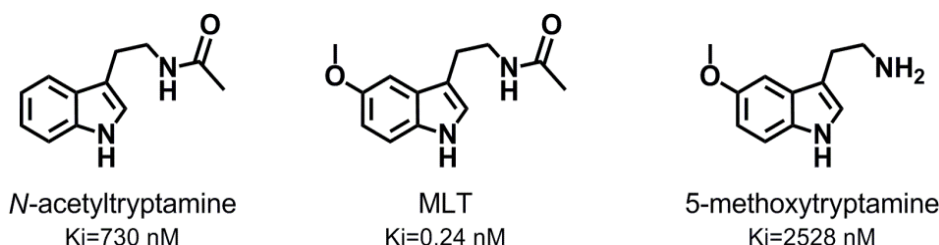
Figure 28. Binding affinities of *N*-acetyltryptamine, MLT and 5-methoxytryptamine for ML_1 (high affinity binding sites for melatonin; MT_1/MT_2) in chicken brain.[73]

- *Alkyl chain attached to the amide function*

The variation in the length of the alkyl substituent attached to the amide carbonyl group of MLT has been thoroughly studied and it was concluded that the enlargement of CH_3 to C_3H_7 enhances the affinity for the receptors but that any larger increase or branching leads to decreased binding affinity, suggesting a small hydrophobic receptor pocket in this area.[73]

- *Linker between central scaffold and amide function: Aliphatic chain of variable length*

At the time of carrying out this bibliographic revision, it was established that the MT_1/MT_2 receptors agonists have to contain an aliphatic chain of variable length substituted over the central core and linked to an amide function. In MLT, this aliphatic chain forms a two-methylene

linker, and quite recently, it has been considered critical for this chain to have two methylenes in order for MLT to adopt the flexible conformation required to fit in the different cavities of both receptors.[40] In fact, as can be observed in table 2, the majority of the molecules shown contain a linker of at least two methylenes, a requirement that should now be included in every future pharmacophore.
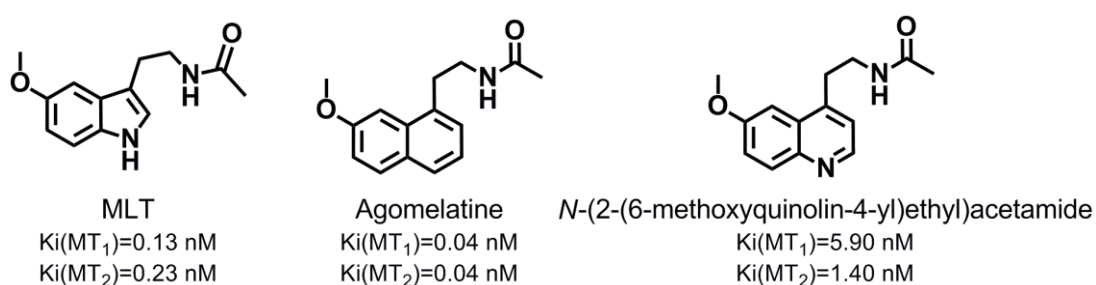
In MLT, the aliphatic chain creates a six-atom distance between the oxygen atom of the alkoxy group and the nitrogen atom of the amide function. This distance appears to be optimum for these critical groups of MLT to reach the required places inside the receptor in order to form the hydrogen bonds needed to bind with them.[73] Most of the published melatoninergic agonists, minus a few exceptions, also present this six-atom distance. In conclusion, it can be said that the vital structural requirements of a melatoninergic agonist include the presence of a methoxy group and an amide function that have to bind in specific positions inside the receptor, and in most cases, these positions are reached due to the existence of this six-atom distance.

- Central nucleus

The indole ring found in MLT as central nucleus is not essential for binding to the receptor. Substitution of the indole ring with a naphthalene ring leads to agomelatine, which presents higher affinity for MT_1/MT_2 than MLT, as can be observed in figure 29. In addition, the substitution of the naphthalene ring with a quinoline ring also leads to compounds of similar affinity.[73, 91]

MLT
Ki(MT_1)=0.13 nM
Ki(MT_2)=0.23 nM

Agomelatine
Ki(MT_1)=0.04 nM
Ki(MT_2)=0.04 nM

N-(2-(6-methoxyquinolin-4-yl)ethyl)acetamide
Ki(MT_1)=5.90 nM
Ki(MT_2)=1.40 nM

Figure 29. Binding affinities of MLT, agomelatine and *N*-(2-(6-methoxyquinolin-4-yl)ethyl)acetamide measures in ^{125}I-MLT displacement assays for binding with human MT_1/MT_2 receptors expressed in CHO cells.[73, 91]

Replacement of the indole scaffold of MLT by tetraline ring afforded ligands with binding affinities generally comparable to that of naphthalenes. Other bioisosteric replacements of the indole ring with a benzofuran and benzothiophene moiety resulted in a slight loss of affinity. However, substitution of the indole ring with benzimidazole dramatically reduced binding affinity. A nucleus of phenylalkyl amide has demonstrated remarkably high affinity in chicken brain despite its simplicity, so it has been set up as the minimal structural requirement for receptor recognition.[73]

III.HYPOTHESIS AND OBJECTIVES

4. HYPOTHESIS

The hypothesis on which this work is based is that the synthesis of new derivatives containing the following structural requirements could result in new potential agonists of the melatonin receptors MT_1 and MT_2 with potential activity for the prevention of insomnia and different alterations of circadian rhythms (figure 30):

- A central core constituted by an aromatic ring.

- A methoxy group linked to the central scaffold.

- A lateral aliphatic chain of variable length substituted on the central ring, joined to different alkylamide functions or derivatives.

Figure 30. Structure requirements of MT_1 and MT_2 receptor agonists.

5. OBJECTIVES AND STRATEGIES

The principal aim of this work is to synthesize new molecules as agonists of the melatonin receptors MT_1 and MT_2 with potential activity for the treatment of insomnia and different sleep disorders.

As can be observed in figure 31, the strategy followed during this project is based on an iterative process in which, after having carried out an extensive bibliographical review, some initial compounds are designed. Later, after obtaining the biological activity values, a SAR study is performed and the starting hypothesis is then redefined, followed by optimization of the initially designed structures. Therefore, a structural refinement is achieved, which allows us to obtain new hits as candidates for future drugs. In addition, when enough data has been collected, alternative molecular modeling techniques can be included in this cycle that allows the prediction of active compounds and consequently, the discrimination of inactive molecules in stages previous to their synthesis and biological evaluation, saving a lot of time and money.
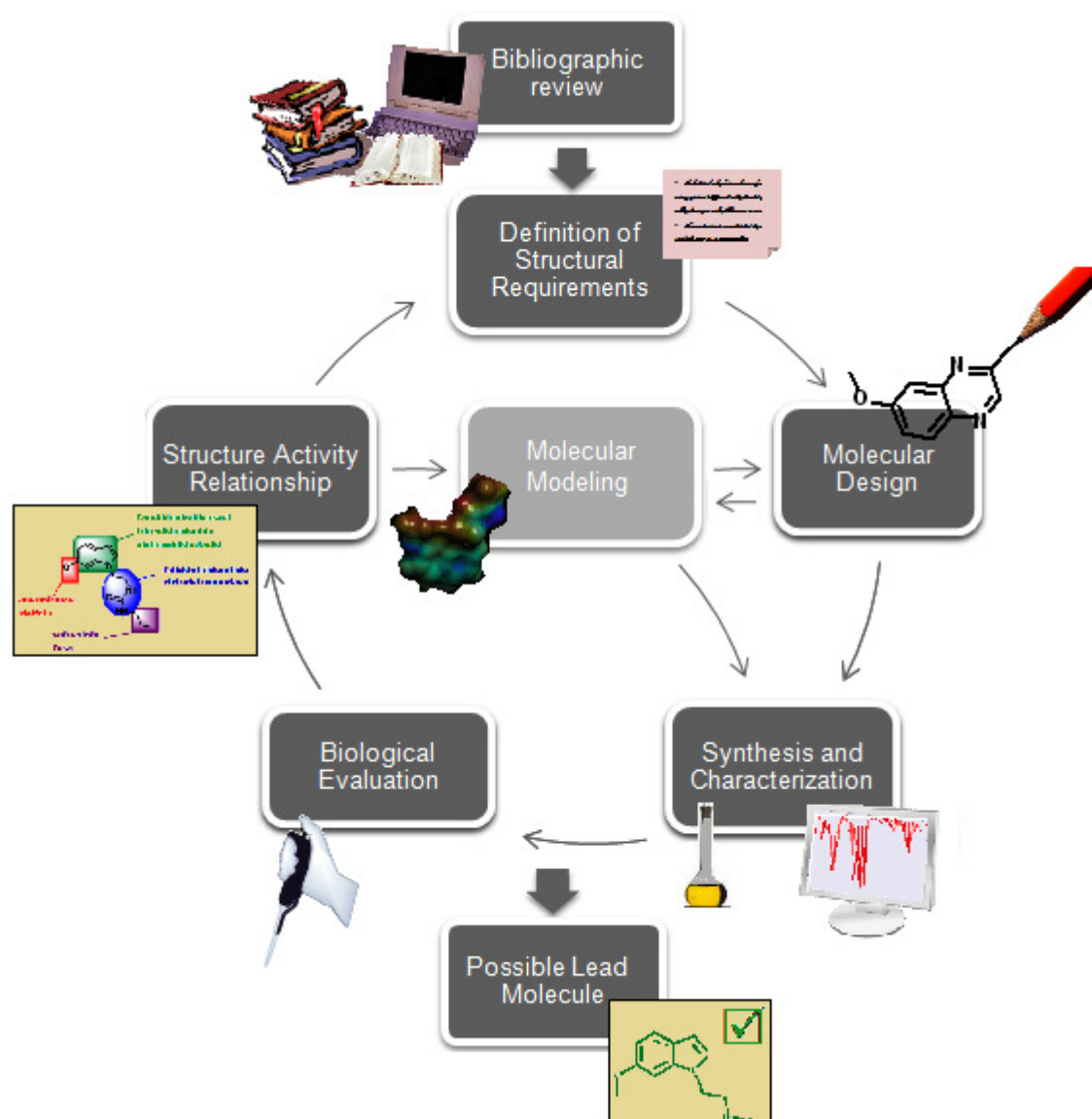
Figure 31. Iterative cycle that represents the strategy followed for drug discovery in this project.

The specific objectives proposed in order to achieve the main aim of this project are as follows:

1. Continuous bibliographic review of the melatonin receptor (MT_1/MT_2) agonists published

2. Definition of the structural requirements of molecules for binding to MT_1/MT_2 receptors

3. Design of new molecules

4. Synthesis and structural characterization of designed molecules

5. Biological evaluation of designed compounds

6. Performance of SAR study and feedback for the definition of more accurate structural requirements and molecular design

7. Use of molecular modeling for creating a model capable of predicting the melatoninergic binding affinity of new designed melatonin receptor agonists before their synthesis

8. Selection of new hit compound (if possible)

IV. MOLECULAR MODELING

6. APPROXIMATION TO MOLECULAR MODELING

6.1. Molecular modeling and drug design

For many years the strategy in drug discovery has consisted in the chemical modulation of a lead structure in order to find analog molecules exhibiting the desired biological properties. This process involves several trial and error cycles where the experience and chemical intuition of the scientists are also necessary in order to ultimately select an analog as a candidate for further development. Therefore, the entire process is laborious and expensive. However, the undeniable fact is that this process has provided most of the existing drugs used today. The traditional methods of drug discovery are now being supplemented by more direct approaches, based on the understanding of the molecular processes involved in each disease. [92]

In the seventies, pure samples of protein targets were isolated, and the X-ray crystallography technique made it possible to learn how three-dimensional (3D) structures control the different chemical processes in the body. Therefore, great interest in this rational approach emerged in the pharmaceutical industry, contributing to the rapid development of molecular modeling as a full discipline.[92]

Molecular modeling can be simply considered a range of computational techniques based on theoretical chemistry methods and experimental data that can be used either to analyze molecules and molecular systems or to predict molecular and biological properties.[92] In this way, computational chemistry is a useful tool for enhancing chemical knowledge, helpful for characterizing and predicting the structure and stability of chemical systems, estimating energy differences between diverse states, and explaining reaction pathways and mechanisms at the atomic level.[93]

Rational drug design is based on the principle that the biological properties of molecules are related to their structural features. Molecular modeling has opened the road to the discovery of lead structures by means of a rational approach. "Direct" and "indirect" designs are the two major modeling strategies currently used. In the first approach, the 3D features of a known receptor site are directly considered, whereas in the second approach, the design is based on the comparative analysis of the structural features of known active and inactive molecules; lead structures are designed on the basis of the pharmacophoric model obtained by said analyses (figure 32).[92]
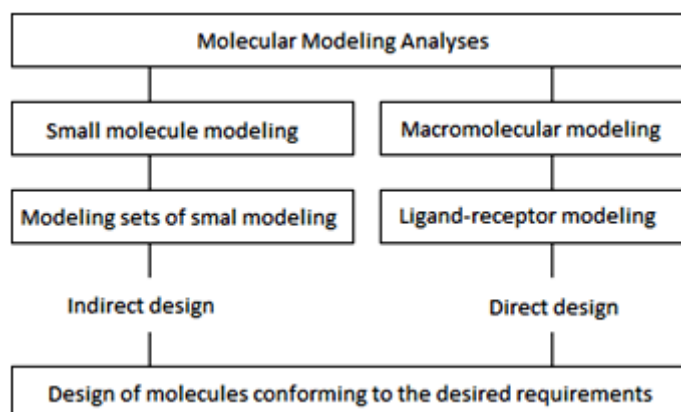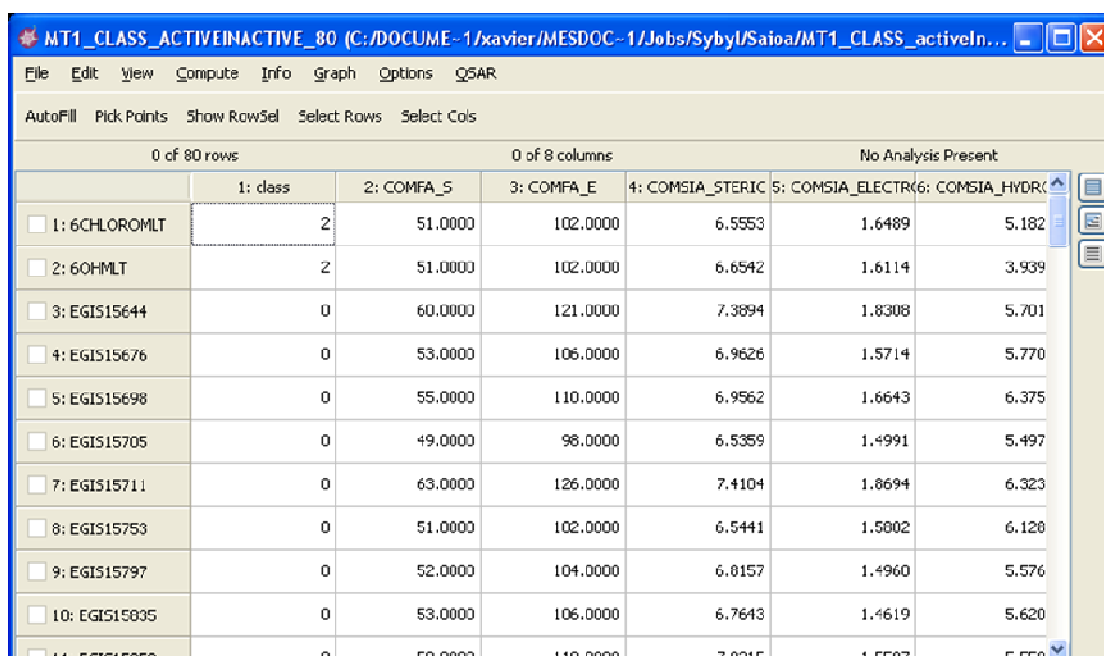

Figure 32. Conceptual frame used in molecular modeling and drug design [Figure adapted from *Cohen, N.C.*].[92]

A pharmacophore may be defined as the essential geometric arrangement of atoms or functional groups necessary for producing a given biological response. Pharmacophores have been described as connectivity-based structural 2D fragments and 3D geometric patterns.[92]

6.2. Quantitative structure activity relationship (QSAR)

QSAR is a technique that quantifies the relationship between structure and biological data and it can be quite useful for chemical optimization of the groups that modulate the potency of a known molecule. This technique has been traditionally used for the rational design of active compounds.[94]

A typical QSAR contains a data table in which each row represents the properties of a single compound and each column shows a different physical or biological molecular property (hydrophobicity, lipophilicity, electronic nature, steric factors,…) of the compounds (figure 33).

Figure 33. An example of a data table or spreadsheet used in Sybyl program.

All these descriptors are then included in a multiple regression analysis in order to obtain the best fitting equation that can provide different information:[94]

• Mechanism of action of the compounds.

• Atomic environment of the biological binding site.

• Indications regarding the interactions between a named substituent in a known position and certain regions of the biological target.

• Estimation of the highest potency expected of a molecule in a series.

• Provide information on whether or not all parts of the molecule are in close contact with the binding site.

Adjusting the descriptors to be used, the mathematical form of the relationship between the potency and the descriptors, and the statistical method used to analyze the relationship are part of the "art" of QSAR practice.[94]

6.2.1. *Modeling and molecule optimization*

The first step of every project always consists in carrying out extensive bibliographic research in order to create the initial database for the study containing a suitable structural variability as well as a broad activity range.

Before creating any model, molecular structures can be generated first, then their molecular geometries can be optimized, and finally, their energies and charges can be calculated. There are three major theoretical computational methods:[94]

• Molecular mechanics (MM) methods

• *Ab initio* methods (pure Quantum mechanics (QM) method)

• Semi-empirical methods (a mixture between QM and experimental data)

During the optimization process, different conformations have to be generated. It should be pointed out that most molecules exist in different conformations. Therefore, it is necessary to deduce which conformation is the one that describes the expected property. In addition, it is important to remember that the bioactive conformation of a ligand bound to its biological target is not necessarily the conformation of minimum energy calculated. What is important is the binding energy of receptor-ligand complex. Indeed, as described in the formula below (equation 1) and according to thermodynamic's principle, the better the receptor-ligand affinity is, the lower the binding free enthalpy is.[94]

$$\Delta G_{binding} = \Delta G_{complex} - \Delta G_{receptor} - \Delta G_{ligand}$$

Equation 1. Formula of the binding energy of receptor-ligand complex.

Finally, it should also be pointed out that the computation of molecular properties requires the 3D superimposition of molecules within a homologous series (figure 34). When molecules possess a rigid common substructure, their alignment is relatively easy but when the molecules are more diverse with a higher flexibility, their superimposition becomes the critical step of the process.[94]
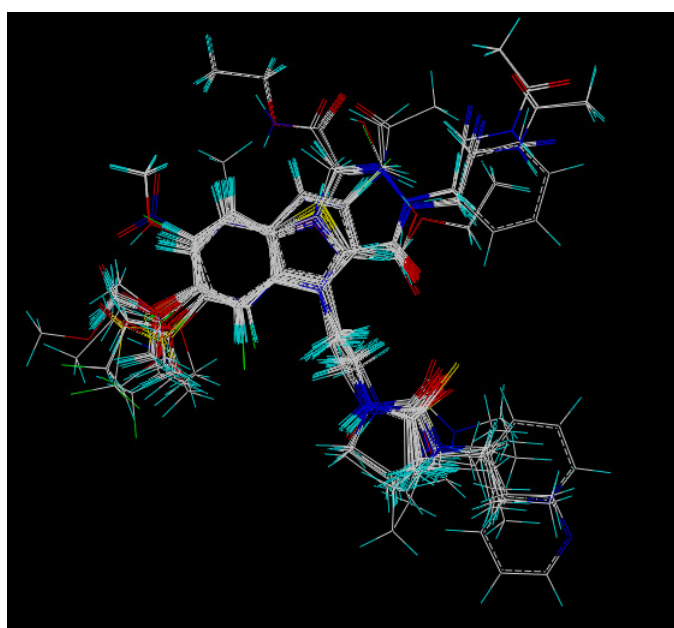
Figure 34. An example of molecular superimposition.

6.2.1.1. MOLECULAR MECHANICS METHODS (MM)

MM methods are the easiest and fastest methods to deal with very large systems such as biological macromolecules. In the MM methods, the molecule is viewed as a collection of points (atoms) connected by springs (bonds) with different elasticities (force constants). This method uses fundamental principles of vibrational spectroscopy that define the energy potential function of a molecule depending on the variation of bonded energy terms and the nonbonded energy terms. The bonded energy terms represent the geometry of the bond distance between two atoms (stretching force), the angle value between three connected atoms (bending force) and the angle value between four connected atoms (torsion force). The nonbonded energy terms render the electrostatic interaction between two nonbonded atoms (Coulomb force) and the contact between two nonbonded atoms (van der Waals force). This potential energy function depicts favorable conformations observed from experimental structural data and permits setting up significant penalties for unobserved conformations. The computations use the Born-Oppenheimer approximation, which describes the energy potential function of the molecule in terms of the nuclear positions. This is the so-called potential energy surface or Born-Oppenheimer surface.[94]

MM calculations then use an empirically derived set of equations for the Born-Oppenheimer surface. This set of potential functions is called the force field and it contains adjustable parameters that are optimized in order to obtain the best fit of calculated and experimental properties of molecules, such as geometries, conformational energies, heats of formation, etc. The assumption made here is that corresponding parameters and force constants are transferable from one molecule to another. The general MM force field is based on the Westheimer method and includes functions for bond stretching, angle bending, torsion, and vdW interactions as follows (equation 2):

$$V_{potential} = V_{stretch} + V_{bend} + V_{torsion} \quad \text{(Bonded energy terms)}$$
$$+ V_{Coulomb} + V_{VdW} \quad \text{(Nonbonded energy terms)}$$

Equation 2. MM force field.

Each component of MM ford field may be represented by its own potential function as shown in equation 3.

$$V_{potential} = \sum k_b (r - r_0)^2 + \sum k_\theta (\theta - \theta_0)^2 + \sum |k_\emptyset| - k_\emptyset \cos(n\emptyset) \quad \text{(Bonded energy terms)}$$

$$+ \sum \frac{q_i q_j}{\varepsilon 4\pi\varepsilon_0 r_{ij}} + \sum \left(\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^6} \right) \quad \text{(Nonbonded energy terms)}$$

Equation 3. Equation Potential function of each component of MM force field.

The sum extends over all bonds, bond angles, torsion angles, and nonbonded interactions between all atoms not bound to each other or to a common atom. More elaborated force fields may include Urey-Bradley terms (1,3-nonbonded interactions), cross interaction terms and electrostatic terms.

MM energy minimization involves successive iterative computations where an initial conformation is submitted to full geometry optimization. All parameters defining the geometry of the system are modified by small increments until the overall structural energy reaches a local minimum. However, the local minimum may not be the global minimum. Searching methods are then used in order to find other conformations. These methods may involve systematic search, which increases all rotatable bonds to explore the complete conformation space of the molecule, distance geometry and other random sampling approaches.[94]

- MM calculations in Sybyl; geometry optimization and energy/charges calculation

Several molecular mechanics force fields can be used within Sybyl. An approximation for determining which force field generally works best for which structures is shown in table 3.[95]

Table 3. Force fields of Sybyl and their application in the study of different structure types

Force field	Small molecule	Large molecule	Protein-ligand complex	Metalloproteins
Tripos	2	2	2	2
AMBER/Kollman		1		1a
MMFF94/MMFF94s	2	2	1	1b

1: The force field was parameterized for this type of molecule, and therefore, it is an excellent first choice for that particular class.

2: The force field is also an acceptable choice and should be considered if the first choice lacks parameters or in order to compare results between the different force fields.

1a/1b: The AMBER force field should be the first choice if the parameters for the metal in the complex are available. Otherwise, use MMFF94 as a first choice.

Every force field contains adjustable parameters that are optimized in order to obtain the best fit between calculated and experimental molecular properties. The nature of a force field depends on the potential functions and parameters used.[95]

Several minimization procedures for finding a local minimum of the energy function are provided. Two types of line searches are available in Sybyl: [95]

- The conventional line search looks for a minimum until the variance between the current minimum and the previous minimum is within a pre-defined value.

- The Wolfe line search seeks a position that satisfies the so-called Wolfe conditions. The Wolfe line search is roughly three times faster than the conventional line search.

There are several methods for locating the minimum of an energy function. The methods can be classified using no derivatives of the function, using first derivatives only or using first and second derivatives. In Sybyl, "MAXIMIN2" is invoked in order to minimize the strain energy of a molecule. MAXIMIN2 uses a combination of the first and non-derivative methods. The procedures available within the MAXIMIN2 command are explained below and in table 4, guidelines for using the different energy minimizing methods are reported.[95]

Table 4. Guidelines for using the different energy minimizing methods.

Method	Molecule size	Relative speed	Notes
Powell	Large or small	Fastest	Most efficient
Conjugate gradient	Small	Slow	Use when Powell fails to converge
Steepest descent	Small with poor geometry	Fast at high gradients	Very large steps
BFGS	Small and with few degrees of freedom	Slowest	Very small steps

• <u>Steepest descent:</u> the line search direction is the derivative of the function at the current position. No information from previous iterations is used. This procedure works well for distorted systems, where the direction of maximum energy change varies greatly from one iteration to the next. The overall convergence properties of this method are quite poor. This method requires storage proportional to the number of variables.

• <u>Conjugate gradient:</u> accumulates information about the energy function, from one iteration to the next. Its convergence properties are superior to steepest descent. Storage requirements are greater than steepest descent but still linear with the number of variables.

• <u>Powell:</u> belongs to the conjugate gradient family of minimization methods, but uses more advanced rules than conjugate gradient for determining the descent direction. It is also more tolerant to inexact line searches. As a result, it is over three times faster than the conjugate gradient method and is well suited for a wide variety of problems.

• <u>BFGS:</u> approximates the inverse of the Hessian matrix (i.e., the second derivative matrix) by accumulating information from first derivatives at each iteration. BFGS has superior convergence properties in comparison with conjugate gradients. This matrix requires the highest storage requirements. In this case, the storage is proportional to the square of the number of variables.

Another important point is the calculation of electrostatic charges on the atoms in a molecule. In each case, Sybyl calculates the formal charge on each atom in the molecule and reports non-zero ones. This calculation is based on the number of free valences of each atom type and the number of bonds currently ascribed to the individual atoms. Charge calculation methods available in Sybyl are:[95]

• <u>Del-Re charges:</u> This is a simple quantum chemical method using the concept of localized bond orbitals. The Del-Re method is intended for molecules containing σ bonds only, but the results for molecules with π bonds are usually satisfactory.

• <u>Gasteiger and Marsili charges:</u> The σ charges are calculated according to the method described by Gasteiger and Marsili and the π charges are calculated according to a method adapted from the former one.

• <u>Hückel charges:</u> The Hückel method is a simple quantum method applicable to π electron systems

• <u>Gasteiger-Hückel charges:</u> This method is a combination of two other charge computation methods; the Gasteiger-Marsili method for calculating the σ component of the atomic charge and the Hückel method for calculating the π component of the atomic charge. The total charge is the sum of the charges calculated by the two methods. Formal charges on atoms included in the

π system are assumed to be delocalized over the entire π system. For this reason, the Hückel portion of the charges is calculated first and these charges are used as the basis for the Gasteiger-Marsili charge calculation.

- Pullman charges: The method of Pullman and Berthod is a combination of two methods: the Del-Re method for calculating the σ component of the atomic charge and the Hückel method for calculating the π component of the atomic charge. The total charge is the sum of the charges calculated by the two methods. As in the Gasteiger-Hückel method, the π charges are calculated first and the σ charges are calculated second. The Pullman method has been very successful in predicting dipole moments and atomic charges for nucleic acids and proteins.

6.2.1.2. QUANTUM MECHANICS (QM) METHODS

- Semi-empirical methods

These types of calculations are basically quantum mechanics but some experimental values are used as well. Below it is explained that the solution of the Schrödinger equation without approximations is the basis of the *ab initio* method, whereas the solution of the Schrödinger equation with approximations is the basis of semi-empirical calculations.[94]

Unlike the molecular mechanics methods, semi-empirical methods use mathematical formulations of the wave functions, which describe hydrogen-like orbitals. The types of wave functions used include the Slater-type orbitals (STO) and Gaussian-type orbitals (GTO). Semi-empirical methods treat the linear combination of orbitals by iterative computations which establish a self-consistent field (SCF) and minimize the energy of the system. Semi-empirical calculations differ in the approximations that are made regarding repulsions between electrons in different orbitals. The approximations are adjusted by setting values, which correspond to either *ab initio* data or available experimental data.[94]

- Ab initio methods

The *ab initio* methods are the highest level of quantum mechanical calculation and they are used for the purpose of finding the solution of the Schrödinger equation. There are two different approaches to *ab initio* methods:[94]

- The calibrated approach used the full exact equations without approximations. The basis set is fixed in a semi-empirical way by calibrating calculations on a broad range of molecules. The error in a new application is estimated based on the average error obtained compared with experimental data for the calibrating molecules.

- The converged approach is used for small molecules and for test cases. In this method, a sequence of calculations improving basis sets is carried out regarding one molecule until convergence is reached. The error is estimated from the sensitivity of the result to further refinements in the basis set.

Ab initio methods are most useful for cases where there is no experimental data to draw from, but suffer from the disadvantage that a great deal of computer power is needed and therefore, the method is not routinely useful for systems with more than 50 heavy atoms (all the atoms except H).[94]

In general, *ab initio* calculations are iterative procedures based on self-consistent field (SCF) methods. Normally, calculations are approached by the Hartree-Fock closed shell

approximation, which treats a single electron at a time interacting with an aggregate of all the other electrons. Self-consistency is achieved by a procedure in which a set of orbitals is assumed and the electron-electron repulsion is calculated. This energy is then used to calculate a new set of orbitals and these, in turn, are used to calculate a new repulsive energy. The process is continued until convergence occurs and self-consistency is achieved. The sophistication of the *ab initio* calculation is dependent on the number of basis set orbitals used. In general, the more orbitals that are used, the more accurate, but also slower, the calculation is. The individual *ab initio* calculations are further identified by abbreviations for the basis set of orbitals that are used, e.g., STO-3G, 4-3 l-G, 6-31G, and so forth. Standard *ab initio* programs provide consistent results for computation of molecular geometry, energy and charge distribution in the ground state for reasonable sized molecules. These computational approaches are also performed in order to highlight reaction processes as well as transition states and intermediates.[94]

- QM calculations in Spartan; energy/charges calculation

Semi-empirical models are the simplest methods, based on quantum mechanics that can be applied to molecules containing 100-200 atoms. The MNDO, AM1, RM1 and PM3 models generally provide geometries coinciding with experimental structures. This makes them suitable for evaluating properties such as polar surface area which depends solely on geometry. However, none of the existing semi-empirical models is suitable, for example, for obtaining the relative energy of the degree of exothermicness/endothermicness of a chemical reaction, or for determining whether or not an isomeric product of a reaction is likely to be more or less stable than another one.

Among the *ab initio* methods, Hartree-Fock models resolve the Schrödinger equation considering the electrons as independent particles. The Hartree-Fock approximation renders the motions of electrons in molecules (molecular orbitals) as the sum of the motions of electrons in atoms (atomic orbitals).[96]

A second approximation, termed the LCAO (Linear Combinations of Atomic Orbital), distinguishes different Hartree-Fock models that use different basis sets; different numbers and types of atomic orbitals. For example, the model with the 3-21G split-valence basis set uses one atomic orbital to describe each non-valence electron and two atomic orbitals to describe each valence (core) electron; it comprises one 1s atomic orbital and two sets of 2s and 2p atomic orbitals.[96]

<div align="center">
1s
2s, 2px, 2py, 2pz
2s´, 2px´, 2py´, 2pz´
</div>
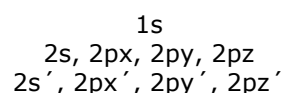
The 6-31G* polarization basis set increases the flexibility by including d type atomic orbitals that, while not occupied in the atom, are used in molecules.

<div align="center">
1s
2s, 2px, 2py, 2pz
2s´, 2px´, 2py´, 2pz´
dxx,dxy, dxz, dyy, dyz, dzz
</div>
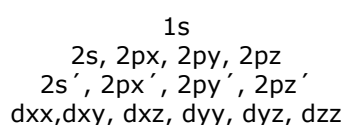
Larger basis sets involving even higher splitting of valence shells, including higher-order (f and g type) functions as well as diffuse functions, are also available. 3-21G and 6-31G* and larger basis set Hartree-Fock models generally provide good accounts of the geometries of organic molecules.[96]

However, the Hartree-Fock models do not allow for properly taking into account the bond-dissociation energy. Indeed, the Hartree-Fock approximation considers interactions between electrons as interactions between one electron and the average field created by all other electrons. As a result, electrons "get in each other's way" to a greater degree than they should. This leads to overestimation of the electron-electron repulsion energy. Termed correlated models allow us to avoid these effects, and these models are divided into two broad categories, density functional models and wave-function based models.[96]

- Density functional models explicitly introduce an approximate correlation term between electrons. In terms of computation time, they are not much longer than the Hartree-Fock models.

- Wave-function based models extend the flexibility of Hartree-Fock models by mixing ground-state and excited-state wave-functions. They take up more computation time than the Hartree-Fock and density functional models.

6.2.2.   3D-QSAR approaches

Among the various 3D-QSAR approaches, CoMFA (comparative molecular field analysis) is the most popular method. CoMFA is based on the assumption that changes in binding affinities of ligands are related to changes in molecular properties, represented by 3D fields. More accurately, the underlying idea in CoMFA is that differences in a target property are often related to differences in the shapes of the non-covalent fields surrounding the tested molecules. When the shape of a molecular field is placed into a QSAR table (molecular spreadsheet), the magnitudes of its steric and electrostatic fields are sampled at regular intervals throughout a defined region. While there are many adjustable parameters in CoMFA, the most important is the relative 3D superimposition of the individual molecules. Properly aligned molecules have a comparable conformation and a similar orientation in cartesian space.[95]

The QSAR is then generated by a PLS (partial least squares) analysis of the data contained in the molecular spreadsheet. A PLS is a general and powerful tool for deriving (multi)linear relationships among columns of data. PLS results are evaluated for their ability to predict new information, using cross-validation or leave-one-out validation. Most often PLS is performed in two stages. The first step is carried out with cross-validation in order to determine the number of components to be used in the model. The consistency of the resulting QSAR can be determined through the value of the cross-validated R^2 reported by the partial least squares regression. If acceptable, a second PLS is run without cross-validation in order to specify the optimum number of components and thus, establish the single model which best represents the data. This final CoMFA-QSAR non-crossvalidated model can most easily be manipulated using various graphic techniques. Once an acceptable QSAR has been derived, edition of the target property value for a new molecule is particularly straightforward.[95]

Another 3D-QSAR approach has currently become very popular; CoMSIA (comparative molecular similarity indices analysis), an extension of the CoMFA methodology, is a method for comparing molecular structures among a group of structures brought into a common alignment. This technique is most commonly used in drug discovery in order to find the common features that are important in binding to the biologically relevant receptor. This technique differs from CoMFA only in the implementation of the fields.[95]

If these quantitative methods do not provide some suitable models, qualitative methods could be used to build model predicting activity classes. SIMCA (soft independent modeling of class analogy) is a technique used for this purpose, associating column data to known categories in a table and computing a separated principal component for each category. This technique produces a mathematical description of the differences between rows of different categories, based on columns of explanatory properties. In terms of concepts, SIMCA is the tool to use in place of PLS when the target property is categorical (nominal) rather than a continuous variable (pKi). This situation arises for both the scales that are arbitrary (such as one of sweet, sour or bitter) and the ordered scales (inactive, marginally active, and active). The immediate output is a summary of how well the final model is able to classify the rows in the categories. Next, a report can be generated in which the "distances between categories" are specified by providing a rms distance of each category's rows from each category. This data indicates whether or not a good separation between the categories has been derived. In the same report, the predicted versus actual category prediction are also given, enabling misclassified rows to be located.[95]

6.3. 2D-Bayesian model and fingerprints

6.3.1. Fingerprints

The fingerprints are numeral 2D molecular descriptors that can be interpreted by different methods for the purpose of building models that are able to predict relevant properties.

A fingerprint is a descriptor used to characterize 2D molecular structures in a binary format. It is based on a molecular abstraction that represents the existing connectivity between atom types or functional groups present in the molecule. All the heavy atoms of the molecule are included in the fingerprints.

6.3.1.1. FINGERPRINTS IN DISCOVERY STUDIO

A variety of molecular fingerprints can be calculated for the input molecules using one of the following algorithms for calculating fingerprints:
- SciTegic extended-connectivity fingerprints
- Daylight-style path fingerprints
- Atom environment fingerprints
- MDL public key fingerprints

For the fingerprint construction, several decisions may be made. First, the maximum path distance (such as the number of bonds) used for indexing an individual fragment has to be

chosen. Next, the best parameters for the molecular fingerprints component should be selected from a large range of options.

All these parameters are used to encode the different fingerprint type. The different fingerprints are encoded by 4 letters and a number as follows:

- The 1st letter is determined by the *atom abstraction* parameter (table 5).

Table 5. Code for atom abstraction parameter.

1st Letter	atom abstraction
F	Functional class
E	Atom type
L	A LogP code
S	Sybyl

- Functional class: uses the rapid functional role codes which are a combination of a hydrogen-bond acceptor, hydrogen-bond donor, positively ionized or positively ionizable, negatively ionized or negatively ionizable, aromatic and halogen parameters.

- Atom type: uses a code derived from the number of connections to an atom, the element type, the charge, and the atomic mass.

- ALogP code: uses a code from the 120 atom types used in the calculation of ALogP.

- Sybyl: uses the Sybyl atom types used in the Tripos™ Mol2 file format.

- The 2nd letter is determined by the *type* parameter, and the type of fingerprint to be calculated is encoded (table 6).

Table 6. Code for type parameter.

2nd Letter	Type
C	Extended connectivity
P	Path
E	Atom environment

- Extended connectivity: generates extended connectivity fingerprints in which each feature represents the presence of a structural (not substructural) unit.

- Path: generates Daylight-style path-based fingerprints. These fingerprints are generated by detecting all paths up to a given length, and then, generating a feature that represents those paths. The union of all the different features present in a molecule is the path-based fingerprint for that molecule.

- Atom environment: generates higher order features from atom types. This creates a string fingerprint.

- The 3rd letter is always F

- The 4rd letter is determined by the *output type* parameter which controls the way the fingerprint is presented (table 7).

Table 7. Code for output type parameter.

4th Letter	Output type
C	Counts
P	Fingerprints

- Counts: a list of the features present in the molecule, with duplicates retained. If a feature occurs more than once in a molecule, that bit value is included more than once in the output list.

- Fingerprints: a list of the features present in the molecule, with duplicates removed.

• The number is determined by the *max distance* parameter. For extended connectivity fingerprints, this is a maximum diameter (in bond lengths) of the largest structure represented by the fingerprint. For path fingerprints, this is the maximum length of the path.

Example 1

For example, a functional class extended connectivity fingerprint of maximum diameter 6 generates a property named "FCFP_6": 1st letter: F (functional class); 2nd letter: C (extended connectivity); 3rd letter: F (always); 4th letter: P (fingerprints); Number: 6 (maximum diameter 6).

Thus, a small molecule will generate different codes depending on the property used (fingerprint type, atom abstraction, maximum diameter…).

Example 2

For example, the initial FCFP atom code for benzamide is shown below. It should be pointed out that the codes are additives.

At iteration 0 (before iterating), it only has information about the atom itself, encoded into its initial atom code (figure 35, table 8).

Table 8. Code for functional class atom abstraction

Functional class	Code
Has lone pairs	1
Is H-bond donor	2
Is negative ionizable	4
Is positive ionizable	8
Is aromatic	16
Is halogen	32
Another	0
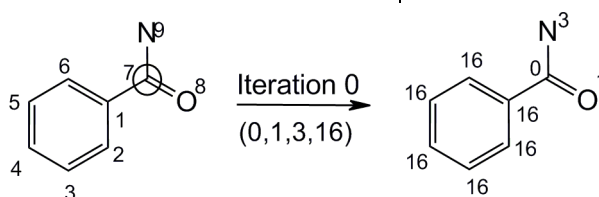
Figure 35. Benzamide and its FCFP_0.

FCFP_0: N (lone pair + H-bond donor)=1+2=3
C_7 (another)=0
O (lone pair)=0
$C_1=C_2=C_3=C_4=C_5=C_6$ (aromatic)=16 (This number will appear just once due to the fact that this fingerprint belongs to a "P" output type. Thus, the duplicate features are removed)
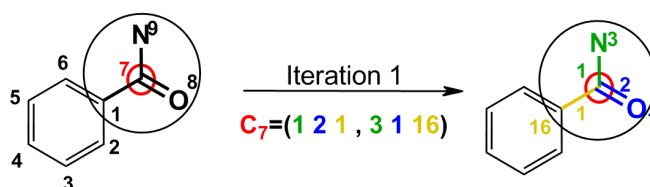
Example 3

Continuing with the same example, upon doing the first iteration, information is collected from all immediate neighbor atoms and a new code is generated. This new code represents the presence of a molecular structure incorporating four atoms: the core atom and its immediate neighbors. This process is performed for each atom in the molecule (figure 36, table 9).
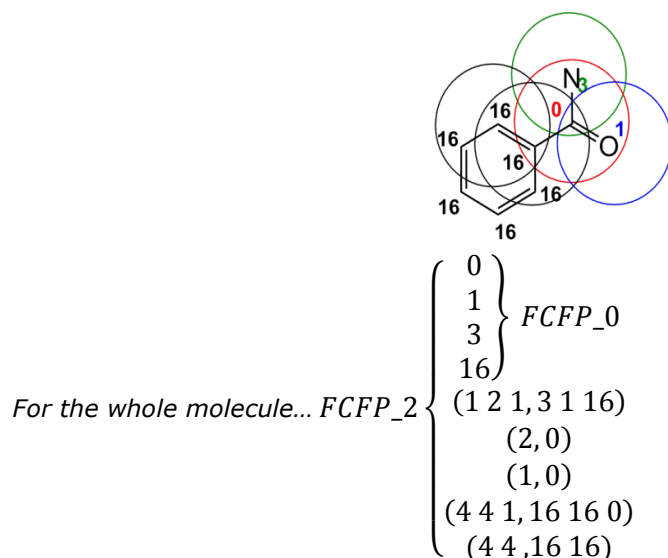
Table 9. Code for bond type.

Bond	Code
Simple	1
Double	2
Triple	3
Aromatic	4

For the atom C_7... FCFP_2: C_7 Bond with simple bond to N = 3
Bond with double bond to O = 1
ond with simple bond to C_6 = 16

Figure 36. Benzamide and its FCFP_2 for C_7 atom.

So for FCFP_2 benzamide will be coded as follows (figure 37);

$$\text{For the whole molecule... } FCFP\_2 \begin{cases} \left.\begin{array}{l} 0 \\ 1 \\ 3 \\ 16 \end{array}\right\} FCFP\_0 \\ (1\ 2\ 1, 3\ 1\ 16) \\ (2, 0) \\ (1, 0) \\ (4\ 4\ 1, 16\ 16\ 0) \\ (4\ 4\ ,16\ 16) \end{cases}$$

Figure 37. Benzamide and its FCFP_2

Example 4

For the second iteration (FCFP_4), the process of collecting information from the neighbors is repeated and a new code is generated. But this time, instead of using the initial atom codes for the atom and its neighbors, it uses the updated codes from iteration 1. The code generated from this step represents an even larger structure around the core atom, in this case, all atoms within two bonds of the core atom.

Finally, a hashing function is applied to all these numbers, and a single number is generated in order to transform the data into bits, allowing it to be indexed and interpreted.

6.3.2.   Bayesian models

Bayesian categorization model is simply a probabilistic classification model which uses probabilities to classify objects into one of various sets of categories. In chemistry, one common categorization is whether a molecule is active or inactive. These models are based on Bayes theorem shown in equation 4.[97]

$$p(h|d) = \frac{p(d|h)p(h)}{p(d)}$$

Equation 4. Bayes theorem

-*h* is the hypothesis or model

-*d* is the observed data

-*p(h)* is the prior belief (probability of hypothesis h before observing any data)

-*p(d)* is the data evidence (marginal probability of the data)

-*p(d|h)* is the likelihood (probability of data d if hypothesis h is true)

-*p(h|d)* is the posterior probability (probability of hypothesis h being true given the observed data d)

This theorem permits inverting the probabilities; if the consequences of a cause are known, the observation of the effects permits returning to the initial cause. This fact can be named as "bottom-up" induction effect. Moreover, the "top-down" induction effect aids in deducing the consequences of a known cause through the interpretation of the theorem.[97]

6.3.2.1.   BAYESIAN MODELS IN DISCOVERY STUDIO

In the implementation of the Bayesian modeling in the Accelrys® applications, the learned models are created with a "learn by example" paradigm: the user marks the sample data that is of interest (good), and then, the system learns to distinguish them from background data. No tuning parameters are required beyond the selection of the input descriptors from which to learn.[97]

The learning process generates a large set of Boolean features from the input descriptors. Then, the frequency of occurrence of each feature is collected in the good subset and in all data samples. To apply the model to a particular sample, the features of the sample are generated, and a weight is calculated for each feature. The weights are added up together in order to provide an estimated probability, which is a relative predictor of the likelihood of that sample being from the good subset. The Bayesian statistics not only consider the likelihood of a model, they also take the complexity of the model into consideration. As a result, it automatically picks the simplest model that can explain the observed data.[97]

Finally, the results of the Bayesian model are represented by the ROC (receiver operating characteristic) curves, which are an objective, quantitative measure of whether or not the model discriminates between two populations (active/inactive compounds).[97]

The curve is created plotting specificity on the x axis against the sensitivity on y axis (figure 38). A random test that cannot distinguish between the groups will provide a straight line from (0,0) to (1,1). As the accuracy of the test improves, the curve moves further towards the ideal situation, where both sensitivity and specificity is 1 (0,1).[97]
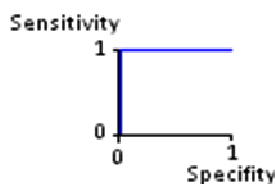
Figure 38. ROC curve representation.

ROC curves consider the four possible outcomes of making a prediction using a model with two populations:
- TP = true positives, actives that are correctly predicted active
- TN = true negatives, inactives that are correctly predicted inactive
- FP = false positives, inactives that are incorrectly predicted active
- FN = false negatives, actives that are incorrectly predicted inactive

• Sensitivity is defined as the power of a model to identify positives, for example, how well the model can pick out active molecules from inactive, or the fraction of your predicted actives that are actually active:

$$Sensitivity = \%TP = TP / (TP + FN)$$

• Specificity is conversely the ability of a model to determine negatives, or the fraction of predicted inactives that are correctly classified:

$$Specificity = \%FP = FP / (FP + TN)$$

The molecules predicted as active will have fewer inactives among them. However, the specificity tends to go down as the molecules predicted as being inactive contain more active molecules. ROC plots are a measure of both sensitivity and specificity across all possible thresholds in your data, giving an overall score of a predictive model.[97]

The accuracy of the test is assessed by measuring the AUC (area under the curve) (figure 39).
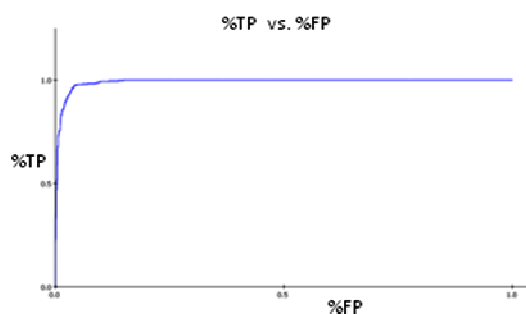
Figure 39. Accuracy of the model or AUC.

Typically, the results are presented with the AUC ranging from 0.5 (a random model) to 1.0 (a perfect model) (table 10).

Table 10. Typical ranging of AUC.

| AUC	Quality
0.9 - 1.0	Excellent
0.8 - 0.9	Good
0.7 - 0.8	Fair
0.6 - 0.7	Poor

6.4. Kohonen Neural Network (KNN)

KNN is a technique used to perform an unsupervised learning in order to generate a projection of objects from a high dimensional space onto a 2D space. Thus, this method offers the possibility of decreasing the dimensionality while conserving the topology of information as much as possible. When the learning process is finished, the objects that were close to each other in multidimensional descriptors remain neighbors in KNN maps. Thus, the KNN can be used for multivariate data classification and visualization.[98]

KNN consists in an input layer, which contains m neurons (input vectors) corresponding to *"m"* molecular descriptors and an output layer that is a 2D geometrical arrangement of "n" neurons. Every "m" neuron in the input layer is connected to each "n" neuron in the output layer by weight vectors, as shown in the figure 40.[98]

The network is trained by the modification of weight vectors in two phases; the competitive learning and self organization phases. The value of weight vector is initialized by a random number. Euclidean distance is then calculated between the input vector and the weight vector of output layer, and the node, having the shortest distance to the input vector, is referred to as winner. In the following step, the winner node and the neighboring nodes are updated in order to make these nodes even closer to the sample. Finally, the same procedure is repeated for all samples.[98]
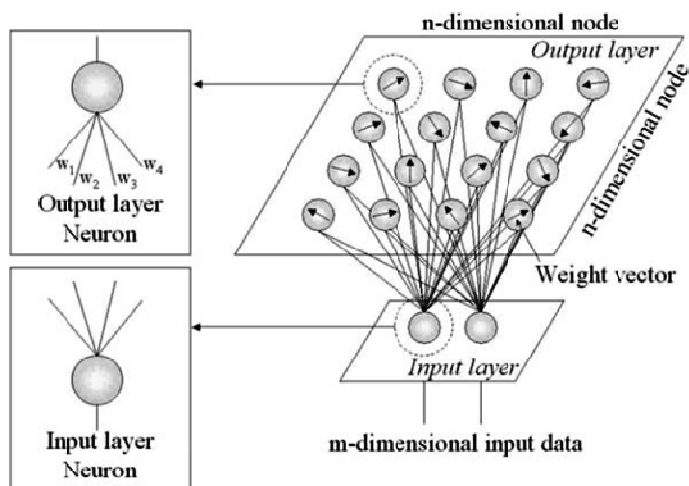
Figure 40. Architecture of KNN [Figure adapted from *Hasegawa K. et. al.*].[98]

More specifically, this KNN algorithm has been used to transform 3D vdW surfaces of the molecules into 2D maps without changing the topological relations and maintaining the possibility to reconstruct original 3D distribution from 2D Kohonen maps. Since molecular interactions between the ligand and receptor mainly occur near the vdW surface of ligand, the parameterization of the molecular electrostatic potential (MEP) gives the option of performing more accurate and realistic analysis. For MEP parameterization, KNN was used through SOM-PAK software. These new approaches are expected to create a 2D grid mapping chemical compounds on the basis of 3D vdW surface.[98]

6.4.1. *Self-organizing map (SOM)*

The SOM algorithm is a "non-linear projection" of the probability density function of the high-dimensional input data onto the 2D display. The program represents the input data space *"R"* onto a regular 2D matrix of nodes. To every node *"i"*, a vector *"m_i"* ($m_i \in R^n$) is associated.

An input vector $x \in R^n$ is compared with $m_i \in R^n$ and the best response is placed on its location in the matrix (figure 41). Every $x \in R^n$ input data vector may be compared with all $m_i \in R^n$ in any metric.[99]
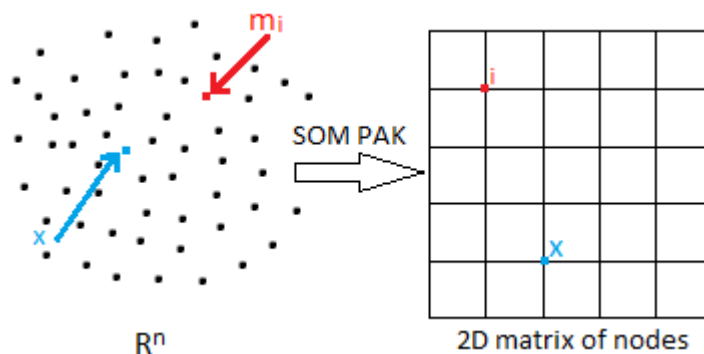
Figure 41. Simplified scheme of 2D mapping of high-dimensional input data.

The array and the location of the response on it are obtained as numerical matrix output data that can be presented as a graphic display creating an image of input data. In practical applications, when x is compared with m_i, the smallest of the Euclidean distances (equation 5) is usually made to define the best matching node, expressed by the subscript c:[99]

$$\|x - m_c\| = \min_i\{\|x - m_i\|\} \quad or \quad c = arg \min_i\{\|x - m_i\|\}$$

Equation 5. Euclidean distances and the best matching node.

Therefore, x is mapped onto the node c relative to the parameter values m_i. An "optimal" mapping would be one that maps the probability density function $p(x)$ in the most "faithful" way, trying to preserve at least the local structures of $p(x)$. In SOM package, the m_i values are defined by the "stochastic approximation type" derivation that defines the original form of the SOM learning procedure. During learning, these nodes that are topographically close together in the matrix (up to a certain distance that can be defined in each study) will activate each other to learn from the same input. [99]

The program states that useful values of the m_i can be found as convergence limits of the following learning process (equation 6).

$$m_i(t + 1) = m_i(t) + h_{ci}(t)[x(t) - m_i(t)]$$

Equation 6.

… where the initial values of the $m_i(0)$ are randomly defined and where t is an integer number; the discrete time coordinate and $h_{ci}(t)$ the neighborhood kernel; a function defined over the lattice points (equation 7).

$$h_{ci}(t) = h(\|r_c - r_i\|, t)$$

Equation 7.

…where $r_c \in R^2$ and $r_i \in R^2$ are the radius vectors in the array of the nodes c and i respectively. Increasing $\|r_c - r_i\|$, $h_{ci} \to 0$. The average width and form of h_{ci} defines the stiffness or rigidity of the "elastic surface" to be fitted to the data points. Notice that it is not usually interesting to describe the exact form of $p(x)$, (especially if x is very high dimensional). It is more important to be able to automatically find those dimensions and domains in the signal space, where x has significant amounts of sample values.

The SOM package contains two options for the definition of $h_{ci}(t)$. The simplest of them refers to a neighborhood set of array points around node c and it is denoted as N_c. Notice that N_C is a function time $N_{C=} N_C(t)$, where if $i \in N_C$, $h_{Ci}(t)=\alpha(t)$ and if $i \notin N_C$, $h_{Ci}=0$, knowing that $\alpha(t)$ is some monotonically decreasing function of time ($0< \alpha(t) <1$). This kind of kernel is nicknamed "bubble".[99]

Another widely applied neighborhood kernel can be written in terms of the Gaussian function (equation 8).

$$h_{ci} = \alpha(t) \cdot e^{\left(-\frac{\|r_c - r_i\|^2}{2\sigma^2(t)}\right)}$$

Equation 8.

…where $\alpha(t)$ is another scalar valued "learning rate", and the parameter $\sigma(t)$ defines the width of the kernel. Both, $\alpha(t)$ and $\sigma(t)$ are some monotonically decreasing functions of time, and their exact forms are not critical. The latter corresponds to the radius of N_C above.

The next step is the calibration of the map in order to be able to locate images of different input data items on it.[99]

V. WORK PLAN

7. WORK PLAN

7.1. Bibliographic review and definition of structural requirements

As a first step in this work, a thorough review has been carried out regarding the MT_1/MT_2 receptor agonists published in the years leading up to the beginning of this project. The next step was to define the main structural requirements that these molecules should contain for binding to the receptors. A summary has been previously outlined (s*ee Ch1 – 3*).

7.2. Design of new MT₁/MT₂ receptor agonists

All the compounds have been designed according to the structural requirements described in the initial hypothesis:

- A central core constituted by a heteroaromatic ring. In this work the series designed can be divided into two main groups according to this heteroaromatic nucleus; series of quinoxalines and series of indoles.

- A methoxy group placed on the central scaffold.

- A lateral aliphatic chain of variable length substituted over the central ring, joined to different alkylamide functions or derivatives such as an alkylurea, an alkylthiourea or a methylsulfonamide.

All the structures synthesized in this project can be summarized in the following general structure (figure 42).
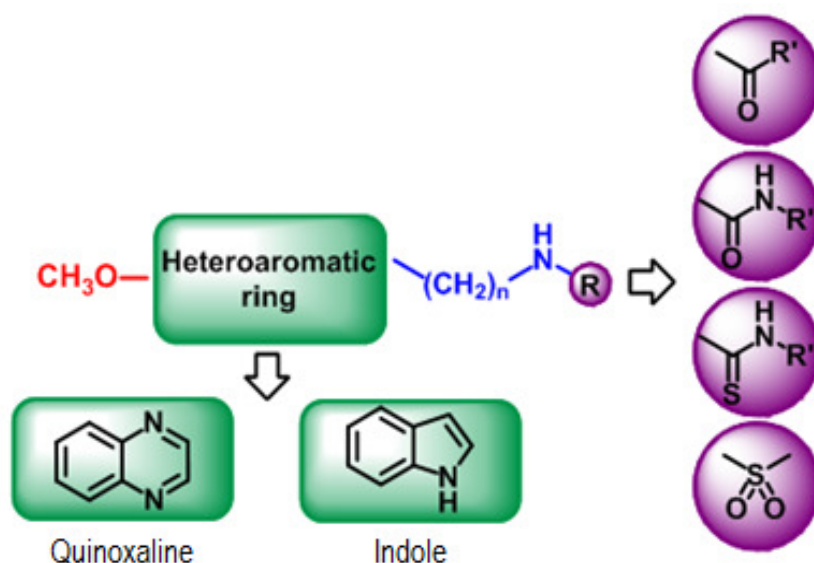
Figure 42. General structure of the compounds synthesized in this project.

As explained previously, the design of these compounds has suffered different changes due to the repetition of the iterative cyclical strategy which is the basis of this work (*see Ch1 – 5*). Therefore, the substituents linked to the amide function vary from the series of quinoxalines to the series of indoles.

7.2.1. *Quinoxaline ring as central core*

The initial series contain a quinoxaline ring as the central nucleus of the structure. As previously explained, the replacement of the indole ring found in MLT with a naphthalene ring or a quinoline ring leads to compounds of MLT-like affinity for MT_1 and MT_2 receptors. Since, chemically speaking, the substitution of the indole ring or the naphthalene ring with a quinoline ring is a clear example of structural approach, the next step in this project is to introduce the quinoxaline ring, a bioisostere of quinoline, and naphthalene rings as the central nucleus of future melatoninergic agonists (figure 43). Therefore, in this work, it is assumed that the quinoxaline ring can be considered the central nucleus of novel MT_1/MT_2 receptors agonists.
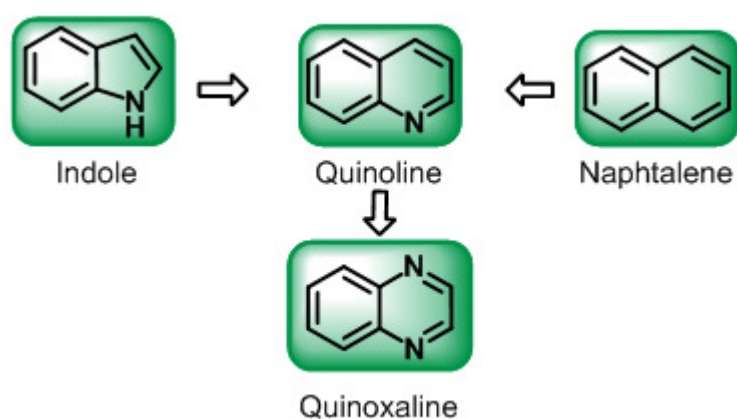
Figure 43. Structural approach of the indole ring and the naphthalene ring to quinoline ring and to quinoxaline ring.

The quinoxaline ring, which is described as a bioisoster of the rings of quinoline, naphthalene, benzothiophene, pyridine and pyrazine, represents a privileged chemical structure due to the diverse biological properties that different quinoxaline derivatives possess with applications in diverse therapeutic areas. These structures have been studied in several fields, such as anticancerous,[100] antiinflammatories/antioxidants,[101] antibacterials,[102, 103] antiprotozoals,[104] sedatives, anticonvulsants, for anxiety treatments or for psychosis or memory loss[105]… etc, and they constitute a very relevant line of research in the "Drug R&D Unit" of the University of Navarra. For all the aforementioned reasons, and based on our expertise regarding quinoxaline chemistry, the possibility of introducing the quinoxaline ring as the central core of some future melatonin receptor agonists arises.

Due to the lack of previous quinoxaline derivatives such as MT_1/MT_2 receptor agonists, it has not been possible to define important facts such as on which side of the quinoxaline the methoxy group might be introduced or whether or not the existing distance between the methoxy group and the nitrogen atom of the aliphatic chain should have a six-atom distance such as in MLT. Thus, a great structural variety has been included in the initial quinoxaline series, not only in order to procure high quality data for SAR study, which will allow the development of a new and more detailed pharmacophore, but also in order to obtain an active compound as a starting point of new quinoxaline as MT_1/MT_2 receptor agonists. Thus, the series

shown in figure 44 have been designed containing all the structural requirements previously defined:

- A central core constituted by a heteroaromatic ring; a quinoxaline ring.

- A methoxy group placed on the central scaffold. This group has been substituted on different carbons of the quinoxaline ring in order to study its influence.

- A lateral aliphatic chain of variable length substituted on the central ring, joined to different alkylamides or alkylurea functions. The length of the chain attached to the amide carbonyl group has been branched and changed in length in order to study the effect that it causes on the binding affinities of the compounds, and to attempt to reach some SAR conclusions regarding the small hydrophobic pocket that is thought to be in this area of the receptor (*see Ch1 - 3*). The existing distance between the aforementioned methoxy and NH groups varies among the series so that it can be determined whether or not there should be a six-atom distance, as in MLT, when the central scaffold is formed by the quinoxaline ring.

Series Q1   Series Q2   Series Q3

$$R \Rightarrow \begin{array}{l} CH_3 \\ CH_2CH_3 \\ CH_2CH_2CH_3 \\ CH(CH_3)_2 \\ Ph \\ NHCH_2CH_3 \\ NHCH_2CH_2CH_3 \\ NHCH(CH_3)_2 \\ NHPh \end{array}$$
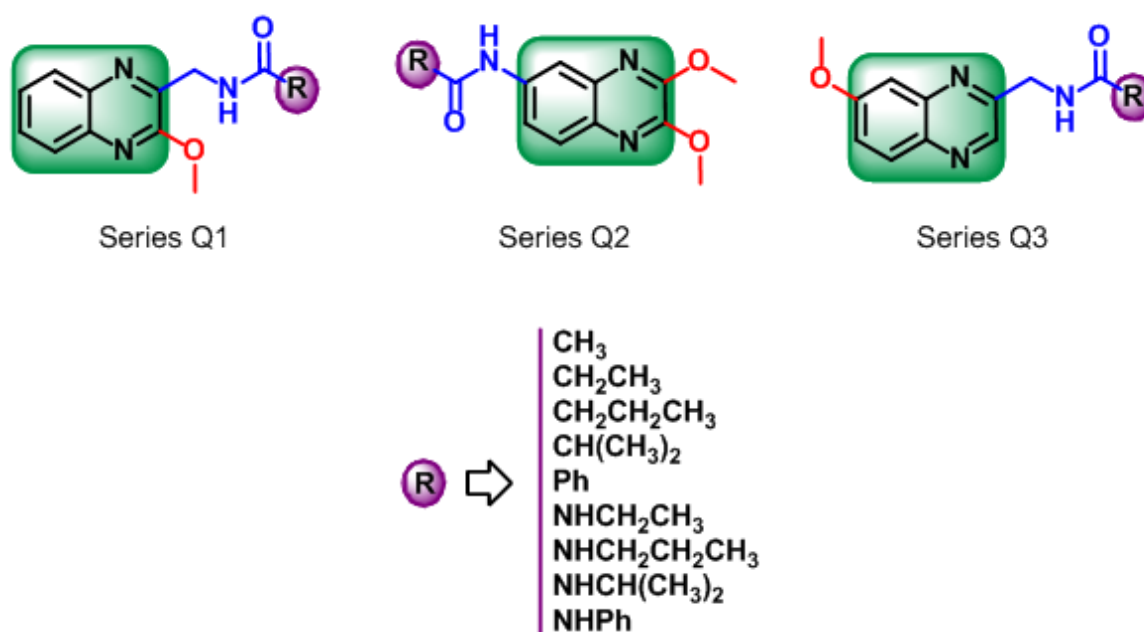
Figure 44. Series designed containing a quinoxaline ring as central nucleus of the structure.

7.2.2. Indole ring as central core

In contrast to the innovative series of quinoxalines, a more classical approach has been designed in this case, where the central nucleus of the structures is formed by an indole ring. These series contain all the structural requirements previously described, including a heteroaromatic ring, an indole ring, as central scaffold, a methoxy group substituted over the indole ring and an aliphatic chain introduced in the central nucleus and linked to an amide function or derivative. The main changes in this series, shown in figure 45, correspond to the aliphatic chain that is replaced over the *N*-atom of the indole ring and the variation in the functions linked to this chain that vary, now including alkylamides, alkylureas, alkylthioureas and sulfonamides.

In this case, the existing distance between the aforementioned methoxy and NH groups varies among the two series. The first series, In1, presents a MLT-like structure because it contains a six-atom distance and a two-methylene linker. In the other series, In2, some variations have been made in order to study the importance of this newly defined structural requirement. Thus, series In2 contains a two-methylene linker; the existing space between the methoxy and the NH groups is reduced to a five-atom distance.
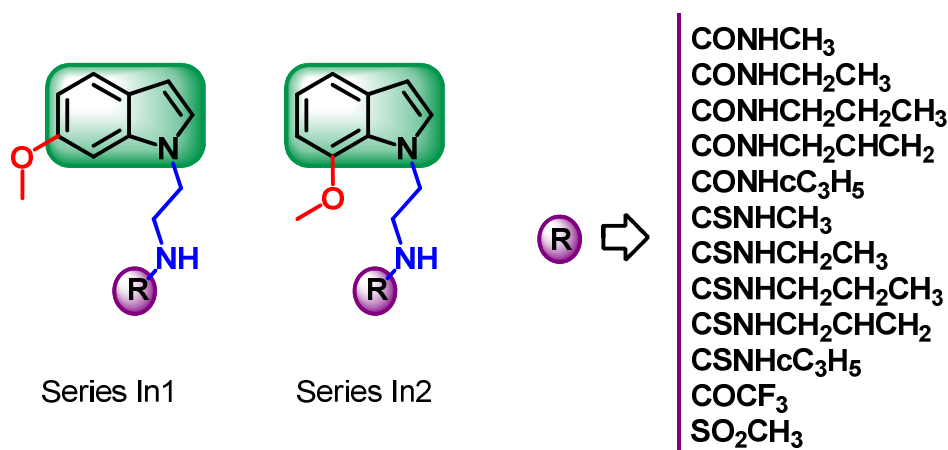
Figure 45. Series designed containing an indole ring as central nucleus of the structure.

7.3. Synthesis and characterization of the designed compounds

Once the molecules have been designed, the synthetic routes that would lead to their obtainment must be defined and optimized. It must be pointed out that although many attempts and changes have been made while attempting to obtain synthetic schemes summarized in this thesis (*see Ch2 – 8 and 9)*, due to the impossibility of agglutinating all the unsuccessful routes and trials, only the final routes have been described.

Next, the synthesized compounds have been chemically characterized by thin layer chromatography (TLC), infrared spectroscopy (IR), proton nuclear magnetic resonance (^{1}H NMR), elemental microanalyses of carbon, hydrogen and nitrogen (CHN) and high performance liquid chromatography (HPLC) (*see Ch3 – 15, 16, 17, 18, 19).*

7.4. Biological evaluation of the synthesized compounds

All the compounds synthesized in this project have been biologically evaluated in the *"Institute de Recherches Servier"* in France. First, ^{125}I-MLT has been used as radioligand in competitive binding assays to measure the affinity of new synthesized drugs for MT_1/MT_2 receptors. Next, the agonism/antagonism profile of the molecules that presented an interesting binding affinity in the previous assay is measured via [^{35}S]GTPγS binding assay.

7.5. Study of structure-activity relationship (SAR)

After obtaining the biological activity values, a SAR study is performed in order to redefine the starting hypothesis and the initially designed structures. Thus, a structural refinement is achieved, which would allow the obtainment of new hits as candidates for future drugs.

7.6. Generation of predicting models with molecular modeling

On this project, several computational chemistry techniques have been used because they provide an unequalable tool for obtaining a suitable model for predicting the binding affinity of new designed MT_1/MT_2 agonists. The following models have been considered:

- 3D-QSAR model
- 2D-Bayesian model
- Kohonen neural network map (KNN map)

7.7. Search for new hits

A joint study of the biological results and the chemical structure of the new synthesized molecules will permit the selection of new hits.

Chapter 2

Material and methods

VI. CHEMICAL SYNTHESIS

8. SYNTHETIC SCHEMES AND REACTIONS – QUINOXALINE DERIVATIVES

8.1. Synthetic scheme and reactions - Series Q1

Q1-3	R
.I	CH₃
.II	CH₂CH₃
.III	CH₂CH₂CH₃
.IV	CH(CH₃)₂
.V	Ph

Q1-3	R
.VI	CH₂CH₃
.VII	CH₂CH₂CH₃
.VIII	CH(CH₃)₂
.IX	Ph
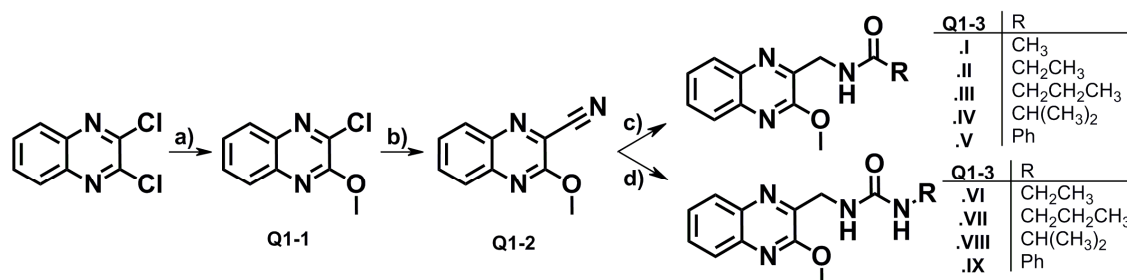
Reagents and conditions:
a)NaOCH₃/THF b)Et₄N⁺CN⁻/AcCN c)H₂/Raney®-Nickel/THF/(RCO)₂O d)H₂/Raney®-Nickel/THF/RNCO

Scheme 1. Designed synthetic route for the synthesis of compounds of series Q1.

8.1.1. Substitution of an aryl chloride by a methoxy group (a)

In order to obtain compound **Q1-1**, an aryl chloride is converted into aryl ether by the substitution of the chlorine atom with a methoxy group. This reaction generally requires activated reagents. Without them, side reactions predominate, though some aryl methyl ethers have been prepared from unactivated chlorides by treatment with methanolate (MeO⁻) in different solvents.[106] Although HMPA (hexamethylphosphoramide) and ammonia have been established as good solvents for this reaction, in this case, THF is used as described by *Beavers M. et al.,*[107] obtaining a very good yield.

This reaction appears to occur by means of a radical nucleophilic aromatic substitution mechanism ($S_{RN}1$).[108] This mechanism involves three types of reactive intermediates; aryl radical [Ar]· and radical anions [ArX]·⁻ and [ArNu]·⁻. The generation of any of them will provide the entry into the propagation cycle. As shown in figure 46, the initial aryl halide accepts an electron from a radical initiator forming the radical anion [ArX]·⁻. In the first step of the propagation cycle, the radical anion [ArX]·⁻ breaks down to form an aryl radical [Ar]· and a halide anion X⁻. Then, in step 2, the aryl radical [Ar]· reacts with the nucleophile anion Nu⁻ to form a radical anion [ArNu]·⁻. Finally, in the third step of the cycle, the radical anion [ArNu]·⁻ transfers its electron to a new aryl halide, forming the final substituted product and restarting the propagation cycle.[109]

$$Initiation: \quad ArX \xrightarrow{-e^-} [ArX]^{\cdot-}$$

$$Propagation: \begin{cases} Step \ 1: [ArX]^{\cdot-} \ \rightarrow \ [Ar]^\cdot + X^- \\ Step \ 2: [Ar]^\cdot + \ Nu^- \ \rightarrow \ [ArNu]^{\cdot-} \\ Step \ 3: [ArNu]^{\cdot-} \ + ArX \ \rightarrow \ ArNu + \ [ArX]^{\cdot-} \end{cases}$$
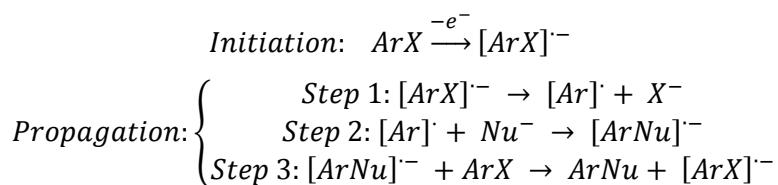
Figure 46. General scheme of $S_{RN}1$ mechanism.

At the time of performing this reaction in the laboratory, and with the aim of substituting one of the two chlorine atoms present in 2,3-dichloroquinoxaline, it has been observed that this reaction is selective through one of the chlorine atoms. If equimolar quantities of initial reagents

are added, just a single atom is substituted, while, if an excess of sodium methoxide is added, the second chlorine atom begins to also be replaced.

8.1.2. *Substitution of an aryl chloride by nitrile (b)*

A cyano-de-halogenation reaction is carried out in order to obtain compound **Q1-2**. Although this is not the case, the typical reaction used for the substitution of the halogen atom of an aryl halide is the Rosenmund von Braun reaction where cuprous cyanide is used.[110]

The substitution of an aryl chloride by a nitrile is usually quite complicated; typical cyanides such as potassium cyanide or sodium cyanide are not able to react with aryl halides even when these are activated. One reason is that these reagents do not mix each other due to the insolubility of the reactive in water and other polar solvents, and the insolubility of nucleophile anion in organic solvents. This problem can be overcome by the use of phase-transfer catalyst. In contrast to most of the inorganic cyanides, quaternary ammonium or phosphonium salts with sufficiently large alkyl chains (R) are poorly dissolved in water and they can be used as phase-transfer catalyst in order to introduce the anion into organic phase and allow it to react with the reagent as shown in figure 47.[111]

Aqueous phase $B^+CN^- + R_4N^+Cl^- \longrightarrow R_4N^+CN^- + B^+Cl^-$

Organic phase $R_4N^+CN^- + \boxed{RCl} \longrightarrow \boxed{RCN} + R_4N^+Cl^-$
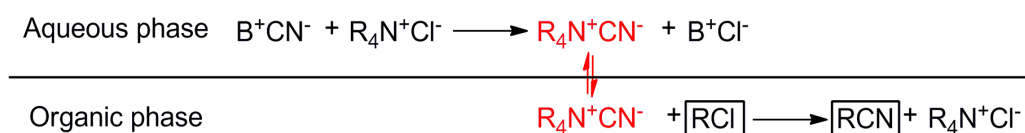
Figure 47. Scheme of the equilibrium formed in reactions catalyzed by ammonium salts where B^+ is an inorganic cation and R is an aliphatic chain [Figure adapted from *Smith M.B. et al.*].[111]

Herrmann K. et al. succeed in removing some activated chlorine atoms of 2-chloroquinoxaline at room temperature (rt) or relatively low temperatures with quite good yield through the use of tetraethylammonium cyanide ($Et_4N^+CN^-$) in DMSO.[112, 113] This reaction appears to be a modification of the nucleophilic substitution reactions catalyzed by a phase-transfer catalyst, explained above. The strategy appears to be the same; ammonium salt is used in order to get the anion to dissolve in DMSO and allow it to collide with the reagent so that the reaction will then take place. At the time of performing this reaction in our laboratory, the DMSO has been replaced by acetonitrile, obtaining similar yields and offering the possibility of removing the solvent easily if needed.

However, no bibliographic data has been found which unequivocally defines this reaction mechanism. The typical aromatic nucleophilic substitution mechanism (S_NAr) has been proposed as the most probable one (figure 48). These reactions are accelerated by the presence of nitrogens as heteroatoms, especially those at α and γ positions to the leaving group.[114] Thus, compound **Q1-1** should be a sufficiently activated reactive in order to suffer a nucleophilic attack. The S_NAr mechanism consists in two steps; first, the nucleophile forms a bond with the reactive, generating an intermediate, and then, the leaving group departs.
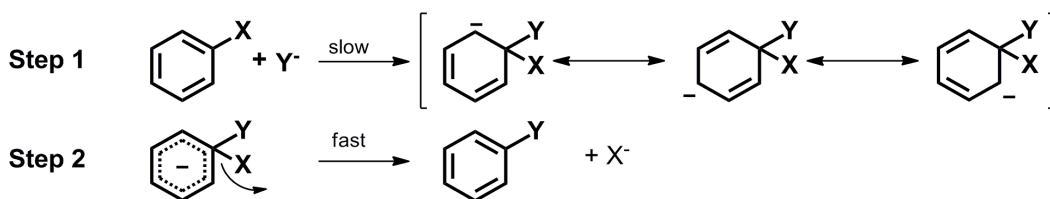
Figure 48. Scheme of S$_N$Ar mechanism [Figure adapted from *Smith M.B. et al.*].[115]

8.1.3. One-pot catalytic nitrile reduction and N-acylation reaction (c)

In order to obtain amide derivatives of series Q1, **[Q1-3.(I-V)]**, one-pot reaction is carried out in which two reactions are achieved: a catalytic hydrogenation and an acylation.

- Nitrile reduction

In the first reaction the nitrile group of **Q1-2** is reduced to a primary amine by using H$_2$ as reducing agent and Raney®-Nickel as catalyst. This reaction occurs through a heterogeneous catalysis in which both the reactive and the molecular hydrogen appear to be adsorbed onto the catalyst surface. In the process of being adsorbed, the molecular hydrogen suffers a dissociation into atomic hydrogens, which are then mobile on the support, allowing approximation to the reactive in order to permit the electrophilic addition reaction.[116, 117]

Typically, nitriles can bind to metal centers through the lone pair of electrons on the nitrogen atom or via the triple bond π-electrons or by way of the lone pair of electrons and the triple bond π-electrons at the same time (figure 49).
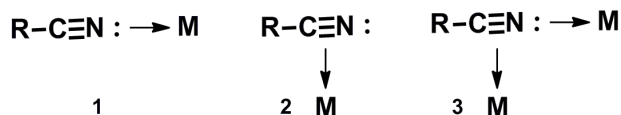
Figure 49. Bonding of nitriles to metal centers [Figure adapted from *Garcia-Alonso, F.J. et al.*].[118]

As shown in figure 50, the hydrogenation reaction often stops at the first step, affording complexes (II) and (III) where a single hydrogen atom has been transferred. Subsequent stepwise additions of hydrogen may lead to a number of intermediates: nitrene (IV), imine (V) or aminoalkylcarbene species (VI). Further hydrogenation yields the surface-adsorbed products (VII) and (VIII), respectively, and finally, after addition of the fourth hydrogen atom, the primary amine (IX) is formed (figure 50).[119]
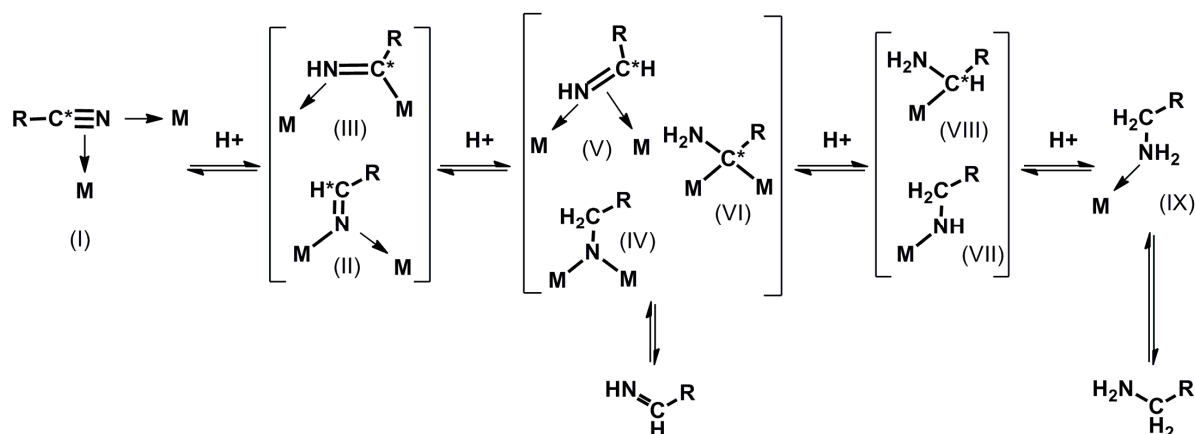
Figure 50. Heterogeneous catalytic hydrogenation of nitriles.

C*: Unsaturated carbon atom

[Figure adapted from *Chojecki A.*].[119]

It is well known that during catalytic hydrogenation of nitriles, undesired secondary and tertiary amines use to be formed mixed with primary amines. It should be pointed out that the majority of surface-adsorbed species are susceptible to a nucleophilic attack because of the presence of an unsaturated carbon atom (marked as C*). The main side reaction in the hydrogenation of nitriles may occur among the partially hydrogenated intermediates. Only nitrene species, (IV) and (VII), are unsusceptible to a nucleophilic attack because the carbon atom is saturated. In figure 51, the proposed mechanism for the formation of aforementioned secondary and tertiary amines is shown.[119]

As can be observed, primary amines are formed in two steps. First, an imine is formed as intermediate and then, the desired primary amine is achieved. Subsequently, secondary amines are obtained; when an imine and a primary amine interact, 1-aminoalkylamine first and then a secondary imine are formed as intermediate. Finally, hydrogenation of the secondary imine affords the formation of secondary amine. For the formation of tertiary amines, it has been proposed that this takes place when a secondary amine is added to the initial imine resulting in gem-diamine and an enamine formation. Finally, the hydrogenation of the enamine leads to tertiary amine.

The formation of these secondary and tertiary amines can be avoided by adding a compound such an anhydride, which reacts with the primary amine as soon as it is formed, or by the use of excess ammonia to displace the equilibrium towards the stabilization of the primary amine (figure 51).[120]
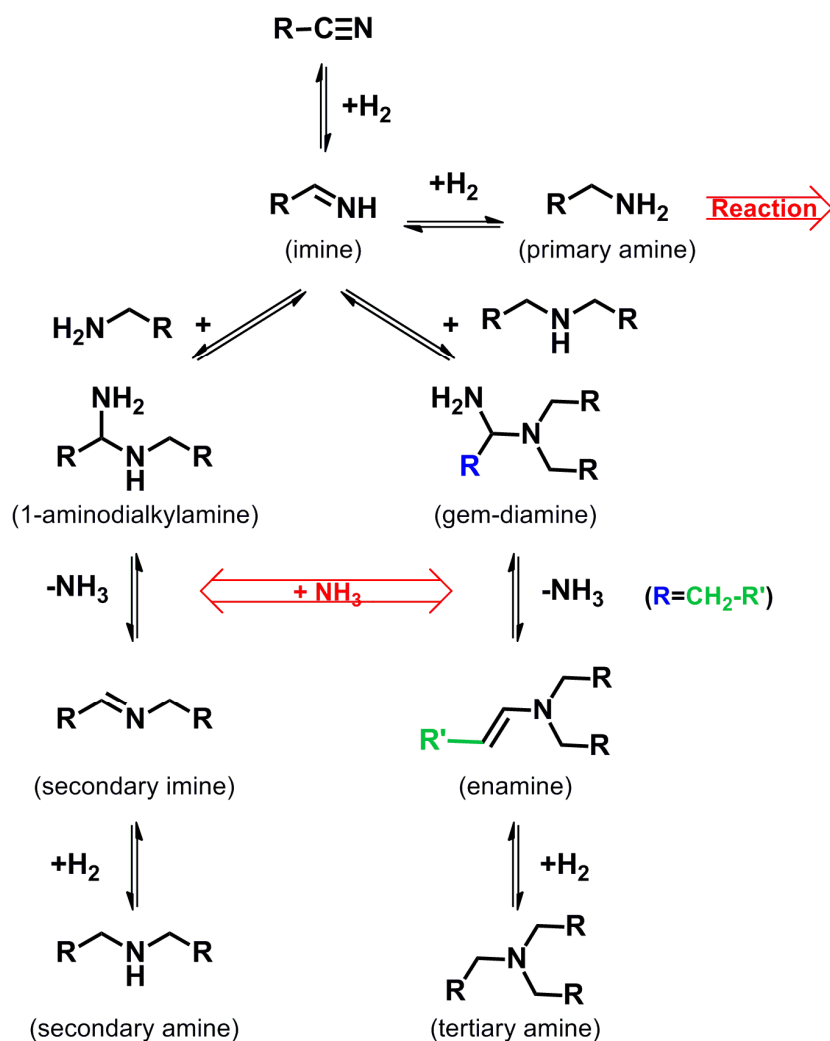
Figure 51. Proposed mechanism for the formation of secondary and tertiary amines in the hydrogenation of nitriles [Figure adapted from *Barrault et al.*].[121]

- *N*-Acylation of amine

In the second phase of this one-pot reaction, the amine is acylated with different anhydrides and the corresponding amide derivatives are achieved.[122] The only inconvenience of this strategy lies in the disgusting smell of some of the acid anhydrides used as well as their high boiling point which hinders the final compound purification.

With acid anhydrides, a rapid acylation of the majority of amines is achieved. It should be taken in account that only one of the two acyl groups is converted into amide due to the fact that the other one leaves in the final step of the nucleophilic acyl substitution reaction mechanism. This reaction mechanism differs from common nucleophilic substitution in the fact that when the nucleophile is added to the electrophilic carbon, the nucleophile itself becomes a tetrahedral intermediate and then, the leaving group departs.

It is important to point out that in the proposed mechanism shown in figure 52, two amine equivalents are needed. The first one acts as nucleophile and the second one as Brønsted base. This is an important point because this second amine is the molecule that gains a proton from the amine involved in the tetrahedral intermediate before the loss of the carboxylate ion.

It is suggested that the reaction starts when the lone pair on the nitrogen atom of the first amine molecule makes a nucleophilic attack to the partially positive carbon of the carbonyl group. After the attack, π electrons of the carbonyl group move onto one of the oxygen atoms of the acid anhydride, resulting in a negative formal charge. Finally, the two electrons of the negatively charged oxygen atom return to the π system of the first carbonyl group. The electronic rearrangement results in the elimination of the ammonium carboxylate salt and formation of the final amide.[123, 124]
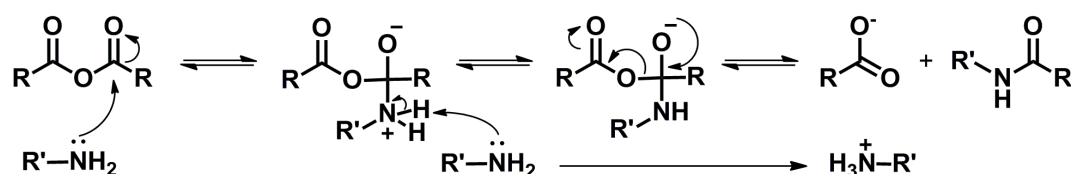
Figure 52. Proposed mechanism for the *N*-acylation reaction between two amine equivalents and acid anhydride.

8.1.4. *One-pot catalytic nitrile reduction and urea formation (d)*

In the case of the synthesis of urea derivatives **[Q1-3.(VI-IX)]**, several trials and strategies have been carried out. Finally, one-pot reaction has been carried out where two different reactions are achieved; a catalytic hydrogenation of nitrile and urea formation.

At the beginning, a classical strategy was designed where two different reactions were drawn up; the first was a catalytic hydrogenation of nitrile, adding an excess of ammonium in the reaction flask in order to avoid extent of the side reactions (*see Ch2 - 8.1.3*) (figure 51), while the second one was the urea formation, through the reaction of primary amine with the corresponding isocyanate.[125]

Due to the unsuccessfulness of this strategy, an unusual reaction has been performed; an isocyanate has been added to the hydrogenation tube, in order to force the amine to react as soon as it is formed (figure 51) (*see Ch2 -8.1.3*). As expected, the result has been that the desired compound was achieved mixed with the hydrolyzed isocyanate.[126]

Considering the electronic structure of the isocyanate group (figure 53) it can be observed that the carbon is the atom with a major positive charge density. Thus, reactions between isocyanates and active hydrogen compounds involve an attack upon the electrophilic carbon of the isocyanate with a nucleophilic center. [127]

Figure 53. Electronic structure of isocyanate.

As mechanism of this reaction, it has been proposed that the lone pair of electrons on nitrogen atom of the amine makes a nucleophilic attack on the C atom of the isocyanate. As a result, π electrons of the carbonyl group move onto the oxygen atom. Then, a reorganization takes place, resulting in the formation of the imidic acid, tautomer of urea, which being less

stable than urea itself, suffers an electronic rearrangement to finally form the desired urea derivative (figure 54).
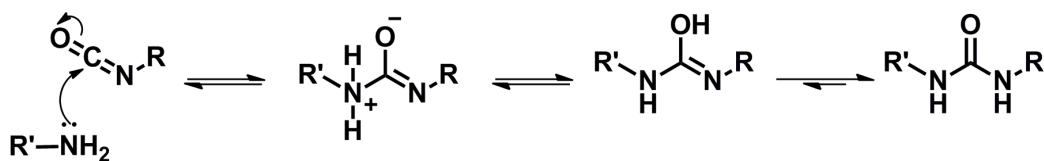
Figure 54. Proposed mechanism for urea formation through the reaction of an amine and an isocyanate.

8.2. Synthetic scheme and reactions - Series Q2

Reagents and conditions: a)NaOCH$_3$/THF b)N$_2$H$_4$·OH/Pd-C/MeOH c)Et$_3$N/THF/(RCO)$_2$O d)N$_2$/dry DCM/RNCO
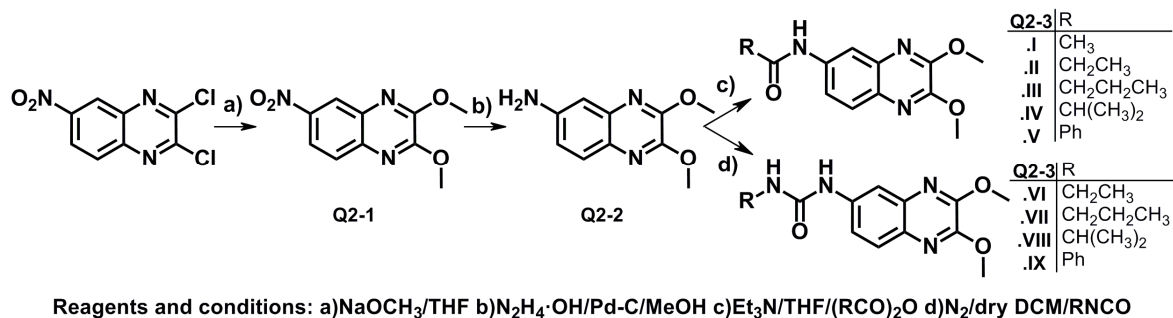
Scheme 2. Designed synthetic route for synthesis of compounds of series Q2.

8.2.1. *Substitution of aryl chloride atoms by methoxy groups (a)*

To obtain compound **Q2-1**, a substitution of the two chlorine atoms of the 2,3-dichloro-6-nitroquinoxaline by two methoxy groups is carried out as explained before (*see Ch2 - 8.1.1*), in this case using two equivalents of sodium methoxide.

8.2.2. *Catalytic reduction of nitro group (b)*

Among all possible reduction conditions, in order to obtain **Q2-2**, hydrazinium hydroxide, as reducing agent, and palladium on carbon (Pd-C), as catalyst, have been selected for the purpose of reducing the nitro group of compound **Q2-1** based on the previous experience of the research group.[128] In the mechanism proposed for this type of reaction, it is assumed that nitroso compounds and hydroxylamines are formed leading to final amines.[129] In the first step of the proposed mechanism (figure 55), the nitrogen atom takes up an electron and the π electrons involved in N=O bond move on, to allow the oxygen atom to accept a proton.

Subsequently, an electrochemical rearrangement takes place, where the nitroso compound is formed as intermediate. Next, the nitrogen atom of nitroso compound gains another electron and the π electrons involved in N=O bond move onto the oxygen atom to permit generating another nucleophilic attack over another proton. Later, a negative charge is formed when a second electron is accepted by the nitrogen atom, which makes a nucleophilic attack, resulting in hydroxylamine formation. This new intermediate takes up another electron, the hydroxyl group departs and finally, the desired amine is formed.
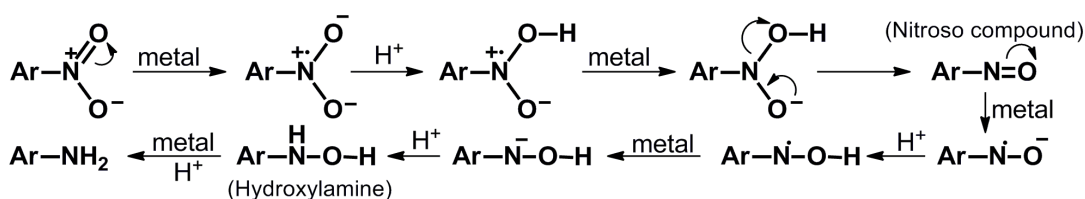
Figure 55. Proposed reaction mechanism for the reduction of nitro compounds.

8.2.3. N-acylation of the primary amine (c)

An acylation is performed in order to obtain the amide derivatives of series Q2 **[Q2-3.(I-V)]** by the treatment of compound **Q2-2** with different acid anhydrides in the presence of triethylamine (Et$_3$N).

This reaction has been explained before *(see Ch2 - 8.1.3)*. The only difference lies in the fact that in the former case, two amine equivalents were needed; the first one acted as a nucleophile and the second one as Brønsted base. However, in this case, just a single amine equivalent is needed due to the fact that Et$_3$N acts as base catalyst (figure 56).
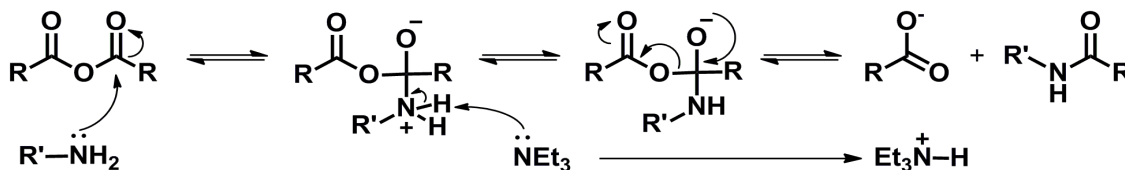
Figure 56. Proposed mechanism for the *N*-acylation of one amine equivalent with different anhydrides and triethylamine.

8.2.4. Urea formation reaction (d)

To synthesize the urea derivatives of series Q2 **[Q2-3.(VI-IX)]** several isocyanates are forced to react with **Q2-2**. This reaction mechanism has been previously explained for the synthesis of urea derivatives of series Q1 (*see Ch2 – 8.1.4*).

8.3. Synthetic scheme and reactions - Series Q3

Reagents and conditions: a)CNCH$_2$CN/*N,N*-DMF/Et$_3$N **b)**(CH$_3$)$_3$CNO$_2$/*N,N*-DMF/N$_2$
c)Na$_2$S$_2$O$_4$/H$_2$O/MeOH/AcOEt **d)**H$_2$/Raney®-Nickel/THF/(RCO)$_2$O

Q3-4	R
.I	CH$_3$
.II	CH$_2$CH$_3$
.III	CH$_2$CH$_2$CH$_3$
.IV	CH(CH$_3$)$_2$
.V	Ph
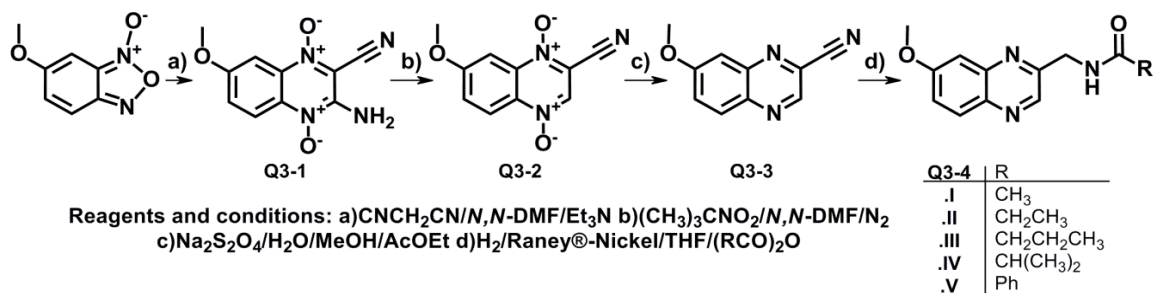
Scheme 3. Designed synthetic route for the synthesis of compounds of series Q3.

8.3.1. *Beirut reaction. Condensation of a benzofuroxan and malononitrile (a)*

Beirut reaction was described at the beginning as one step condensation between a benzofuroxan (BFX) and an enamine to give a quinoxaline-1,4-di-*N*-oxide.[130] Subsequently, it was reported that BFXs can also react with ketones or aldehydes in the presence of ammonia or secondary amines through *in situ* formed enamine intermediates.[131] Currently, the general definition of the Beirut reaction has been extended and it actually involves any type of condensation between a BFX and a ketocarboxylic acid derivative or similar, in a base catalyzed reaction and with or without formation of corresponding enamine intermediates.[102]

Among these variations, the condensation between BFX and the malononitrile has been of vital importance for obtaining compound **Q3-1**, where 6-methoxybenzofuroxan reacts with malononitrile using *N,N*-dimethylformamide (*N,N*-DMF) as solvent and triethylamine as condensing base catalyst.[104]

As explained below, when mono-substituted BFXs are used as starting reagents, Beirut reaction results in an isomeric mixture of 6/7-substituted quinoxaline-1,4-di-*N*-oxides. The fact is that mono-substituted BFXs normally appear as a mixture of 5/6-isomers, presumably interconverting via the *ortho*-nitroso intermediate (figure 57). It has also been observed that if the substituent is a methoxy, a chloro or an acetoxy group, 5-isomer is the most stable. In contrast, if R is a carboxyl or an ethyl carboxyl group, the stability is reversed.

(6-isomer)   (5-isomer)
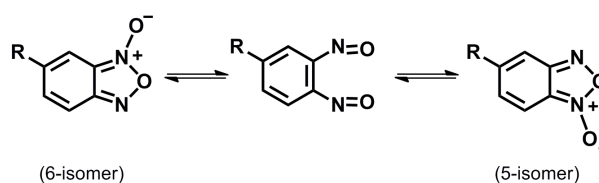
Figure 57. Mixture of isomer present in mono-substituted BFX derivatives.

Although adding the condensation step determines the structure of final compounds and the amount of isomeric mixture formed, in the reference literature, it is not clear which is the nitrogen that suffers the nucleophilic attack. While some studies suggest that the *N*-3 of the most stable isomer is the site where the nucleophilic attack takes place, later works suggest

that the addition might occur on the *o*-dinitroso intermediate.[130] In the reaction mechanism proposed in this work, it is assumed that the nucleophilic attack takes place over the aforementioned *o*-dinitroso intermediate (figure 58).[132]
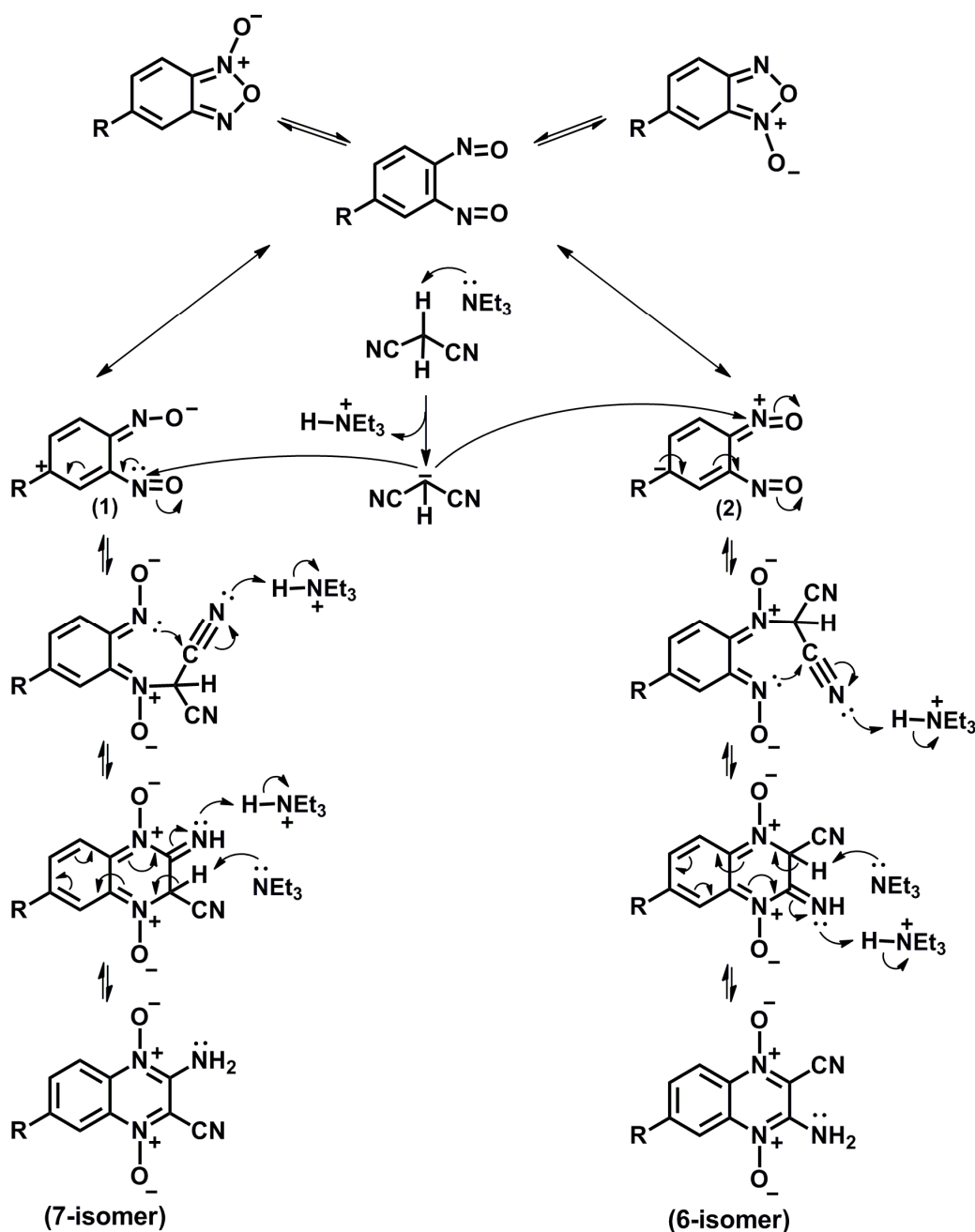
Figure 58. Mechanism proposed for the condensation of mono-substituted BFXs with malononitrile.

The *o*-dinitroso intermediate contains different resonance forms that would be stabilized differently, depending on the substituent present in the BFX ring. An electron-releasing substituent would stabilize the resonance form **(1)** over **(2)**, while an electron-accepting substituent would have the opposite effect. Thus, when BFX is 5/6-substituted with an electron-donating group, theoretically, an isomeric mixture of 7/6-substituted 2-amino-3-cyano-quinoxaline-1,4-di-*N*-oxide should be achieved where 7-isomer prevails over 6-isomer but in

practice, the workup and purification of the isomeric mixture affords only 7-isomer with the 6-isomer being discarded. Moreover, in the case of methoxy substituent, only 7-isomer has been detected.[133]

With regard to the reaction mechanism, and as can be observed in figure 58, it is proposed that triethylamine removes a proton from the activated methylene of malononitrile, resulting in a carbanion which makes a nucleophilic attack over one of the nitrogen atoms of the o-dinitroso intermediate. After an electronic rearrangement, the lone pair of electrons on the neutral nitrogen atom makes a nucleophilic attack over one of the cyano groups, forcing the ring to close and the cyano group to accept the first proton. Next, the catalyst removes the remaining proton on position 2/3 of the quinoxaline ring and after another electronic reorganization, the final compound is achieved.

8.3.2. *Deamination reaction (b)*

The aromatic amine of compound **Q3-1** is removed, leading to compound **Q3-2**. In general, this reaction used to be carried out in two different steps.

First, the amine is usually diazotized by the treatment of the compound with sodium nitrite or nitrous acid, hydrochloric acid and water at low temperature (0-5ºC). As can be observed in figure 59, in the mechanism proposed for this reaction, a nitroso ion (NO$^+$) is formed in the first step which is highly reactive and forces dehydration of the amine and diazonium salt formation. Normally, diazonium ions are not very stable and can break down easily under their formation conditions. However, aryl diazonium ions can be stabilized in aqueous solution at 0-5ºC for a reasonable period of time.[134]
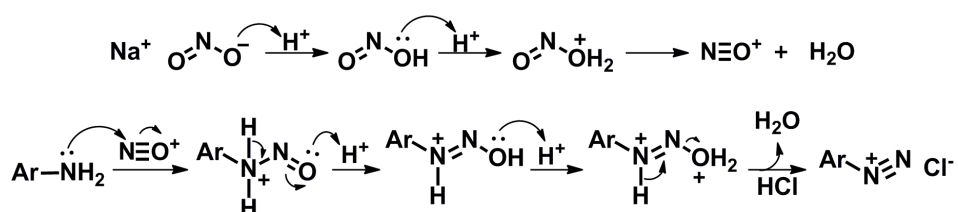
Figure 59. Reaction mechanism proposed for diazonium salt formation.

Continuing with deamination process, in a second reaction, the molecular nitrogen acts as a leaving group and is replaced by a hydrogen atom through the use of some hydrogen donors as the hypophosphorous acid (H$_3$PO$_2$) that react via free-radical reaction mechanism or ethanol, which presents a S$_N$1 mechanism.[135-137]

Aromatic amines can also be deaminated in one step by treatment with an alkyl nitrite in *N,N*-DMF or boiling THF, through the formation of diazonium salt as intermediate.[136] Some studies state that performing these reactions using *N,N*-DMF instead of THF or dioxane, effectively minimizes the formation of secondary products. It is also suggested that these reactions occur through free radical reaction mechanism.[137] Thus, in this case, one-pot reaction is performed using *tert*-butyl nitrite and *N,N*-DMF as reaction conditions.[138]

The mechanism of the diazotization of arylamines in *N,N*-DMF is not fully clear. Before describing the reaction mechanism, two main points have to be taken into account. First, when a diazonium salt is formed in an alkaline solution as in this case, since *N,N*-DMF is a relatively strong basic solvent, the diazonium salt can couple with another aromatic ring.[139] Secondly, *Ek et al.* suggest that the mechanism for *the Gomberg-Bachmann (GB) reaction*, proposed by *Rüchhardt et a.l* is also probably applicable to aprotic diazotization of arylamines.[140] Thus, the reaction mechanism in this case has been suggested as follows (figure 60):
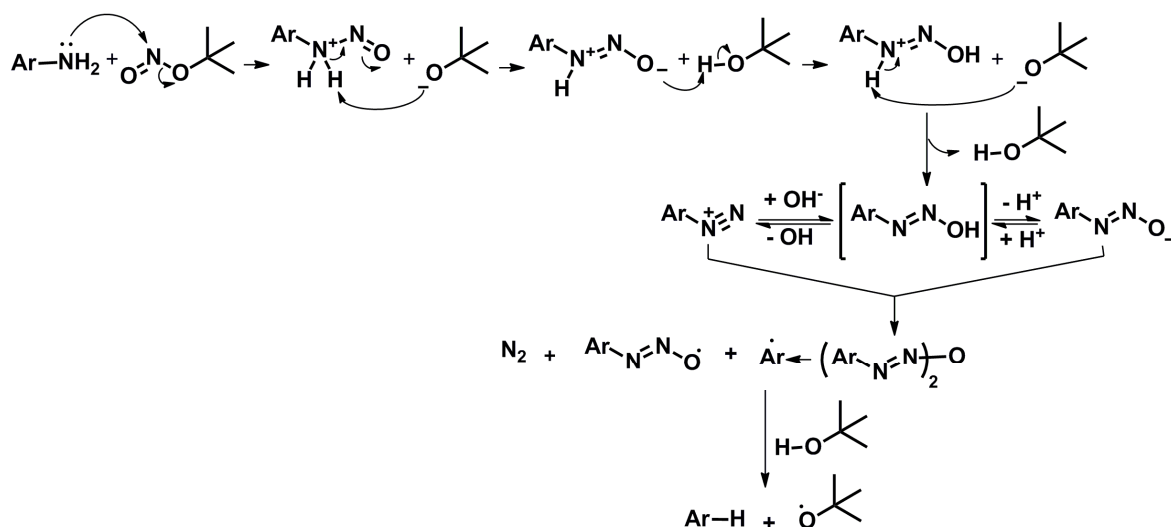
Figure 60. Reaction mechanism suggested for an arylamine deamination reaction.

As can be observed in figure 60, when the aryl amine is treated with *tert*-butyl nitrite in *N,N*-DMF, a hydroxydiazene intermediate is formed first. This intermediate is involved in a pH-dependent equilibrium with the corresponding diazonium salt and diazenolate anion that react then together to form the oxybis(diazene) which, in turn, breaks down and yields nitrogen gas, aryl radical and the long-life diazenoxyl radical. When the aryl radical abstracts a hydrogen atom, the final deamination is achieved.

8.3.3. Reduction of N-oxide groups (c)

Sodium dithionite is a readily available, inexpensive reducing agent which is capable of reducing a wide variety of functional groups. In addition, the use of $Na_2S_2O_4/H_2O$ for the purpose of reducing the *N*-oxide groups of several quinoxaline-1,4-di-*N*-oxide derivatives is a method widely backed up by group's experience.[101, 141] However, its mode of action on the reduction of quinoxaline *N*-oxides is not well established yet.

In 1974, a first mechanism was suggested by *Haddadin et al.* where the first step involved the nucleophilic attack by dithionite anion ($[S_2O_4]^{2-}$) at position 2/3 of the quinoxaline-1,4-di-*N*-oxide. The resulting intermediate was postulated to lose SO_2 to give the sulfinate anion, which breaks down to form the product (figure 61).[138, 142] As both *N*-oxides are reduced, it can be assumed that this reaction would take place twice.
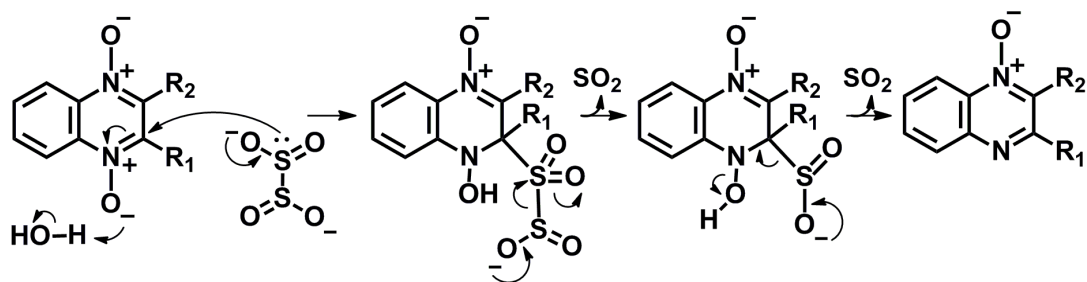
Figure 61. Reaction mechanism proposed by *Haddadin et al.* in 1974 for *N*-oxide reduction.

Currently, it is known that although $[S_2O_4]^{2-}$ itself is occasionally important as a reducing agent, numerous kinetic studies have shown that the sulfur dioxide radical anion $(SO_2)^{\cdot-}$ usually formed from dissociation of $[S_2O_4]^{2-}$ (figure 62) is the active specie in "dithionite reductions" even though it is present in only small amounts.[143, 144]
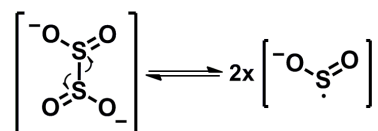
Figure 62. Dissociation of sodium dithionite into sulfur dioxide radical anion.

As the reactivity of sodium dithionite to reduce *N*-oxide groups has been explained in terms of the addition of the $(SO_2)^{\cdot-}$ anion radical to form a sulfinate intermediate,[145] it is possible to think that the reduction of quinoxaline-1,4-di-*N*-oxide would most likely occur through this route. Taking into account the previous information, Haddadin's mechanism can be rewritten as shown below (figure 63). In this case, just as before, it can be assumed that this reaction would take place twice.
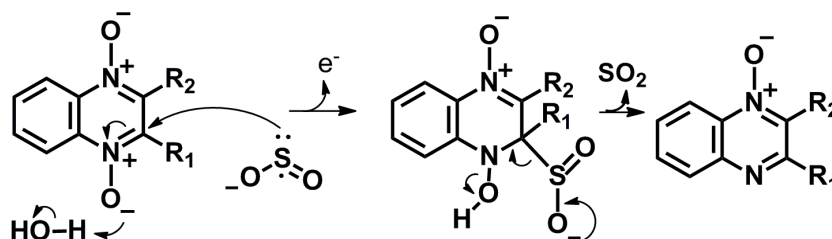
Figure 63. Reaction mechanism proposed for *N*-oxide reduction on quinoxaline-1,4-di-*N*-oxide derivatives.

8.3.4. *One-pot catalytic nitrile reduction and acylation reaction (d)*

In order to obtain amide derivatives of series Q3, **[Q3-4.(I-V)]**, one-pot hydrogenation and acylation reactions are performed as previously described *(see Ch2 – 8.1.3).*

9. SYNTHETIC SCHEMES AND REACTIONS – INDOLE DERIVATIVES

9.1. Synthetic scheme and reactions – Series In1 and In2
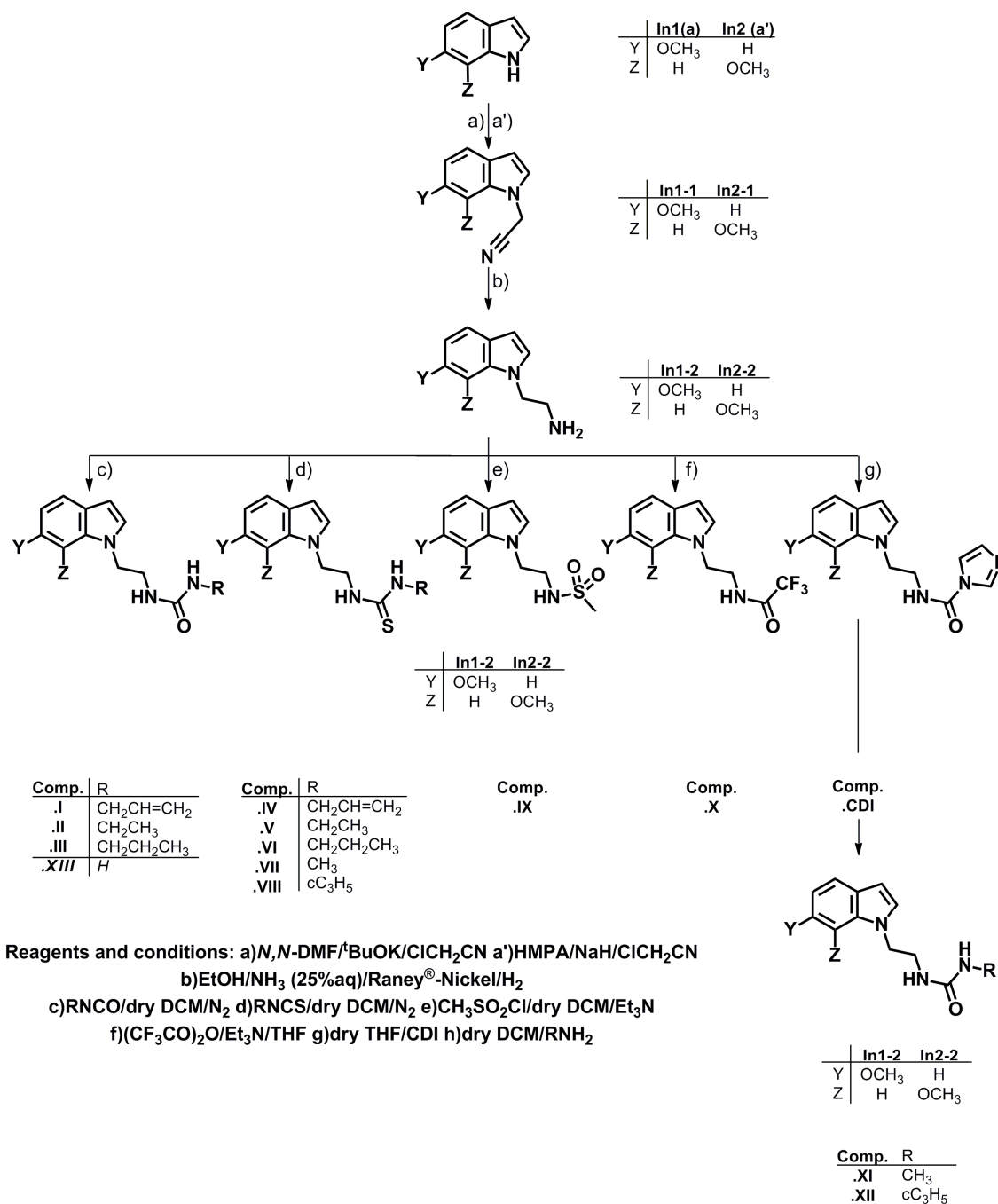
Scheme 4. Designed synthetic route for the synthesis of compounds of series In1 and In2
Compound **In1-3.XIII** was not in the initial work plan; it was designed after having obtained the results in the molecular modeling study (*see Ch3 - 27*).

9.1.1. *Aryl N-alkylation reaction (a/a')*

The first reaction for the synthesis of series In1 and In2 is an *N*-alkylation of the NH group of 6/7-methoxyindole with chloroacetonitrile and a base in order to obtain compounds **In1-1** and **In2-1**. The *N*-alkylation of indoles and related heterocyclic compounds containing an acidic hydrogen atom attached to nitrogen takes place through the treatment of these compounds with an appropriate base that results in an indolide anion which is the intermediate that generates the nucleophilic attack over the alkyl halide in order to afford the *N*-alkylated compound. However, when indoles are alkylated, this reaction can also occur at C-3 atom so substantial quantities of 3-alkylindole may be obtained combined with the *N*-alkylated product. This fact can be explained by the two main isomeric structures that are present in the indolyl anion (figure 64).[146]
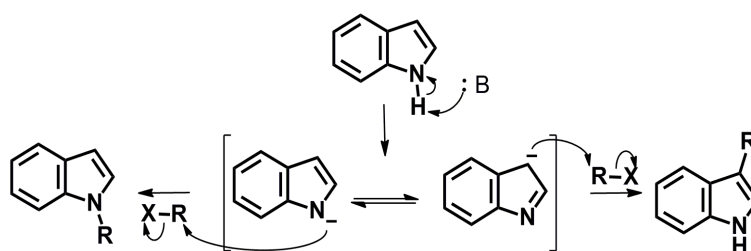
Figure 64. Indolide anion formation and indole alkylation.

The amount of *N*-alkylated and C-alkylated product obtained depends on a number of factors, including the base used for the deprotonation of indole or the solvent. With regard to the bases, when salts are formed by a sodium or potassium cation, *N*-alkylation reactions are usually predominant. With harder cations such as lithium or magnesium, which are tightly bound to nitrogen, the amount of C-alkylation reactions is dramatically increased. The solvent can also influence the ratio of *N*/C-alkylation; if polar aprotic solvents are used, mainly *N*-alkylated products are achieved.[147]

After several trials in which different reaction conditions have been tested, potassium *tert*-butoxide in *N,N*-DMF has been selected as the reaction condition for *N*-alkylation of 6-methoxyindole (series In1) and sodium hydride in HMPA for the *N*-alkylation of 7-methoxyindole (series In2). Although different bases and solvents have been selected, it should be pointed out that both bases contain a cation that favors the formation of the *N*-alkylated compound and that both solvents are polar aprotic solvents for the same reason.

9.1.2. *Catalytic nitrile reduction (b)*

The nitrile groups of compounds **In1-1** and **In2-1** are reduced to a primary amine by the use of H_2 as reducing agent and Raney®-Nickel as catalyst, leading to the obtainment of compounds **In1-2** and **In2-2**. These reactions have been discussed before when describing the synthesis of quinoxaline derivatives *(see Ch2 – 8.1.3)*.

9.1.3.   Urea/thiourea formation (c/d)

The corresponding amines **In1-2** and **In2-2** are treated with iso(thio)cyanates in order to synthesize some urea/thiourea derivatives of series In1 and In2 **[In1-3.(I-VIII)** and **In2-3.(I-VIII)]**. The urea formation reaction has been explained before in the section corresponding to quinoxaline derivatives *(see Ch2 – 8.1.4)*. It is important to point out that the mechanism of the thiourea formation reaction is not explained because it is almost the same as that discussed in the obtainment of urea derivatives.

9.1.4.   Sulfonamide formation (e)

In this reaction, the corresponding amines, **In1-2** and **In2-2**, are treated with methanesulfonyl chloride in dry conditions in order to obtain compounds **In1-3.IX** and **In2-3.IX**.

Nucleophilic substitutions at sulfonyl halides (RSO_2X) are essentially the same as those that take place at carbonyl halides (RCOX). The main difference is that the tetrahedral intermediate formed in traditional S_N2 reactions, in the case of the sulfonyl halides, is proposed as a trigonal bipyramidal intermediate where five groups are distributed around the central nucleus (figure 65).[148]
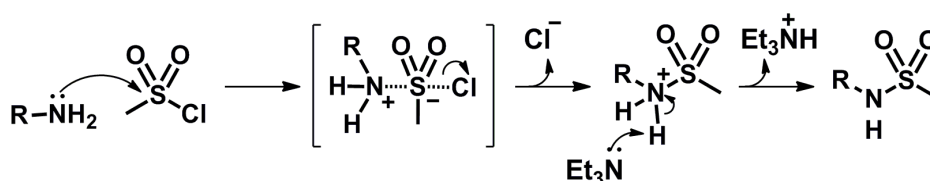
Figure 65. Proposed mechanism for nucleophilic substitutions of methanesulfonyl chloride with an amine.

9.1.5.   N-acylation of a primary amine (f/g/h)

For the synthesis of compounds **In1-3.X** and **In2-3.X**, an acylation reaction has been performed between the corresponding amine derivatives, **In1-2** and **In2-2**, and trifluoroacetic anhydride in the presence of triethylamine. No more reaction details are given because this reaction has already been explained *(see Ch2 -8.2.3)*.

For the synthesis of compounds **In1-3.CDI** and **In2-3.CDI**, another acylation is performed. In this case, the 1,1'-carbonyldiimidazole has been used as both catalyst and reagent, as can be observed in figure 66.[149]
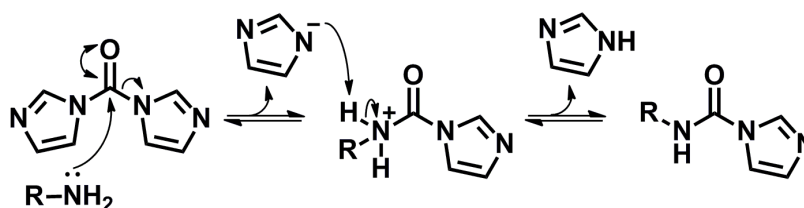
Figure 66. Proposed mechanism for the *N*-acylation of one amine equivalent with 1,1'-carbonyldiimidazole.

For the purpose of synthesizing the urea derivatives that are lacking for the completion of series In1/In2/In3 **[In1-4.(XI-XII)** and **In2-4.(XI-XII)]** the last acylation reactions are performed. In this case, just as before, the imidazole derivative plays a double role; it acts as reagent as well as catalyst, as can be observed in figure 67.[149]
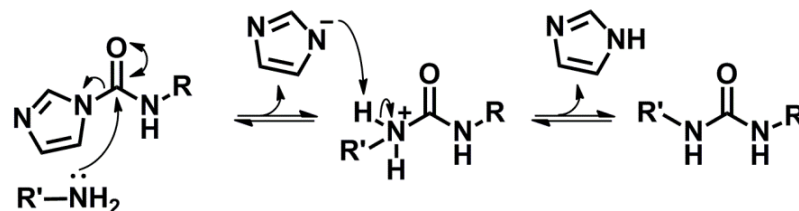
Figure 67. Proposed mechanism for the *N*-acylation of one amine equivalent with carbonylimidazole derivative.

10. CHEMICALS AND INSTRUMENTS - GENERAL REMARKS

10.1. Chemicals

Chemicals have been purchased from Panreac Química S.A. (Montcada i Reixac, Spain), Sigma-Aldrich-Fluka Química S.A. (Alcobendas, Spain), Alfa Aesar Avocado, GmbH & Co. KG (Karlsruhe, Germany) and E. Merck (Darmstadt, Germany).

10.2. Instruments

All of the final compounds have been chemically characterized by thin layer chromatography (TLC), infrared spectroscopy (IR), proton nuclear magnetic resonance (^{1}H NMR), elemental microanalyses of carbon, hydrogen and nitrogen (CHN), melting point (MP) and high performance liquid chromatography (HPLC).

10.2.1. Thin layer chromatography (TLC)

Alugram® SIL G/UV254 (layer 0.2 mm) (Macherey-Nagel GmbH & Co. KG., Düren, Germany) has been used for thin layer chromatography. TLCs have been studied under UV wavelength at 254 nm and 365 nm. As mobile phase, dichloromethane/methanol (DCM/MeOH) and hexane/ethyl acetate (n-Hex/AcOEt) in different gradients have been used and also toluene/dioxane/acetic acid (TDA) in a proportion of 90:25:4.

10.2.2. Column chromatography (CC)

Column Chromatography has been performed using glass columns and silica gel 60 (0.040-0.063 mm) as stationary phase. As mobile phase DCM/MeOH and n-Hex/AcOEt in different gradients have been used.

Flash column chromatography has been developed on an automated flash chromatography system CombiFlash® R$_f$ (Teledyne Isco, Lincoln, USA) instrument. Redisep® R$_f$ 12 g silica columns have been used which contain an average particle size from 35 to 70 microns and an average pore size of 60 Å. As mobile phase DCM/MeOH and n-Hex/AcOEt in different gradients have been used.

10.2.3. Infrared spectroscopy (IR)

The IR spectra have been recorded on a Thermo Nicolet FTIR Nexus Euro (Madhas beenon, USA) with OMNIC 6.0 software, using KBr pellets for solid samples and NaCl pellets for liquid samples. Frequencies are expressed in cm^{-1} and signal intensities are represented as: w (weak), m (medium), s (strong) and vs (very strong).

10.2.4. Melting point (MP)

Melting points have been determined with a Mettler FP82+FP80 apparatus (Greifense, Switzerland). Melting points are expressed in degree centigrades (ºC).

10.2.5. Nuclear magnetic resonance (NMR)

The NMR spectra have been recorded on a Bruker 400 Ultrashield[TM] instrument (Bruker BioSpin GmbH, Rheinstetten, Germany), using tetramethylsilane (TMS) as internal standard and with DMSO-d_6 or $CDCl_3$ as solvents. The chemical shifts (δ) are reported in ppm and coupling constant (J) values are given in Hertz (Hz). Signal multiplicities are represented as: s (singlet), bs (broad singlet), d (doublet), dd (doublet of doublets), ddd (doublet of doublet of doublets), dddd (doublet of doublet of doublet of doublets), t (triplet), dt (doublet of triplets), td (triplet of doublets), tdd (triplet of doublets of doublets), tt (triplet of triplets); ttd (triplet of triplets of doublets); q (quartet); dq (doublet of quartets) and m (multiplet).

10.2.6. Elemental microanalyses of carbon, hydrogen and nitrogen (C.H.N.)

Elemental microanalyses have been obtained on an elemental analyzer LECO CHN-900 (Tres Cantos, Spain) from vacuum-dried samples. The analytical results for carbon, hydrogen and nitrogen, are within ±0.5 of the theoretical values indicating a purity of >95%.

10.2.7. High performance liquid chromatography (HPLC)

HPLC experiments have been developed on an Ultimate 3000 chromatograph (Dionex) with Chromoleon v.6.8 software. The measurements have been performed using an RP 18 column (Lichrospher 100 RP 18 E.C. 5 μm; 10x0.46; Teknokroma) as the stationary phase, at a flow of 1 mL/min and with methanol/water (70:30 and 60:40) as the mobile phase. The retention times (t_R) are expressed in minutes and the reference wavelength is set at 254 nm.

11. EXPERIMENTAL SYNTHETIC METHODS

11.1. Experimental synthetic methods - Series Q1

Synthesis of 2-chloro-3-methoxyquinoxaline [Q1-1].
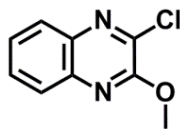
2.30 g (10.04 mmol) of 2,3-dichloroquinoxaline (90%) and 20.00 mL of THF are placed in a flask. The mixture is cooled until 0ºC. A solution of sodium methoxide (NaOMe) is made through a reaction between Na (0.50 g) and MeOH (5.00 mL). 2.34 mL (10.04 mmol) of NaOMe are added dropwise over the first suspension. The color of the mixture changed from purple to yellow. The mixture is stirred during 30' at 0ºC and during 1h at rt. Next, 80 mL of DCM are added and the mixture is quenched with brine. The DCM solution is dried with Na_2SO_4 and filtered. The solvent is removed under reduced pressure in order to obtain an oily residue.

Synthesis of 3-methoxyquinoxaline-2-carbonitrile [Q1-2].
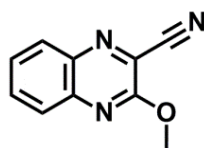
1.59 g (8.17 mmol) of **Q1-1** and 2.55 g (16.32 mmol) of tetraethylammonium cyanide are dissolved in 60.00 mL of AcCN. The mixture is heated between 50-60ºC during 2h 30'. The solution is added onto ice and a black precipitate is formed, which is purified by silica gel column chromatography using DCM as mobile phase. The solvent is removed under reduced pressure.

Synthesis of N-(3-methoxyquinoxalin-2-ylmethyl)-alkylamide [Q1-3.(I-V)].
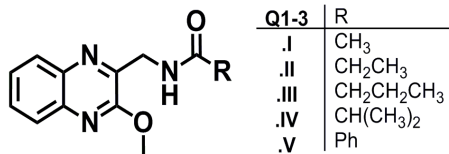
Q1-3	R
.I	CH$_3$
.II	CH$_2$CH$_3$
.III	CH$_2$CH$_2$CH$_3$
.IV	CH(CH$_3$)$_2$
.V	Ph

Q1-2, THF and the corresponding acid anhydride in proportion 1:2 are placed in the hydrogenation flask. A spatula of Raney®-nickel is added and the mixture is hydrogenated during 7h at 50ºC and 50 psi. Next, DCM is added and the mixture is decanted in order to remove the catalyst. Next, the mixture is quenched with water, dried with Na_2SO_4 and filtered. The solvent is removed under reduced pressure to obtain brown oil. Finally, n-hexane is added and a white solid precipitates, which is filtered and washed with diethyl ether and water. If necessary, it is purified by silica gel column chromatography using DCM/MeOH as mobile phase.

Synthesis of 1-alkyl-3-(3-methoxyquinoxalin-2-ylmethyl)-urea [Q1-3.(VI-IX)].

Q1-3	R
.VI	CH_2CH_3
.VII	$CH_2CH_2CH_3$
.VIII	$CH(CH_3)_2$
.IX	Ph
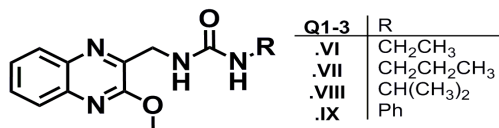
Q1-2, THF and the corresponding isocyanate are placed in the hydrogenation tube. A spatula of Raney®-nickel is added and the reaction is carried out during 7h at rt and 50 psi. Next, DCM is added and the mixture is decanted with the aim of leaving the catalyst off. Next, the mixture is quenched with water, dried with Na_2SO_4 and filtered. The solvent is removed under reduced pressure and the orange oil obtained is purified by flash column chromatography employing DCM/MeOH as mobile phase. The solid is dissolved in water/diethyl ether with vigorous stirring and it is then quenched. The organic layer is dried with Na_2SO_4 and filtered. The solvent is removed under reduced pressure.

11.2. Experimental synthetic methods - Series Q2

Synthesis of 2,3-dimethoxy-6-nitroquinoxaline [Q2-1].
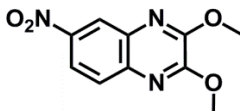
100.00 mL of THF and 5.12 g (20.35 mmol) of 2,3-dichloro-6-nitroquinoxaline are mixed and cooled until 0ºC. 9.60 mL (41.00 mmol) of just made solution of NaOMe [Na (1.20 g) and MeOH (12.00 mL)] are added dropwise. The mixture color changes from yellow to orange. The mixture is stirred during 30' at 0ºC and during 1h at rt. Then, DCM is added and the mixture is quenched with brine. The DCM solution is dried with Na_2SO_4 and filtered. The solvent is removed under reduced pressure and a yellow solid is obtained.

Synthesis of 6-amino-2,3-dimethoxyquinoxaline [Q2-2].
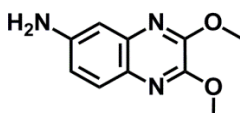
2.50 g (10.64 mmol) of **Q2-1** are dissolved in 150.00 mL of MeOH and the mixture is cooled until 0ºC. Then a tip of spatula of Pd-C and great excess of hydrazinium hydroxide are added. After stirring 30' at 0ºC the mixture is stirred at rt for 15h. The mixture is filtered to eliminate the Pd-C and the MeOH is removed under reduced pressure. Then, DCM is added and the mixture is quenched with brine. The DCM solution is dried with Na_2SO_4 and filtered. The solvent is removed under reduced pressure.

Synthesis of N-(2,3-dimethoxyquinoxalin-6-yl)alkylamide [Q2-3.(I-V)].

Q2-3	R
.I	CH$_3$
.II	CH$_2$CH$_3$
.III	CH$_2$CH$_2$CH$_3$
.IV	CH(CH$_3$)$_2$
.V	Ph
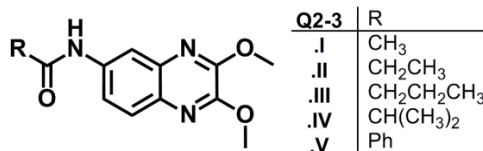
The compound **Q2-2** is placed in suspension in THF. Next, the corresponding acid anhydride in proportion 1:3 and a catalytic amount of Et$_3$N are added dropwise. The reaction is carried out during 1h 30' at rt. Then, THF is removed under reduced pressure. Subsequently, the obtained residue is treated with *n*-hexane and a solid precipitate. The solid is then filtered and washed with water.

Synthesis of 1-(2,3-dimehtoxyquinoxaline-6-yl)-3-alkylurea [Q2-3.(VI-IX)].

Q2-3	R
.VI	CH_2CH_3
.VII	$CH_2CH_2CH_3$
.VIII	$CH(CH_3)_2$
.IX	Ph
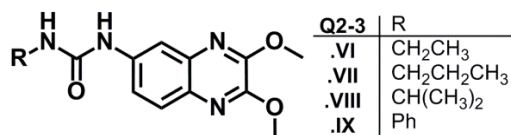
The compound **Q2-2** is dissolved in dry DCM and placed under N_2 atmosphere. The mixture is cooled until 0ºC and the corresponding isocyanate is added in proportion 1:2. The reaction is carried out during 30' at 0ºC and 3h at rt. Next, the mixture was quenched with brine. The DCM solution was dried with Na_2SO_4 and filtered. The solvent was removed under reduced pressure. Finally, the solid was precipitated with diethyl ether and filtered.

11.3. Experimental synthetic methods - Series Q3

Synthesis of 2-amino-3-cyano-6-methoxyquinoxaline-1,4-di-N-oxide [Q3-1].
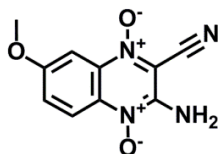
2.50 g (15.06 mmol) of 5-methoxybenzofuroxan, 1.80 mL (18.00 mmol) of malononitrile and 5.00 mL of *N,N*-DMF are placed in a 50.00 mL flask. Next, the mixture is cooled until 0ºC and 1.50 mL of Et₃N are added dropwise. The reaction is carried out during 6h at rt. Cool diethyl ether and MeOH are added and the mixture is filtered. The obtained compound is used without further purification.

Synthesis of 2-cyano-7-methoxyquinoxaline-1,4-di-N-oxide [Q3-2].
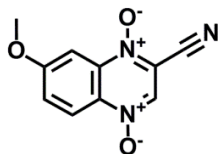
1.60 g (6.88 mmol) of **Q3-2** and 30.00 mL of *N,N*-DMF are placed in a 100.00 mL flask, under N₂ atmosphere. Next, the mixture is heated until 65ºC and 2.80 mL (23.57 mmol) of *tert*-butyl nitrite are added dropwise. Effervescence is observed. After stirring 10' at 65ºC, another 2.10 mL (17.68 mmol) of *tert*-butyl nitrite are added and the mixture is stirred for 1h. The solvent is removed under reduced pressure. Finally, the solid is purified by column chromatography using DCM as mobile phase.

Synthesis of 7-methoxyquinoxaline-2-carbonitrile [Q3-3].
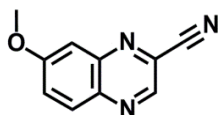
A mixture of 0.91 g (0.97 mmol) of **Q3-2** and 40.00 mL of AcOEt/MeOH (1:1) is heated until 65ºC. Next, a solution of sodium dithionite in water is "*in situ*" prepared (1.01 g of Na₂S₂O₄ in 20.00 mL of water) and added over the mixture. The reaction is carried out during 10' at 65ºC. Next, the mixture is filtered with the aim of eliminating the sodium dithionite and the solvent is removed under reduced pressure. Finally, the compound is purified by flash column chromatography.

Synthesis of N-[(7-methoxyquinoxalin-2-yl)methyl]alkylamide [Q3-4.(I-V)].

Q3-4	R
.I	CH_3
.II	CH_2CH_3
.III	$CH_2CH_2CH_3$
.IV	$CH(CH_3)_2$
.V	Ph
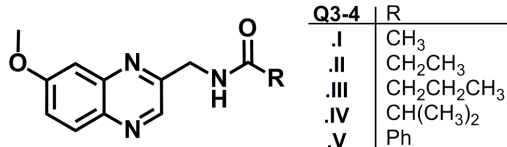
The previously obtained **Q3-3**, THF and the corresponding acid anhydride in proportion 1:2 are placed in the hydrogenation flask. A spatula of Raney®-nickel is added and the mixture is hydrogenated during 7h at 50ºC and 50 psi. Next, DCM is added and the mixture is decanted with the aim of leaving the catalyst off. Next, the mixture is quenched with water, dried with Na_2SO_4 and filtered. The solvent is removed under reduced pressure and the orange oil obtained is purified by flash column chromatography.

11.4. Experimental synthetic methods - Series In1 and In2

Synthesis of 2-(6-methoxyindolyl)acetonitrile [In1-1].
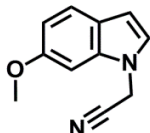
600 mg (4.08 mmol) of 6-methoxyindole are dissolved on 15.00 mL of dry *N,N*-DMF and the mixture is placed under N_2 atmosphere. Next, 800 mg (4.92 mmol) of $(CH_3)_3COK$ are added and the mixture is stirred at rt for 40'. The mixture is heated until 65ºC, 0.80 mL (12.68 mmol) of chloroacetonitrile is added dropwise and the mixture is stirred for 30'. Finally, the mixture is stirred at rt for 20h. The solution is poured onto ice and the formed precipitate is filtered. The obtained solid is used without further purification.

Synthesis of 2-(7-methoxyindolyl)acetonitrile [In2-1].
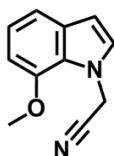
0.30 g (2.04 mmol) of 7-methoxyindole are dissolved in 20.00 mL of HMPA and the mixture is cooled to 0ºC. Next 0.20 g (4.90 mmol) of NaH (60%) is added, N_2 atmosphere is placed and the mixture is stirred at rt for 4h. Next, 0.39 mL (6.12 mmol) of chloroacetonitrile is added and the reaction is stirred for 10' and then stirred again at rt for 20h. The mixture is poured into ice and the formed precipitate is filtered. Finally, the obtained product is purified by flash column chromatography.

Synthesis of 2-(6/7-methoxyindolyl)ethanamine [In1-2 and In2-2].

	In1-2	In2-2
Y	OCH₃	H
Z	H	OCH₃
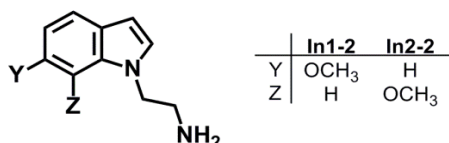
The corresponding nitrile (**In1-1** or **In2-1**) is placed in the hydrogenation flask, where 120.00 mL of EtOH and 12.00 mL of NH_3 (25% aq) are added. A spatula of Raney®-nickel is also added and the reaction is carried out during 7h at rt and 50 psi. Next, the mixture is decanted with the aim of leaving the catalyst off. DCM is added to the mixture and it is quenched with water in order to eliminate the ammonia and remained residues of Raney®-nickel. The organic layer is dried with Na_2SO_4 and filtered. The solvent is removed under reduced pressure and the oily amine is used without further purification.

Synthesis of N-(2-(6/7-methoxyindol)ethyl)-1-imidazolecarboxamide [In1-3.CDI and In2-3.CDI]

	In1-3.CDI	In2-3.CDI
Y	OCH$_3$	H
Z	H	OCH$_3$
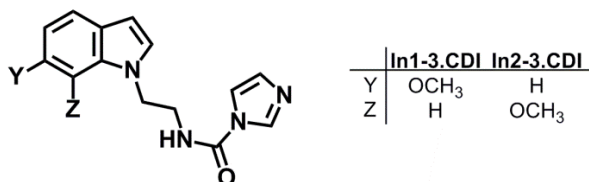
Previously obtained amine (**In1-2** or **In2-2**) is dissolved in 10.00 mL of dry THF and 1,1'-carbonyldiimidazole is added dropwise at 50ºC, in proportion 1:1.5. The reaction is stirred at 50ºC for 3h. Next, the solvent is eliminated under reduced pressure and the obtained crude product is purified by flash column chromatography.

Synthesis of 1-alkyl(2-(6/7-methoxyindolyl)ethyl)-3-urea [In1-3.(I-III), In2-3.(I-III) and In1-3.XIII]

	In1-3	In2-3
Y	OCH$_3$	H
Z	H	OCH$_3$

Comp.	R
.I	CH$_2$CH=CH$_2$
.II	CH$_2$CH$_3$
.III	CH$_2$CH$_2$CH$_3$
.XIII	H
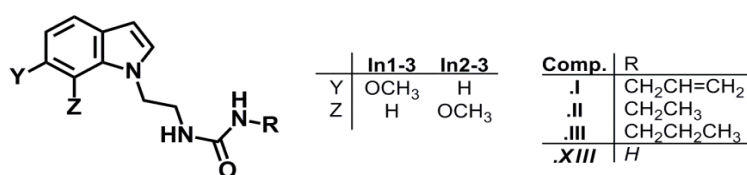
To a solution of the corresponding amine (**In1-2** or **In2-2**) in dry DCM (10.00 mL) the adequate isocyanate (KNCO when R=H) is added in proportion 1:1.5, at 0ºC and under N$_2$ atmosphere. The reaction is stirred at rt for 2h. Then, the reaction mixture is concentrated under reduced pressure and the obtained crude product is purified by flash column chromatography.

Synthesis of 1-alkyl(2-(6/7-methoxyindolyl)ethyl)-3-thiourea [In1-3.(IV-VIII) and In2-3.(IV-VIII)]

	In1-3	In2-3
Y	OCH$_3$	H
Z	H	OCH$_3$

Comp.	R$_1$
.IV	CH$_2$CH=CH$_2$
.V	CH$_2$CH$_3$
.VI	CH$_2$CH$_2$CH$_3$
.VII	CH$_3$
.VIII	cC$_3$H$_5$
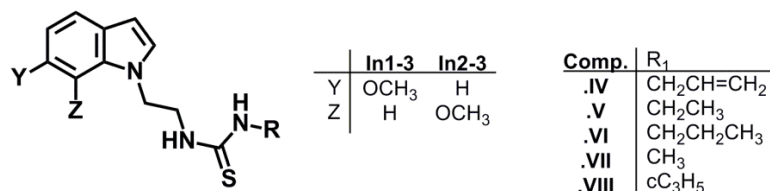
The corresponding amine (**In1-2** or **In2-2**) is dissolved in dry DCM (10.00 mL) at 0ºC and under N$_2$ atmosphere. Then, the corresponding isothiocyanate is added in proportion 1:1.5. The reaction is stirred at rt for 2h and the solvent is eliminated under reduced pressure. The obtained crude product is purified by flash column chromatography.

Synthesis of N-(2-(6/7-methoxyindolyl)ethyl)methanesulfonamide [In1-3.IX and In2-3.IX]

	In1-3.IX	In2-3.IX
Y	OCH$_3$	H
Z	H	OCH$_3$
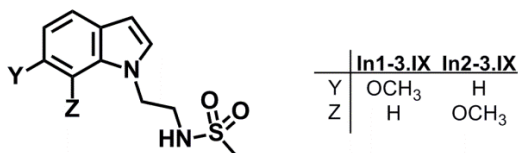
To a solution of the corresponding amine (**In1-2** or **In2-2**) in 10.00 mL of dry DCM, methanesulfonyl chloride in proportion 1:1.5 and catalytic amount of Et$_3$N are added dropwise. Next, the reaction is stirred for 7h at rt. Then, the solvent is eliminated under reduced pressure and the obtained solid is purified by flash column chromatography.

Synthesis of 2,2,2-trifluoro-N-(2-(6/7-methoxyindolyl)ethyl)acetamide [In1-3.X and In2-3.X]

	In1-3.X	In2-3.X
Y	OCH$_3$	H
Z	H	OCH$_3$
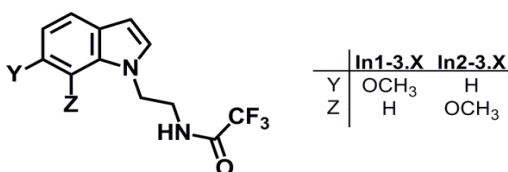
The corresponding amine (**In1-2** or **In2-2**) is dissolved in 10.00 mL of THF and placed at 0ºC. Then, trifluoroacetic anhydride in proportion 1:1.5 and catalytic amount of Et$_3$N are added dropwise. The mixture is stirred at 0ºC for 2h and at rt for 24h. The solvent is removed under reduced pressure. The residue is purified by flash column chromatography.

Synthesis of 1-alkyl-3-(2-(6/7-methoxyindolyl)ethyl) [In1-4.(XI-XII) and In2-4.(XI-XII)]

	In1-4	In2-4
Y	OCH$_3$	H
Z	H	OCH$_3$

Comp.	R
.XI	CH$_3$
.XII	cC$_3$H$_5$
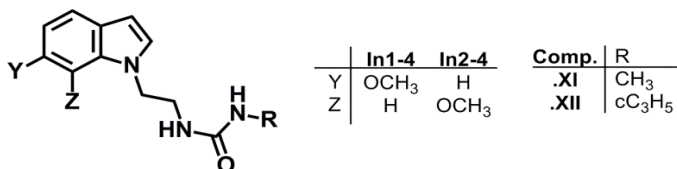
To a solution of corresponding carbonylimidazole derivative (**In1-3.CDI** or **In2-3.CDI**) in dry DCM (10.00 mL) corresponding amines are added in proportion 1:5. The reaction is stirred 2h at rt and 4h at T$_{reflux}$. Next, the solvent is eliminated under reduced pressure and the obtained crude product is purified by flash column chromatography.

113

VII.   BIOLOGICAL EVALUATION

12. PHARMACOLOGICAL ASSAYS

All compounds synthesized in this project have been biologically evaluated in the "*Institute the Recherches Servier*" in France.

12.1. Basis of pharmacological assays

The actions of ligand at receptor depend on two fundamental events; first, the ligand must bind to the receptor. In other words, it might have affinity for the receptor. Second, the ligand may have effects on the receptor and its associated signaling systems. This second attribute has been termed efficacy. Thus, agonist compounds are said to have positive efficacy, inverse agonists or antagonists are said to have negative efficacy and in this hypothetical scale, neutral antagonists should have zero efficacy.

12.1.1. Affinity

The affinity of drugs is tested through radioligand binding assays with membrane preparations. Membranes containing the receptor of interest are incubated with the radioligand for an appropriate period of time and then receptor-bound radioactivity is measured. First, a saturation experiment is carried out by maintaining the receptor quantity constant and varying the radioligand concentration, thus permitting a saturation curve to be generated. From this type of experiments, both the receptor density (B_{max}) and dissociation constant for the radioligand (K_D) can be estimated. Next, inhibition experiments are performed. In this case, the amount of a competing non-radioactive drug included in the incubation is varied while both radioligand and receptor amounts are held constant. Finally, the inhibition constant (Ki) of drugs for the receptor are calculated.

Data from inhibition experiments is then analyzed using non-linear regression techniques. Data is visualized as sigmoid inhibition curves of bound radioligand expressed as percent of maximum response vs. the logarithm of the inhibitor concentrations. The concentration of the inhibitor that reduces bound radioligand by half is the IC_{50} (inhibitory concentration 50%) or EC_{50} (effective concentration 50%) and can be estimated by inhibition curve. The GraphPad Prism software fits inhibition curves using a non-linear regression analysis, directly calculates $logEC_{50}$, and derives K_i using the equation of Cheng and Prusoff (equation 9), where [L] is the concentration of ligands.[150]

$$Ki = \frac{IC_{50}}{1 + \frac{[L]}{K_D}}$$

Equation 9. Cheng-Prusoff equation.

12.1.2. *Efficacy*

The term efficacy has been used to describe the variation of response produced by a drug. Efficacy is measured with the aim of setting up the agonism/antagonism profile of drugs that present an interesting binding affinity.

A molecule that binds to a receptor causing activation is termed as agonist. A full agonist produces a maximum response even when occupying a small proportion of the available receptors, whereas, a partial agonist cannot fully activate the receptors and thus, cannot exert a maximal response. Finally, an inverse agonist or an antagonist is a compound that binds to a receptor but produces the opposite effect.[151]

Whereas binding affinity can be estimated by equilibrium binding studies using the radiolabeled ligand, the evaluation of drug's efficacy requires suitable functional assays that allow a change in the biological state to be measured. The level of G protein activation can be measured by determining the binding of a radiolabeled non-hydrolyzable analog of GTP, [^{35}S]GTPγS, to Gα subunits.[152]

Structurally, GPCRs consist of an extracellular *N*-terminus, seven transmembrane domains and an intracellular C terminus bonded to a heterotrimeric G protein, which is formed by a Gα subunit and a dimer comprising Gβ and Gγ subunits. In the inactive state, the G protein exists as the Gαβγ heterotrimer with GDP bound to the Gα subunit.[153] Agonist binding to receptor causes the dissociation of GDP from Gα, allowing GTP to bind to Gα. This fact leads to the dissociation of the Gα-GTP and Gβγ subunits that are then able to interact with effector systems. The GTPase activity of the Gα subunit hydrolyses GTP to GDP, forming Gα-GDP and thus, allowing Gα and Gβγ subunits to re-associate and the system to turn off (figure 68).[153]
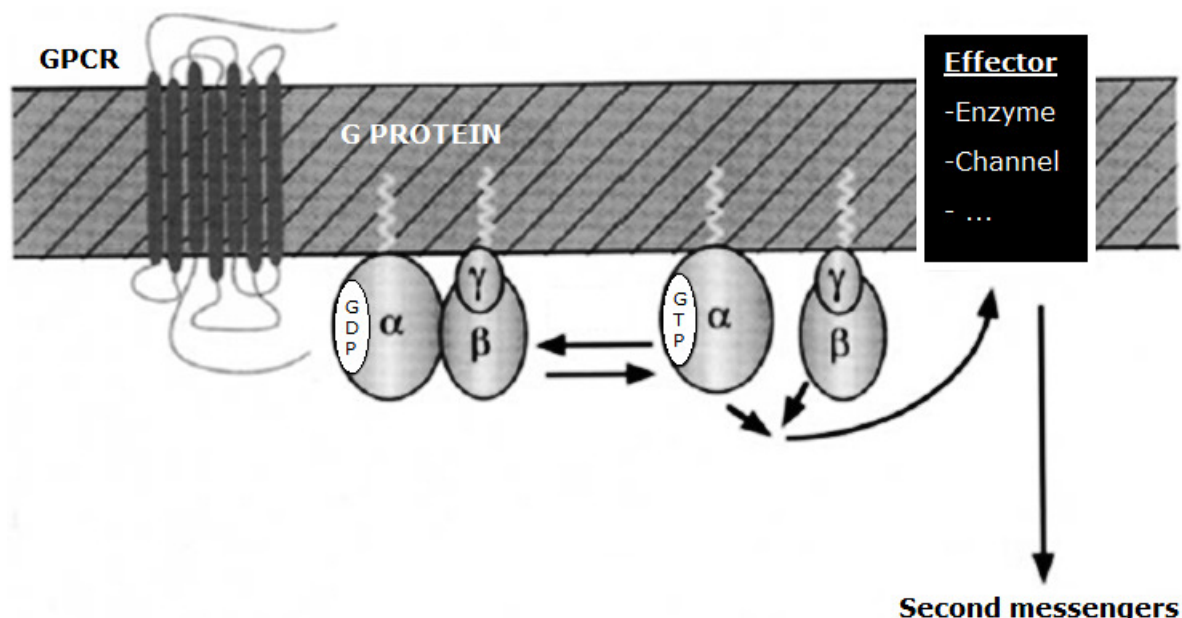
Figure 68. G protein activation and deactivation process [Figure adapted from *De Vries, L. et al.*][154]

The [^{35}S]GTPγS binding assay measures the level of G protein activation generated due to the binding of an agonist to GPCR. In the assay, [^{35}S]GTPγS replaces endogenous GTP and binds to the Gα subunit giving the Gα-[^{35}S]GTPγS species, which is resistant to hydrolysis by the GTPase of Gα. As result, G protein is prevented from reforming as a heterotrimer and, thus, Gα-[^{35}S]GTPγS subunits accumulate and can be measured by counting the amount of [^{35}S]-label incorporated. The amount of accumulated Gα-[^{35}S]GTPγS is a consequence of the action of an activated receptor and thus can be used to determine the degree of agonism and the potency of compounds acting at a particular GPCR. This assay allows the concentration-effect curves construction and therefore, potency (EC_{50}) and relative efficacy (E_{max}) measures.[153] The curve of concentration–effect relationship commonly follows a hyperbola. If this relationship is re-expressed as the percent of maximum response vs. the logarithm of agonist concentration, a sigmoid curve is achieved (figure 69).
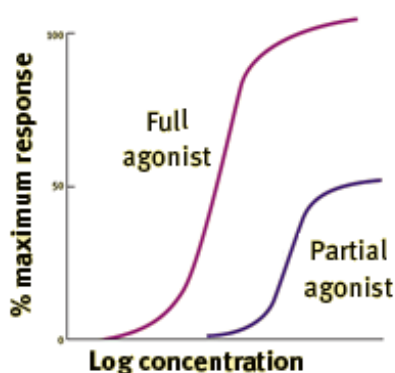
Figure 69. The curve corresponding to the percentage of maximum response vs. the logarithm of agonist concentration.[Figure adapted from *Pleuvry, B.J.]*[151]

12.2. Chemicals, reagents and cell cultures of binding assays

2-[^{125}I]Iodomelatonin (2200 Ci/mmol) has been purchased from NEN (Boston, MA). Other drugs and chemicals have been purchased from Sigma-Aldrich (Saint Quentin, France).

HEK (provided by A.D. Strosberg, Paris, France) and CHO cell lines stably expressing the human melatonin MT_1 or MT_2 receptors are grown in DMEM medium supplemented with 10% fetal calf serum, 2 mM glutamine, 100 IU/mL penicillin and 100 µg/mL streptomycin. Grown at confluence at 37ºC (95% O_2 / 5% CO_2), they are harvested in PBS containing EDTA 2 mM and centrifuged at 1000 g for 5 min (4ºC). The resulting pellet is suspended in Tris 5 mM (pH 7.5), containing EDTA 2 mM and homogenized using a Kinematica polytron. The homogenate is then centrifuged (95,000 g, 30 min, 4ºC) and the resulting pellet suspended in 75 mM Tris (pH 7.5), 12.5 mM $MgCl_2$ and 2 mM EDTA. Aliquots of membrane preparations are stored at -80ºC until use.

12.3. Binding assay protocols

12.3.1. MT_1/MT_2 binding assays

12.3.1.1. MT_1/MT_2 AFFINITY ASSAY (2-[^{125}I]Iodomelatonin)

2-[^{125}I]Iodomelatonin binding assay conditions are essentially as previously described.[155] Briefly, binding is initiated by addition of membrane preparations from stable transfected HEK or CHO cells diluted in binding buffer (50 mM Tris-HCl buffer, pH 7.4, containing 5 mM MgCl$_2$) to 2-[^{125}I]iodomelatonin (25 or 200 pM for MT_1 and MT_2 receptors, respectively, expressed in HEK cells or 20 pM for MT_1 and MT_2 receptors expressed in CHO cells) and the tested drug. Non-specific binding is defined in the presence of 1 µM melatonin. After 120 min incubation at 37ºC, reaction is stopped by rapid filtration through GF/B filters presoaked in 0.5% (v/v) polyethylenimine. Filters are washed three times with 1 mL of ice-cold 50 mM Tris-HCl buffer, pH 7.4. Data from dose-response curves (seven concentrations in duplicate) is analyzed using the program PRISM (Graph Pad Software Inc., San Diego, CA) to yield IC$_{50}$ (inhibitory concentration 50). Results are expressed as Ki using the Cheng and Prusoff equation (equation 9).

12.3.1.2. MT_1/MT_2 EFFICACY ASSAY ([^{35}S]GTPγS)

[^{35}S]GTPγS binding assay is performed according to published methodology.[155] Briefly, membranes from transfected CHO cells expressing MT_1 and MT_2 receptor subtypes and compounds are diluted in binding buffer (20 mM HEPES, pH 7.4, 100 mM NaCl, 3 µM GDP, 3 mM MgCl$_2$, and 20 µg/mL saponin). Incubation is started by the addition of 0.2 nM [^{35}S]GTPγS to membranes (20 µg/mL) and drugs, and continued for 1h at rt. For experiments with antagonists, membranes are pre-incubated with both, the melatonin (3 nM) and the antagonist, for 30 min prior the addition of [^{35}S]GTPγS. Non-specific binding is defined using cold GTPγS (10 µM). Reaction is stopped by rapid filtration through GF/B filters followed by three successive washings with ice-cold buffer. Usual levels of [^{35}S]GTPγS binding (expressed in dpm) are for CHO-MT_2 membranes: 2000 for basal activity, 8000 in the presence of melatonin 1 µM and 180 in the presence of GTPγS 10 µM which defined the non-specific binding. Data from the dose-response curves (seven concentrations in duplicate) are analyzed by using the program PRISM (Graph Pad Software Inc., San Diego, CA) to yield EC$_{50}$ (effective concentration 50%) and E$_{max}$ (maximal effect) for agonists. Antagonist potencies are expressed as K$_B$ using equation 10 where IC$_{50}$ is the inhibitory concentration of antagonist that gives 50% inhibition of [^{35}S]GTPγS binding in the presence of a fixed concentration of melatonin ([Ago]) and EC$_{50}$ ago is the EC$_{50}$ of the molecule when tested alone. I$_{max}$ (maximal inhibitory effect) is expressed as a percentage of that observed with melatonin at 3 nM for MT_2 receptor.

$$K_B = \frac{IC_{50}}{1 + \left(\frac{[Ago]}{EC_{50}ago}\right)}$$

Equation 10. Equation to calculate antagonist potencies.

VIII. MOLECULAR MODELING

13. MATERIALS FOR MOLECULAR MODELING

13.1. Hardware

HP xw6400 workstation, 4 CPUs (central processing units).

13.2. Software

13.2.1. MDL® ISIS/Draw™ (Accelrys®)

ISIS/Draw is a chemically intelligent drawing program that understands the fundamentals of chemistry such as valence limits, bond angles and aromatic ring systems. This understanding enables users to create chemical graphics intuitively, which can be inserted into documents or used to build and search databases of 2D and 3D molecules, polymers and reactions.[156]

13.2.2. Pipeline Pilot™ (Accelrys®)

Pipeline Pilot is a scientific informatic platform that provides immediate access to a large volume of research data, automates the scientific analysis of that data and enables to rapidly explore, visualize and report research results. Pipeline Pilot and its integrated set of applications is a tool that helps to organize the modeling, simulation, informatics and scientific needs of research projects. A powerful key of the platform is the ability to manage and synthesize scientific information in a way that allows rapidly adjusted workflows and optimized research cycles.[157]

13.2.3. Discovery Studio® (Accelrys®)

Discovery Studio is a type of software used to provide molecular design solutions to computational chemistry. Discovery Studio makes it easier to examine the properties of large and small molecules, study systems, identify leads and optimize candidates. Furthermore, built on Pipeline Pilot technology, Discovery Studio rapidly automates routine tasks and integrates third party applications.[158]

13.2.4. Sybyl™ (Tripos®)

Sybyl is a program created to provide comprehensive tools for molecular modeling, optimization, comparison, data association and visualization of structures. Sybyl permits performance of routine tasks in a fast and easy way and its discovery workflows have been enhanced from obtaining data to deliver easy and understanding results.[95]

13.2.5. Spartan™ (Wavefunction®)

Spartan is a tool used to explore chemical compounds through molecular mechanics and quantum chemical calculations. This software is implemented with an array of graphical models that allows useful management of the results of these calculations.[159]

13.2.6. SOM_PAK (Self Organizing Map Program Package)

SOM_PAK is a program capable of performing "non-linear projection" of the probability density function of the high dimensional input data onto the 2D display. Thus, it preserves the majority of the information contained in high dimension data in 2D format and as a consequence, the management of the data becomes easier and more understandable.[99]

13.3. Ligands

All molecules used in this work belong to the database of the *"Institute de Recherches Servier"* including the molecules synthesized on this project. Thus, all these compounds have been tested by Servier, using the same protocols in every case *(see Ch2 – 11.3)*.

MLT is taken as reference to set the 3D conformation of the compounds used in this work *(see Ch1 – 1.2.4)*

14. PROTOCOLS

14.1. 3D-QSAR model generation protocol

14.1.1. 3D modeling of ligands

- Draw molecules in 2D

- Convert the molecules to 3D

- Refinement of a correct geometry with MM energy minimization runs:

  ❖ Conjugated gradient minimizer

  ❖ 300 iterations

  ❖ 0.05 Kcal/(Mol*A)

  ❖ Force field: Tripos

  ❖ Charges: Gasteiger-Hückel

- Refine an appropriate conformation of the molecule to be superimposed with the MT_1-bound conformation of the reference MLT (from a 3D model of MT_1 validated to preferentially bind MT agonists)

- Charges and energy computation of the molecules with QM

  ❖ Calculate energy at ground state

  ❖ With density functional/B3LYP/6-31G*

  ❖ In water

  ❖ Print atomic charges

14.1.2. Generation 3D-QSAR models

- Create a quantitative 3D-QSAR model for activity prediction:

  ❖ Create CoMFA models with PLS

  ❖ Create CoMSIA models with PLS

  ❖ Validate the best models and analyze the results

- Create a qualitative 3D-QSAR model for a two-class (active/inactive) prediction model:

  ❖ Create CoMFA models with SIMCA

  ❖ Create CoMSIA models with SIMCA

  ❖ Validate the best models and analyze the results

14.1.3. Activity prediction of molecules without known biological data

- Create the new molecules and align them with those of the training set.

- Predict their activity with the best model.

14.2. Bayesian model generation method

14.2.1. Database generation

- Select a large amount of MT_1 agonists with known biological data and obtain them in 2D in a single ".sd" file.

- Select a range to set the active/inactive class of the molecules in the database.

- Prepare a training set file and a validation set file.

14.2.2. Molecular features description and model construction

- Obtain the fingerprint types that best characterized the 2D molecular structures in the training set and generate the model with those fingerprints.

- Check the results of the generated models:

   ❖   ROC curve

   ❖   AUC value

- Make combinations between fingerprint types to find the models that best represent the molecules in the training set.

- Compare the different models and choose the best ones.

- Use the best models to predict the validation sets.

- Analyze the ROC curves and AUC values obtained with the different models

- Select the best Bayesian model.

14.2.3. Activity prediction of molecules without known biological data

- Predict the activity of the molecules using the best model.

14.3. Kohonen neural network (KNN)

14.3.1. Database generation

- Make a Sybyl database with the molecules converted into 3D with QM computations.

- Calculate the vdW surface of each molecule.

   ❖   183.9 mol/Å2 density of dots

   ❖   Conolly surface

- Prepare a text file per molecule with the coordination and potential value of each dot.

- Create a single file placing all the text files with dot information in the same file to form the test file.

14.3.2. Selection of parameters

- Create a map using MLT as reference file and also as test file.

- Change the parameters and create new maps.

- Select the parameters (**P0**) of the best map (**MAP_0**) as starting point.

14.3.3. Input file generation

- Create a map using MLT as reference file, the generated test file as test file and the parameters P0.

- Modify the parameters and select the best map (**MAP_1**) and its parameters (**P1**).

- Make the visualization of each molecule in **MAP_1** in order to obtain one potential map per molecule.

- Make a matrix with the potential map of each molecule in one line and use it as input file.

14.3.4. Map generation

- Create a map with the generated input file as reference file and test as test file, using the parameters **P1**.

- Change the parameters and select the best map (**MAP_2**).

14.3.5. Molecule classification

- Make the visualization of input file in **MAP_2** and obtain the distribution of each test molecule in the same map.

14.3.6. Activity prediction of molecules without known biological data

- Calculate the vdW surface of the molecules and calculate their potential map.

- Visualize the molecules in **MAP_1** and create a matrix with all the information.

- Create a new test file by combination of the two previous matrixes.

- Make the visualization of the new test file in **MAP_2** and obtain the distribution of each test molecule in the same map.

Chapter 3

Results and discussion

IX. COMPOUND CHARACTERIZATION

15. COMPOUND CHARACTERIZATION – SERIES Q1

15.1. Series Q1 - Intermediates

2-chloro-3-methoxyquinoxaline [Q1-1]
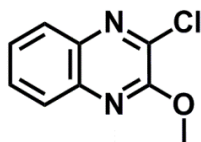
Appearance: Beige solid
MF: $C_9H_7N_2OCl$
MW: 194.5 g/mol
Yield: 86%

IR (KBr cm⁻¹): 3032 (w, aromatic C-H); 2987 and 2936 (w, aliphatic C-H).

¹H RMN (CDCl₃, 400 MHz) δ: 7.96 (ddd, 1H, **H₈**, J_{8-7}=8.3 Hz, J_{8-6}=1.5 Hz, J_{8-5}=0.5 Hz); 7.88 (ddd, 1H, **H₅**, J_{5-6}=8.3 Hz, J_{5-7}=1.4 Hz, J_{5-8}=0.5 Hz); 7.71 (ddd, 1H, **H₆**, J_{6-5}=8.4 Hz, J_{6-7}=7.1 Hz, J_{6-8}=1.5 Hz); 7.60 (ddd, 1H, **H₇**, J_{7-8}=8.4 Hz, J_{7-6}=7.1 Hz, J_{7-5}=1.5 Hz); 4.20 (s, 3H, OC**H₃**) ppm.

CHN	Calculated	Found
C	55.53%	55.23%
H	3.60%	3.34%
N	14.40%	14.41%

3-methoxyquinoxaline-2-carbonitrile [Q1-2]
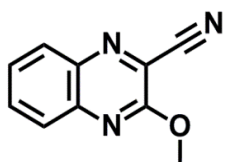
Appearance: Beige solid
MF: $C_{10}H_7N_3O$
MW: 185 g/mol
Yield: 20%

IR (KBr cm⁻¹): 3026 (w, aromatic C-H); 2943 (w, aliphatic C-H); 2232 (s, C≡N).

¹H RMN (CDCl₃, 400 MHz) δ: 8.08 (ddd, 1H, **H₈**, J_{8-7}=8.4 Hz, J_{8-6}=1.5 Hz, J_{8-5}=0.5 Hz); 7.88 (ddd, 1H, **H₅**, J_{5-6}=8.3 Hz, J_{5-7}=1.4 Hz, J_{5-8}=0.5 Hz); 7.71 (ddd, 1H, **H₆**, J_{6-5}=8.4 Hz, J_{6-7}=6.9 Hz, J_{6-8}=1.4 Hz); 7.60 (ddd, 1H, **H₇**, J_{7-8}=8.4 Hz, J_{7-6}=6.9 Hz, J_{7-5}=1.5 Hz); 4.20 (s, 3H, OC**H₃**) ppm.

CHN	Calculated	Found
C	64.86%	65.23%
H	3.78%	3.61%
N	22.70%	22.70%

15.2. Series Q1 - Final compounds

N-[(3-methoxyquinoxalin-2-yl)methyl]acetamide [Q1-3.I]
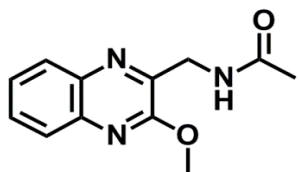
Appearance: White solid
*MF: $C_{12}H_{13}N_3O_2 \cdot \frac{1}{4} H_2O$*
MW: 235.5 g/mol
Yield: 7%

IR (KBr cm^{-1}): 3314 (s, NH); 3084 (w, aromatic C-H); 2937 (w, aliphatic C-H); 1648 (vs, C=O).

^{1}H RMN (CDCl$_3$, 400 MHz) δ: 8.11 (dd, 1H, **H$_8$**, $J_{8\text{-}7}$=8.1 Hz, $J_{8\text{-}6}$=1.1 Hz); 7.91 (ddd, 1H, **H$_5$**, $J_{5\text{-}6}$=8.3 Hz, $J_{5\text{-}7}$=1.3 Hz); 7.71 (ddd, 1H, **H$_6$**, $J_{6\text{-}5}$=8.3 Hz, $J_{6\text{-}7}$=7.1 Hz, $J_{6\text{-}8}$=1.3 Hz); 7.62 (ddd, 1H, **H$_7$**, $J_{7\text{-}8}$=8.3 Hz, $J_{7\text{-}6}$=7.1 Hz, $J_{7\text{-}5}$=1.4 Hz); 4.76 (d, 2H, C**H$_2$**, $J_{CH2\text{-}NH}$=4.3 Hz); 4.17 (s, 3H, OC**H$_3$**); 2.21 (s, 3H, C**H$_3$**) ppm.

CHN	Calculated	Found
C	61.14%	61.26%
H	5.73%	5.47%
N	17.83%	17.65%

MP: 231-235ºC
HPLC: N.D

N-[(3-methoxyquinoxalin-2-yl)methyl]propionamide [Q1-3.II]
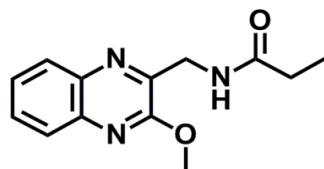
Appearance: White solid
MF: $C_{13}H_{15}N_3O_2 \cdot \frac{1}{10} H_2O$
MW: 246.8 g/mol
Yield: 27%

**IR (KBr cm^{-1}):** 3308 (s, NH); 3078 (w, aromatic C-H); 2969 and 2943 (w, aliphatic C-H); 1638 (vs, C=O).

^{1}H RMN (DMSO-d_6, 400 MHz) δ: 8.27 (t, 1H, N**H**, $J_{NH\text{-}CH2}$=5.3 Hz); 7.96 (dd, 1H, **H$_8$**, $J_{8\text{-}7}$=8.2 Hz, $J_{8\text{-}6}$=1.2 Hz); 7.85 (dd, 1H, **H$_5$**, $J_{5\text{-}6}$=8.3 Hz, $J_{5\text{-}7}$=1.0 Hz); 7.72 (ddd, 1H, **H$_6$**, $J_{6\text{-}5}$=8.3 Hz, $J_{6\text{-}7}$=7.0 Hz, $J_{6\text{-}8}$=1.5 Hz); 7.63 (ddd, 1H, **H$_7$**, $J_{7\text{-}8}$=8.3 Hz, $J_{7\text{-}6}$=7.0 Hz, $J_{7\text{-}5}$=1.5 Hz); 4.53 (d, 2H, C**H$_2$**NH, $J_{CH2\text{-}NH}$=5.6 Hz); 4.07 (s, 3H, OC**H$_3$**); 2.23 (q, 2H, C**H$_2$**CH$_3$, $J_{CH2\text{-}CH3}$=7.6 Hz); 1.06 (t, 3H, CH$_2$C**H$_3$**, $J_{CH3\text{-}CH2}$=7.6 Hz) ppm.

CHN	Calculated	Found
C	63.21%	63.44%
H	6.16%	6.08%
N	17.02%	16.65%

MP: 164ºC
HPLC: N.D

N-[(3-methoxyquinoxalin-2-yl)methyl]butyramide [Q1-3.III]
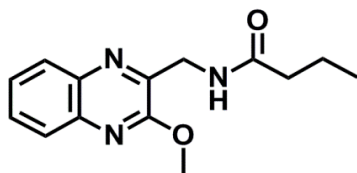
Appearance: White solid
MF: $C_{14}H_{17}N_3O_2 \cdot \frac{1}{8} H_2O$
MW: 261.25 g/mol
Yield: 9%

IR (KBr cm⁻¹): 3309 (s, NH); 3065 (w, aromatic C-H); 2951 (w, aliphatic C-H); 1641 (vs, C=O).

¹H RMN (DMSO-*d₆*, 400 MHz) δ: 8.30 (bs, 1H, N**H**); 7.93 (d, 1H, **H₈**, J_{8-7}=7.9 Hz); 7.84 (d, 1H, **H₅**, J_{5-6}=8.2 Hz); 7.70 (dd, 1H, **H₆**, J_{6-5}=8.2 Hz, J_{6-7}=6.7 Hz); 7.61 (dd, 1H, **H₇**, J_{7-8}=8.1 Hz, J_{7-6}=6.9 Hz); 4.52 (d, 2H, C**H₂**NH, J_{CH2-NH}=5.4 Hz); 4.07 (s, 3H, OC**H₃**); 2.18 (t, 2H, C**H₂**CH₂CH₃, $J_{CH2-CH2}$=7.2 Hz); 1.56 (dd, 2H, CH₂C**H₂**CH₃, $J_{CH2-CH2}$=7.2 Hz, $J_{CH2-CH3}$=7.2 Hz, J_{gem}=14.6 Hz); 0.91 (t, C**H₃**, $J_{CH3-CH2}$=7.3 Hz) ppm.

CHN	Calculated	Found
C	64.31%	64.34%
H	6.60%	6.65%
N	16.08%	15.93%

MP: 259ºC
HPLC: 95.449% / t_R=2.647'

N-[(3-methoxyquinoxalin-2-yl)methyl]isobutyramide [Q1-3.IV]
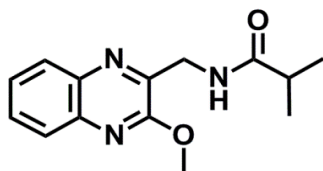
Appearance: White solid
MF: $C_{14}H_{17}N_3O_2$
MW: 259 g/mol
Yield: 28%

IR (KBr cm⁻¹): 3312 (s, NH); 3065 (w, aromatic C-H); 2968 and 2937 (w, aliphatic C-H); 1644 (vs, C=O).

¹H RMN (DMSO-*d₆*, 400 MHz) δ: 8.23 (t, 1H, N**H**, J_{NH-CH2}=5.6 Hz); 7.93 (dd, 1H, **H₈**, J_{8-7}=8.2 Hz, J_{8-6}=1.0 Hz); 7.85 (dd, 1H, **H₅**, J_{5-6}=8.2 Hz, J_{5-7}=0.9 Hz); 7.72 (ddd, 1H, **H₆**, J_{6-5}=8.3 Hz, J_{6-7}=7.1 Hz, J_{6-8}=1.5 Hz); 7.62 (ddd, 1H, **H₇**, J_{7-8}=8.4 Hz, J_{7-6}=7.1 Hz, J_{7-5}=1.5 Hz); 4.51 (d, 2H, C**H₂**NH, J_{CH2-NH}=5.6 Hz); 4.07 (s, 3H, OC**H₃**); 2.55 (q, 1H, C**H**(CH₃)₂, $J_{CH-(CH3)2}$=6.8 Hz); 1.08 (d, 6H, CH(C**H₃**)₂, $J_{(CH3)2-CH}$=6.8 Hz) ppm.

CHN	Calculated	Found
C	64.86%	64.55 %
H	6.56%	6.79%
N	16.22%	15.91%

MP: 142-144ºC
HPLC: N.D.

N-[(3-methoxyquinoxalin-2-yl)methyl]benzamide [Q1-3.V]
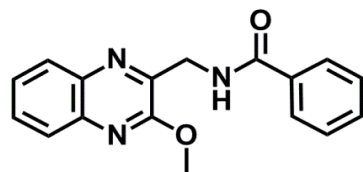
Appearance: White solid
MF: $C_{17}H_{15}N_3O_2 \cdot \frac{1}{8} H_2O$
MW: 295.25 g/mol
Yield: 38%

IR (KBr cm^{-1}): 3391 (s, NH); 3071 and 3026 (w, aromatic C-H); 2943 and 2904 (w, aliphatic C-H); 1659 (vs, C=O).

^{1}H RMN (DMSO-d_6, 400 MHz) δ: 9.00 (t, 1H, N**H**, J_{NH-CH2}=5.4 Hz); 7.93 (m, 3H, **H$_8$+H$_{2'}$+H$_{6'}$**); 7.86 (d, 1H, **H$_5$**, J_{5-6}=8.3 Hz); 7.71 (dd, 1H, **H$_6$**, J_{6-5}=8.1 Hz, J_{6-7}=7.2 Hz); 7.54 (m, 4H, **H$_7$+H$_{3'}$+H$_{4'}$+H$_{5'}$**); 4.74 (d, 2H, C**H$_2$**, J_{CH2-NH}=5.4 Hz); 4.10 (s, 3H, OC**H$_3$**) ppm.

CHN	Calculated	Found
C	69.09%	69.13%
H	5.17%	5.14%
N	14.23%	13.99%

MP: 137ºC
HPLC: 96.063% / t_R=3.516'

1-ethyl-3-[(3-methoxyquinoxalin-2-yl)methyl]urea [Q1-3.VI]
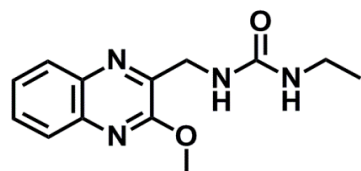
Appearance: Beige solid
MF: $C_{13}H_{16}N_4O_2$
MW: 260 g/mol
Yield: 3%

IR (KBr cm^{-1}): 3334 (m, NH); 3061 (w, aromatic C-H); 2978 and 2939 (w, aliphatic C-H); 1629 (vs, C=O).

^{1}H RMN (DMSO-d_6, 400 MHz) δ: 7.96 (dd, 1H, **H$_8$**, J_{8-7}=8.2 Hz, J_{8-6}=1.2 Hz); 7.85 (dd, 1H, **H$_5$**, J_{5-6}=8.2 Hz, J_{5-7}=1.2 Hz); 7.71 (ddd, 1H, **H$_6$**, J_{6-5}=8.4 Hz, J_{6-7}=7.0 Hz, J_{6-8}=1.4 Hz); 7.63 (ddd, 1H, **H$_7$**, J_{7-8}=8.4 Hz, J_{7-6}=7.1 Hz, J_{7-5}=1.4 Hz); 6.41 (t, 1H, ArCH$_2$N**H**, J_{NH-CH2}=5.0 Hz); 6.33 (t, 1H, N**H**CH$_2$CH$_3$, J_{NH-CH2}=5.3 Hz); 4.50 (d, 2H, ArC**H$_2$**NH, J_{CH2-NH}=5.4 Hz); 4.07 (s, 3H, OC**H$_3$**); 3.05 (dq, 2H, C**H$_2$**CH$_3$, $J_{CH2-CH3}$=7.2 Hz, J_{CH2-NH}=5.8 Hz); 1.02 (dt, 3H, C**H$_3$**, $J_{CH3-CH2}$=7.2 Hz, J_{CH3-NH}=1.3 Hz) ppm.

CHN	Calculated	Found
C	60.00%	59.91%
H	6.15%	6.12%
N	21.54%	21.14%

MP: 160.5-160.8ºC
HPLC: 95.664% / t_R=2.169'

1-[(3-methoxyquinoxalin-2-yl)methyl]-3-propylurea [Q1-3.VII]
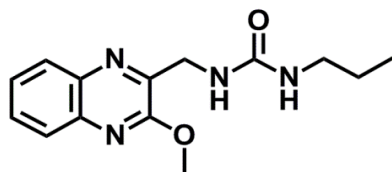
Appearance: White solid

MF: C$_{14}$H$_{18}$N$_4$O$_2$

MW: 274 g/mol

Yield: 9%

IR (KBr cm^{-1}): 3330 (s, NH); 3061 (w, aromatic C-H); 2960 (w, aliphatic C-H); 1633 (vs, C=O).

^{1}H RMN (DMSO-*d$_6$*, 400 MHz) δ: 7.95 (d, 1H, **H$_8$**, $J_{8\text{-}7}$=8.1 Hz); 7.85 (d, 1H, **H$_5$**, $J_{5\text{-}6}$=8.2 Hz); 7.71 (dd, 1H, **H$_6$**, $J_{6\text{-}5}$=8.0 Hz, $J_{6\text{-}7}$=7.3 Hz); 7.63 (dd, 1H, **H$_7$**, $J_{7\text{-}8}$=8.0 Hz, $J_{7\text{-}6}$=7.1 Hz); 6.41 (t, 1H, ArCH$_2$N**H**, $J_{NH\text{-}CH2}$=5.5 Hz); 6.37 (t, 1H, N**H**CH$_2$CH$_2$CH$_3$, $J_{NH\text{-}CH2}$=6.0 Hz); 4.50 (d, 2H, ArC**H$_2$**NH, $J_{CH2\text{-}NH}$=5.2 Hz); 4.07 (s, 3H, OC**H$_3$**); 2.98 (dd, 2H, NHC**H$_2$**CH$_2$CH$_3$, $J_{CH2\text{-}NH}$=6.4 Hz, $J_{CH2\text{-}CH2}$=6.7 Hz); 1.40 (tq, 2H, NHCH$_2$C**H$_2$**CH$_3$, $J_{CH2\text{-}CH2}$=6.8 Hz, $J_{CH2\text{-}CH3}$=7.1 Hz, J_{gem}=13.5 Hz); 0.86 (t, 3H, C**H$_3$**, $J_{CH3\text{-}CH2}$=7.4 Hz) ppm.

CHN	Calculated	Found
C	61.31%	60.99%
H	6.57%	6.45%
N	20.44%	20.31%

MP: 115.3-116ºC

HPLC: 95.193% / t$_R$=2.593'

1-isopropyl-3-[(3-methoxyquinoxalin-2-yl)methyl]urea [Q1-3.VIII]
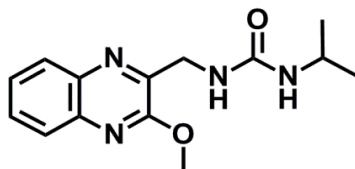
Appearance: White solid

MF: C$_{14}$H$_{18}$N$_4$O$_2$

MW: 274 g/mol

Yield: 23%

IR (KBr cm^{-1}): 3363 and 3283 (s, NH); 3100 (w, aromatic C-H); 2971 (w, aliphatic C-H); 1620 (vs, C=O).

^{1}H RMN (DMSO-*d$_6$*, 400 MHz) δ: 7.95 (dd, 1H, **H$_8$**, $J_{8\text{-}7}$=8.1 Hz, $J_{8\text{-}6}$=1.1 Hz); 7.85 (dd, 1H, **H$_5$**, $J_{5\text{-}6}$=8.2 Hz, $J_{5\text{-}7}$=0.9 Hz); 7.71 (ddd, 1H, **H$_6$**, $J_{6\text{-}5}$=8.2 Hz, $J_{6\text{-}7}$=7.7 Hz, $J_{6\text{-}8}$=1.2 Hz); 7.63 (ddd, 1H, **H$_7$**, $J_{7\text{-}8}$=8.0 Hz, $J_{7\text{-}6}$=7.4 Hz, $J_{7\text{-}5}$=1.1 Hz); 6.34 (t, 1H, ArCH$_2$N**H**, $J_{NH\text{-}CH2}$=5.1 Hz); 6.26 (d, 1H, N**H**CH(CH$_3$)$_2$, $J_{NH\text{-}CH}$=7.6 Hz); 4.49 (d, 2H, ArC**H$_2$**NH, $J_{CH2\text{-}NH}$=5.3 Hz); 4.07 (s, 3H, OC**H$_3$**); 3.71-3.67 (m, 1H, C**H**(CH$_3$)$_2$); 1.06 (d, 6H, CH(C**H$_3$**)$_2$, $J_{(CH3)2\text{-}CH}$=6.5 Hz) ppm.

CHN	Calculated	Found
C	61.31%	61.33%
H	6.57%	6.58%
N	20.44%	20.43%

MP: 171.2ºC

HPLC: 99.820% / t$_R$=2.551'

1-[(3-methoxyquinoxalin-2-yl)metal]-3-phenylurea [Q1-3.IX]
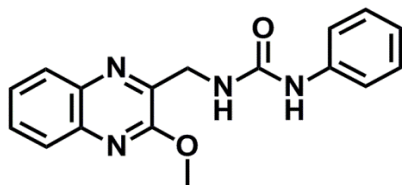
Appearance: White solid
MF: $C_{17}H_{16}N_4O_2 \cdot {}^1/_8 H_2O$
*MW: 310.25 g/mol*
Yield: 4%

IR (KBr cm^{-1}): 3310 (s, NH); 3177 and 3023 (w, aromatic C-H); 2952 and 2907 (w, aliphatic C-H); 1640 (vs, C=O).

^{1}H RMN (DMSO-d_6, 400 MHz) δ: 9.01 (s, 1H, CON**H**-Ph); 7.99 (dd, 1H, **H$_8$**, J_{8-7}=8.2 Hz, J_{8-6}=1.2 Hz); 7.87 (dd, 1H, **H$_5$**, J_{5-6}=8.3 Hz, J_{5-7}=1.1 Hz); 7.73 (ddd, 1H, **H$_6$**, J_{6-5}=8.4 Hz, J_{6-7}=7.1 Hz, J_{6-8}=1.5 Hz); 7.64 (ddd, 1H, **H$_7$**, J_{7-8}=8.4 Hz, J_{7-6}=7.1 Hz, J_{7-5}=1.5 Hz); 7.43 (dd, 2H, **H$_{2'}$**+**H$_{6'}$**, $J_{2'-3'}$=$J_{6'-5'}$=8.6 Hz, $J_{2'-4'}$=$J_{6'-4'}$=1.1 Hz); 7.24 (dd, 2H, **H$_{3'}$**+**H$_{5'}$**, $J_{3'-2'}$=$J_{5'-6'}$=8.4 Hz, $J_{3'-4'}$=$J_{5'-4'}$=7.5 Hz); 6.91 (tt, 1H, **H$_{4'}$**, $J_{4'-3'}$=$J_{4'-5'}$=7.5 Hz, $J_{4'-2'}$=$J_{4'-6'}$=1.1Hz); 6.83 (t, 1H, ArCH$_2$N**H**, J_{NH-CH2}=5.2 Hz); 4.61 (d, 2H, ArC**H$_2$**NH, J_{CH2-NH}=5.2 Hz); 4.10 (s, 3H, OC**H$_3$**) ppm.

CHN	Calculated	Found
C	65.75%	65.76%
H	5.24%	5.32%
N	18.05%	18.12%

MP: 201-201.5ºC
HPLC: 95.854% / t_R=3.604'

16. COMPOUND CHARACTERIZATION - SERIES Q2

16.1. Series Q2 - Intermediates

2,3-dimethoxy-6-nitroquinoxaline [Q2-1]
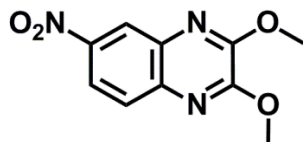
Appearance: Beige solid
MF: $C_{10}H_9N_3O_4$
MW: 235 g/mol
Yield: 52%

IR (KBr cm⁻¹): 3116 and 3014 (w, aromatic C-H); 2956 (w, aliphatic C-H).

¹H RMN (DMSO-d_6, 400 MHz) δ: 8.50 (d, 1H, **H₅**, J_{5-7}=2.6 Hz); 8.30 (dd, 1H, **H₇**, J_{7-8}=9.0 Hz, J_{7-5}=2.6 Hz); 7.93 (d, 1H, **H₈**, J_{8-7}=9.0 Hz); 4.11 (s, 3H, OC**H₃**-C2); 4.10 (s, 3H, OC**H₃**-C3) ppm.

CHN	Calculated	Found
C	51.06%	50.99%
H	3.83%	3.79%
N	17.87%	17.40%

2,3-dimethoxyquinoxaline-6-amine [Q2-2]
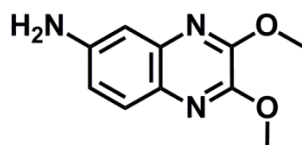
Appearance: White solid
MF: $C_{10}H_{11}N_3O_2$
MW: 205 g/mol
Yield: 96%

IR (KBr cm⁻¹): 3442 and 3347 (w, NH₂); 3219 (w, aromatic C-H); 2937 (w, aliphatic C-H).

¹H RMN (DMSO-d_6, 400 MHz) δ: 7.42 (d, 1H, **H₈**, J_{8-7}=8.6 Hz); 6.87 (dd, 1H, **H₇**, J_{7-8}=8.6 Hz, J_{7-5}=2.4 Hz); 6.78 (d, 1H, **H₅**, J_{5-7}=2.5 Hz); 5.43 (s, 2H, N**H₂**); 3.97 (s, 3H, OC**H₃**-C2); 3.94 (s, 3H, OC**H₃**-C3) ppm.

CHN	Calculated	Found
C	58.54%	58.08%
H	5.37%	5.37%
N	20.49%	20.41%

16.2. Series Q2 - Final compounds

N-(2,3-dimethoxyquinoxalin-6-yl)acetamide [Q2-3.I]
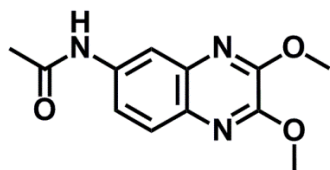
Appearance: White solid
MF: $C_{12}H_{13}N_3O_3 \cdot \frac{1}{2} H_2O$
MW: 256 g/mol
Yield: 13%

IR (KBr cm^{-1}): 3276 (s, N-H); 3116 (w, aromatic C-H); 2943 (w, aliphatic C-H); 1664 (s, C=O).

^{1}H RMN (DMSO-d_6, 400 MHz) δ: 10.19 (s, 1H, N**H**); 8.21 (d, 1H, **H$_5$**, J_{5-7}=2.2 Hz); 7.67 (d, 1H, **H$_8$**, J_{8-7}=8.8 Hz); 7.57 (dd, 1H, **H$_7$**, J_{7-8}=8.9 Hz, J_{7-5}=2.2 Hz); 4.03 (s, 3H, OC**H$_3$**$_{-C2}$); 4.01 (s, 3H, OC**H$_3$**$_{-C3}$); 2.10 (s, 3H, C**H$_3$**) ppm.

CHN	Calculated	Found
C	56.25%	56.23%
H	5.47%	5.42%
N	16.41%	16.04%

MP: 290ºC
HPLC: 99.568% / t_R=2.019'

N-(2,3-dimethoxyquinoxalin-6-yl)propionamide [Q2-3.II]
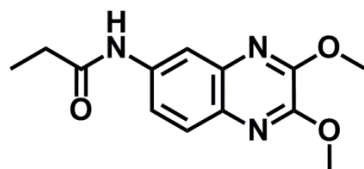
Appearance: White solid
MF: $C_{13}H_{15}N_3O_3 \cdot \frac{1}{2} H_2O$
*MW: 270 g/mol*
Yield: 68%

IR (KBr cm^{-1}): 3284 (s, NH); 3175 and 3105 (w, aromatic C-H); 2987 and 2943 (w, aliphatic C-H); 1663 (vs, C=O).

^{1}H RMN (DMSO-d_6, 400 MHz) δ: 10.11 (s, 1H, N**H**); 8.21 (s, 1H, **H$_5$**); 7.67 (dd, 1H, **H$_8$**, J_{8-7}=8.8 Hz, J_{8-5}=1.2 Hz); 7.60 (ddd, 1H, **H$_7$**, J_{7-8}=8.9 Hz, J_{7-5}=2.2 Hz, J_{H7-NH}=1.5 Hz); 4.03 (d, 3H, OC**H$_3$**$_{-C2}$, $J_{OCH3-OCH3}$=1.4 Hz); 4.01 (d, 3H, OC**H$_3$**$_{-C3}$, $J_{OCH3-OCH3}$=1.4 Hz); 2.38 (dq, 2H, C**H$_2$**, $J_{CH2-CH3}$=7.5 Hz, J_{CH2-NH}=1.3 Hz); 2.12 (dt, 3H, C**H$_3$**, $J_{CH3-CH2}$=7.5 Hz, J_{CH3-NH}=1.4 Hz) ppm.

CHN	Calculated	Found
C	57.78%	58.15%
H	5.93%	5.92%
N	15.56%	15.61%

MP: 182-183ºC
HPLC: 99.658% / t_R=2.359'

NOTE! In all the compounds of this series, the signals corresponding to the two methoxy groups present in quinoxaline ring show a coupling between them but instead of the expected quadruplet, a doublet is obtained. A 2D ^{1}H-^{1}H COSY experiment was carried out in order to confirm the existence of this coupling and avoid a possibly poorly resolved spectrum which would generate confusion (see Ch3 -21.2).

N-(2,3-dimethoxyquinoxalin-6-yl)butyramide [Q2-3.III]
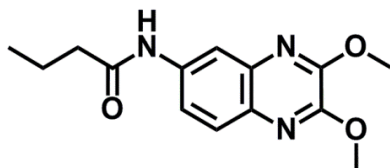
Appearance: White solid

MF: $C_{14}H_{17}N_3O_3$

MW: 275 g/mol

Yield: 26%

IR (KBr cm^{-1}): 3289 (m, NH); 2949 (w, aliphatic C-H); 1656 (vs, C=O).

^{1}H RMN (DMSO-d_6, 400 MHz) δ: 10.11 (s, 1H, N**H**); 8.21 (s, 1H, **H$_5$**); 7.66 (dd, 1H, **H$_7$**, J_{7-8}=8.4 Hz, J_{7-5}=2.9 Hz); 7.59 (d, 1H, **H$_8$**, J_{8-7}=8.7 Hz); 4.03 (d, 3H, OC**H$_3$**$_{-C2}$, $J_{OCH3-OCH3}$=2.6 Hz); 4.01 (d, 3H, OC**H$_3$**$_{-C3}$, $J_{OCH3-OCH3}$=2.6 Hz); 2.36-2.32 (m, 2H, COC**H$_2$**CH$_2$CH$_3$); 1.66-1.62 (m, 2H, COCH$_2$C**H$_2$**CH$_3$); 0.96-0.92 (m, 3H, C**H$_3$**) ppm.

CHN	Calculated	Found
C	61.09%	60.78%
H	6.18%	6.32%
N	15.27%	15.59%

MP: 136-140ºC

HPLC: 99.831% / t_R=2.879'

N-(2,3-dimethoxyquinoxalin-6-yl)isobutyramide [Q2-3.IV]
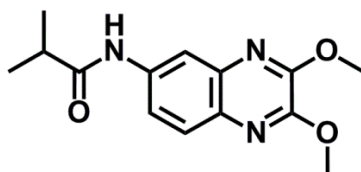
Appearance: White solid

MF: $C_{14}H_{17}N_3O_3 \cdot {}^1/_8 H_2O$

*MW: 277.25*

Yield: 72%

IR (KBr cm^{-1}): 3253 (s, NH); 3026 (w, aromatic C-H); 2969 (w, aliphatic C-H); 1653 (vs, C=O).

^{1}H RMN (DMSO-d_6, 400 MHz) δ: 10.07 (s, 1H, N**H**); 8.21 (d, 1H, **H$_5$**, J_{5-7}=2.2 Hz); 7.67 (d, 1H, **H$_8$**, J_{8-7}=8.9 Hz); 7.62 (dd, 1H, **H$_7$**, J_{7-8}=8.9 Hz, J_{7-5}=2.3 Hz); 4.03 (d, 3H, OC**H$_3$**$_{-C2}$, $J_{OCH3-OCH3}$=1.2 Hz); 4.01 (d, 3H, OC**H$_3$**$_{-C3}$, $J_{OCH3-OCH3}$=1.3 Hz); 2.64 (dq, 1H, C**H**(CH$_3$)$_2$, $J_{CH-(CH3)2}$=6.7 Hz); 1.14 (d, 6H, CH(C**H$_3$**)$_2$, $J_{(CH3)2-CH}$=6.8 Hz) ppm.

CHN	Calculated	Found
C	60.60%	60.60%
H	6.22%	6.07%
N	15.15%	15.12%

MP: 201-203ºC

HPLC: 98.374% / t_R=2.692'

N-(2,3-dimethoxyquinoxalin-6-yl)benzamide [Q2-3.V]
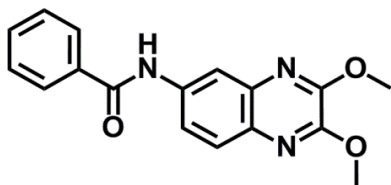
Appearance: Beige solid
*MF: $C_{17}H_{15}N_3O_3 \cdot {}^1/_8 H_2O$*
MW: 311.25 g/mol
Yield: 42%

IR (KBr cm^{-1}): 3285 (s, NH); 3061 (w, aromatic C-H); 2984 and 2939 (w, aliphatic C-H); 1651 (vs, C=O).

^{1}H RMN (DMSO-d_6, 400 MHz) δ: 10.49 (s, 1H, N**H**); 8.37 (d, 1H, **H$_5$**, J_{5-7}=2.2 Hz); 8.00-7.98 (m, 2H, **H$_{2'}$+H$_{6'}$**); 7.87 (dd, 1H, **H$_7$**, J_{7-8}=8.9 Hz, J_{7-5}=2.4 Hz); 7.73 (d, 1H, **H$_8$**, J_{8-7}=8.8 Hz); 7.65-7.53 (m, 3H, **H$_{3'}$+H$_{4'}$+H$_{5'}$**); 4.05 (s, 3H, OC**H$_3$**-C2, $J_{OCH3-OCH3}$=0.5 Hz); 4.03 (s, 3H, OC**H$_3$**-C3, $J_{OCH3-OCH3}$=0.5 Hz) ppm.

CHN	Calculated	Found
C	65.54%	65.52%
H	4.90%	4.99%
N	13.49%	13.40%

MP: 165-170ºC
HPLC: 99.721% / t_R=3.518'

1-(2,3-dimethoxyquinoxalin-6-yl)-3-ethylurea [Q2-3.VI]
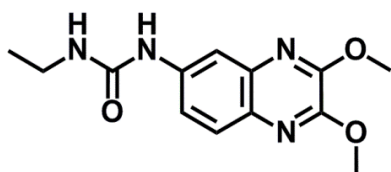
Appearance: White solid
*MF: $C_{13}H_{16}N_4O_3 \cdot {}^1/_4 H_2O$*
MW: 280.5 g/mol
Yield: 24%

IR (KBr cm^{-1}): 3303 (s, NH); 3103 (w, aromatic C-H); 2981 (w, aliphatic C-H); 1633 (vs, C=O).

^{1}H RMN (DMSO-d_6, 400 MHz) δ: 8.69 (s, 1H, CON**H**Ar); 7.97 (d, 1H, **H$_5$**, J_{5-7}=2.2 Hz); 7.60 (d, 1H, **H$_8$**, J_{8-7}=8.8 Hz); 7.41 (dd, 1H, **H$_7$**, J_{7-8}=8.8 Hz, J_{7-5}=2.5 Hz); 6.20 (t, 1H, CON**H**CH$_2$, J_{NH-CH2}=5.4 Hz); 4.02 (d, 3H, OC**H$_3$**-C2, $J_{OCH3-OCH3}$=0.9 Hz); 3.99 (s, 3H, OC**H$_3$**-C3, $J_{OCH3-OCH3}$=0.9 Hz); 3.14 (dq, 2H, C**H$_2$**CH$_3$, J_{CH2-NH}=5.9 Hz, $J_{CH2-CH3}$=7.2 Hz); 1.08 (t, 3H, CH$_2$C**H$_3$**, $J_{CH3-CH2}$=7.3 Hz) ppm.

CHN	Calculated	Found
C	55.61%	55.57%
H	5.88%	5.66%
N	19.96%	19.94%

MP: >300ºC
HPLC: 99.797% / t_R=2.117'

1-(2,3-dimethoxyquinoxalin-6-yl)-3-propylurea [Q2-3.VII]
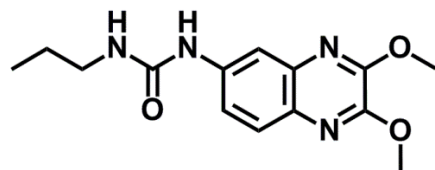
Appearance: White solid
MF: $C_{14}H_{18}N_4O_3$
MW: 290 g/mol
Yield: 29%

IR (KBr cm^{-1}): 3289 (m, NH); 2949 (w, aliphatic C-H); 1656 (vs, C=O).

^{1}H RMN (DMSO-d_6, 400 MHz) δ: 8.67 (s, 1H, CON**H**Ar); 7.96 (dd, 1H, **H$_5$**, J_{5-7}=2.2 Hz, J_{5-8}=0.6 Hz); 7.60 (dd, 1H, **H$_8$**, J_{8-7}=8.8 Hz, J_{8-5}=0.6 Hz); 7.40 (ddd, 1H, **H$_7$**, J_{7-8}=8.9 Hz, J_{7-5}=2.4 Hz, J_{H7-NH}=0.8 Hz); 6.24 (t, 1H, CON**H**CH$_2$, J_{NH-CH2}=5.8 Hz); 4.02 (s, 3H, OC**H$_3$**-C2, $J_{OCH3-OCH3}$=0.9 Hz); 3.99 (s, 3H, OC**H$_3$**-C3, $J_{OCH3-OCH3}$=0.9 Hz); 3.08 (dd, 2H, NHC**H$_2$**CH$_2$CH$_3$, $J_{CH2-CH2}$=6.4 Hz, J_{CH2-NH}=6.3 Hz); 1.47 (tq, 2H, NHCH$_2$C**H$_2$**CH$_3$, $J_{CH2-CH3}$=7.3 Hz, $J_{CH2-CH2}$=7.2 Hz, J_{gem}=14.4 Hz); 0.89 (t, 3H, C**H$_3$**, $J_{CH3-CH2}$=7.4 Hz) ppm.

CHN	Calculated	Found
C	57.93%	57.78%
H	6.21%	6.11%
N	19.31%	19.31%

MP: 122-123ºC
HPLC: 97.767% / t_R=2.549'

1-(2,3-dimethoxyquinoxalin-6-yl)-3-isopropylurea [Q2-3.VIII]
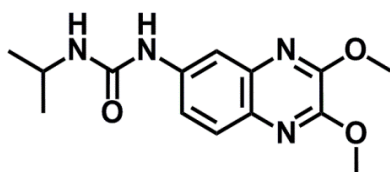
Appearance: White solid
MF: $C_{14}H_{18}N_4O_3 \cdot {}^1/_4 H_2O$
MW: 294.5 g/mol
Yield: 10%

IR (KBr cm^{-1}): 3334 (s, NH); 3297 (s, NH); 3097 (w, aromatic C-H); 2962 (w, aliphatic C-H); 1637 (vs, C=O).

^{1}H RMN (DMSO-d_6, 400 MHz) δ: 8.55 (s, 1H, CON**H**Ar); 7.95 (dd, 1H, **H$_5$**, J_{5-7}=2.3 Hz, J_{5-8}=1.2 Hz); 7.60 (dd, 1H, **H$_8$**, J_{8-7}=8.9 Hz, J_{8-5}=1.0 Hz); 7.39 (ddd, 1H, **H$_7$**, J_{7-8}=8.9 Hz, J_{7-5}=2.4 Hz, J_{H7-NH}=1.1 Hz); 6.10 (d, 1H, N**H**CH(CH$_3$)$_2$, J_{NH-CH}=7.7 Hz); 4.02 (s, 3H, OC**H$_3$**-C2, $J_{OCH3-OCH3}$=1.3 Hz); 3.99 (s, 3H, OC**H$_3$**-C3, $J_{OCH3-OCH3}$=1.3 Hz); 3.84-3.75 (m, 1H, C**H**); 1.12 (dd, 6H, CH(C**H$_3$**)$_2$, $J_{(CH3)2-CH}$=6.5 Hz, $J_{(NH-(CH3)2)}$=1.2 Hz) ppm.

CHN	Calculated	Found
C	57.05%	57.12%
H	6.28%	6.24%
N	19.02%	18.88%

MP: 250ºC
HPLC: 98.752% / t_R=2.506'

1-(2,3-dimethoxyquinoxalin-6-yl)-3-benzamide [Q2-3.IX]
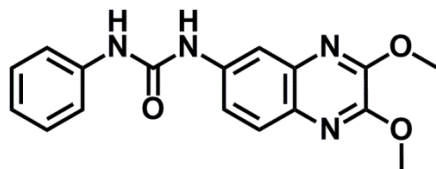
Appearance: Beige solid
MF: $C_{17}H_{16}N_4O_3$
MW: 324 g/mol
Yield: 55%

IR (KBr cm^{-1}): 3275 (s, NH); 3068 and 3030 (w, aromatic C-H); 2997 and 2939 (w, aliphatic C-H); 1639 (vs, C=O).

^{1}H RMN (DMSO-d_6, 400 MHz) δ: 8.94 (s, 1H, N**H**); 8.77 (s, 1H, N**H**); 8.02 (d, 1H, **H$_5$**, J_{5-7}=2.1 Hz); 7.66 (d, 1H, **H$_8$**, J_{8-7}=8.8 Hz); 7.54-7.43 (m, 3H, **H$_7$**+**H$_{2'}$**+**H$_{6'}$**); 7.30 (dd, 2H, **H$_{3'}$**+**H$_{5'}$**, $J_{3'-2'}$=$J_{5'-6'}$=8.1 Hz, $J_{3'-4'}$=$J_{5'-4'}$=7.7 Hz); 6.99 (t, 1H, **H$_{4'}$**, $J_{4'-3'}$=$J_{4'-5'}$=7.4 Hz); 4.03 (s, 3H, OC**H$_3$**$_{-C2}$); 4.00 (s, 3H, OC**H$_3$**$_{-C3}$) ppm.

CHN	Calculated	Found
C	62.96%	63.30%
H	4.94%	5.23%
N	17.28%	17.56%

MP: 260ºC

HPLC: 99.052% / t$_R$=3.796'

17. COMPOUND CHARACTERIZATION - SERIES Q3

17.1. Series Q3 - Intermediates

2-amino-3-cyano-1,4-di-N-oxido-6-methoxyquinoxaline [Q3-1]
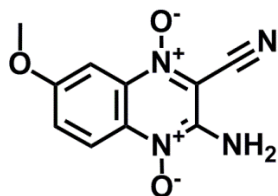
Appearance: Orange solid
MF: $C_{10}H_8N_4O_3$
MW: 232 g/mol
Yield: 45%

IR (KBr cm^{-1}): 3336 and 3267 (m, NH$_2$); 3107 (w, aromatic C-H); 2210 (m, C≡N); 1342 (s, N$^+$-O$^-$).

^{1}H RMN (DMSO-d_6, 400 MHz) δ: 8.22 (d, 1H, **H$_8$**, $J_{8\text{-}7}$=9.2 Hz); 7.85 (s, 2H, N**H$_2$**); 7.60-7.55 (m, 1H, **H$_7$**); 7.59 (s, 1H, **H$_5$**); 3.93 (s, 3H, OC**H$_3$**) ppm.

CHN	Calculated	Found
C	51.72%	N.D.
H	3.45%	N.D.
N	24.14%	N.D.

NOTE! This compound has been previously described by Monge et al.[138] Although theoretically, when synthesizing compound Q3-1 a mixture of two positional isomers (6/7-methoxyquinoxaline-2-carbonitrile) could be formed, Zarranz et al. stated that only one, 7-methoxyquinoxaline-2-carbonitrile, is obtained.[160]

2-cyano-1,4-di-N-oxido-7-methoxyquinoxaline [Q3-2]
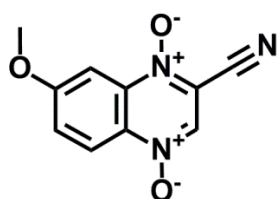
Appearance: Yellow solid
MF: $C_{10}H_7N_3O_3 \cdot \frac{1}{8} H_2O$
MW: 219.25 g/mol
Yield: 31%

IR (KBr cm^{-1}): 3090 (w, aromatic C-H); 2975 (w, aliphatic C-H); 2238 (w, C≡N); 1367 (s, N$^+$O$^-$).

^{1}H RMN (DMSO-d_6, 400 MHz) δ: 9.19 (s, 1H, **H$_3$**); 8.38 (d, 1H, **H$_5$**, $J_{5\text{-}6}$=9.4 Hz); 7.74 (d, 1H, **H$_8$**, $J_{8\text{-}6}$=2.7 Hz); 7.69 (dd, 1H, **H$_6$**, $J_{6\text{-}5}$=9.4 Hz, $J_{6\text{-}8}$=2.8 Hz); 4.02 (s, 3H, OC**H$_3$**) ppm.

CHN	Calculated	Found
C	54.73%	54.86%
H	3.31%	3.30%
N	19.16%	18.85%

7-methoxyquinoxaline-2-carbonitrile [Q3-3]
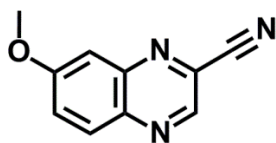
Appearance: White solid
MF: $C_{10}H_7N_3O \cdot \frac{1}{8} H_2O$
MW: 187.25 g/mol
Yield: 39%

IR (KBr cm^{-1}): 3090 and 3039 (w, aromatic C-H); 2975 (w, aliphatic C-H); 2226 (w, C≡N).

^{1}H RMN (DMSO-d_6, 400 MHz) δ: 9.22 (s, 1H, **H$_3$**); 8.13 (d, 1H, **H$_5$**, J_{5-6}=9.3 Hz); 7.73 (dd, 1H, **H$_6$**, J_{6-5}=9.3 Hz, J_{6-8}=2.8 Hz); 7.58 (d, 1H, **H$_8$**, J_{8-6}=2.8 Hz); 4.00 (s, 3H, OC**H$_3$**) ppm.

CHN	Calculated	Found
C	64.09%	64.36%
H	3.87%	3.73%
N	22.43%	22.29%

17.2. Series Q3 - Final compounds

N-[(7-methoxyquinoxalin-2-yl)methyl]acetamide [Q3-4.I]
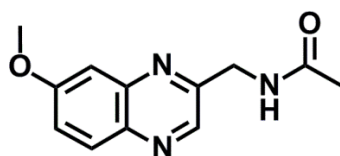
Appearance: Beige solid
*MF: $C_{12}H_{13}N_3O_2 \cdot \frac{1}{8} H_2O$*
MW: 233.25 g/mol
Yield: 6%

IR (KBr cm^{-1}): 3289 (s, NH); 3068 (w, aromatic C-H); 2965 and 2920 (w, aliphatic C-H); 1648 (vs, C=O).

^{1}H RMN (DMSO-d_6, 400 MHz) δ: 8.71 (s, 1H, **H$_3$**); 8.62 (t, 1H, N**H**, J_{NH-CH2}=5.3 Hz); 7.97 (d, 1H, **H$_5$**, J_{5-6}=9.1 Hz); 7.62 (dd, 1H, **H$_6$**, J_{6-5}=9.1 Hz, J_{6-8}=2.8 Hz); 7.40 (d, 1H, **H$_8$**, J_{8-6}=2.6 Hz); 4.56 (d, 2H, C**H$_2$**NH, J_{CH2-NH}=5.8 Hz); 3.95 (s, 3H, OC**H$_3$**); 1.93 (s, 3H, C**H$_3$**) ppm.

CHN	Calculated	Found
C	61.74%	61.99%
H	5.68%	5.69%
N	18.01%	17.67%

MP: 123.9-126.9ºC
HPLC: 97.333% / t_R=2.589'

N-[(7-methoxyquinoxalin-2-yl)methyl]propionamide [Q3-4.II]
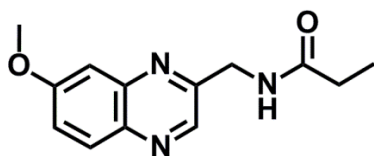
Appearance: Beige solid
*MF: $C_{13}H_{15}N_3O_2 \cdot \frac{1}{4} H_2O$*
MW: 249.5 g/mol
Yield: 24%

IR (KBr cm^{-1}): 3280 (s, NH); 3068 (w, aromatic C-H); 2971 and 2913 (w, aliphatic C-H); 1643 (vs, C=O).

^{1}H RMN (DMSO-d_6, 400 MHz) δ: 8.70 (s, 1H, **H$_3$**); 8.54 (t, 1H, N**H**, J_{CH2-NH}=5.6 Hz); 7.97 (d, 1H, **H$_5$**, J_{5-6}=9.1 Hz); 7.46 (dd, 1H, **H$_6$**, J_{6-5}=9.1 Hz, J_{6-8}=2.8 Hz); 7.40 (d, 1H, **H$_8$**, J_{8-6}=2.8 Hz); 4.56 (d, 2H, C**H$_2$**NH, J=5.8 Hz); 3.95 (s, 1H, OC**H$_3$**); 2.21 (q, 2H, C**H$_2$**CH$_3$, $J_{CH2-CH3}$=7.6 Hz); 1.04 (t, 3H, C**H$_3$**, $J_{CH3-CH2}$=7.6 Hz) ppm.

CHN	Calculated	Found
C	62.53%	62.92%
H	6.21%	6.23%
N	16.83%	16.85%

MP: 143-145ºC.
HPLC: 97.814% / t$_R$=3.267'

N-[(7-methoxyquinoxalin-2-yl)methyl]butyramide [Q3-4.III]
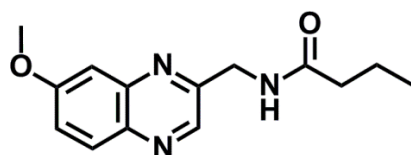
Appearance: Greyish solid
MF: $C_{14}H_{17}N_3O_2 \cdot \frac{1}{2} H_2O$
MW: 268 g/mol
Yield: 24%

IR (KBr cm^{-1}): 3282 (s, NH); 3081 (w, aromatic C-H); 2965 (w, aliphatic C-H); 1643 (vs, C=O).

^{1}H RMN (DMSO-d_6, 400 MHz) δ: 8.69 (s, 1H, **H$_3$**); 8.58 (s, 1H, N**H**); 7.96 (d, 1H, **H$_5$**, J_{5-6}=8.7 Hz); 7.45 (d, 1H, **H$_6$**, J_{6-5}=8.4 Hz); 7.38 (s, 1H, **H$_8$**); 4.56 (d, 2H, C**H$_2$**NH, J_{CH2-NH}=4.8 Hz); 3.94 (s, 3H, OC**H$_3$**); 2.17 (t, 2H, C**H$_2$**CH$_2$CH$_3$, $J_{CH2-CH2}$=6.5 Hz); 1.56 (m, 2H, CH$_2$C**H$_2$**CH$_3$); 0.88 (t, 3H, C**H$_3$**, $J_{CH3-CH2}$=6.6 Hz) ppm.

CHN	Calculated	Found
C	62.69%	62.49%
H	6.72%	6.55%
N	15.67%	15.53%

MP: 128-129ºC.
HPLC: 98.927% / t$_R$=4.589'

NOTE! Although theoretically, when synthesizing compound Q3-1 a mixture of two positional isomers (6/7-methoxyquinoxaline-2-carbonitrile) could be formed, Zarranz et al. stated that only one, 7-methoxyquinoxaline-2-carbonitrile, is isolated.[160] Different experiments were carried out, ^{13}C RMN, HMQC and HMBC, in order to confirm that the obtained compound was the one with methoxy group in position 7 (see Ch3 -21.3).

N-[(3-methoxyquinoxalin-2-yl)methyl]isobutyramide [Q3-4.IV]
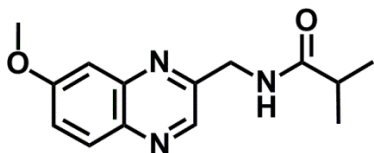
Appearance: White solid

MF: $C_{14}H_{17}N_3O_2$

MW: 259 g/mol

Yield: 62%

IR (KBr cm^{-1}): 3280 (s, NH); 3074 (w, aromatic C-H); 2971 and 2926 (w, aliphatic C-H); 1648 (vs, C=O).

^{1}H RMN (DMSO-d_6, 400 MHz) δ: 8.67 (s, 1H, **H$_3$**); 8.53 (t, 1H, N**H**, J_{NH-CH2}=5.3 Hz); 7.97 (d, 1H, **H$_5$**, J_{5-6}=9.2 Hz); 7.46 (dd, 1H, **H$_6$**, J_{6-5}=9.2 Hz, J_{6-8}=2.8 Hz); 7.39 (d, 1H, **H$_8$**, J_{8-6}=2.7 Hz); 4.55 (d, 2H, C**H$_2$**NH, J_{CH2-NH}=5.8 Hz); 3.95 (s, 3H, OC**H$_3$**); 2.50 (m, 1H, C**H**); 1.07 (d, 6H, C**H$_3$**, $J_{(CH3)2-CH}$=6.9 Hz) ppm.

CHN	Calculated	Found
C	64.86%	64.76 %
H	6.56%	6.67%
N	16.22%	16.27%

MP: 133-133.5ºC

HPLC: 99.581% / t$_R$=4.396'

N-[(7-methoxyquinoxalin-2-yl)methyl]benzamide [Q3-4.V]
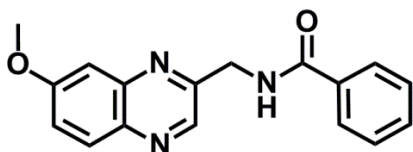
Appearance: Beige solid

MF: $C_{17}H_{15}N_3O_2 \cdot {}^1/_4 H_2O$

MW: 297.5 g/mol

Yield: 29%

IR (KBr cm^{-1}): 3299 (s, NH); 3061 and 3010 (w, aromatic C-Hs); 2965 and 2920 (w, aliphatic C-H); 1640 (vs, C=O).

^{1}H RMN (DMSO-d_6, 400 MHz) δ: 9.28 (t, 1H, N**H**, J_{NH-CH2}=5.4 Hz); 8.79 (s, 1H, **H$_3$**); 7.98 (d, 1H, **H$_5$**, J_{5-6}=9.1 Hz); 7.94 (dd, 2H, **H$_{2'}$**+**H$_{6'}$**, $J_{2'-3'}$=$J_{6'-5'}$=7.1 Hz, $J_{2'-4'}$=$J_{6'-4'}$=1.2 Hz); 7.59-7.44 (m, 4H, **H$_6$**+**H$_{3'}$**+**H$_{4'}$**+**H$_{5'}$**); 7.41 (d, 1H, **H$_8$**, J_{8-6}=2.7 Hz); 4.79 (d, 2H, C**H$_2$**NH, J_{CH2-NH}=5.7 Hz); 3.94 (s, 3H, OC**H$_3$**) ppm.

CHN	Calculated	Found
C	68.57%	68.66%
H	5.21%	5.35%
N	14.12%	13.95%

MP: 142-143.9ºC

HPLC: 95.653% / t$_R$=7.780'

18. COMPOUND CHARACTERIZATION - SERIES In1

18.1. Series In1 – Intermediates

2-(6-methoxyindolyl)acetonitrile [In1-1]
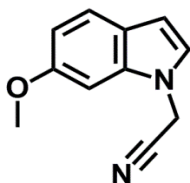
Appearance: Brown oil
MF: $C_{11}H_{10}N_2O$. $^1/_3$ H_2O
MW: 192 g/mol
Yield: 45%

IR (KBr cm⁻¹): 3106 and 3074 (w, aromatic C-H); 2997 and 2965 (w, aliphatic C-H); 2245 (vw, C≡N).

¹H RMN (DMSO-*d₆*, 400 MHz) δ: 7.46 (d, 1H, **H₄**, J_{4-5}=8.6 Hz); 7.28 (d, 1H, **H₂**, J_{2-3}=3.2 Hz); 7.17 (s, 1H, **H₇**); 6.77 (dd, 1H, **H₅**, J_{5-4}=8.6 Hz, J_{5-7}=2.1 Hz); 6.46 (d, 1H, **H₃**, J_{3-2}=3.2 Hz); 5.48 (s, 2H, C**H₂**); 3.82 (s, 3H, OC**H₃**) ppm.

CHN	Calculated	Found
C	68.75%	68.75%
H	5.56%	5.40%
N	14.58%	14.76%

MP: 101.9-102.5ºC

2-(6-methoxyindolyl)ethanamine [In1-2]
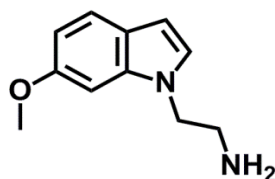
Appearance: Brown-yellow oil
MF: $C_{11}H_{14}N_2O$
MW: 190 g/mol

IR (KBr cm⁻¹): N.D.

¹H RMN (DMSO-*d₆*, 400 MHz) δ: 7.39 (d, 1H, **H₄**, J_{4-5}=8.6 Hz); 7.20 (d, 1H, **H₂**, J_{2-3}=3.1 Hz); 7.02 (d, 1H, **H₇**, J_{7-5}=2.3 Hz); 6.66 (dd, 1H, **H₅**, J_{5-4}=8.6 Hz, J_{5-7}=2.3 Hz); 6.32 (dd, 1H, **H₃**, J_{3-2}=3.2 Hz, J_{3-7}=0.6 Hz); 4.08 (t, 2H, ArC**H₂**CH₂NH, $J_{CH2-CH2}$=6.6 Hz); 3.78 (s, 3H, OC**H₃**); 3.30 (bs, 2H, N**H₂**); 2.87 (t, 2H, ArCH₂C**H₂**NH₂, $J_{CH2-CH2}$=6.6 Hz) ppm.

CHN	Calculated	Found
C	69.47%	N.D.
H	7.37%	N.D.
N	14.74%	N.D.

N-(2-(6-methoxyindolyl)ethyl) -1-imidazolecarboxamide [In1-3.CDI]
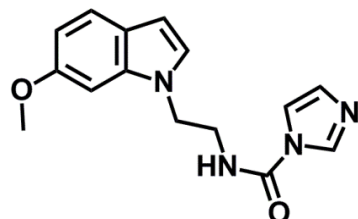
Appearance: Brown oil
MF: $C_{15}H_{16}N_4O_2$
MW: 284 g/mol
Yield: 25%

IR (KBr cm^{-1}): N.D.
^{1}H RMN (DMSO-d_6, 400 MHz) δ: N.D.

The oily product was used without further purification and it was not possible to isolate a fraction clear enough for its characterization

CHN	Calculated	Found
C	63.38%	N.D.
H	5.63%	N.D.
N	19.72%	N.D.

18.2. Series In1 - Final compounds

1-allyl-3-(2-(6-methoxyindolyl)ethyl)urea [In1-3.I]
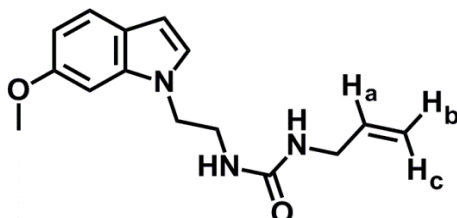
Appearance: White solid
MF: $C_{15}H_{19}N_3O_2$
MW: 273 g/mol
Yield: 8%

IR (KBr cm^{-1}): 3351 and 3317 (s, NH); 3132 (w, aromatic C-H); 2965 and 2926 (w, aliphatic C-H); 1628 (vs, C=O).

^{1}H RMN (DMSO-d_6. 400 MHz) δ: 7.39 (d, 1H, **H$_4$**, J_{4-5}=8.6 Hz); 7.15 (d, 1H, **H$_2$**, J_{2-3}=3.1 Hz); 7.02 (d, 1H, **H$_7$**, J_{7-5}=1.9 Hz); 6.66 (dd, 1H, **H$_5$**, J_{5-4}=8.6 Hz, J_{5-7}=2.2 Hz); 6.33 (d, 1H, **H$_3$**, J_{3-2}=3.1 Hz); 6.07 (t, 1H, N**H**CH$_2$CH=CH$_2$, J_{NH-CH2}=5.9 Hz); 5.97 (t, 1H, Ar(CH$_2$)$_2$N**H**, J_{NH-CH2}=5.6 Hz); 5.79 (tdd, 1H, **Ha**, J_{a-c}=17.1 Hz, J_{a-b}=10.1 Hz, J_{a-CH2}=5.0 Hz); 5.08 (ddd, 1H, **Hc**, J_{c-a}=17.3 Hz, J_{c-CH2}=3.2 Hz, J_{c-b}=1.7 Hz); 5.01 (ddd, 1H, **Hb**, J_{b-a}=10.3 Hz, J_{b-CH2}=3.5 Hz, J_{b-c}=1.8 Hz); 4.14 (t, 2H, ArC**H$_2$**CH$_2$NH, $J_{CH2-CH2}$=6.4 Hz); 3.78 (s, 3H, OC**H$_3$**); 3.64 (t, 2H, C**H$_2$**CH=CH$_2$, J_{CH2-NH}=5.4 Hz); 3.36-3.31 (m, 2H, ArCH$_2$C**H$_2$**NH) ppm.

CHN	Calculated	Found
C	65.93%	65.63%
H	6.96%	6.84%
N	15.38%	15.25%

MP: 110-110.3ºC
HPLC: 99.090% / t_R=3.201'

1-ethyl-3-(2-(6-methoxyindolyl)ethyl)urea [In1-3.II]
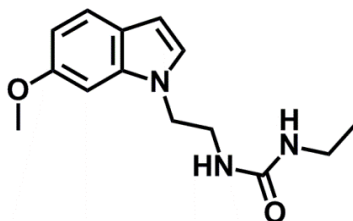
Appearance: White solid

MF: $C_{14}H_{19}N_3O_2$

*MW: 261 g/mol*

Yield: 20%

IR (KBr cm^{-1}): 3335 (s, NH); 3132 (w, aromatic C-H); 2977 and 2939 (w, aliphatic C-H); 1625 (vs, C=O).

^{1}H RMN (DMSO-d_6, 400 MHz) δ: 7,39 (d, 1H, **H$_4$**, J_{4-5}=8,6 Hz); 7,15 (d, 1H, **H$_2$**, J_{2-3}=3,1 Hz); 7,02 (d, 1H, **H$_7$**, J_{7-5}=2,1 Hz); 6,66 (dd, 1H, **H$_5$**, J_{5-4}=8,6 Hz, J_{5-7}=2,3 Hz); 6,33 (dd, 1H, **H$_3$**, J_{3-2}=3,1 Hz, J_{3-7}=0,6 Hz); 5,88 (dd, 2H, N**H**+N**H**, J_{NH-CH2}=5,5 Hz, J_{NH-CH2}=5,5 Hz); 4,13 (t, 2H, ArC**H$_2$**CH$_2$NH, $J_{CH2-CH2}$=6,4 Hz); 3,79 (s, 3H, OC**H$_3$**); 3,34-3,29 (m, 2H, ArCH$_2$C**H$_2$**NH); 3,01 (dq, 2H, C**H$_2$**CH$_3$, $J_{CH2-CH3}$=7,1 Hz, J_{CH2-NH}=5,8 Hz); 0,82 (t, 3H, C**H$_3$**, $J_{CH3-CH2}$=7,2 Hz) ppm.

CHN	Calculated	Found
C	64.37%	63.94%
H	7.28%	7.14%
N	16.09%	15.87%

MP: 148.3-149.3ºC

HPLC: 98.212% / t$_R$ 2.846'

1-(2-(6-methoxyindolyl)ethyl)-3-propylurea [In1-3.III]
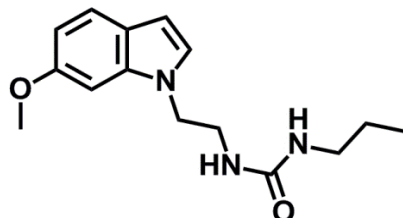
Appearance: White solid

MF: $C_{15}H_{21}N_3O_2 \cdot {}^1/_4\, H_2O$

MW: 279.5 g/mol

*Yield: 36%*

IR (KBr cm^{-1}): 3331 (s, NH); 3132 (w, aromatic C-H); 2965 and 2933 (w, aliphatic C-H); 1623 (vs, C=O).

^{1}H RMN (DMSO-d_6, 400 MHz) δ: 7.39 (d, **H$_4$**, 1H, J_{4-5}=8.6 Hz); 7.15 (d, 1H, **H$_2$**, J_{2-3}=3.1 Hz); 7.02 (d, 1H, **H$_7$**, J_{7-5}=1.9 Hz); 6.66 (dd, 1H, **H$_5$**, J_{5-4}=8.6 Hz, J_{5-7}=2.2 Hz); 6.33 (dd, 1H, **H$_3$**, J_{3-2}=3.1 Hz, J_{3-7}=0.7 Hz); 5.93 (t, 1H, N**H**(CH$_2$)$_2$CH$_3$, J_{NH-CH2}=5.7 Hz); 5.86 (t, 1H, Ar(CH$_2$)$_2$N**H**, J_{NH-CH2}=5.6 Hz); 4.13 (t, 2H, ArC**H$_2$**CH$_2$NH, $J_{CH2-CH2}$=6.3 Hz); 3.79 (s, 3H, OC**H$_3$**); 3.34-3.29 (m, 2H, ArCH$_2$C**H$_2$**NH); 2.94 (m, 2H, C**H$_2$**CH$_2$CH$_3$); 1.35 (m, 2H, CH$_2$C**H$_2$**CH$_3$); 0.82 (t, 3H, C**H$_3$**, $J_{CH3-CH2}$=7.4 Hz) ppm.

CHN	Calculated	Found
C	64.40%	64.17%
H	7.69%	7.82%
N	15.03%	14.99%

MP: 118,7-119ºC

HPLC: 99.037% / t$_R$=3.734'

1-allyl-3-(2-(6-methoxyindolyl)ethyl)thiourea [In1-3.IV]
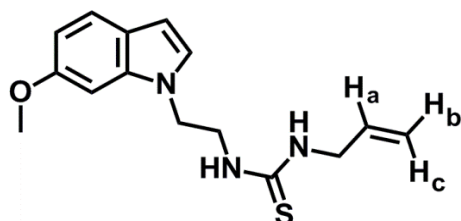
Appearance: Brown oil
MF: $C_{15}H_{19}N_3OS \cdot {}^1/_4 H_2O$
MW: 293.5 g/mol
Yield: 17%

IR (KBr cm^{-1}): 3357 and 3261 (m, NH); 3055 (w, aromatic C-H); 2932 (w, aliphatic C-H).

^{1}H RMN (DMSO-d_6, 400 MHz) δ: 7.66 (bs, 1H, Ar(CH$_2$)$_2$N**H**); 7.48 (bs, 1H, N**H**CH$_2$CH=CH$_2$); 7.39 (dd, 1H, **H$_4$**, J_{4-5}=8.6 Hz, J_{4-7}=1.1 Hz); 7.18-7.14 (m, 2H, **H$_2$**+**H$_7$**); 6.97 (dd, 1H, **H$_5$**, J_{5-4}=8.6 Hz, J_{5-7}=2.1 Hz); 6.35 (d, 1H, **H$_3$**, J_{3-2}=3.7 Hz); 5.81 (bs, 1H, **H**a); 5.11 (d, 1H, **H**c, J_{c-a}=17.3 Hz); 5.05 (d, 1H, **H**b, J_{b-a}=10.3 Hz); 4.30 (t, 2H, C**H**$_2$CH=CH$_2$, $J_{CH2-CH2}$=6.4 Hz); 4.04 (bs, 2H, ArC**H**$_2$CH$_2$NH); 3.78 (s, 3H, OC**H$_3$**); 3.75 (bs, 2H, ArCH$_2$C**H$_2$**NH) ppm.

CHN	Calculated	Found
C	61.33%	61.30%
H	6.64%	6.98%
N	14.31%	14.08%

HPLC: N.D.

1-ethyl-3-(2-(6-methoxyindolyl)ethyl)thiourea [In1-3.V]
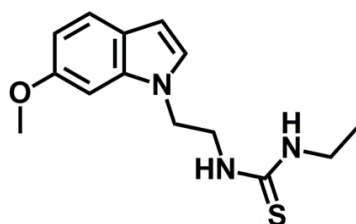
Appearance: Brown oil
MF: $C_{14}H_{19}N_3OS$
MW: 277 g/mol
Yield: 44%

IR (KBr cm^{-1}): 3344 and 3260 (m, NH); 3067 (w, aromatic C-H); 2971 and 2932 (m, aliphatic C-H).

^{1}H RMN (DMSO-d_6, 400 MHz) δ: 7.53 (bs, 1H, Ar(CH$_2$)$_2$N**H**); 7.39 (d, 1H, **H$_4$**, J_{4-5}=8.6 Hz); 7.38 (bs, 1H, N**H**CH$_2$CH$_3$); 7.16 (m, 2H, **H$_2$**+**H$_7$**); 6.66(dd, 1H, **H$_5$**, J_{5-4}=8.6 Hz, J_{5-7}=2.2 Hz); 6.34 (d, 1H, **H$_3$**, J_{3-2}=3.1 Hz); 4.29 (t, 2H, ArC**H$_2$**CH$_2$NH, $J_{CH2-CH2}$=6.3 Hz); 3.78 (s, 3H, OC**H$_3$**); 3.72 (bs, 2H, ArCH$_2$C**H$_2$**NH); 3.33 (bs, 2H, C**H$_2$**CH$_3$); 1.02 (t, 3H, C**H$_3$**, $J_{CH2-CH3}$=6.0 Hz) ppm.

CHN	Calculated	Found
C	60.65%	61.13%
H	6.86%	6.82%
N	15.16%	14.97%

HPLC: N.D.

1-(2-(6-methoxyindolyl)ethyl)-3-propylthiourea [In1-3.VI]
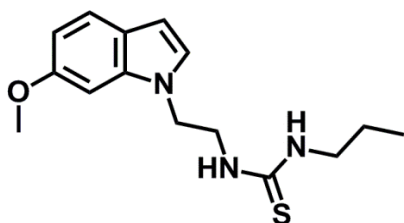
Appearance: Brown oil
MF: $C_{15}H_{21}N_3OS$
MW: 291 g/mol
Yield: 15%

IR (KBr cm⁻¹): 3364 and 3267 (m, NH); 3055 (w, aromatic C-H); 2959 and 2926 (m, aliphatic C-H).

¹H RMN (DMSO-d_6, 400 MHz) δ: 7.53 (bs, 1H, Ar(CH₂)₂N**H**); 7.39 (dd, 1H, **H₄**, J_{4-5}=8.6 Hz, J_{4-7}=1.1 Hz); 7.38 (bs, 1H, N**H**(CH₂)₂CH₃); 7.18-7.14 (m, 2H, **H₂**+**H₇**); 6.66 (td, 1H, **H₅**, J_{5-4}=8.6 Hz, J_{5-7}=1.8 Hz); 6.34 (dd, 1H, **H₃**, J_{3-2}=3.1 Hz, J_{3-7}=0.7 Hz); 4.30 (t, 2H, ArC**H₂**CH₂NH, $J_{CH2-CH2}$=6.2 Hz); 3.78 (s, 3H, OC**H₃**); 3.76-3.70 (m, 2H, ArCH₂C**H₂**NH); 3.10-3.40 (bs, 2H, C**H₂**CH₂CH₃); 1.44 (bs, 2H, CH₂C**H₂**CH₃); 0.80 (t, 3H, CH3, $J_{CH2-CH3}$=7.3 Hz) ppm.

CHN	Calculated	Found
C	61.86%	61.47%
H	7.22%	7.10%
N	14.43%	14.13%

HPLC: 99.037% / t_R=3.734'

1-(2-(6-methoxyindolyl)ethyl)-3-methylthiourea [In1-3.VII]
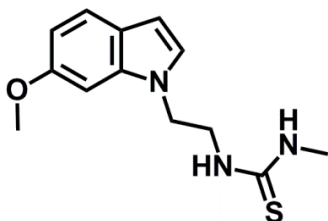
Appearance: Brown oil
MF: $C_{13}H_{17}N_3OS \cdot {}^1/_4 H_2O$
MW: 267.5 g/mol
Yield: 14%

IR (KBr cm⁻¹): 3368 and 3274 (m, NH); 2939 (w, aliphatic C-H).

¹H RMN (DMSO-d_6, 400 MHz) δ: 7,50 (bs, 2H, N**H**+N**H**); 7,39 (d, 1H, **H₄**, J_{4-5}=8,6 Hz); 7,16 (m, 2H, **H₂**+**H₇**); 6,66 (dd, 1H, **H₅**, J_{5-4}=8,6 Hz, J_{5-7}=2,1 Hz); 6.35 (d, 1H, **H₃**, J_{3-2}=3.0 Hz); 4,30 (t, 2H, ArC**H₂**CH₂NH, $J_{CH2-CH2}$=6.4 Hz); 3,78 (s, 3H, OC**H₃**); 3,73 (bs, 2H, ArCH₂C**H₂**NH); 2,81 (t, 3H, C**H₃**) ppm.

CHN	Calculated	Found
C	58.32%	58.40%
H	6.54%	6.62%
N	15.70%	15.86%

HPLC: N.D.

1-cyclopropyl-3-(2-(6-methoxyindolyl)ethyl)thiourea [In1-3.VIII]
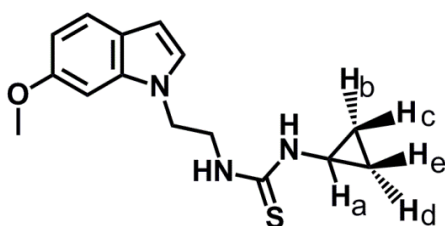
Appearance: Beige solid
MF: $C_{15}H_{19}N_3OS \cdot {}^{1}/_{4} H_2O$
MW: 293.5 g/mol
Yield: 11%

IR (KBr cm⁻¹): 3267 and 3215 (s, NH); 2997 and 2932 (w, aliphatic C-H).

¹H RMN (DMSO-d_6, 400 MHz) δ: 8.06 (bs, 1H, Ar(CH₂)₂N**H**); 7.55 (bs, 1H, N**H**-cC₃H₅); 7.39 (d, 1H, **H₄**, J_{4-5}=8.6 Hz); 7.18-7.14 (m, 2H, **H₂+H₇**); 6.66 (dd, 1H, **H₅**, J_{5-4}=8.5 Hz, J_{5-7}=1.9 Hz); 6.35 (d, 1H, **H₃**, J_{3-2}=2.8 Hz); 4.34 (t, 2H, ArC**H₂**CH₂NH, $J_{CH2-CH2}$=6.3 Hz); 3.78 (s, 3H, OC**H₃**); 3.34-3.30 (m, 2H, ArCH₂C**H₂**NH); 2.33 (bs, 1H, **Ha**); 0.58 (bs, 2H, **Hb+Hd**); 0.35 (bs, 2H, **Hc+He**) ppm.

CHN	Calculated	Found
C	61.33%	61.15%
H	6.64%	6.67%
N	14.31%	14.23%

MP: 116.5ºC
HPLC: 97.845% / t_R=3.861'

N-(2-(6-methoxyindolyl)ethyl)methanesulfonamide [In1-3.IX]
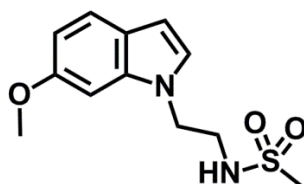
Appearance: Beige solid
MF: $C_{12}H_{16}N_2O_3S \cdot {}^{1}/_{4} H_2O$
MW: 272.5 g/mol
Yield: 6%

IR (KBr cm⁻¹): 3327 (m, NH); 3001 (w, aromatic C-H); 2948 (w, aliphatic C-H).

¹H RMN (DMSO-d_6, 400 MHz) δ: 7.40 (d, 1H, **H₄**, J_{4-5}=8.6 Hz); 7.26 (d, 1H, N**H**, J_{NH-CH2}=5.8 Hz); 7.23 (d, 1H, **H₂**, J_{2-3}=3.2 Hz); 7.05 (s, 1H, **H₇**, J_{7-5}=2.0 Hz); 6.68 (dd, 1H, **H₅**, J_{5-4}=8.6 Hz, J_{5-7}=2.2 Hz); 6.35 (d, 1H, **H₃**, J_{3-2}=3.1 Hz); 4.22 (t, 2H, ArC**H₂**CH₂NH, $J_{CH2-CH2}$=6.4 Hz); 3.80 (s, 3H, OC**H₃**); 3.33-3.27 (m, 2H, CH₂C**H₂**NH); 2.71 (s, 3H, C**H₃**) ppm.

CHN	Calculated	Found
C	52.84%	53.14%
H	6.06%	6.12%
N	10.27%	10.18%

MP: 80.0-80.4ºC
HPLC: N.D.

2,2,2-trifluoro-N-(2-(6-methoxyindolyl)ethyl)acetamide [In1-3.X]
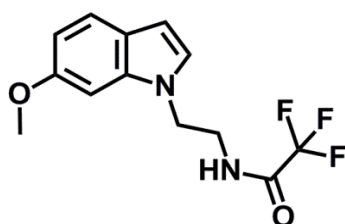
Appearance: White solid
*MF: $C_{13}H_{13}N_2O_2F_3 \cdot {}^1/_4 H_2O$*
MW: 290.5 g/mol
Yield: 6%

IR (KBr cm^{-1}): 3313 (m, NH); 3004 (w, aromatic C-H); 2952 and 2894 (w, aliphatic C-H); 1706 (vs, C=O).

^{1}H RMN (DMSO-d_6, 400 MHz) δ: 9.56 (t, 1H, N**H**, J_{NH-CH2}=5.6 Hz); 7.39 (d, 1H, **H$_4$**, J_{4-5}=8.6 Hz); 7.15 (d, 1H, **H$_2$**, J_{2-3}=3.2 Hz); 7.03 (d, 1H, **H$_7$**, J_{7-5}=2.1 Hz); 6.67 (dd, 1H, **H$_5$**, J_{5-7}=2.3 Hz, J_{5-4}=8.6 Hz); 6.34 (dd, 1H, **H$_3$**, J_{3-7}=0.6 Hz, J_{3-2}=3.2 Hz); 4.29 (t, 2H, ArC**H$_2$**CH$_2$NH, $J_{CH2-CH2}$=6.1 Hz); 3.79 (s, 3H, OC**H$_3$**); 3.55 (dd, 2H, ArCH$_2$C**H$_2$**NH, J_{CH2-NH}=5.8 Hz, $J_{CH2-CH2}$=6.1 Hz) ppm.

CHN	Calculated	Found
C	53.70%	53.96%
H	4.65%	4.35%
N	9.64%	9.64%

MP: 97.8-98ºC
HPLC: 96.765% / t_R=3.788'

1-(2-(6-methoxyindolyl)ethyl)-3-methylurea [In1-4.XI]
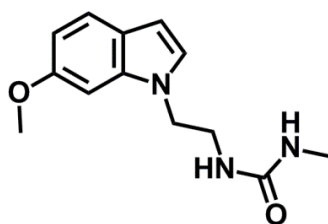
Appearance: White-pink solid
MF: $C_{13}H_{17}N_3O_2 \cdot H_2O$
MW: 265 g/mol
Yield: 23%

IR (KBr cm^{-1}): 3339 (s, NH); 3150 (w, aromatic C-H); 2995 and 2939 (w, aliphatic C-H); 1631 (vs, C=O).

^{1}H RMN (DMSO-d_6, 400 MHz) δ: 7.39 (d, 1H, **H$_4$**, J_{4-5}=8.6 Hz); 7.15 (d, 1H, **H$_2$**, J_{2-3}=3.1 Hz); 7.03 (d, 1H, **H$_7$**, J_{7-5}=2.2 Hz); 6.66 (dd, 1H, **H$_5$**, J_{5-4}=8.6 Hz, J_{5-7}=2.3 Hz); 6.33 (d, 1H, **H$_3$**, J_{3-2}=3.1 Hz); 5.97 (t, 1H, ArCH$_2$CH$_2$N**H**, J_{NH-CH2}=5.8 Hz); 5.81 (q, 1H, CON**H**CH$_3$, J_{NH-CH3}=5.3 Hz); 4.13 (t, 2H, ArC**H$_2$**CH$_2$NH, $J_{CH2-CH2}$=6.4 Hz); 3.79 (s, 3H, OC**H$_3$**); 3.33-3.27 (m, 2H, CH$_2$C**H$_2$**NH); 2.54 (d, 3H, C**H$_3$**, J_{CH3-NH}=4.7 Hz) ppm.

CHN	Calculated	Found
C	58.87%	58.93%
H	7.17%	6.71%
N	15.85%	15.61%

MP: 158.2ºC
HPLC: 97.798% / t_R=2.383'

1-cyclopropyl-3-(2-(6-methoxyindolyl)ethyl)urea [In1-4.XII]
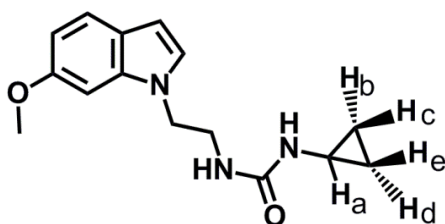
Appearance: White-pink solid

MF: $C_{15}H_{19}N_3O_2 \cdot \frac{1}{4} H_2O$

MW: 277.5 g/mol

Yield: 28%

IR (KBr cm^{-1}): 3345 and 3315 (s, NH); 3127 (w, aromatic C-H); 2995 and 2932 (w, aliphatic C-H); 1631 (vs, C=O).

^{1}H RMN (DMSO-*d$_6$*, 400 MHz) δ: 7.39 (d, 1H, **H$_4$**, J_{4-5}=8.6 Hz); 7.15 (d, 1H, **H$_2$**, J_{2-3}=3.1 Hz); 7.04 (d, 1H, **H$_7$**, J_{7-5}=2.0 Hz); 6.67 (dd, 1H, **H$_5$**, J_{5-4}=8.6 Hz, J_{5-7}=2.2 Hz); 6.33 (d, 1H, **H$_3$**, J_{3-2}=3.1 Hz); 6.24 (bs, 1H, CON**H**CH); 5.94 (t, 1H, ArCH$_2$CH$_2$N**H**, J_{NH-CH2}=4.8 Hz); 4.16 (t, 2H, ArC**H$_2$**CH$_2$NH, $J_{CH2-CH2}$=6.4 Hz); 3.79 (s, 3H, OC**H$_3$**); 3.33-3.27 (m, 2H, ArCH$_2$C**H$_2$**NH); 2.33-2.37-2.30 (m, 1H, **Ha**); 0.53 (qd, 2H, **Hb**+**Hd**, $Jb-c=Jd-e$=4.7 Hz, $Jb-e=Jd-c$=4.7 Hz, $Jb-a=Jd-a$=6.7 Hz, $Jb-d$=6.7 Hz); 0.27 (qd, 2H, **Hc**+**He**, $Jc-b=Je-d$=4.4 Hz, $Jc-a$=4.4 Hz, $Jc-d=Je-b$=4.6 Hz, $Jc-e$=6.9 Hz) ppm.

CHN	Calculated	Found
C	64.86%	64.87%
H	7.03%	7.00%
N	15.14%	15.03%

MP: 133-134ºC

HPLC: 98.756% / t_R=3.095'

1-(2-(6-methoxyindolyl)ethyl)urea [In1-3.XIII]
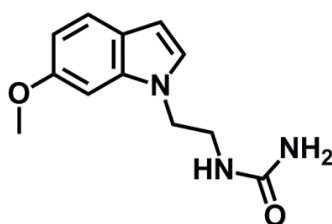
Appearance: White solid

MF: $C_{12}H_{15}N_3O_2 \cdot \frac{1}{2} H_2O$

MW: 242 g/mol

Yield: 2%

IR (KBr cm^{-1}): 3457, 3401 and 3359 (s, N-H); 3193 and 3098 (w, aromatic C-H); 2934 (w, aliphatic C-H); 1654 (vs, C=O).

^{1}H RMN (DMSO-*d$_6$*, 400 MHz) δ: 7.39 (d, 1H, **H$_4$**, J_{4-5}= 8.8 Hz); 7.16 (d, 1H, **H$_2$**, J_{2-3}=3.3 Hz); 7.06 (s, 1H, **H$_7$**); 6.66 (dd, 1H, **H$_5$**, J_{5-4}= 8.6 Hz, J_{5-7}= 2.2 Hz); 6.33 (d, 1H, **H$_3$**, J_{3-2}= 3.3 Hz); 6.00 (m, 1H, CH$_2$N**H**CO); 5.53 (s, 2H, N**H$_2$**); 4.13 (t, 2H, ArC**H$_2$**CH$_2$NH, $J_{CH2-CH2}$= 6.6 Hz); 3.79 (s, 3H, OC**H$_3$**); 3.29 (m, 2H, CH$_2$C**H$_2$**NHCO) ppm.

CHN	Calculated	Found
C	59.50%	59.20%
H	6.61%	6.64%
N	17.36%	17.05%

MP: 136.6ºC

HPLC: 95.678% / t_R=10.263'

19. COMPOUND CHARACTERIZATION - SERIES In2

19.1. Series In2 - Intermediates

2-(7-methoxyindolyl)acetonitrile [In2-1]
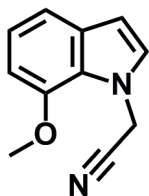
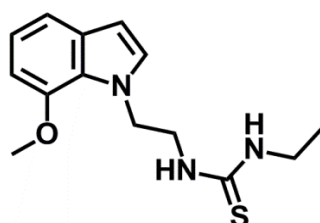*Appearance: Beige solid*
MF: $C_{11}H_{10}N_2O$
MW: 186 g/mol
Yield: 79%

IR (KBr cm^{-1}): 3345 (s, NH); 3101 (m, aromatic C-H); 2967 and 2937 (w, aliphatic C-H).

^{1}H RMN (DMSO-d_6, 400 MHz) δ: 7.33 (d, 1H, **H$_2$**, J_{2-3}=3.1 Hz); 7.16 (d, 1H, **H$_4$**, J_{4-5}=7.9 Hz); 7.01 (t, 1H, **H$_5$**, J_{5-4}=J_{5-6}=7.8 Hz); 6.76 (d, 1H, **H$_6$**, J_{6-5}=7.8 Hz); 6.49 (d, 1H, **H$_3$**, J_{3-2}=3.1 Hz); 5.54 (s, 2H, C**H$_2$**CN); 3.93 (s, 3H, OC**H$_3$**) ppm.

CHN	Calculated	Found
C	70.97%	ND
H	5.38%	ND
N	15.05%	ND

2-(7-methoxyindolyl)ethanamine [In2-2]
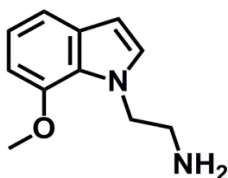
Appearance: Beige solid
MF: $C_{11}H_{14}N_2O$
MW: 190 g/mol
Yield: 67%

IR (KBr cm^{-1}): N.D.
^{1}H RMN (DMSO-d_6, 400 MHz) δ: N.D.

The oily product has been used without further purification and it was not possible to isolate a fraction that was not clear enough for its characterization

CHN	Calculated	Found
C	69.47%	N.D.
H	7.37%	N.D.
N	14.74%	N.D.

N-(2-(7-methoxyindolyl)ethyl)-imidazole-1-carboxamide [In2-3.CDI]
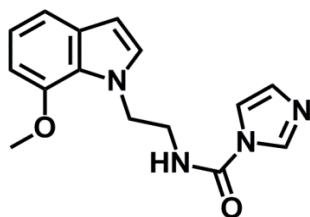
Appearance: White solid

MF: $C_{15}H_{16}N_4O_2$

MW: 284 g/mol

Yield: 8%

IR (KBr cm⁻¹): 3160 (w, NH); 3117 (m, aromatic C-H); 2990 and 2945 (w, aliphatic C-H); 1716 (vs, C=O).

¹H RMN (DMSO-*d_6*, 400 MHz) δ: 8.53 (t, 1H, N**H**, J_{NH-CH2}=5.4 Hz); 8.13 (s, 1H, **H$_{2'}$**); 7.55 (d, 1H, **H$_{5'}$**, $J_{5'-4'}$=1.2 Hz); 7.20 (d, 1H, **H$_2$**, J_{2-3}=2.9 Hz); 7.11 (d, 1H, **H$_4$**, J_{4-5}=7.9 Hz); 7.00 (d, 1H, **H$_{4'}$**, $J_{4'-5'}$=0.7 Hz); 6.90 (t, 1H, **H$_5$**, J_{5-4}=J_{5-6}=8.0 Hz); 6.62 (d, 1H, **H$_6$**, J_{6-5}=7.7 Hz); 6.37 (d, 1H, **H$_3$**, J_{3-2}=3.0 Hz); 4.54 (t, 2H, ArC**H$_2$**CH$_2$NH, $J_{CH2-CH2}$=5.7 Hz); 3.78 (s, 3H, OC**H$_3$**); 3.59 (dd, 2H, ArCH$_2$C**H$_2$**NH, $J_{CH2-CH2}$=6.0 Hz, J_{CH2-NH}=5.5 Hz) ppm.

CHN	Calculated	Found
C	63.38%	N.D.
H	5.63%	N.D.
N	19.72%	N.D.

19.2. Series In2 - Final compounds

1-allyl-3-(2-(7-methoxyindolyl)ethyl)urea [In2-3.I]
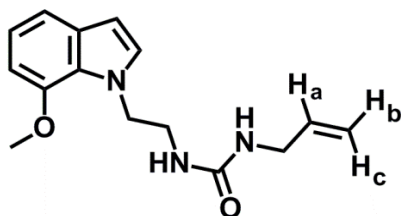
Appearance: Pale pink solid

MF: $C_{15}H_{19}N_3O_2$

MW: 273 g/mol

Yield: 28%

IR (KBr cm⁻¹): 3337 (m, NH); 3042 (w, aromatic C-H); 2961 and 2907 (m, aliphatic C-H); 1625 (vs, C=O).

¹H RMN (DMSO-*d_6*, 400 MHz) δ: 7.13 (d, 1H, **H$_2$**, J_{2-3}=2.8 Hz); 7.11 (dd, 1H, **H$_4$**, J_{4-5}=7.9 Hz, J_{4-6}=0.6 Hz); 6.91 (dt, 1H, **H$_5$**, J_{5-4}=7.9 Hz, J_{5-6}=7.8 Hz, J_{5-3}=2.2 Hz); 6.65 (d, 1H, **H$_6$**, J_{6-5}=7.7 Hz); 6.36 (d, 1H, **H$_3$**, J_{3-2}=2.6 Hz); 5.99 (t, 1H, ArCH$_2$CH$_2$N**H**, J_{NH-CH2}=5.7 Hz); 5.91 (t, 1H, N**H**CH$_2$CH=CH$_2$, J_{NH-CH2}=5.6 Hz); 5.80 (dddd, 1H, **Ha**, J_{a-c}=17.1 Hz, J_{a-b}=10.2 Hz, J_{a-CH2}=5.1 Hz, J_{a-NH}=3.5 Hz); 5.09 (dd, 1H, **Hc**, J_{c-a}=17.2 Hz, J_{c-b}=1.8 Hz); 5.01 (d, 1H, **Hb**, J_{b-a}=10.3 Hz, J_{b-c}=1.6 Hz); 4.37 (t, 2H, ArC**H$_2$**CH$_2$NH, $J_{CH2-CH2}$=6.1 Hz); 3.88 (s, 3H, OC**H$_3$**); 3.63 (dt, 2H, NHC**H$_2$**CH=CH$_2$, J_{CH2-CH}=7.2 Hz, J_{CH2-NH}=3.5 Hz); 3.34 (m, 2H, ArCH$_2$C**H$_2$**NH) ppm.

CHN	Calculated	Found
C	65.93%	65.75%
H	6.96%	6.91%
N	15.38%	15.20%

MP: 104-107.7ºC

HPLC: 98.669% / t_R=4.418'

1-ethyl-3-(2-(7-methoxyindolyl)ethyl)urea [In2-3.II]
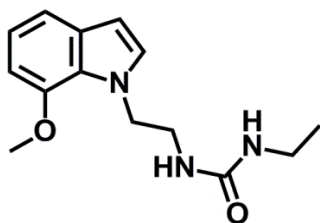
Appearance: White solid
MF: $C_{14}H_{19}N_3O_2$
MW: 261 g/mol
Yield: 67%

IR (KBr cm^{-1}): 3345 (s, NH); 3045 (w, aromatic C-H); 2965 and 2904 (m, aliphatic C-H); 1626 (vs, C=O).

^{1}H RMN (DMSO-d_6, 400 MHz) δ: 7.13 (d, 1H, **H$_2$**, $J_{2\text{-}3}$=3.1 Hz); 7.10 (dd, 1H, **H$_4$**, $J_{4\text{-}5}$=7.9 Hz, $J_{4\text{-}6}$=0.6 Hz); 6.90 (t, 1H, **H$_5$**, $J_{5\text{-}4}$=$J_{5\text{-}6}$=7.8 Hz); 6.65 (d, 1H, **H$_6$**, $J_{6\text{-}5}$=7.6 Hz); 6.35 (d, 1H, **H$_3$**, $J_{3\text{-}2}$=3.0 Hz); 5.99 (t, 1H, ArCH$_2$CH$_2$N**H**, $J_{NH\text{-}CH2}$=5.7 Hz); 5.80 (t, 1H, CON**H**CH$_2$CH$_3$, $J_{NH\text{-}CH2}$=5.6 Hz); 4.36 (t, 2H, ArC**H$_2$**CH$_2$NH, $J_{CH2\text{-}CH2}$=6.2 Hz); 3.89 (s, 3H, OC**H$_3$**); 3.34 (m, 2H, NHC**H$_2$**CH$_3$); 2.99 (dq, 2H, ArCH$_2$C**H$_2$**NH, $J_{CH2\text{-}CH3}$=7.2 Hz, $J_{CH2\text{-}NH}$=5.6 Hz); 0.97 (t, 3H, C**H$_3$**, $J_{CH2\text{-}CH3}$=7.2 Hz) ppm.

CHN	Calculated	Found
C	64.37%	64.04%
H	7.28%	7.14%
N	16.09%	15.88%

MP: 103.2-104.1ºC
HPLC: 99.894% / t$_R$=3.837' /

1-(2-(7-methoxyindolyl)ethyl)-3-propylurea [In2-3.III]
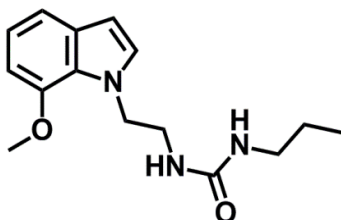
Appearance: White solid
MF: $C_{15}H_{21}N_3O_2$
MW: 275 g/mol
Yield: 28%

IR (KBr cm^{-1}): 3330 (m, NH); 2959 and 2906 (w, aliphatic C-H); 1625 (vs, C=O).

^{1}H RMN (DMSO-d_6, 400 MHz) δ: 7.13 (d, 1H, **H$_2$**, $J_{2\text{-}3}$=3.1 Hz); 7.10 (d, 1H, **H$_4$**, $J_{4\text{-}5}$=8.0 Hz); 6.90 (t, 1H, **H$_5$**, $J_{5\text{-}4}$=$J_{5\text{-}6}$=7.8 Hz); 6.65 (d, 1H, **H$_6$**, $J_{6\text{-}5}$=7.7 Hz); 6.35 (d, 1H, **H$_3$**, $J_{3\text{-}2}$=3.0 Hz); 5.85 (t, 1H, ArCH$_2$CH$_2$N**H**, $J_{NH\text{-}CH2}$=5.6 Hz); 5.79 (t, 1H, CON**H**CH$_2$CH$_3$, $J_{NH\text{-}CH2}$=5.6 Hz); 4.36 (t, 2H, ArC**H$_2$**CH$_2$NH, $J_{CH2\text{-}CH2}$=6.1 Hz); 3.88 (s, 3H, OC**H$_3$**); 3.60-3.29 (m, 2H, ArCH$_2$C**H$_2$**NH); 2.92 (td, 2H, NHC**H$_2$**CH$_2$CH$_3$, $J_{CH2\text{-}CH2}$=6.4 Hz, $J_{CH2\text{-}NH}$=6.4 Hz); 1.35 (ttd, 2H, NHCH$_2$C**H$_2$**CH$_3$, J_{gem}=11.5 Hz, $J_{CH2\text{-}CH3}$=7.2 Hz, $J_{CH2\text{-}CH2}$=5.7 Hz); 0.82 (t, 3H, C**H$_3$**, $J_{CH3\text{-}CH2}$=7.4 Hz) ppm.

CHN	Calculated	Found
C	65.45%	65.51%
H	7.64%	7.94%
N	15.27%	15.19%

MP: 103.2-104.1ºC
HPLC: 99.596% / t$_R$=5.279'

1-allyl-3-(2-(7-methoxyindolyl)ethyl)thiourea [In2-3.IV]
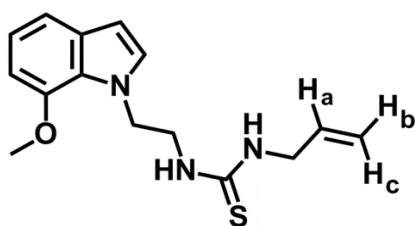
Appearance: Light beige solid
MF: $C_{15}H_{19}N_3OS$
MW: 289 g/mol
Yield: 46%

IR (KBr cm^{-1}): 3245 (s, NH); 3065 (w, aromatic C-H); 2965 71 and 2927 (w, aliphatic C-H).

^{1}H RMN (DMSO-d_6, 400 MHz) δ: 7.56 (bs, 1H, ArCH$_2$CH$_2$N**H**); 7.41 (bs, 1H, CSN**H**CH$_2$CHCH$_2$); 7.14 (bs, 1H, **H$_2$**); 7.11 (d, 1H, **H$_4$**, J_{4-5}=7.9 Hz); 6.91 (t, 1H, **H$_5$**, J_{5-4}=J_{5-6}=7.8 Hz); 6.66 (d, 1H, **H$_6$**, J_{6-5}=7.7 Hz); 6.37 (d, 1H, **H$_3$**, J_{3-2}=3.0 Hz); 5.80 (m, 1H, **Ha**); 5.11 (dd, 1H, **Hc**, J_{c-a}=17.3 Hz, J_{c-b}=1.3 Hz); 5.05 (dd, 1H, **Hb**, J_{b-a}=10.3 Hz, J_{b-c}=1.3 Hz); 4.51 (t, 2H, ArC**H$_2$**CH$_2$NH, $J_{CH2-CH2}$=5.9 Hz); 4.02 (bs, 2H, NHC**H$_2$**CH=CH$_2$); 3.89 (s, 3H, OC**H$_3$**); 3.74 (bs, 2H, ArCH$_2$C**H$_2$**NH) ppm.

CHN	Calculated	Found
C	62.28%	62.22%
H	6.57%	7.01%
N	14.53%	14.62%

MP: 69.5-71.0ºC
HPLC: 99.401% / t_R=6.324′

1-ethyl-3-(2-(7-methoxyindolyl)ethyl)thiourea [In2-3.V]

Appearance: : Light beige solid
MF: $C_{14}H_{19}N_3OS$
MW: 277 g/mol
Yield: 71%

IR (KBr cm^{-1}): 3350 and 3223 (s, NH); 3087 and 3010 (w, aromatic C-H); 2970 and 2912 (w, aliphatic C-H).

^{1}H RMN (DMSO-d_6, 400 MHz) δ: 7.43 (bs, 1H, ArCH$_2$CH$_2$N**H**); 7.32 (bs, 1H, N**H**CH$_2$CH$_3$); 7.14 (bs, 1H, **H$_2$**); 7.11 (dd, 1H, **H$_4$**, J_{4-5}=7.9 Hz, J_{4-6}=0.8 Hz); 6.91 (dt, 1H, **H$_5$**, J_{5-4}=J_{5-6}=7.8 Hz, J_{5-3}=1.8 Hz); 6.66 (d, 1H, **H$_6$**, J_{6-5}=7.7 Hz); 6.37 (dd, 1H, **H$_3$**, J_{3-2}=3.0 Hz, J_{3-5}=1.9 Hz); 4.51 (t, 2H, ArC**H$_2$**CH$_2$NH, $J_{CH2-CH2}$=5.9 Hz); 3.89 (s, 3H, OC**H$_3$**); 3.70 (bs, 2H, C**H$_2$**CH$_3$); 3.30 (bs, 2H, ArCH$_2$C**H$_2$**NH); 1.03 (t, 3H, C**H$_3$**, $J_{CH3-CH2}$=7.2 Hz) ppm.

CHN	Calculated	Found
C	60.65%	60.54%
H	6.86%	7.18%
N	15.16%	15.03%

MP: 76.8-77ºC
HPLC: 99.676% / t_R=5.159′/

1-(2-(7-methoxyindolyl)ethyl)-3-propylthiourea [In2-3.VI]
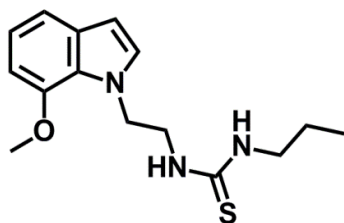
Appearance: White solid
MF: $C_{15}H_{21}N_3OS$
MW: 291 g/mol
Yield: 65%

IR (KBr cm^{-1}): 3354 and 3230 (s, NH); 3071 and 3007 (w, aromatic C-H); 2963 and 2932 (w, aliphatic C-H).

^{1}H RMN (DMSO-d_6, 400 MHz) δ: 7.45 (bs, 1H, ArCH$_2$CH$_2$N**H**); 7.32 (bs, 1H, N**H**CH$_2$CH$_2$CH$_3$); 7.14 (bs, 1H, **H$_2$**); 7.11 (d, 1H, **H$_4$**, J_{4-5}=7.9 Hz); 6.91 (t, 1H, **H$_5$**, $J_{5-4}=J_{5-6}$=7.8 Hz); 6.66 (d, 1H, **H$_6$**, J_{6-5}=7.7 Hz); 6.37 (d, 1H, **H$_3$**, J_{3-2}=3.0 Hz); 4.51 (t, 2H, ArC**H$_2$**CH$_2$NH, $J_{CH2-CH2}$=5.9 Hz); 3.89 (s, 3H, OC**H$_3$**); 3.72 (bs, 2H, ArCH$_2$C**H$_2$**NH); 3.26 (bs, 2H, C**H$_2$**CH$_2$CH$_3$); 1.45 (tt, 2H, CH$_2$C**H$_2$**CH$_3$, J_{gem}=14.3 Hz, $J_{CH2-CH3}$=7.2 Hz); 0.83 (t, 3H, CH$_2$CH$_2$C**H$_3$**, $J_{CH3-CH2}$=7.3 Hz) ppm.

CHN	Calculated	Found
C	61.86%	61.48%
H	7.22%	7.22%
N	14.43%	14.24%

MP: 101.8-102ºC
HPLC: 97.497% / t_R= 7.830'

1-(2-(7-methoxyindolyl)ethyl)-3-methylthiourea [In2-3.VII]
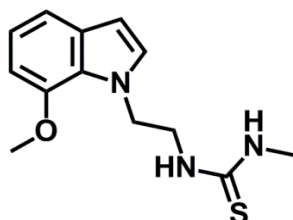
Appearance: White solid
MF: $C_{13}H_{17}N_3OS$
MW: 263 g/mol
Yield: 18%

IR (KBr cm^{-1}): 3281 and 3237 (s, NH); 3058 (w, aromatic C-H); 2924 (w, aliphatic C-H).

^{1}H RMN (DMSO-d_6, 400 MHz) δ: 7.41 (bs, 2H, N**H**+N**H**); 7.14 (bs, 1H, **H$_2$**); 7.10 (d, 1H, **H$_4$**, J_{4-5}=7.9 Hz); 6.91 (dt, 1H, **H$_5$**, $J_{5-4}=J_{5-6}$=7.8 Hz, J_{5-3}=1.9 Hz); 6.66 (d, 1H, **H$_6$**, J_{6-5}=7.6 Hz); 6.37 (d, 1H, **H$_3$**, J_{3-2}=2.9 Hz); 4.52 (bs, 2H, ArC**H$_2$**CH$_2$NH); 3.89 (s, 3H, OC**H$_3$**); 3.71 (bs, 2H, ArCH$_2$C**H$_2$**NH); 2.78 (bs, 3H, C**H$_3$**) ppm.

CHN	Calculated	Found
C	59.32%	58.90%
H	6.46%	6.49%
N	15.97%	15.76%

MP: 97.9-99.7ºC
HPLC: 99.763% / t_R=3.932'

1-cyclopropyl-3-(2-(7-methoxyindolyl)ethyl)thiourea [In2-3.VIII]
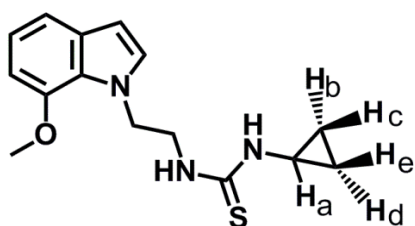
Appearance: White solid
MF: $C_{15}H_{19}N_3OS$
MW: 289 g/mol
Yield: 49%

IR (KBr cm^{-1}): 3369 and 3204 (s, NH); 3083 (w, aromatic C-H); 2990 (w, aliphatic C-H).

^{1}H RMN (DMSO-d_6, 400 MHz) δ: 7.95 (bs, 1H, N**H**-cC$_3$H$_5$); 7.41 (bs, 1H, ArCH$_2$CH$_2$N**H**); 7.14 (d, 1H, **H$_2$**, J_{2-3}=3.1 Hz); 7.11 (d, 1H, **H$_4$**, J_{4-5}=7.9 Hz); 6.91 (t, 1H, **H$_5$**, $J_{5-4}=J_{5-6}$=7.8 Hz); 6.66 (d, 1H, **H$_6$**, J_{6-5}=7.7 Hz); 6.37 (d, 1H, **H$_3$**, J_{3-2}=3.0 Hz); 4.57 (t, 2H, ArC**H$_2$**CH$_2$NH, $J_{CH2-CH2}$=5.5 Hz); 3.90 (s, 3H, OC**H$_3$**); 3.81 (bs, 2H, ArCH$_2$C**H$_2$**NH); 2.38 (bs, 1H, **Ha**); 0.59 (qd, 2H, **Hb**+**Hd**, $Jb-c=Jd-e$=4.6 Hz, $Jb-e=Jd-c$=4.8 Hz, $Jb-a$=4.8, Hz $Jb-d$=6.5 Hz); 0.41 (tdd, 2H, **Hc**+**He**, $Jc-b=Je-d$=4.5 Hz, $Jc-a$=4.5 Hz, $Jc-d=Je-b$=5.0 Hz, $Jc-e$=7.3 Hz) ppm.

CHN	Calculated	Found
C	62.28%	62.46%
H	6.57%	6.92%
N	14.53%	14.76%

MP: 129-131ºC
HPLC: 99.912% / t$_R$=5.407'

N-(2-(7-methoxyindolyl)ethyl)methanesulfonamide [In2-3.IX]
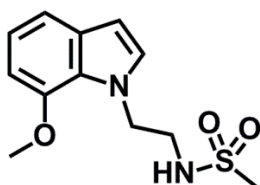
Appearance: Dark beige solid
MF: $C_{12}H_{16}N_2O_3S$
MW: 268 g/mol
Yield: 30%

IR (KBr cm^{-1}): 3271 (s, NH); 2938 (w, aliphatic C-H).

^{1}H RMN (DMSO-d_6, 400 MHz) δ: 7.27-7.20 (m, 1H, N**H**); 7.22 (d, 1H, **H$_2$**, J_{2-3}=3.0 Hz); 7.11 (d, 1H, **H$_4$**, J_{4-5}=7.9 Hz); 6.92 (d, 1H, **H$_5$**, $J_{5-4}=J_{5-6}$=7.8 Hz, J_{5-3}=3.3 Hz); 6.67 (d, 1H, **H$_6$**, J_{6-5}=7.7 Hz) 6.36 (d, 1H, **H$_3$**, J_{3-2}=3.1Hz); 4.41 (t, 2H, ArC**H$_2$**CH$_2$NH, $J_{CH2-CH2}$=4.9 Hz); 3.89 (s, 3H, OC**H$_3$**); 3.35-3.27 (m, 2H, ArCH$_2$C**H$_2$**NH); 2.73 (s, 3H, C**H$_3$**) ppm.

CHN	Calculated	Found
C	53.73%	53.63 %
H	5.97%	6.06%
N	10.44%	9.99%

MP: 86-88ºC
HPLC: 96.149% / t$_R$=3.070'

2,2,2-trifluoro-N-(2-(7-methoxyindolyl)ethyl)acetamide [In2-3.X]
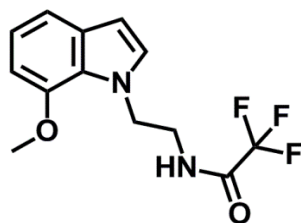
Appearance: Pale rose solid

MF: $C_{13}H_{13}N_2O_2F_3$

MW: 286 g/mol

Yield: 25%

IR (KBr cm^{-1}): 3302 (s, NH); 3006 (w, aromatic C-H); 2960 (w, aliphatic C-H); 1701 (vs, C=O).

^{1}H RMN (DMSO-d_6, 400 MHz) δ: 9.46 (t, 1H, N**H**, $J_{NH\text{-}CH2}$=5.0 Hz); 7.14 (d, 1H, **H$_2$**, $J_{2\text{-}3}$=3.1 Hz); 7.11 (dd, 1H, **H$_4$**, $J_{4\text{-}5}$=7.9 Hz, $J_{4\text{-}6}$=0.7 Hz); 6.92 (t, 1H, **H$_5$**, $J_{5\text{-}4}$=$J_{5\text{-}6}$=7.8 Hz); 6.66 (d, 1H, **H$_6$**, $J_{6\text{-}5}$=7.8 Hz); 6.37 (d, 1H, **H$_3$**, $J_{3\text{-}2}$=3.1 Hz); 4.49 (t, 2H, ArC**H$_2$**CH$_2$NH, $J_{CH2\text{-}CH2}$=5.9 Hz); 3.88 (s, 3H, OC**H$_3$**); 3.55 (q, 2H, ArCH$_2$C**H$_2$**NH, $J_{CH2\text{-}CH2}$=5.7 Hz, $J_{CH2\text{-}NH}$=5.7 Hz) ppm.

CHN	Calculated	Found
C	54.54%	54.36%
H	4.54%	4.39%
N	9.79%	9.64%

MP: 61.0-63.2ºC

HPLC: 99.346% / t_R=5.360'

1-(2-(7-methoxyindolyl)ethyl)-3-methylurea [In2-4.XI]
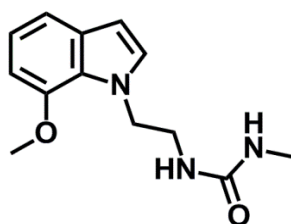
Appearance: White solid

MF: $C_{13}H_{17}N_3O_2$

MW: 247 g/mol

Yield: 12%

IR (KBr cm^{-1}): 3347 (s, NH); 3042 (w, aromatic C-H); 2933 (w, aliphatic C-H); 1631 (vs, C=O).

^{1}H RMN (DMSO-d_6, 400 MHz) δ: 7.14 (d, 1H, **H$_2$**, $J_{2\text{-}3}$=2.9 Hz); 7.10 (d, 1H, **H$_4$**, $J_{4\text{-}5}$=7.9 Hz); 6.90 (t, 1H, **H$_5$**, $J_{5\text{-}4}$=$J_{5\text{-}6}$=7.8 Hz); 6.65 (d, 1H, **H$_6$**, $J_{6\text{-}5}$=7.8 Hz); 6.35 (d, 1H, **H$_3$**, $J_{3\text{-}2}$=2.9 Hz); 5.89 (t, 1H, ArCH$_2$CH$_2$N**H**, $J_{NH\text{-}CH2}$=5.4 Hz); 5.73 (q, 1H, N**H**CH$_3$, $J_{NH\text{-}CH3}$=4.6 Hz); 4.35 (t, 2H, ArC**H$_2$**CH$_2$NH, $J_{CH2\text{-}CH2}$=6.1 Hz); 3.88 (s, 3H, OC**H$_3$**); 3.35-3.33 (m, 2H, ArCH$_2$C**H$_2$**NH); 2.53 (d, 3H, C**H$_3$**, $J_{CH3\text{-}NH}$=4.6 Hz) ppm.

CHN	Calculated	Found
C	63.16%	62.88%
H	6.88%	7.00%
N	17.00%	17.00%

MP: 131-132.8ºC

HPLC: 99.773% / t_R=3.083'

1-cyclopropyl-3-(2-(7-methoxyindolyl)ethyl)urea [In2-4.XII]
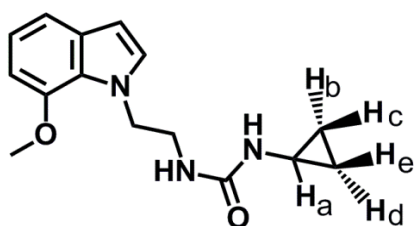
Appearance: White solid
MF: $C_{15}H_{19}N_3O_2 \cdot \frac{1}{6} H_2O$
MW: 276 g/mol
Yield: 28%

IR (KBr cm⁻¹): 3328 (s, NH); 3004 (w, aromatic C-H); 2956 and 2936 (w, aliphatic C-H); 1629 (vs, C=O).

¹H RMN (DMSO-*d₆*, 400 MHz) δ: 7.13 (d, 1H, **H₂**, $J_{2\text{-}3}$=3.1 Hz); 7.10 (d, 1H, **H₄**, $J_{4\text{-}5}$=7.9 Hz); 6.90 (t, 1H, **H₅**, $J_{5\text{-}4}=J_{5\text{-}6}$=7.8 Hz); 6.65 (d, 1H, **H₆**, $J_{6\text{-}5}$=7.7 Hz); 6.35 (d, 1H, **H₃**, $J_{3\text{-}2}$=3.0 Hz); 6.15 (bs, 1H, N**H**-cC₃H₅); 5.83 (t, 1H, ArCH₂CH₂N**H**, $J_{NH\text{-}CH2}$=5.8 Hz); 4.38 (t, 2H, ArC**H₂**CH₂NH, $J_{CH2\text{-}CH2}$=6.1 Hz); 3.89 (s, 3H, OC**H₃**); 3.34 (q, 2H, ArCH₂C**H₂**NH, $J_{CH2\text{-}CH2}=_{CH2\text{-}NH}$=5.5 Hz); 2.36-2.29 (m, 1H, **Ha**); 0.52 (qd, **Hb**+**Hd**, 2H, *Jb-c=Jd-e*=4.6 Hz, *Jb-e=Jd-c*=4.7 Hz, *Jb-a*=4.7 Hz, *Jb-d*=6.7 Hz); 0.28 (tdd, **Hc**+**He**, 2H, *Jc-b=Je-d*=4.3 Hz, *Jc-a*=4.3 Hz, *Jc-d=Je-b*=4.7 Hz, *Jc-e*=7.0 Hz) ppm.

CHN	Calculated	Found
C	65.22%	64.99%
H	7.00%	6.88%
N	15.22%	15.03%

MP: 141-143ºC
HPLC: 99.552% / t_R=4.366'

X. INSTRUMENTAL STRUCTURAL CHARACTERIZATION

20. INFRARED SPECTROSCOPY

Infrared spectroscopy has been used in order to determine the presence of the most prominent functional groups based on the way in which infrared radiation is absorbed by their chemical bonds. The main and most interesting functional groups appearing in the characterization of compounds of this project are listed in table 11, accompanied by their corresponding frequency range.[161, 162]

Table 11. IR of the most interesting functional groups appearing in the characterization of compounds of this project.

Funtional group	Frequency band (cm^{-1})	Observation	Example
(1) Aliphatic C-H stretching	3000-2840 cm^{-1}		(Figure 70)
(2) Heteroaromatic C-H stretching	3100-3000 cm^{-1}		(Figure 70)
(3) Amide NH stretching	3500-3400 cm^{-1}	(A single band in free amines)	
(4) Amide NH stretching	3350-3100 cm^{-1}	(Wider bands in associated amines)	(Figure 70)
(5) Urea NH stretching	3500-3200 cm^{-1}		(Figure 71)
(6) Amide C=O	1740-1630 cm^{-1}		(Figure 70)
(7) Urea C=O	1690-1620 cm^{-1}		(Figure 71)
(8) Aromatic C≡N stretching	2240-2215 cm^{-1}		(Figure 72)
(9) N-CH2-C≡N		(Not observed)	(Figure 72)

As can be observed in figure 70, the frequency of the NH bond of the amide group is in the range of associated amides. This means that, at least in solid state, the hydrogen atom of NH group is involved in a hydrogen bond, probably with the carbonyl group of another amide derivative. Different studies have established that secondary amides in solid state usually form chains via amide-amide hydrogen bonds.[163]
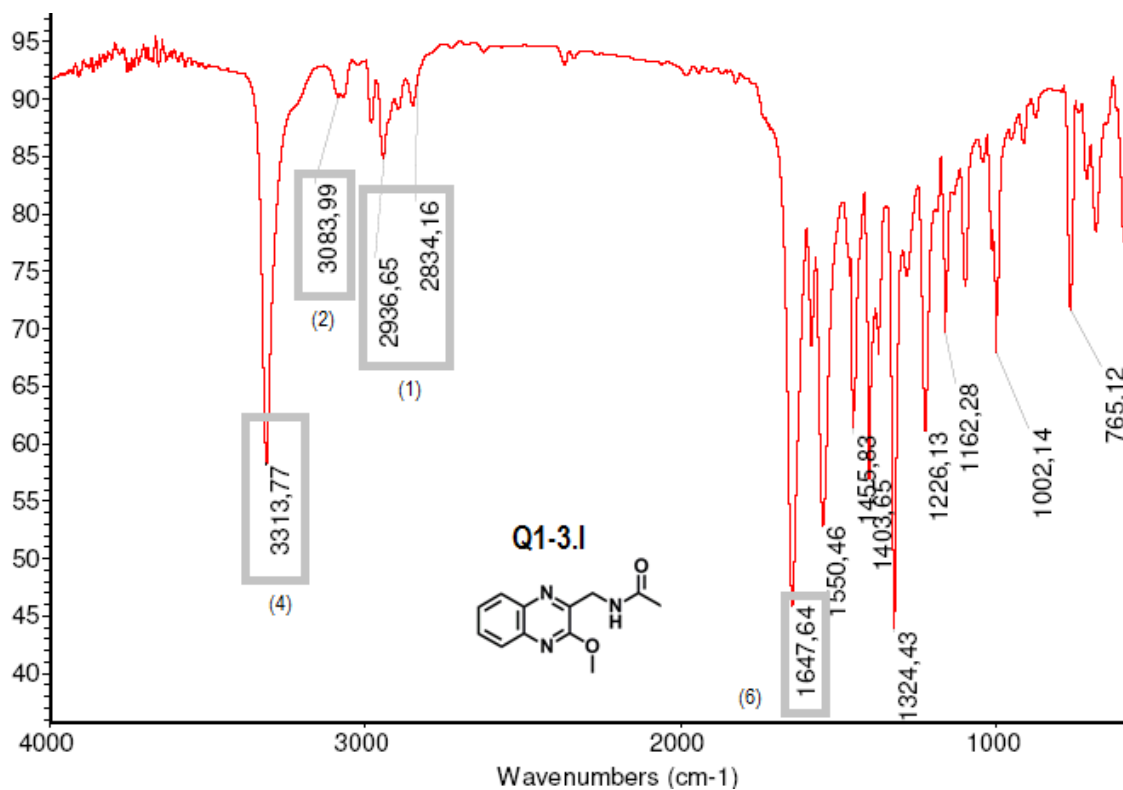
Figure 70. IR spectra of compound **In1-3.I** showing typical bands of aliphatic CH (1), aromatic CH (2), NH of associated amide (4) and CO of amide (6)

Taking into account the fact that a urea can be described as an amide derivative regardless of the second NH being attached to the carbonyl group,[163] it is possible to propose that all final compounds containing an amide or a urea group in their structure, whose NH bonds present a frequency<3350 cm^{-1}, would probably contain a chain structure in solid state (figure 71).

As IR spectra of thiourea derivatives show a signal pattern very close to their homologous urea derivatives and as sulfur atom can form hydrogen bonds even easier than an oxygen atom, it is possible to hypothesize that thiourea-containing compounds would also form thiourea-thiourea hydrogen bonds, resulting in a chain structure in solid state.
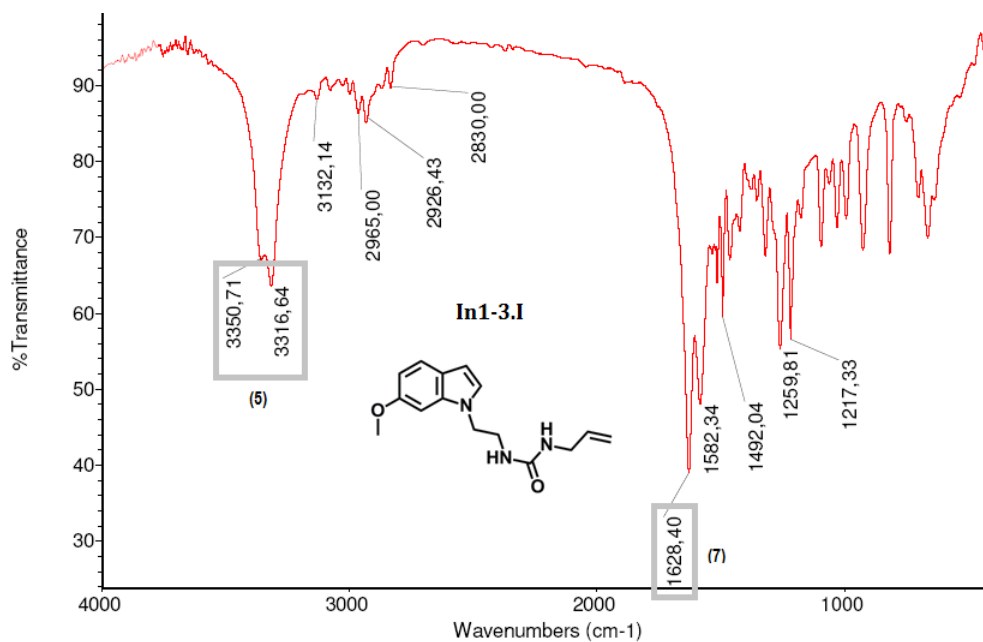
Figure 71. IR spectra of compound **In1-3.I** showing typical bands of urea NH (5) and urea CO (7).

Finally, it is also remarkable that while aromatic C≡N stretching band is a characteristic signal, it disappears at the moment in which a methylene bridge is introduced between the aromatic ring and the nitrile group (figure 72).
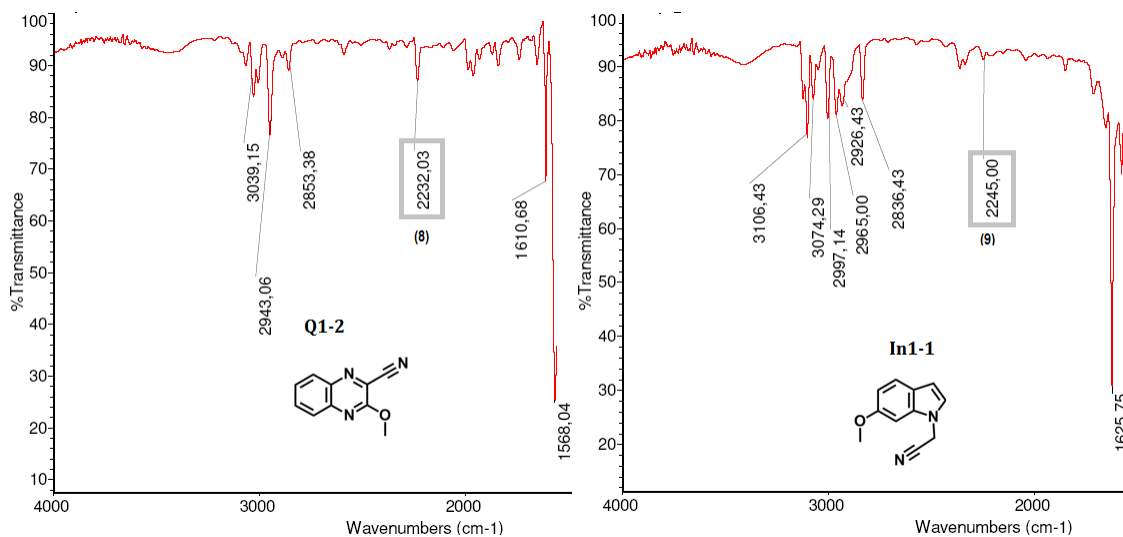
Figure 72. IR spectra of compounds **In1-1** and **Q1-2** showing typical band generated by an aromatic nitrile (8) and the lack of a band in the case of aliphatic nitriles (9)

21. NUCLEAR MAGNETIC RESONANCE SPECTROSCOPY

Nuclear magnetic resonance spectroscopy is a research technique, which provides detailed information regarding the structure, dynamics, reaction state and chemical environment of molecules. This technique is based on magnetic properties of a nucleus possessing spin, with ^1H and ^{13}C being the most used.

^1H NMR experiments have been performed on all molecules synthesized in this project for the purpose of characterizing them. Different experiments such as ^{13}C NMR, 2D ^1H-^{13}C HMQC, 2D ^1H-^{13}C HMBC and 2D ^1H-^1H COSY have also been carried out, if necessary, in order to allow the full description of some compounds.

21.1. ^1H NMR spectroscopy

A ^1H NMR experiment permits quantification of the number of the protons present in a molecule and characterization of the type of proton as well as the chemical environment it presents.

Molecules involved in this work can be divided into two main groups, depending on the aromatic nucleus they contain; quinoxaline or indole. Figure 73 shows the ^1H chemical shift of unsubstituted quinoxaline ring and indole ring and their corresponding coupling constants *(J)* with respect to TMS.

Quinoxaline:
8.07 (8), 8.74 (2), 8.74 (3), 8.07 (4), 7.68 (6), 7.68 (7), 8.07 (1)

$^3J_{2-3}=1.8$
$^3J_{5-6}=^3J_{7-8}=8.4$
$^4J_{5-7}=^4J_{8-6}=1.6$
$^5J_{5-8}=0.6$
$^3J_{6-7}=6.9$

Indole:
7.55 (4), 6.45 (3), 7.00 (5), 7.08 (6), 7.27 (2), 7.40 (7), 10.1 (1)

$^3J_{2-3}=3.1$ $^4J_{4-6}=1.2$
$^5J_{2-4}=^6J_{2-5}=^6J_{2-6}=^5J_{2-7}=0$ $^5J_{4-7}=0.9$
$^3J_{2-1}=2.5$ $^5J_{4-1}=0.8$
$^4J_{3-4}=^5J_{3-5}=^6J_{3-6}=0$ $^3J_{5-6}=7.1$
$^5J_{3-7}=0.7$ $^4J_{5-7}=1.3$
$^4J_{3-1}=2.0$ $^3J_{6-7}=8.1$
$^3J_{4-5}=7.8$ $^6J_{5-1}=^5J_{6-1}=^4J_{7-1}=0$
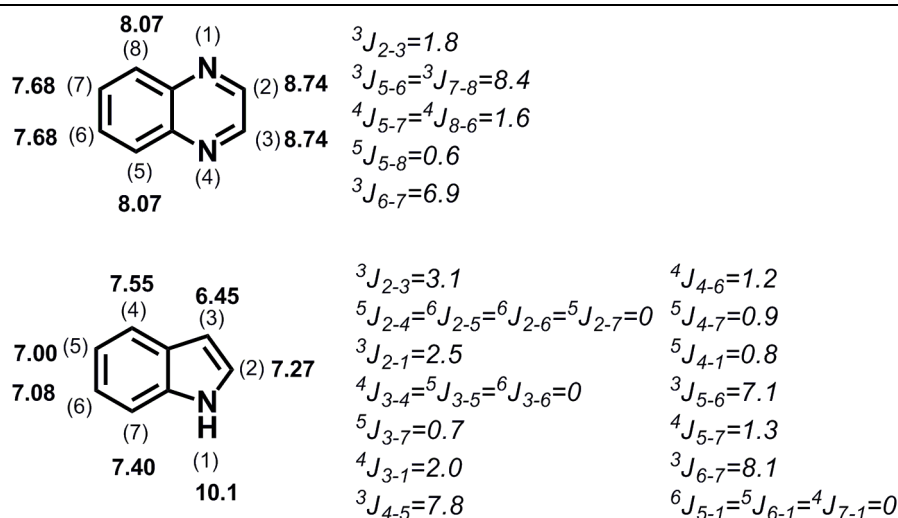
Figure 73. Chemical shift in ppm and coupling constant (*J*) in Hz of quinoxaline and indole rings. [Figure modified from *Pretsch E. et al.*][161, 162]

These displacements and coupling constants *(J)* have been used as the basis for the structural characterization of synthesized compounds, knowing that the effect of substituents would change the values; it is known that electron-withdrawing groups induce a downfield effect in chemical shift and electronegative groups make *J* values smaller.

As can be observed in figure 74, the displacement of peaks linked to the signal's multiplicity and *J* values has permitted the characterization of each proton of the quinoxaline ring. It is remarkable that the order of H_6 and H_8 protons of compound **Q3-3** is changed compared with the unsubstituted quinoxaline, due to the effect caused by methoxy and nitrile groups. The

methoxy group displaces both protons, H_6 and H_8, in a similar way. However, H_8 proton suffers a higher downfield displacement because as a nitrile group is closer to H_8 than to H_6, it induces a higher effect on the proton H_8.
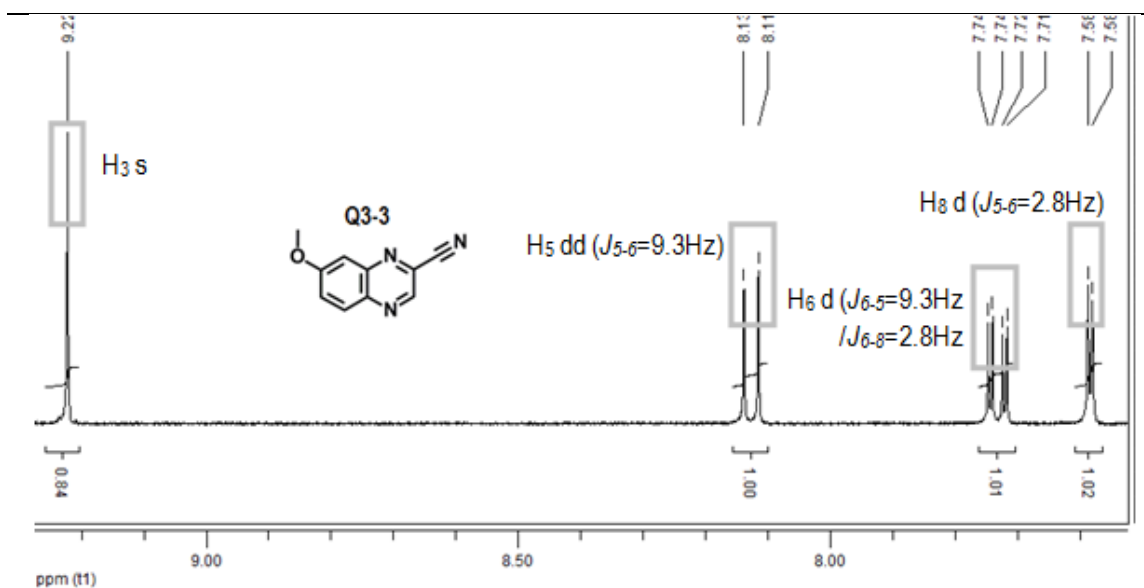
Figure 74. Aromatic region of ^{1}H NMR spectra of compound **Q3-3.**

Figure 75 shows the identification of protons present in indole ring of compound **In2-3.II** as well as their multiplicity and J values. H_5-H_6 protons appear in a different order with respect to protons of unsubstituted indole ring because of the downfield effect induced by methoxy group, which affects the H_6 proton more due to their proximity. In this case, the methoxy group acts as an electron withdrawing group due to the inductive effect of the oxygen atom. The position of H_2-H_4 is also different, probably due to the electron-donating properties of the chain placed in position 1, which has an upfield effect in H_2 proton.
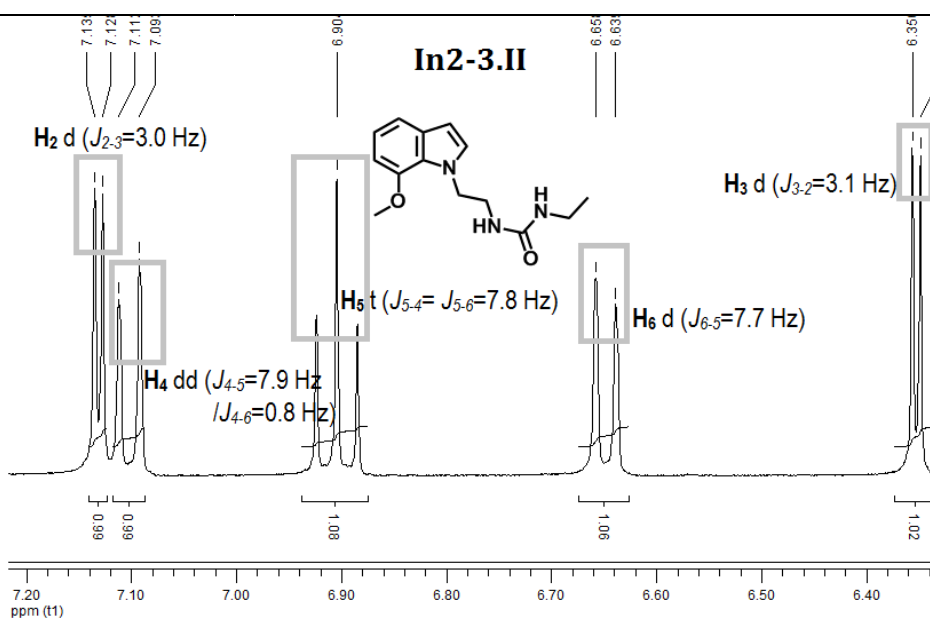
Figure 75. Aromatic region of ^{1}H NMR spectra of compound **In2-3.II.**

In some derivatives, a phenyl group can be found as a substituent. Thus, the chemical shift of 1-substituted phenyl ring is shown in figure 76 as well as a real example from the ^{1}H RMN spectra of compound **Q1-3.IX**.
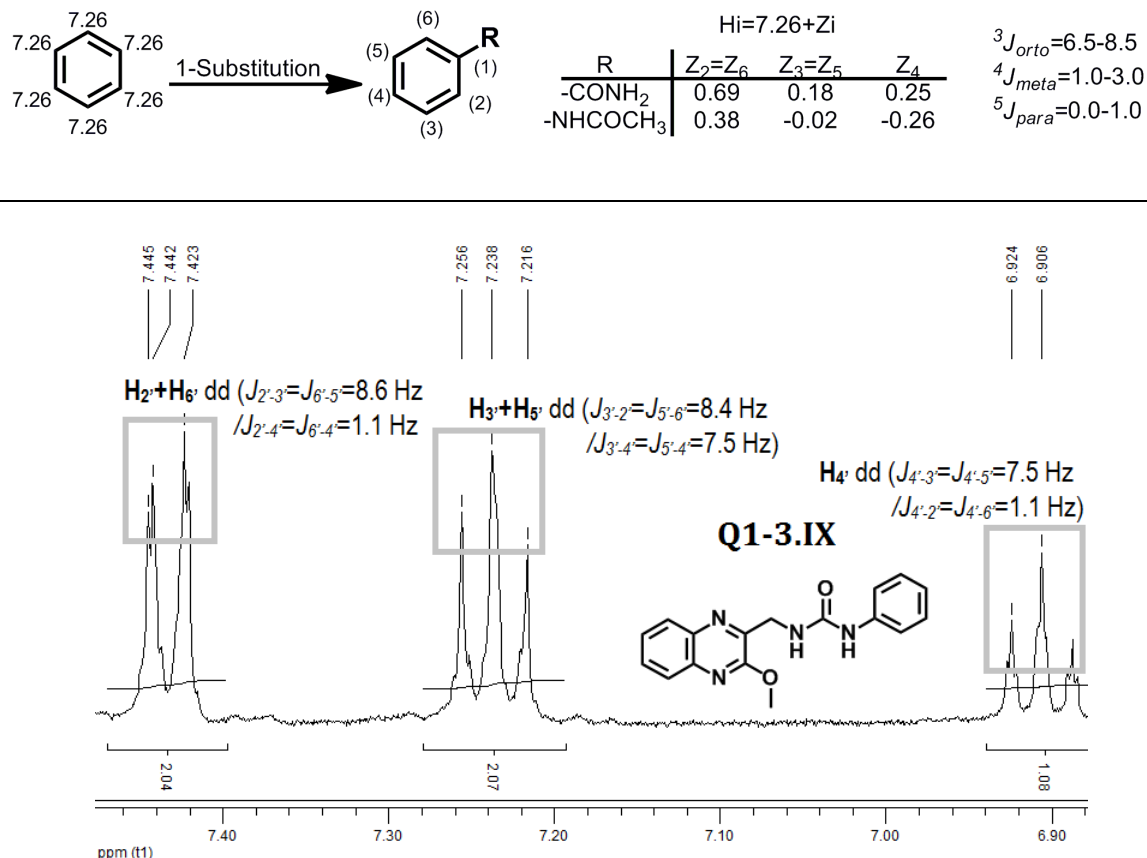
Figure 76. Chemical shift in ppm of protons of 1-substituted phenyl ring (Hi) and their coupling constants (*J*) in Hz [Figure modified from *Pretsch E. et al.*][164], and the aromatic region of ^{1}H NMR spectra of compound **Q1-3.IX** showing signals corresponding to phenyl group.

Allyl rest is another interesting group that can be found as substituent in the aliphatic chain of some derivatives. In this case, as in other substituted ethylenes, geminal and neighboring coupling constants are clearly observed. Normally, Jgem ranges from -4 to 4 Hz, thus, its absolute value is described. In neighboring protons, Jcis is usually smaller than Jtrans and their values range from 4 to 12 Hz for protons in cis position and from 14 to 19 Hz for protons in trans position (figure 77).[165]
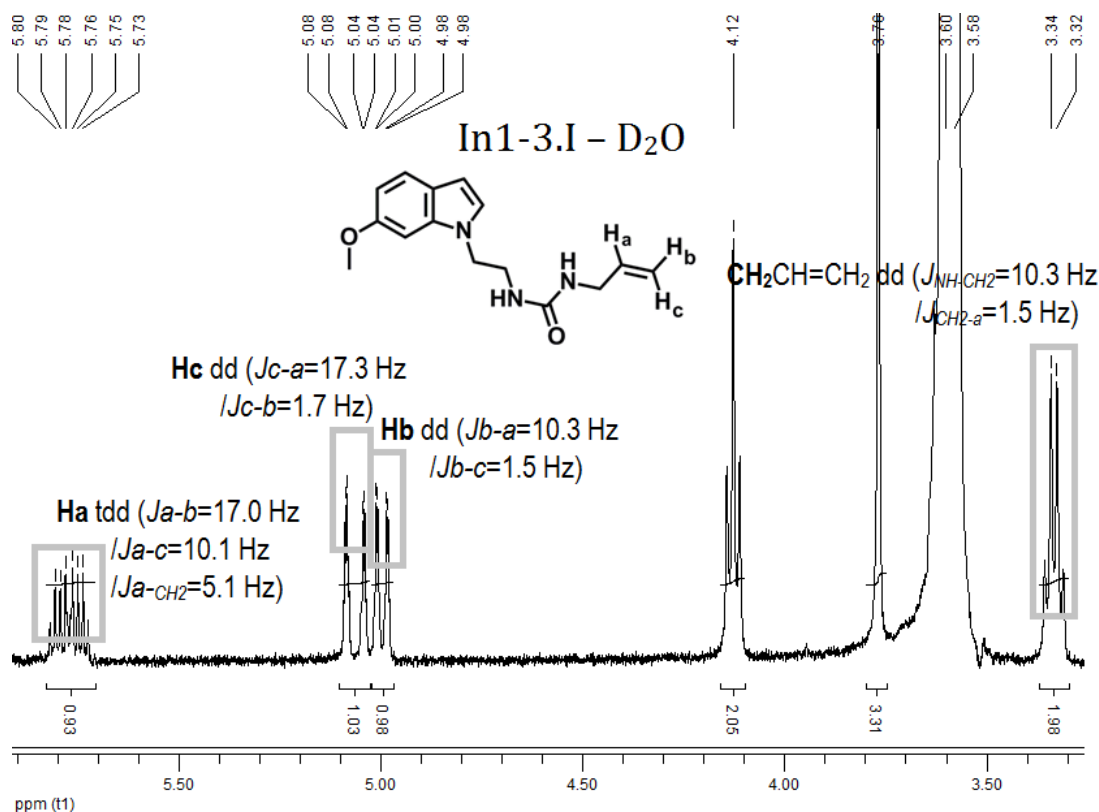
Figure 77. ^{1}H NMR spectra of compound **In1-3.I** showing signals corresponding to allyl group.

Cyclopropyl ring is also an interesting group to describe. In this case, unlike in allyl group, *Jcis* is usually higher than *Jtrans.* In figure 78, the *J* values and chemical shift of protons of 1-substituted cyclopropyl ring can be observed. The ^{1}H NMR spectra of compound **In1-4.XII** showing signals corresponding to cyclopropyl ring is shown in figure 79.
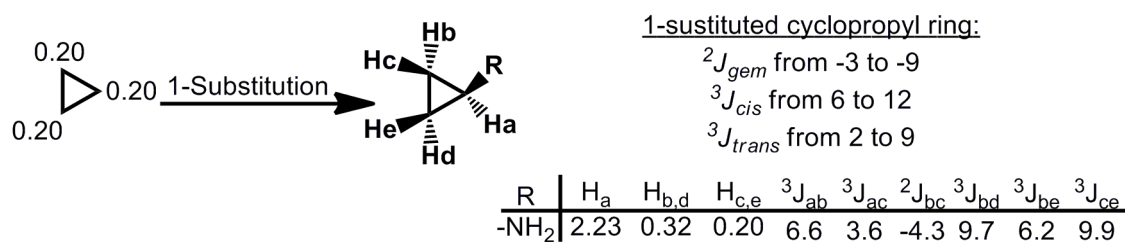
Figure 78. Chemical shift in ppm and coupling constant (*J*) in Hz of free and 1-substituted cyclopropyl ring [Figure modified from *Pretsch E. et al.*].[166]
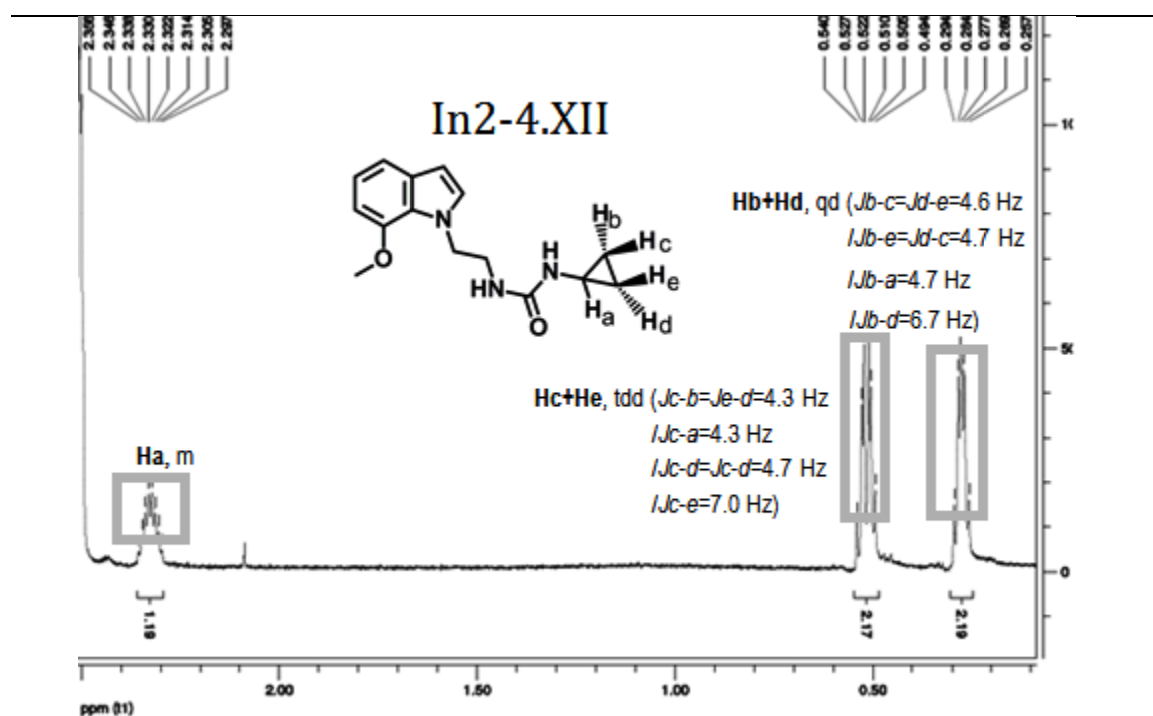
Figure 79. ^{1}H NMR spectra of compound **In1-4.XII** showing signals corresponding to cyclopropyl ring.

21.2. 2D ^{1}H-^{1}H COSY (correlation spectroscopy)

2D ^{1}H-^{1}H COSY is a homonuclear 2D hydrogen-hydrogen correlation spectroscopy technique which allows us to determine the protons which are through-bond coupled to each other. The peaks generated by the protons bonded to the same carbon form a diagonal, while the protons bonded to adjacent carbons generate off-diagonal peaks. Figure 80 shows the COSY spectra of compound **Q2-3.II** carried out in order to confirm the coupling that takes place between the two methoxy groups present in compounds of series Q2.
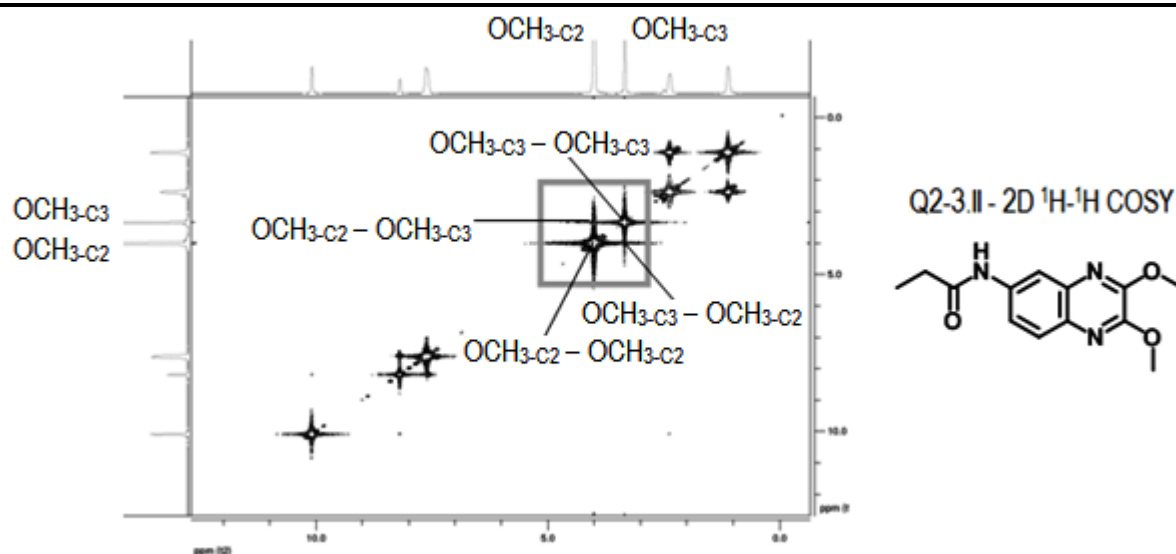
Figure 80. 2D ^{1}H-^{1}H COSY spectra of compound **Q2-3.II**.

21.3. ^{13}C NMR spectroscopy, 2D ^{1}H-^{13}C HMQC (heteronuclear multiple quantum coherence); 2D ^{1}H-^{13}C HMBC (heteronuclear multiple bond coherence)

^{13}C NMR is a monodimensional NMR technique similar to ^{1}H NMR. This experiment provides information regarding the carbons (instead of hydrogens), their chemical shift and the carbon-containing functional groups of the molecule.

2D ^{1}H-^{13}C HMQC and 2D ^{1}H-^{13}C HMBC are heteronuclear 2D carbon-hydrogen spectroscopies that show correlations between the chemical shifts of protons with the chemical shifts of carbons. 2D ^{1}H-^{13}C HMQC is selective for direct C-H coupling and 2D ^{1}H-^{13}C HMBC gives longer range couplings (2-4 bonds).

The obtainment of the final compounds of series Q3 (**Q3-4.(I-V)**) begins with the synthesis of compound **Q3-1** through Beirut reaction. As explained previously, when condensation between a BFX and malononitrile takes place, a mixture of two positional isomers can be achieved (*see Ch2 - 8.3.1);* therefore, these techniques have been used for the purpose of determining whether these compounds (**Q3-4.(I-V)**) are 6-isomer or 7-isomer. First, a ^{13}C NMR has been performed and each peak has been related to a carbon as shown in figure 81. Although this spectrum does not clarify isomerization, it is necessary to analyze the other experiments.

	δ(ppm)	14	20	38	44	57	107	123	131	138	143	144	155	161	173
Carbons	**7-isomer**	C_{16}	C_{15}	C_{14}	C_{12}	C_{11}	C_8	C_6	C_5	C_{10}	C_3	C_9	C_2	C_7	C_{13}
	6-isomer						C_5	C_7	C_8	C_9		C_{10}		C_6	
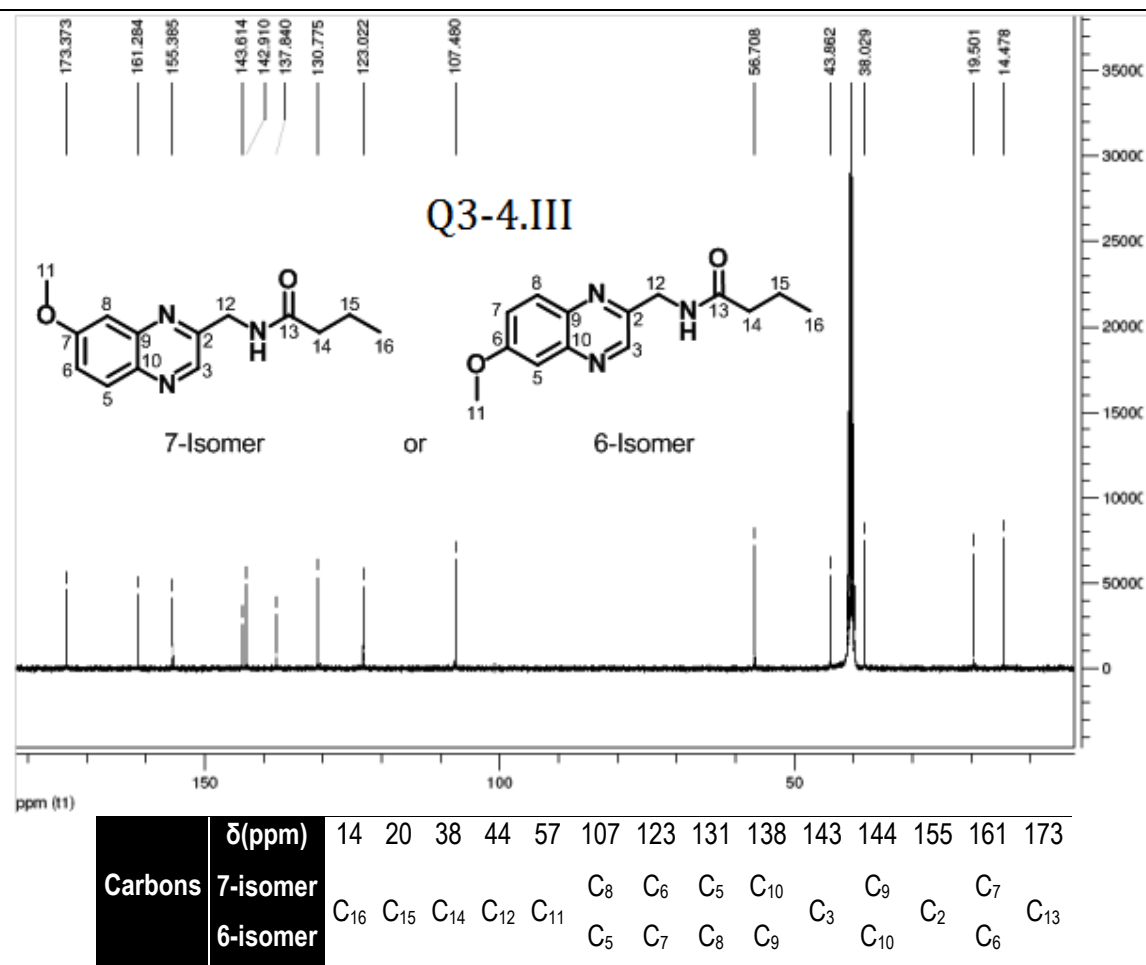
Figure 81. ^{13}C NMR spectra of compound **Q3-4.III** and peak interpretation.

Then 2D ^{1}H-^{13}C HMQC experiment has been performed in order to confirm the peak assignation made upon interpreting ^{13}C NMR spectra (figure 82).

	δ(ppm)	14	20	38	44	57	107	123	131	138	143	144	155	161	173
Carbons	7-isomer	C_{16}	C_{15}	C_{14}	C_{12}	C_{11}	C_8	C_6	C_5	C_{10}	C_3	C_9	C_2	C_7	C_{13}
	6-isomer						C_5	C_7	C_8	C_9		C_{10}		C_6	
Hydrogens	7-isomer	CH$_3$	CH$_2$CH$_3$	COCH$_2$	CH$_2$NH	OCH$_3$	H$_8$	H$_6$	H$_5$	X	H$_3$	X	X	X	X
	6-isomer						H$_5$	H$_7$	H$_8$						
	δ(ppm)	0.88	1.56	2.8	4.56	3.94	7.38	7.45	7.96	X	8.69	X	X	X	X

Figure 82. 2D ^{1}H-^{13}C HMQC spectra of compound **Q3-4.III** and peak interpretation.

Finally, 2D ^{1}H-^{13}C HMBC experiment has been performed in order to determine, which isomer, 7/6, has been obtained. **H$_3$** proton shows a three-bond coupling with C_{10}, whose chemical shift changed from 144 ppm in 6-isomer to 138 ppm in 7-isomer. As can be observed in figure 83, C_{10} appears at 138 ppm; therefore, it can be concluded that the obtained compound is 7-isomer. This fact makes sense as it coincides with previous reports where it was stated that only 7-isomer is obtained upon performing Beirut reaction and using 5-methoxybenzofuroxan as the initial reagent[133] (*see Ch2 - 8.3.1*).

Figure 83. 2D ^{1}H-^{13}C HMBC spectra of compound **Q3-4.III** and peak interpretation
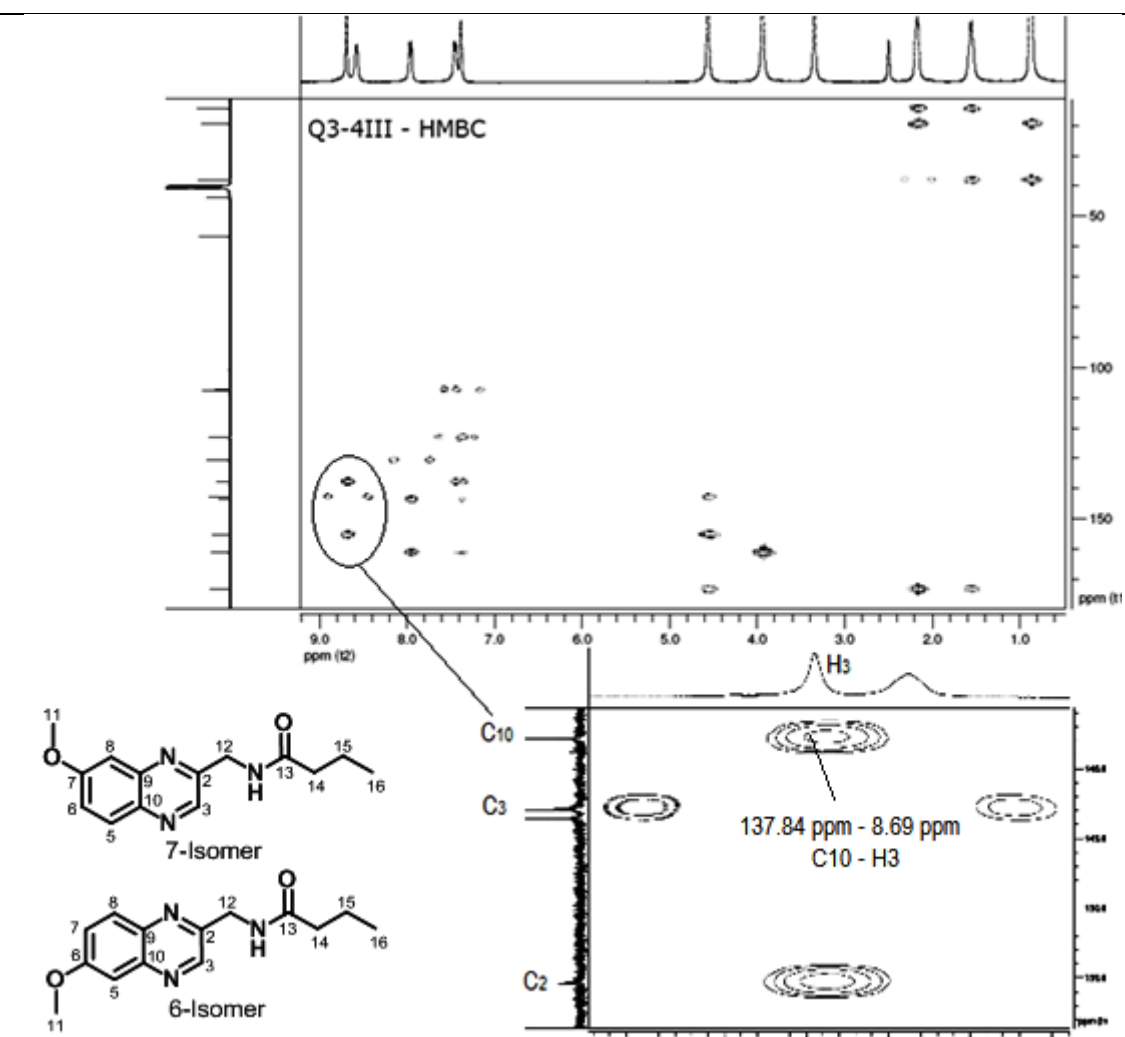
Couplings	Hydrogens	H$_3$	H$_5$	H$_6$	H$_8$	NH	OCH$_3$	C$\underline{H}_2$-NH	CO-CH$_2$	C$\underline{H}_2$-CH$_3$	CH$_3$
Carbons	δ(ppm)	8.69	7.96	7.45	7.38	8.58	3.94	4.56	2.18	1.56	0.88
C2	155.39	O						+			
C3	142.91	X						O			
C5	130.78		X								
C6	123.02			X	O						
C7	161.28		O		+		O				
C8	107.48			O	X						
C9	143.61		O		+						
C10	137.84	O		O	O						
C11	56.71					X					
C12	43.86					X					
C13	173.37							O		O	
C14	38.03								X	+	O
C15	19.50								+	X	+
C16	14.48								O	+	X

Distance: X - 1 bond / + - 2 bonds / O - 3 bonds

XI. BIOLOGICAL EVALUATION

22. BIOLOGICAL EVALUATION
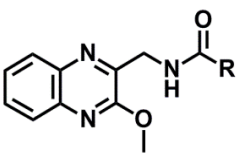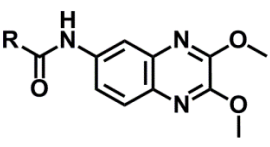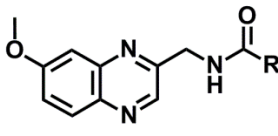
22.1. MT$_1$/MT$_2$ binding affinities. 2-[^{125}I]Iodomelatonin binding assay

2-[^{125}I]Iodomelatonin is an MT$_1$/MT$_2$ receptor agonist used as radioligand in competitive binding experiments for measuring the affinity of new synthesized drugs. In this assay, the unlabeled molecules compete with the radioligand for binding with the receptors. The results are obtained as IC$_{50}$, which is then converted to an absolute inhibition constant Ki.

22.1.1. Binding affinity of quinoxaline derivatives

The Ki values of new synthesized quinoxalines are shown in table 12.

Table 12. Ki value of quinoxalines synthesized on this project.

| Quinoxalines	Name	R	MT$_1$ Ki(µM)±SEM	MT$_2$ Ki(µM)±SEM
	Q1-3.I	CH$_3$	2.980±0.358	0.882±0.298
	Q1-3.II	CH$_2$CH$_3$	2.600±0.232	0.659±0.005
	Q1-3.III	CH$_2$CH$_2$CH$_3$	0.750±0.362	1.100±0.033
	Q1-3.IV	CH(CH$_3$)$_2$	1.230±0.321	0.474±0.020
	Q1-3.V	Ph	I	I
	Q1-3.VI	NHCH$_2$CH$_3$	I	0.395±0.125
	Q1-3.VII	NHCH$_2$CH$_2$CH$_3$	I	0.444±N.D.
	Q1-3.VIII	NHCH(CH$_3$)$_2$	I	I
	Q1-3.IX	NHPh	I	I
	Q2-3.I	CH$_3$	20.000±1.820	0.084±
	Q2-3.II	CH$_2$CH$_3$	17.600±7.810	4.360±1.220
	Q2-3.III	CH$_2$CH$_2$CH$_3$	11.500±2.800	1.350±0.306
	Q2-3.IV	CH(CH$_3$)$_2$	3.400±1.370	10.500±2.960
	Q2-3.V	Ph	I	I
	Q2-3.VI	NHCH$_2$CH$_3$	3.410±1.890	28.800±N.D
	Q2-3.VII	NHCH$_2$CH$_2$CH$_3$	1.630±0.442	0.49±0.065
	Q2-3.VIII	NHCH(CH$_3$)$_2$	I	I
	Q2-3.IX	NHPh	I	I
	Q3-4.I	CH$_3$	I	I
	Q3-4.II	CH$_2$CH$_3$	I	0.338±0.146
	Q3-4.III	CH$_2$CH$_2$CH$_3$	0.205±0.106	0.098±0.010
	Q3-4.IV	CH(CH$_3$)$_2$	0.324±0.039	0.161±0.001
	Q3-4.V	Ph	I	I

I:Inactive

As shown in table 12, series Q3 contains the best affinity results. By comparing the results obtained for compounds **Q1-3.III** and **Q1-3.IV** of series Q1 with those obtained for their analogous **Q3-4.III** and **Q3-4.IV** of series Q3, where the only difference lies in the position of the methoxy group, it can be concluded that the position of this group has a very important role. Indeed, the results suggest that the compounds in series Q3 are more active than their homologous compounds in series Q1 due to the fact that the distance between the methoxy group and the nitrogen atom of the amide/urea group is similar to the distance in the MLT molecule, a distance of six atoms.[73] Derivative **Q3-4.III** is the most active compound against both receptors and therefore, this compound could be selected as a hit in the search for new and more active quinoxalines, MT$_1$/MT$_2$ receptor agonists.

Compound **Q2-3.I** also shows quite interesting biological results because it presents high and selective affinity for the MT_2 receptor. Obtaining selective compounds is of vital importance in order to broaden our knowledge regarding the exact biological functions displayed by each receptor and the physiological effects they activate as well as to obtain new drugs that would allow a more direct and specific treatment against each type of the existing disorders.

With regard to the amide/urea substitution, the results are controversial. Therefore, it cannot be concluded which group (amide/urea) leads to better affinity for MT_1/MT_2 receptors. Comparing the results obtained from series Q1, it could be assumed that the amide derivatives **(Q1-3.(I-V))** generally present better affinity for MT_1 receptor than their homologous urea derivatives **(Q1-3.(VI-IX))**. However, the affinity of urea derivatives with MT_2 is higher than the affinity of amide derivatives, provided that the aliphatic chain linked to the urea/amide group is a lineal aliphatic chain. In contrast, taking into account the results obtained for the compounds from series Q2, it could be concluded that the urea derivatives **(Q2-3.(VI-IX))** present a higher affinity than their homologous amide derivatives **(Q2-3.(I-VI))** for both receptors, provided that the aliphatic chain linked to the urea/amide group is a lineal aliphatic chain. Thus, it can be said that the introduction of a urea function, instead of the amide function existing in the MLT, does not negatively affect the affinity of the ligands for MT_1 and MT_2 receptors. In addition, since it is already well known that urea derivatives usually present better metabolic stability than amide derivatives, the urea derivatives are preferred.

Finally, the introduction of a benzene ring, replacing the aliphatic chain substituted over the amide/urea group, results in a total loss of affinity. By contrast, substitution with an aliphatic lineal chain gives better affinity values, especially when this chain is a propylamide chain, in which case, the affinity values obtained are the best of all.

As explained previously, the SAR study suggested that a six-atom length is the best distance between the methoxy group and the nitrogen atom of the *N*-alkylamide chain in order to obtain high MT_1/MT_2 affinity. Therefore, the initial pharmacophore has been redefined to include a six-atom distance between the methoxy group and *N*-alkylamide/urea group as an essential structural requirement for obtaining future series of quinoxalines such as melatonin receptor agonists. Moreover, it is currently known that the optimum distance should be obtained by a two-methylene linker in order to allow compounds to adopt a MLT-like configuration (figure 84), as explained in *Ch3 - 23.1*.[40]
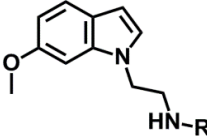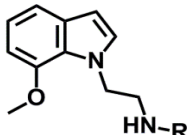

(Six atom distance)
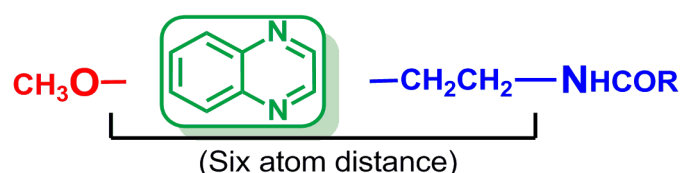
Figure 84. Future pharmacophore of quinoxalines as MT_1/MT_2 receptor agonists.

22.1.1.   *Binding affinity of indole derivatives*

As can be observed in table 13, series In1 contains the best binding affinity results, with three compounds from this series, **In1-3.X**, **In1-3.II** and **In1-3.IX** being the derivatives which present the best MT_1/MT_2 Ki values among all of the synthesized compounds.

Comparing the results obtained for any one compound in series In1 **(In1-3.R)** with its homolog in series In2 **(In2-3.R)**, where the only difference lies in the position of the methoxy group, it has been confirmed that the existing distance between the methoxy group and the first nitrogen atom of the aliphatic chain should be a six-atom distance, as in MLT.

Table 13. Values of binding affinity of indoles to MT_1/MT_2 receptors

Indoles	Name	R	MT_1 Ki(nM)±SEM	MT_2Ki(nM)±SEM
	In1-3.I	$CONHCH_2CH=CH_2$	24	5
	In1-3.II	$CONHCH_2CH_3$	6	1
	In1-3.III	$CONHCH_2CH_2CH_3$	28	2
	In1-3.IV	$CSNHCH_2CH=CH_2$	70	30
	In1-3.V	$CSNHCH_2CH_3$	36	22
	In1-3.VI	$CSNHCH_2CH_2CH_3$	35	22
	In1-3.VII	$CSNHCH_3$	29	30
	In1-3.VIII	$CSNH-cC_3H_5$	280	300
	In1-3.IX	SO_2CH_3	12	4
	In1-3.X	$COCF_3$	1	0.2
	In1-4.XI	$CONHCH_3$	4	20
	In1-4.XII	$CONH-cC_3H_5$	50	4
	In1-3.XIII	$CONH_2$	N.D.	N.D.
	In2-3.I	$CONHCH_2CH=CH_2$	190	17
	In2-3.II	$CONHCH_2CH_3$	100	5
	In2-3.III	$CONHCH_2CH_2CH_3$	330	9
	In2-3.IV	$CSNHCH_2CH=CH_2$	160	44
	In2-3.V	$CSNHCH_2CH_3$	70	13
	In2-3.VI	$CSNHCH_2CH_2CH_3$	270	57
	In2-3.VII	$CSNHCH_3$	120	74
	In2-3.VIII	$CSNH-cC_3H_5$	600	340
	In2-3.IX	SO_2CH_3	16	4
	In2-3.X	$COCF_3$	12	2
	In2-4.XI	$CONHCH_3$	75	16
	In2-4.XII	$CONH-cC_3H_5$	I	93

I: Inactive
N.D.: Non data

Classifying the information above, it can be said that it is an undeniable fact that the pharmacophore structure of future MT_1/MT_2 receptor agonists should display a six-atom distance between the methoxy group and the first nitrogen atom of the aliphatic chain. In addition, the aforementioned space should be formed by a linker containing a two-methylene chain as will now be explained below (*see Ch3 - 23.3*)

Taking into account urea/thiourea variation, when comparing the urea derivatives with their homologous thiourea derivatives in series In1, it can be observed that urea derivatives (**In1-3.I**, **In1-3.II**, **In1-3.III**, **In1-4.XI** and **In1-4.XII**) generally have higher affinity for both receptors, MT_1 and MT_2, than their homologous thioureas (**In1-3.IV**, **In1-3.V**, **In1-3.VI**, **In1-3.VII** and **In1-3.VIII**). Although this reasoning is also confirmed for the MT_2 affinities of

183

urea/thiourea derivatives of series In2, the MT_1 affinity values of these compounds cannot be explained by this hypothesis.

As stated above, the MT_1 binding affinity displayed by the urea/thiourea derivatives of series In2 is an exception to the SAR conclusions defined during this work. It is possible to wonder whether or not these "exceptions" could be explained alternatively. The first point to consider is that the "exceptions" occur in MT_1 receptor binding which, as is known, possesses a more restrictive binding site than MT_2. Secondly, it has to be taken into account that series In2 does not fulfill the six-atom distance that should exist between the methoxy group and the first *N* atom of the aliphatic chain defined in the pharmacophore. Therefore, it would be logical to assume that this fact could possibly be the reason why molecules are prevented from completely filling up the active site of MT_1, resulting in a decrease of their Ki values. Finally, in the case of thiourea derivatives, the loss of activity might be attenuated due to the fact that their sulfur atom could probably be involved in building up some hydrogen bonds with another active rest of the MT_1 receptor.

Among the substituents introduced at the end of the aliphatic chain, it can be said that the trifluoroacetic group leads to the most active compounds because compounds **In1-3.X** and **In2-3.X** show the best binding affinities within their corresponding series. Another interesting group is the methanesulfonyl group which also results in compounds with very good MT_1/MT_2 binding affinity (**In1-3.IX** and **In2-3.IX**). The insertion of a saturated lineal alkyl chain attached to both, urea or thiourea groups, generates higher MT_1/MT_2 activity than the introduction of other alkyl chains such as an allyl group or a cyclopropyl ring. Our results suggest that an ethyl group (**In1-3.II**, **In2-3.II**) is the best alkyl chain to be introduced at this position.

22.2. MT_1/MT_2 binding efficacy. [^{35}S]GTPγS binding assay

Once the binding affinity is known, the efficacy is measured via [^{35}S]GTPγS binding assay for the purpose of setting up the agonism/antagonism of drugs that presents interesting binding affinity. As explained previously, the evaluation of drug efficacy requires an assay that allows a change in the biological state to be measured. In this case, the level of G protein activation is measured by determining the binding of [^{35}S]GTPγS to Gα subunits.[152] The amount of accumulated Gα-[^{35}S]GTPγS is a consequence of the action of an activated receptor and therefore, it is used to determine the degree of agonism and the potency of compounds acting at a particular GPCR.[153] The results are obtained as potency (EC_{50}) and relative efficacy (E_{max}) values.
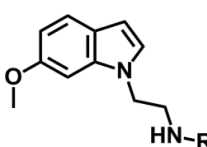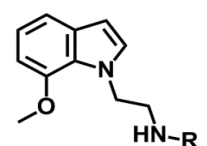
Comparing the different responses that the new derivatives produce in each receptor, it is observed that it is easier to obtain an agonist effect in receptor MT_2 than in MT_1, probably due to the more restrictive binding site that MT_1 receptor presents.

In every series, the urea derivatives present better efficacy than their analogous thiourea derivatives and therefore, the insertion of urea groups is preferred.

The series In1 contains the best efficacy results as can be observed in table 14. These results were previously expected because In1 is the only series that completely fits into the defined MT_1/MT_2 agonist pharmacophore structure. Among this series, the best efficacy at the

time of binding with MT_1/MT_2 receptors is displayed by compounds containing a lineal aliphatic urea chain and methanesulphonyl and trifluoroacetyl groups linked to the first *N* atom of the side chain. Therefore, in combination with their binding values, compounds **In1-3.X**, **In1-3.II** and **In1-3.IX** have been selected as hits in the work.

Table 14. Potency and relative efficacy values of Indoles upon binding to MT_1/MT_2 receptors

Indoles	Name	R	MT_1		MT_2	
			E_{max}	$EC_{50}(nM)$	E_{max}	$EC_{50}(nM)$
	In1-3.I	$CONHCH_2CH=C$	81%	500	94%	40
	In1-3.II	$CONHCH_2CH_3$	86%	200	89%	10
	In1-3.III	$CONHCH_2CH_2CH$	71%	700	107%	30
	In1-3.IV	$CSNHCH_2CH=CH$	50%	600	68%	200
	In1-3.V	$CSNHCH_2CH_3$	73%	400	69%	200
	In1-3.VI	$CSNHCH_2CH_2CH$	40%	200	69%	200
	In1-3.VII	$CSNHCH_3$	63%	200	65%	100
	In1-3.VIII	$CSNH-cC_3H_5$	N.T.	N.T.	N.T.	N.T.
	In1-3.IX	SO_2CH_3	99%	300	89%	40
	In1-3.X	$COCF_3$	103%	10	81%	1
	In1-4.XI	$CONHCH_3$	N.D.	N.D.	N.D.	N.D.
	In1-4.XII	$CONH-cC_3H_5$	N.D.	N.D.	N.D.	N.D.
	In1-3.XIII	$CONH_2$	N.D.	N.D.	N.D.	N.D.
	In2-3.I	$CONHCH_2CH=C$	30%	2000	78%	200
	In2-3.II	$CONHCH_2CH_3$	37%	1000	80%	60
	In2-3.III	$CONHCH_2CH_2CH$	32%	8000	66%	90
	In2-3.IV	$CSNHCH_2CH=CH$	25%	800	72%	100
	In2-3.V	$CSNHCH_2CH_3$	44%	900	71%	200
	In2-3.VI	$CSNHCH_2CH_2CH$	27%	900	70%	200
	In2-3.VII	$CSNHCH_3$	46%	700	53%	300
	In2-3.VIII	$CSNH-cC_3H_5$	N.T.	N.T.	N.T.	N.T.
	In2-3.IX	SO_2CH_3	18%	700	22%	10
	In2-3.X	$COCF_3$	29%	100	58%	20
	In2-4.XI	$CONHCH_3$	41%	600	81%	40
	In2-4.XII	$CONH-cC_3H_5$	N.T.	N.T.	N.T.	N.T.

N.T.: Not tested
N.D.: No data

The structural characteristics that have lead to the obtaining of the hits in the work are the following (figure 85):

- A distance of six atoms between the methoxy group and the first nitrogen of the side chain.

- A two-methylene linker between the central scaffold and the N atom of the side chain.

- The introduction of different functions such as an alkylurea (specially an ethylurea), a trifluoroacethyl and a methanesulfonylamide attached to the two methylene linker.
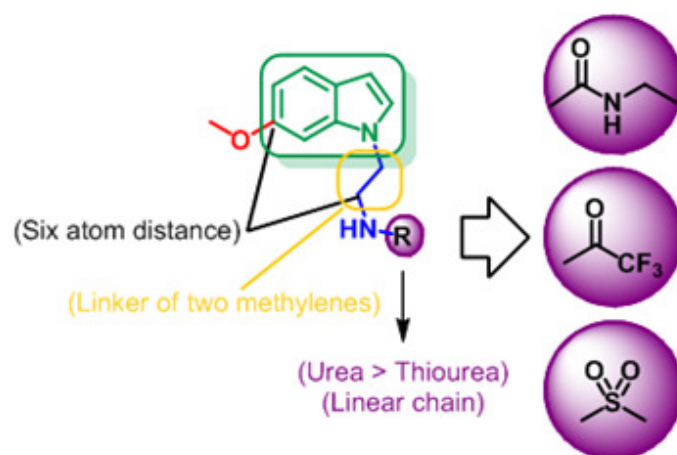
Figure 85. Structural features of the hits in the work.

XII. MOLECULAR MODELING

All the compounds used in this study (including the derivatives synthesized in the first part of this project) belong to the chemical library of the "Institute de Recherches Servier". Therefore, all of the biological data has been obtained through the same protocols (*see Ch2 - 12.3).*

As explained previously, while several MT_2 selective agonists are already known, there is no molecule identified as a MT_1 selective agonist yet. This fact results in this study being mainly centered on building a model capable of predicting the MT_1 binding affinity of newly designed structures.

23. 3D-QSAR MODELS

23.1. 3D Modeling of molecules

First, an extensive revision has been performed in order to select several molecules with great structural variation and different MT_1 binding affinity. Next, their configuration has been corrected, using the conformation of MLT molecule when it is bound to MT_1 receptor as reference (figure 86) ,[40] and they have been superimposed taking into account the central aromatic scaffold and the position of the alkyl-amide chain (or derivative) present in all molecules.
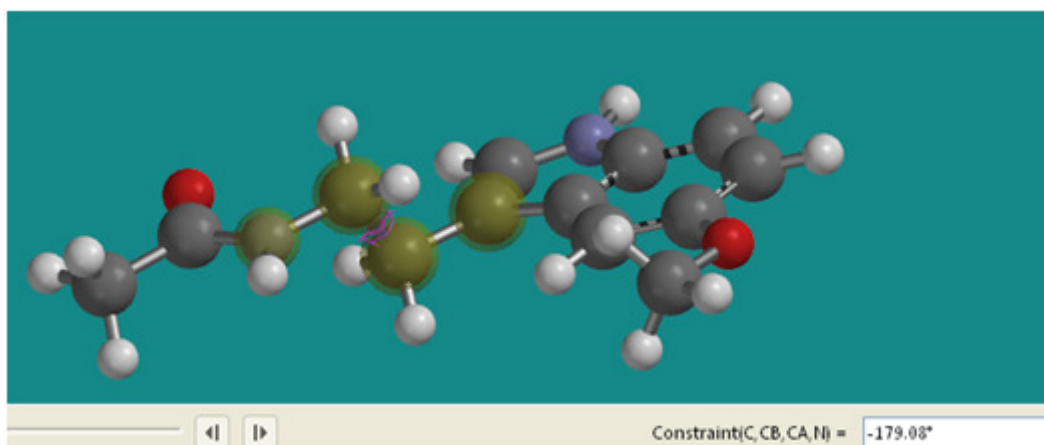
Figure 86. MLT conformation at binding site.

For a correct superposition of the molecules with MT_1-bound conformation of MLT, the ligand must contain a two-methylene linker between the central scaffold and the *N* atom of the side chain. Thus, quinoxaline derivatives synthesized at the beginning of the project have been eliminated from the study because they do not contain a two-methylene linker, and this new structural requirement has been introduced as the basis of the pharmacophore of future MT_1/MT_2 receptor agonists (*see Ch3 - 22.1.1, Ch3 - 23.3, Ch3 - XIII*).

Different sets have been created with them:

- 79_MT$_1$_agoninsts_TrainingSet: 79 molecules have been selected as the first training set of the quantitative 3D-QSAR study.

- 32_MT$_1$_active_UNA_ValSet: 32 molecules (including the indole derivatives synthesized in the first part of this project, which had known MT$_1$ binding affinity data at the time of carrying out this work) have been selected as first validation set.

- 28_MT$_1$_agonists_ValSet: 28 molecules have been selected as second validation set.

- 94_MT$_1$_agoninsts_TrainingSet: 15 molecules from 32_MT$_1$_active_UNA_ValSet have been added to 79_MT1_agoninsts_TrainingSet and used as the second training set of the quantitative 3D-QSAR study.

- 111_MT$_1$_agoninsts_TrSet_3CLASS: 111 molecules have been selected to form the training set of the qualitative 3D-QSAR study with 3 classes. The activity of these compounds has been classified as active, inactive or intermediate.

- 80_MT$_1$_agoninsts_TrSet_2CLASS: The compounds classified as active or inactive in 111_MT$_1$_agoninsts_TrSet_3CLASS have been selected to form the training set of the qualitative 3D-QSAR study with 2 classes, but not those classified as intermediate.

23.2. Create 3D-QSAR models

23.2.1. Create a quantitative 3D-QSAR model for activity prediction

23.2.1.1. MODEL 1: MLT_ALIGNED_79

First, a database has been made with 79_MT$_1$_agoninsts_TrainingSet, and PLS (partial least square) runs have been performed for both the CoMFA and CoMSIA models, obtaining the data shown in table 15.

Table 15. Results of PLS of MLTaligned_79 with CoMFA and CoMSIA models.

MLTaligned_79	Internal Q^2	Comp	R^2	rms
CoMFA	0.328	2	0.596	0.856
CoMSIA	0.371	3	0.667	0.783

-Internal Q^2: predictive power of the model, to predict compounds included in the training set.
-Comp: amount of components used by PLS to correlate all compounds according to descriptors.
-R^2: Correlation between the real value and the value predicted by the model.
-rms: root mean squared deviation.

Consequently, the MLTaligned_79_CoMSIA model has been selected because its predictive power (internal Q^2) has been higher than CoMFA's and the correlation between the real value and the value predicted by the model (R^2) has also been greater.

Next, the compounds have been represented in a graph showing their real and predicted activity values and the residue between them (figure 87). Fifteen outliers have been selected; these compounds displaced more than one point from the middle of the Residue scale.
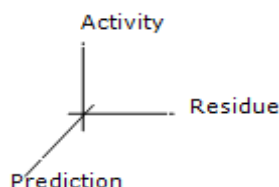
Figure 87. Graph showing the real and predicted activity of the molecules and the residue between them.

The next step uses 2/3 of the compounds to make a model, leaving the other 1/3 of the molecules as validation set in order to search for the external Q^2 value of the model. Several models have been constructed automatically and the values for the best models are shown in table 16:

Table 16. Results of the best models and their external Q^2 created from MLTaligned_79_CoMSIA model.

MLTaligned_79_CoMSIA	Internal Q^2	Comp	R^2	rms	External Q^2
KSL_33	0.446	5	0.842	0.552	0.188
KSL_46	0.383	4	0.802	0.624	0.168
KSL_8	0.459	4	0.826	0.587	0.158

-External Q^2: the predictive power of the model, to predict compounds that have been not included in the training set.

As can be concluded from the external Q^2 value, the models have not been good enough to continue with the study. Thus, the outliers have been taken out of the training set for the purpose of obtaining a better model.

23.2.1.2. MODEL 2: MLT_ALIGNED_64

The 15 outliers found in MLTaligned_79_CoMSIA model have been removed and another PLS analysis has been performed obtaining the following values (table 17):

Table 17. Results of PLS of MLTaligned_64 with CoMFA and CoMSIA models.

MLTaligned_64	Internal Q2	Comp	R^2	rms
CoMFA	0.397	3	0.744	0.661
CoMSIA	0.726	4	0.879	0.459

Taking into account the internal Q^2 and R^2 values, the MLTaligned_64_CoMSIA model has been selected, finding no outliers on it. After calculation of the external Q^2 value (table 18), MLTaligned_79_CoMSIA_ KSL_35 model has been selected.

Table 18. The best models and their external Q2 created from MLTaligned_64_CoMSIA model.

MLTaligned_64_CoMSIA	Internal Q^2	Comp	R^2	rms	External Q^2
KSL_35	0.557	4	0.903	0.416	0.833
KSL_40	0.693	5	0.913	0.394	0.808
KSL_15	0.606	4	0.884	0.420	0.792

A priori and taking into account the values shown in table 18, the MLTaligned_64_CoMSIA_KSL_35 looks like a good, suitable model. Therefore, it should be validated.

A first validation has been made using 32_MT1_active_UNA_ValSet. The predicted data has been compared with the real activity data and represented in the graph 1.
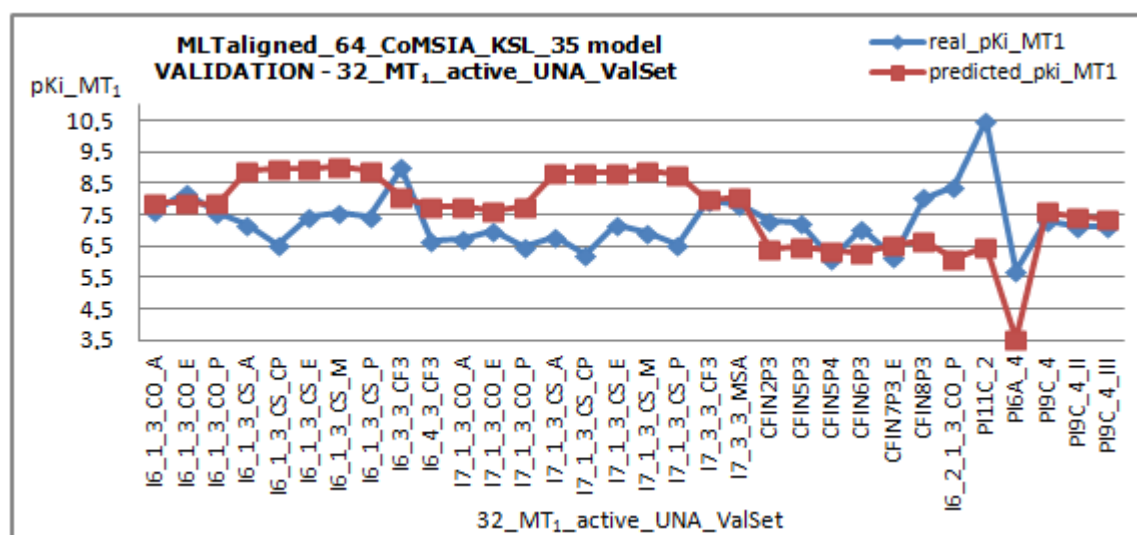
Graph 1. Comparison of the real biological activity of 32_MT$_1$_active_UNA_ValSet and the biological activity predicted with the MLTaligned_64_CoMSIA_KSL_35 model.

As can be observed in the graph, the obtained model has not been good enough to correctly predict the affinity of the ValSet compounds. Thus, the 15 worst predicted compounds have been added to the training set and new models have been performed.

23.2.1.3.   MODEL 3: MLT_ALIGNED_UNA_94

The 94_MT1_agoninsts_TrainingSet has been used to construct the CoMFA and CoMSIA models (table 19).

Table 19. Results of PLS of MLTaligned_UNA_94 with CoMFA and CoMSIA models.

MLTaligned_UNA_94	Internal Q2	Comp	R^2	rms
CoMFA	0.208	5	0.717	0.716
CoMSIA	0.375	3	0.623	0.818

Taking into account the internal Q^2 and R^2 values, MLTaligned_UNA_94_CoMSIA model has been selected as the best model and 18 outliers have been found. No further study has been performed on this model because the already obtained data did not show a good perspective, but another model has been created taking out the new outliers from the training set.

23.2.1.4.   MODEL 4: MLT_ALIGNED_UNA_76

The 18 outliers found in MLTaligned_UNA_94_CoMSIA model have been removed from the table and another PLS analysis has been performed, obtaining the following values (table 20):

Table 20. Results of PLS of MLTaligned_UNA_76 with CoMFA and CoMSIA models.

| MLTaligned_UNA_76	Internal Q^2	Comp	R^2	rms
CoMFA	0.435	6	0.473	0.846
CoMSIA	0.651	4	0.823	0.500

Taking into account the internal Q^2 and R^2 values, MLTaligned_UNA_76_CoMSIA model has been selected, finding no outliers on it. After calculation of the external Q^2 value (table 21), MLTaligned_UNA_76_CoMSIA_KSL_33 model has been selected.

Table 21. Best models and their external Q^2 created from MLTaligned_UNA_76_CoMSIA model.

MLTaligned_UNA_76_CoMSIA	Internal Q^2	Comp	R^2	rms	External Q^2
KSL_33	0.513	4	0.840	0.432	0.772
KSL_22	0.514	4	0.847	0.472	0.765
KSL_29	0.478	6	0.884	0.409	0.740
KSL_32	0.601	6	0.916	0.346	0.730
KSL_5	0.625	4	0.854	0.461	0.727

A priori and taking into account the values shown in table 21, the MLTaligned_UNA_76_CoMSIA_KSL_33 model looks like a good and suitable model. Therefore, it should be validated.

A first validation has been made using the other 17 molecules from 32_MT1_active_UNA_ValSet that have not been used in 94_MT1_agoninsts_TrainingSet, and the obtained data is represented in graph 2.
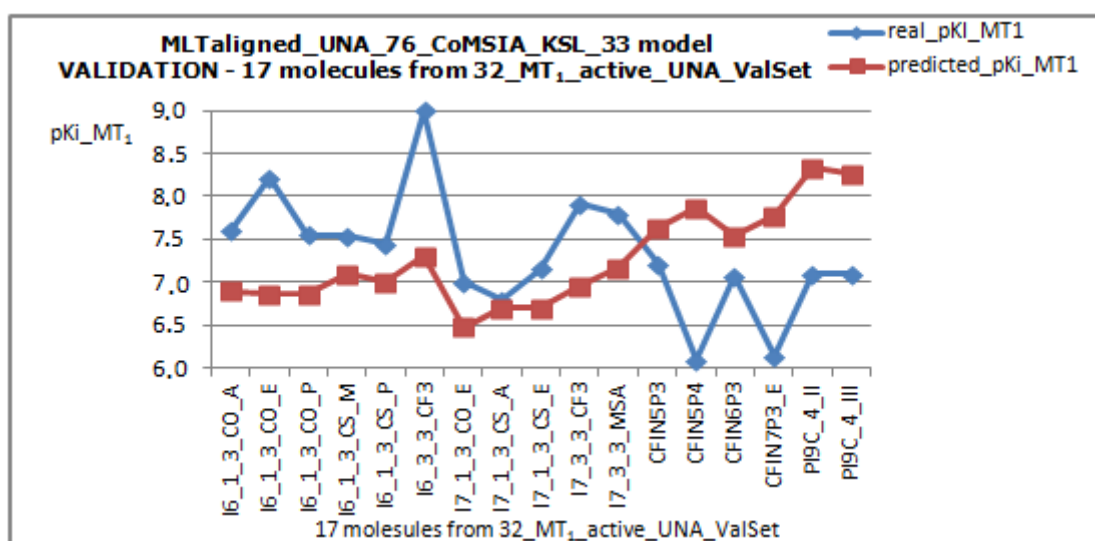
Graph 2. Comparison of the real biological activity of 17 molecules from 32_MT₁_active_UNA_ValSet set and the biological activity predicted with the MLTaligned_UNA_76_CoMSIA_KSL_33 model.

The obtained results are ambiguous due to the fact that the model has been able to predict the variation on the tendency of the affinity but not the accurate values, therefore, another validation has been carried out with the 28_MT1_agonists_ValSet (graph 3).
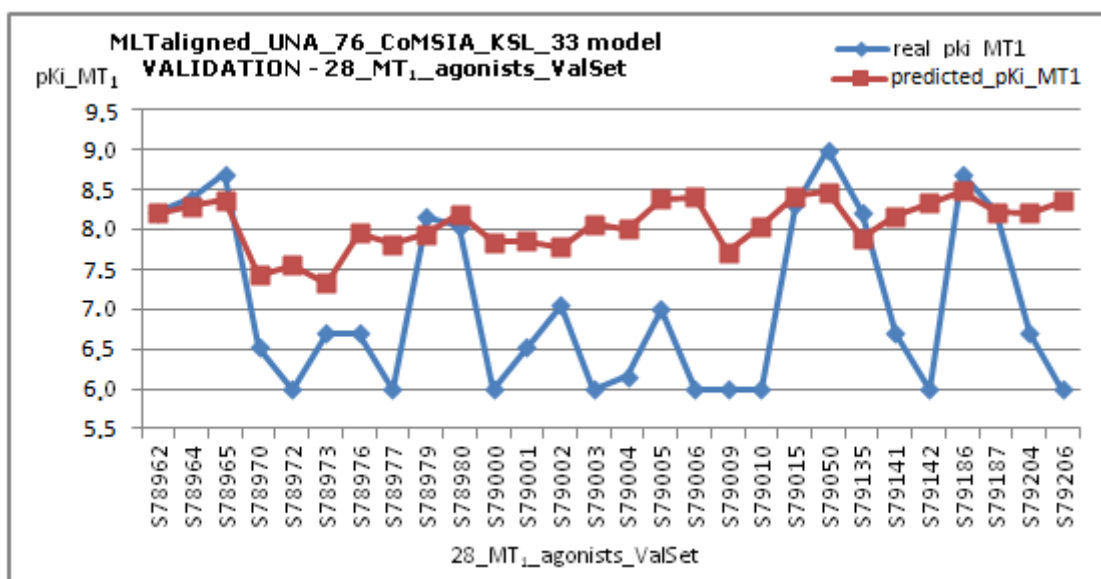
Graph 3. Comparison of the real biological activity of the molecules in 28_MT$_1$_agonists_ValSet and the biological activity predicted with the MLTaligned_UNA_76_CoMSIA_KSL_33 model.

As can be concluded from the graph, the model is not capable of predicting the inactive compounds. This fact could be explained by a lack of information in the training set. The fact is that there is not very much biological data regarding the inactive compounds. Therefore, most of them are not included in the training set, thus, the model is not able to predict this type of compounds. The solution could be to generate a model that is able to first classify compounds as active/inactive and then create another model capable of predicting the biological activity of the active ones.

23.2.2. Create a qualitative 3D-QSAR model for activity class prediction

23.2.2.1. CLASIFFICATION IN 3 CLASSES

23.2.2.1.1. Model 1: C3_MLTaligned_UNA_111_CoMFA_SIMCA

First, 111_MT1_agoninsts_TrSet_3CLASS has been used as training set and their biological data has been placed by class as follows:

-2 (Active) has been defined for the compounds with pki≥8

-0 (Inactive) has been defined for the compounds with pki≤6

-1 (Medium activity) has been defined for the compounds with 6<pki<8

The values of CoMFA regions have then been calculated and a SIMCA analysis has been run.

In table 22 the number of compounds that has been correctly predicted in each class and the number of compounds belonging to that class are compared. As can be observed in the table below, this model predicts the compounds from the intermediate class better than the compounds from the active and inactive classes.

Table 22. Correctly predicted compounds in C3_MLTaligned_UNA_111_CoMFA_SIMCA model.

CLASS	SAMPLE	CORRECT_CoMFA
2	41	28
0	39	29
1	31	31

The table 23 represents the "distance between categories" matrix that shows the differences between the real and predicted values as well as the power of the model to discriminate between different categories. The values of the diagonal of the matrix are expected to be as small as possible meaning that there are small differences between real and predicted values. The other values are expected to be as high as possible meaning that the model can discriminate between different categories. Thus, it can be said that in this case the obtained model has been not good enough.

Table 23. Distances between categories of the model C3_MLTaligned_UNA_111_CoMFA_SIMCA.

	ActualCat2	ActualCat0	ActualCat1
2 ProjectedCat2	10.764	16.163	11.852
0 ProjectedCat0	30.229	19.074	29.754
1 ProjectedCat1	14.705	17.561	12.081

Finally, an automatically generated report where the predicted activity and the real activity of each compound have been described has then been checked. A total of 13 compounds have been predicted in an incorrect category and therefore, marked as outliers.

23.2.2.1.2. Model 2: C3_MLTaligned_UNA_98_CoMFA_SIMCA

The 13 outliers found in C3_MLTaligned_UNA_111_CoMFA_SIMCA model have been removed and another SIMCA analysis has been performed, obtaining the values shown in tables 24 and 25.

Table 24. Correctly predicted compounds in C3_MLTaligned_UNA_98_CoMFA_SIMCA model.

CLASS	SAMPLE	CORRECT_CoMFA
2	35	28
0	32	29
1	31	28

Table 25. Distances between categories of the model C3_MLTaligned_UNA_98_CoMFA_SIMCA.

	ActualCat2	ActualCat0	ActualCat1
2 ProjectedCat2	10.328	18.210	11.855
0 ProjectedCat0	32.603	18.940	32.102
1 ProjectedCat1	14.771	19.300	12.099

As can be observed in table 24, this model predicts all the classes with a similar accuracy; thus, this point shows an improvement in comparison with the previous model. Moreover, it can be said that table 25 also expresses that improvement. Although the values of the diagonal of the matrix that shows the distances between categories are similar to those obtained in the previous model and the distance between class 1 and class 2 shows no significant change, the

distance between class 1 and class 0 has become greater. This is optimistic data with regard to obtaining a model capable of identifying the inactive compounds.

Finally, in this model 3 outliers have been found.

A first validation has been made using 32_MT1_active_UNA_ValSet, and the data obtained has been compared with the real activity data and represented in graph 4.
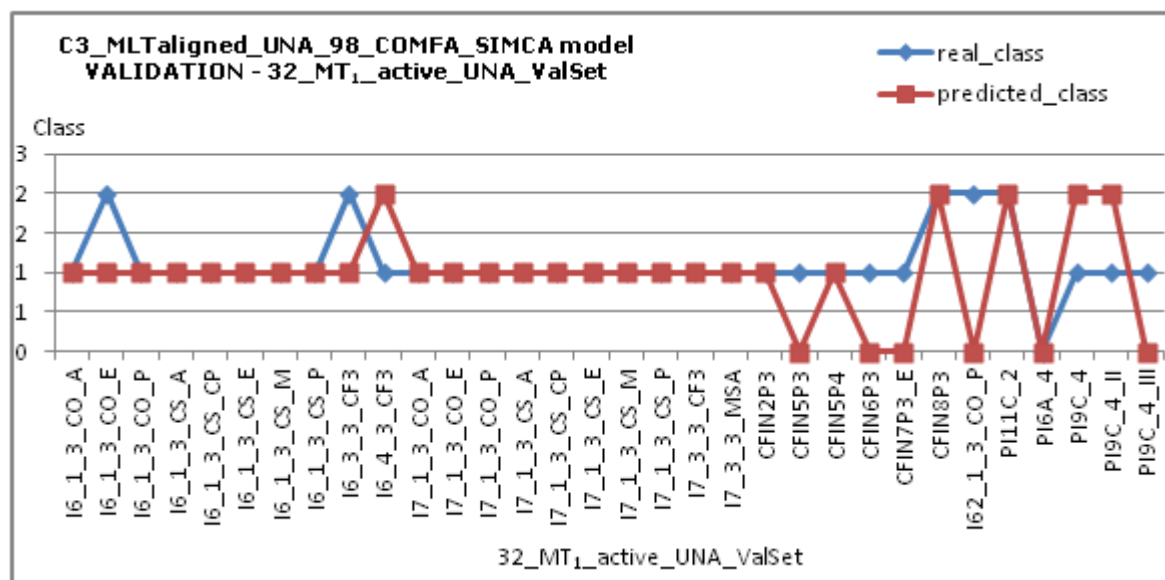
Graph 4. Comparison of the real biological activity of the molecules in 32_MT$_1$_active_UNA_ValSet and the biological activity predicted with the C3_MLTaligned_UNA_98_CoMFA_SIMCA model.

Another validation has been carried out using 28_MT1_agonists_ValSet (graph 5).
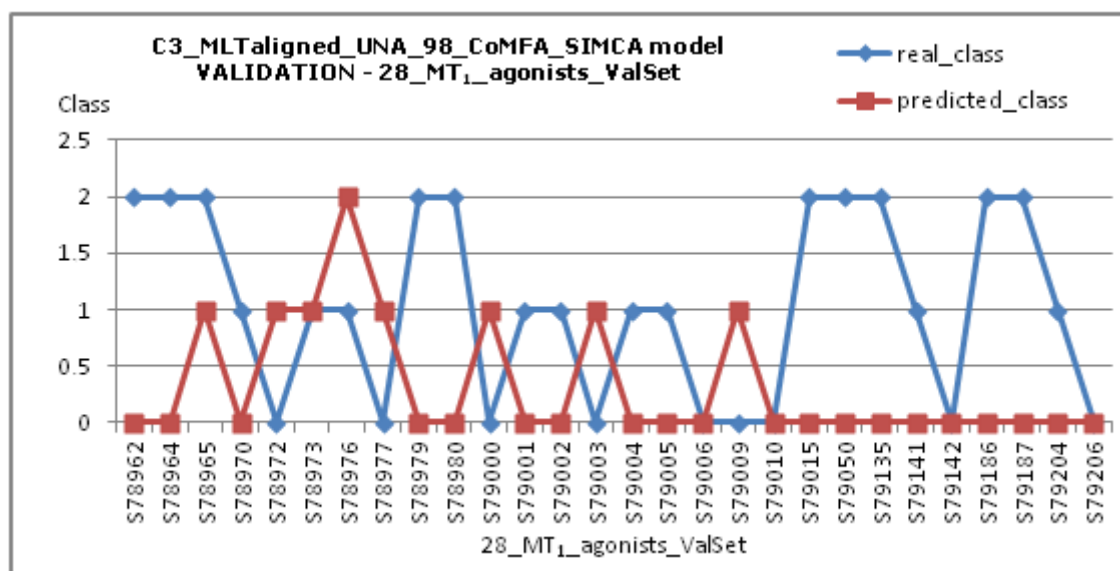
Graph 5. Comparison of the real biological activity of the molecules in 28_MT$_1$_agonists_ValSet and the biological activity predicted with the C3_MLTaligned_UNA_98_CoMFA_SIMCA model.

As can be observed in graphs 4 and 5, this model is not good enough for classifying these compounds by class.

23.2.2.1.3. Model 3: C3_MLTaligned_UNA_95_CoMFA_SIMCA

The 3 outliers found in MLTaligned_UNA_98_CoMFA_SIMCA model have been removed and another SIMCA analysis has been performed, obtaining the following values (tables 26 and 27).

Table 26. Correctly predicted compounds in C3_MLTaligned_UNA_95_CoMFA_SIMCA model.

CLASS	SAMPLE	CORRECT_CoMFA
2	34	29
0	31	29
1	30	25

Table 27. Distances between categories of the model 3_MLTaligned_UNA_95_CoMFA_SIMCA

	ActualCat2	ActualCat0	ActualCat1
2 ProjectedCat2	10.212	18.309	12.169
0 ProjectedCat0	32.887	18.928	32.725
1 ProjectedCat1	14.789	19.591	13.393

Comparing this model with the MLTaligned_UNA_98_CoMFA_SIMCA model, it can be said that the difference between the intermediate class and the other ones has decreased probably due to the "over-learning" of the model.

23.2.2.1.4. Model 4: C3_MLTaligned_UNA_111_CoMSIA_SIMCA

In this model, the process used in MLTaligned_UNA_111_CoMFA_SIMCA has been repeated but in this case, applying the CoMSIA values instead of the CoMFA ones (tables 28 and 29).

Table 28. Correctly predicted compounds in C3_MLTaligned_UNA_111_CoMSIA_SIMCA model.

CLASS	SAMPLE	CORRECT_CoMSIA
2	41	36
0	39	31
1	31	21

Table 29. Distances between categories of the model C3_MLTaligned_UNA_111_CoMSIA_SIMCA

	ActualCat2	ActualCat0	ActualCat1
2 ProjectedCat2	25.456	40.442	31.271
0 ProjectedCat0	59.877	44.891	59.035
1 ProjectedCat1	41.795	47.795	32.400

This new model makes a better prediction for the active compounds than for the inactive and intermediate compounds. Thus, this model does not show a good perspective for obtaining a model capable of identifying the inactive compounds. Moreover, taking into account the diagonal of the "distance between categories" matrix, it can be said that there are big differences between real and predicted values in this model.

23.2.2.2.   CLASSIFICATION IN 2 CLASSES

23.2.2.2.1. Model 5: C2_MLTaligned_UNA_80_CoMFA_

First, 80_MT1_agoninsts_TrSet_2CLASS has been used as training set and its biological data has been placed by class as follows:

-2 (Active) has been defined for the compounds with pki≥8

-0 (Inactive) has been defined for the compounds with pki≤6

The values of CoMFA regions have been then calculated and a SIMCA analysis has been run. The results have been placed in tables 30 and 31.

Table 30. Correctly predicted compounds in C2_MLTaligned_UNA_80_CoMFA_SIMCA model.

CLASS	SAMPLE	CORRECT_CoMFA
2	41	41
0	39	31

Table 31. Distances between categories of the model C2_MLTaligned_UNA_80_CoMSIA_SIMCA

	ActualCat2	ActualCat0
2 ProjectedCat2	11.066	40.442
0 ProjectedCat0	30.988	44.891

In this model, 8 outliers have been found and removed in order to generate a better model. As can be deduced from the tables above, this model has problems for identifying the inactive compounds.

23.2.2.2.2. Model 6: C2_MLTaligned_UNA_72_CoMFA_SIMCA

After removing the 8 outliers found in C2_MLTaligned_UNA_80_CoMFA_SIMCA model, another SIMCA analysis has been run obtaining the results shown in tables 32 and 33.

Table 32. Correctly predicted compounds in C2_MLTaligned_UNA_72_CoMFA_SIMCA model.

CLASS	SAMPLE	CORRECT_CoMFA
2	41	41
0	31	31

Table 33. Distances between categories of the model C2_MLTaligned_UNA_72_CoMSIA_SIMCA

	ActualCat2	ActualCat0
2 ProjectedCat2	11.115	40.442
0 ProjectedCat0	33.887	44.891

In this model, the inactive compounds are identified better than in the previous one and no outliers have been found on it.

28_MT1_agonists_ValSet and 32_MT1_active_UNA_ValSet has been used to validate this model (graph 6) but the desired results have not been achieved; the model made more errors than goals predicting the active and inactive compounds of the validation sets.

Graph 6. Comparison of the real biological activity of the 28 MT_1 agonists from Lille laboratory and 32 MT_1 agonists from UNA laboratory validation set and the biological activity predicted with the C2_MLTaligned_UNA_72_CoMFA_SIMCA model.

23.2.2.2.3. Model 7: C2_MLTaligned_UNA_80_CoMSIA_SIMCA

In this model, the process used in C2_MLTaligned_UNA_80_CoMFA_SIMCA has been repeated, but in this case, the CoMSIA values have been applied instead of the CoMFA ones (tables 34 and 35).

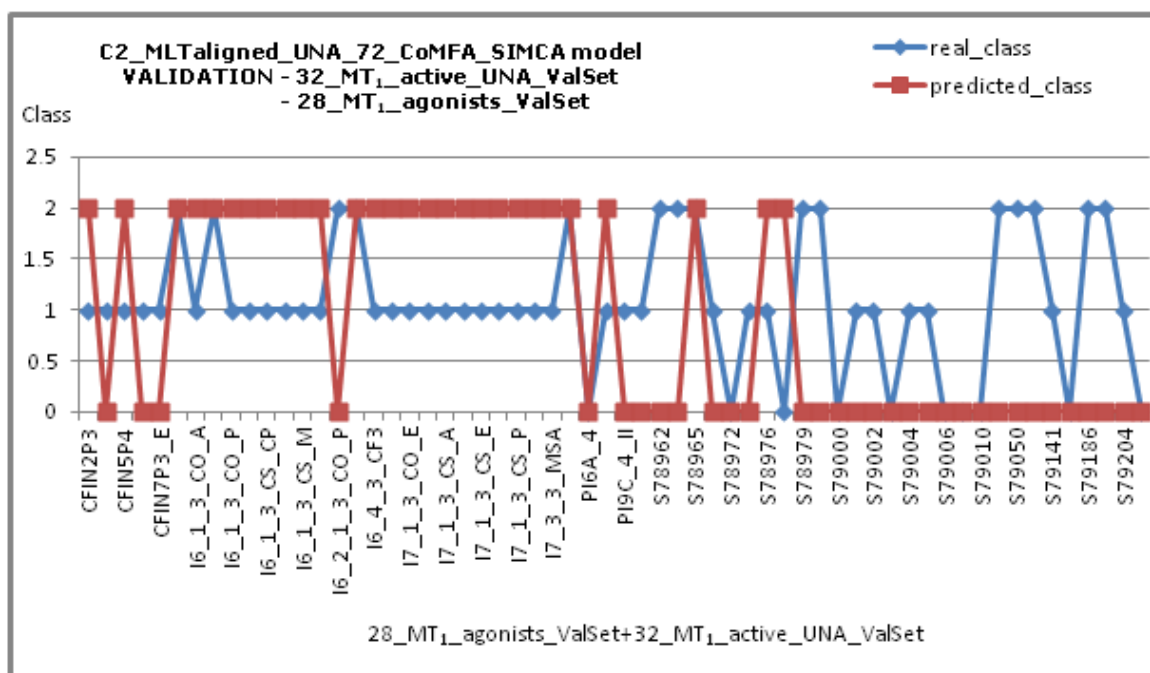**Table 34.** Correctly predicted compounds in C2_MLTaligned_UNA_80_CoMSIA_SIMCA model.

CLASS	SAMPLE	CORRECT_CoMSIA
2	41	41
0	39	31

Table 35. Distances between categories of the model C2_MLTaligned_UNA_80_CoMSIA_SIMCA.

	ActualCat2	ActualCat0
2 ProjectedCat2	25.783	41.037
0 ProjectedCat0	60.816	45.484

As can be observed in the table 35, this model predicts the active compounds better than the inactive ones and the values in the diagonal of "distances between categories" is higher than in C2_MLTaligned_UNA_80_CoMFA_SIMCA model; therefore, this strategy has been discarded and another one has been designed.

23.3. Observations and conclusions

Although this study has not provided any predicting model, three main conclusions have been suggested from the superimposition of the compounds with the MLT molecule:

1. As explained previously (*see Ch3 - 23.1*), quinoxaline derivatives synthesized at the beginning of the project have been eliminated from the study because they do not contain a two-methylene linker and therefore, it is not possible to superimpose them with the MT_1-bound conformation of MLT (figure 88). This fact can also be the reason why quinoxaline derivatives present a very low binding affinity to MT_1/MT_2 receptors, even when they present a six-atom distance between the methoxy group and the first *N* atom of the side chain.
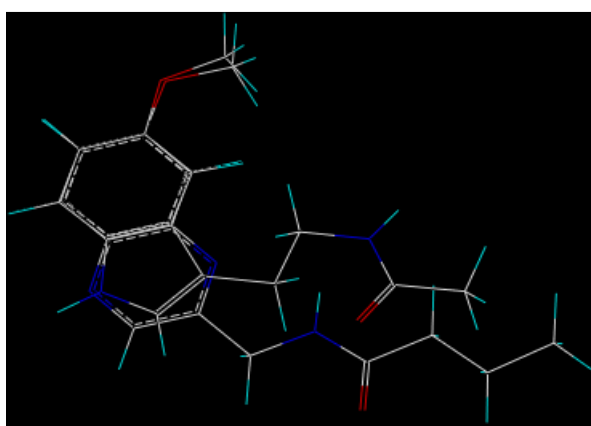
Figure 88. Superimposition of the MLT and compound **Q3-3.I**

2. The methoxy group in series "In2" cannot be correctly superimposed with the MLT due to the fact that its methoxy group is at 7-position of the indole ring, instead of being at 6-position, which is the MLT-like position (figure 89).
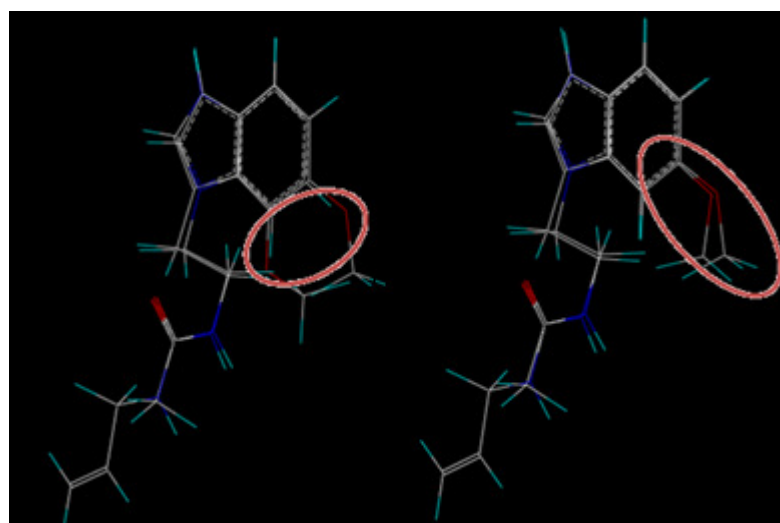
Figure 89. Superimposition of the MLT and compound **In2-3.IV** first and MLT and **In1-3.IV** then.

3. An enlargement or shortening of the two-methylene linker would lead to a displacement and a change on the orientation of the carbonyl group, which is of vital importance in order to form hydrogen bonds with the binding sites of MT_1 and MT_2 receptors and thus, the affinity would decrease (figure 90). Therefore, it has been suggested that a new pharmacophoric structure should contain a two-methylene linker as a structural requirement.
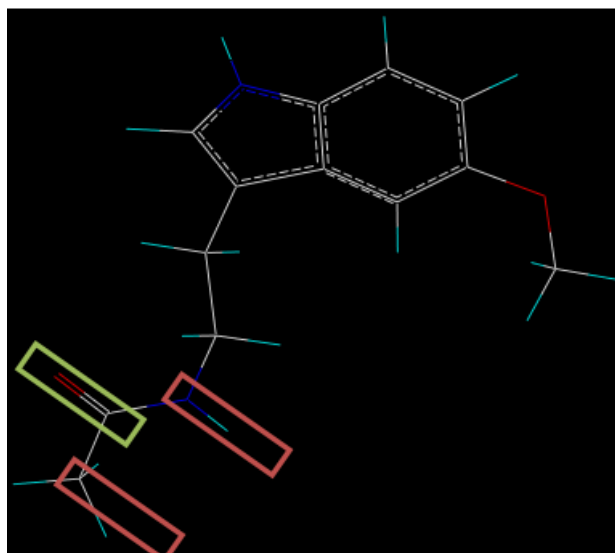
Figure 90. Structure of MLT highlighting the position and orientation of its carbonyl (green) and the position and orientation of this group in molecules with larger or shorter linkers (red).

24. 2D BAYESIAN MODEL – FINGERPRINTS

24.1. Database selection

First, an extensive revision has been carried out at the chemical library of the *"Institute de Recherches Servier"* in order to select several molecules with great structural variation and diverse MT_1 binding affinity. Finally, "MT1initretva.sdf" file has been selected, which contains 3563 MT_1 agonists with different biological activity and already randomly labeled "training/validation" molecules.

All the data contained in "MT1initretva.sdf" file has been divided into two new different files according to the "training/validation" label that every compound possessed.

• MT1_training_Set with 2376 compounds

• MT1_Validation_Set with 1187 compounds

Two other validation sets have been also used:

• 32_MT1_active_UNA_ValSet: 32 molecules (including those indole derivatives which have been synthesized in the first part of this project, and that had a known MT_1 binding affinity data at the time of doing this work) have been selected as first validation set.

• Comp1_ValSet: A file containing 623 MT_1 agonists with a great structural variation and a broad range of measured binding affinities.

Finally, another file has been created, "MT1predictions35", and 35 new compounds have been placed in it. This group contains several molecules already synthesized by UNA group which did not contain a known binding affinity data at the time of doing this experiment. In addition, some other compounds have been virtually designed and added to the file with the aim of synthesizing them in the case of being assigned as active compounds by the predicting model.

This new file has been assayed by the best model of the study for the prediction of their MT_1 binding affinity. Next, all compounds have been classified as active/inactive.

Different trials have been performed changing the classification parameters as follows:

- pKi>6 Active / pKi<6 Inactive
- pKi>7 Active / pKi<7 Inactive
- pKi>8 Active / pKi<8 Inactive
- pKi>7 Active / pKi<6 Inactive
- pKi>8 Active / pKi<6 Inactive
- pKi>8 Active / pKi<7 Inactive

24.2. Description of molecular features and model construction

In the first step of this study, the potency of the different fingerprints has been evaluated according to the ROC scores from the training set. This first approach has been made using all the available fingerprint types one by one and setting the active/inactive class separation at 6 pki (table 36).

Table 36. Characterization of the 2D molecules placed in the training set at pki 6, with different FP types. The best fingerprints of each type are highlighted in grey.

MODELS	ROC AUC	MODELS	ROC AUC	MODELS	ROC AUC	MODELS	ROC AUC
ECFC_20	0.920	ECFP_10	0.918	FCFP_16	0.917	LCFP_20	0.918
ECFC_18	0.920	EPFC_10	0.887	FCFP_14	0.916	LCFP_18	0.918
ECFC_16	0.920	EPFP_10	0.866	FCFP_12	0.915	LCFP_16	0.918
ECFC_14	0.920	FCFC_20	0.916	FPFC_10	0.888	LCFP_14	0.918
ECFC_12	0.919	FCFC_18	0.916	FPFP_10	0.866	LCFP_12	0.916
ECFC_10	0.918	FCFC_16	0.916	LCFC_20	0.919	LCFP_10	0.915
ECFP_20	0.920	FCFC_14	0.916	LCFC_18	0.919	LCFC_20	ERROR
ECFP_18	0.920	FCFC_12	0.915	LCFC_16	0.918	LPFP_20	ERROR
ECFP_16	0.920	FCFC_10	0.913	LCFC_14	0.918	SEFP_20	0.931
ECFP_14	0.920	FCFP_20	0.917	LCFC_12	0.918	SEFP_18	0.930
ECFP_12	0.918	FCFP_18	0.916	LCFC_10	0.917	SEFP_16	0.929

Next, several combinations have been made between the selected fingerprint types in order to determine the best characterization of the molecules in the training set (table 37). Finally, the best models have been selected and used for predicting the validation sets (table 38).

Table 37. Combination of fingerprints that better characterized the molecules in the training set.

BM_act6	AUC	BM_act7	AUC
SEFP_20	0.931	ECFC14_SEFP20	0.929
ECFC14_SEFP20	0.930	FCFC14_SEFP20	0.928
FCFC14_SEFP20	0.930	LCFC18_SEFP20	0.928
BM_act8	**AUC**	**BM_act7_inact6**	**AUC**
ECFP14_SEFP20	0.927	SEFP_20	0.965
ECFP14_FCFP16_SEFP20	0.927	ECFC14_SEFP20	0.965
FCFP16_SEFP20	0.926	FCFP16_SEFP20	0.964
BM_act8_inact7	**AUC**	**BM_act8_inact6**	**AUC**
ECFP14_FCFP16_SEFP20	0.958	ECFP14_LCFP14_SEFP20	0.980
ECFP14_SEFP20	0.959	ECFP14_SEFP20	0.981
FCFP16_SEFP20	0.958	FCFP16_SEFP20	0.980
		SEFP20	0.980

Table 38. Prediction of validation sets with the best models of each range.

BM_act6	AUC			BM_act7	AUC		
BEST MODELS	Val_1	Val_2	Val_3	BEST MODELS	Val_1	Val_2	Val_3
SEFP20	0.917	0.967	0.695	ECFC14_SEFP20	0.914	0.908	0.723
ECFC14_SEFP20	0.918	0.973	0.705	FCFC14_SEFP20	0.913	0.887	0.734
FCFC14_SEFP20	0.917	0.973	0.704	LCFC18_SEFP20	0.915	0.898	0.730
BM_act8	AUC			BM_act7_inact6	AUC		
BEST MODELS	Val_1	Val_2	Val_3	BEST MODELS	Val_1	Val_2	Val_3
ECFP14_SEFP20	0.920	0.949	0.761	SEFP20	0.909	0.930	0.742
ECFP14_FCFP16_SEFP20	0.922	0.943	0.771	ECFC14_SEFP20	0.909	0.953	0.735
FCFP16_SEFP20	0.922	0.938	0.778	FCFP16_SEFP20	0.909	0.936	0.747
BM_act8_inact7	AUC			BM_act8_inact6	AUC		
BEST MODELS	Val_1	Val_2	Val_3	BEST MODELS	Val_1	Val_2	Val_3
ECFP14_FCFP16_SEFP20	0.924	0.946	0.764	ECFP14_LCFP14_SEFP20	0.914	0.959	0.770
ECFP14_SEFP20	0.923	0.946	0.758	ECFP14_SEFP20	0.916	0.951	0.767
FCFP16_SEFP20	0.924	0.951	0.774	FCFP16_SEFP20	0.917	0.943	0.783
				SEFP20	0.917	0.954	0.780

Val_1: MT$_1$_Validation_Set / Val_2: 32_MT$_1$_active_UNA_ValSet / Val_3: Comp1_ValSet.

In general, all the obtained models are quite good. Finally, BM_act8_inact6_SEFP_20 has been selected as the best predictive model because it provides the best results within the validation set "Comp1_ValSet" which contains the greatest molecular variation. Moreover, this model is also a very good predicting model within the two other validation sets, "MT$_1$_Validation_Set" and "32_MT$_1$_active_UNA_ValSet". This model has been used for the prediction of new molecules (table 39). It should be noted that all the obtained models use fingerprints based on "Sybyl atom type" and as reported in the 3D-QSAR experiments (*see Ch3 – 23.3)*, this data is not enough so as to achieve a predicting model. Therefore, every result obtained from these models should be confirmed by another assay, in this case, KNN map.

24.3. Prediction of new proposed molecules

35 molecules placed on the file "MT1predictions35" have been used and tested by the BM_act8_inact6_SEFP_20 model (table 39). Since, in this group, there are several compounds synthesized by UNA group, which did not contain a known binding affinity data at the time of carrying out this experiment, they were a good option for confirming the validity of this predicting model as soon as the biological results arrived. As explained previously, in addition to these molecules, some other compounds have been virtually designed, and added to the file with the aim of synthesizing them in the case of being assigned as active compounds by the predicting model.

As stated above, all information obtained through this model should be confirmed by KNN maps.

Table 39. Prediction of new, virtually designed molecules by BM_act8_inact6_SEFP_20.

| Indoles	Prediction	Indoles	Prediction
In1-3COCF3	Act	**In2-3CSCF3**	inact
In1-4.XII	Act	**In2-3CSH**	inact
In1-3COH	Act	**I743CF3**	act
In1-4.XI	Act	**I753COA**	inact
In1-3CSCF3	Act	**I753COCF3**	inact
In1-3CSH	Inact	**I754COM**	inact
I623	Inact	**I753COE**	inact
I624N	Inact	**I753COH**	inact
I624O	Inact	**I754COM**	act
I624S	Inact	**I753COP**	inact
I625N	Inact	**I753CSA**	inact
I625O	Inact	**I753CSCF3**	inact
I625S	Inact	**I753CScP**	act
In1-3.IX	Act	**I753CSE**	inact
In2-3COCF3	Inact	**I753CSH**	inact
In2-4.XII	Inact	**I753CSM**	inact
In2-3COH	Inact	**I753CSP**	act
In2-4.XI	Act		

25. KOHONEN NEURAL NETWORK (KNN)

Due to the impossibility of obtaining good predicting models based on QSAR and the low reliability of fingerprints, the idea of analyzing the vdW surface of the molecules arose. The great amount of information that should be manipulated forces the use of KNN, even though it is a completely new technique for the group. Thus, in this study, a large number of trials have been carried out in order to learn how to manipulate, parameterize and analyze the application and its results. Therefore, only the most important assays have been reported.

25.1. Assay 1: MOL_11_DIM_50

25.1.1. Database generation

First, 11 MLT analogs have been selected in order to learn how to manipulate the software and analyze the results. Next, the KNN_11_MOL set has been created. The vdW surface of each molecule has been calculated with 183.9 mol/$Å^2$ density of dots (figure 91). Finally, the input file has been generated, placing all the information (coordination and dots) of each molecule in it.
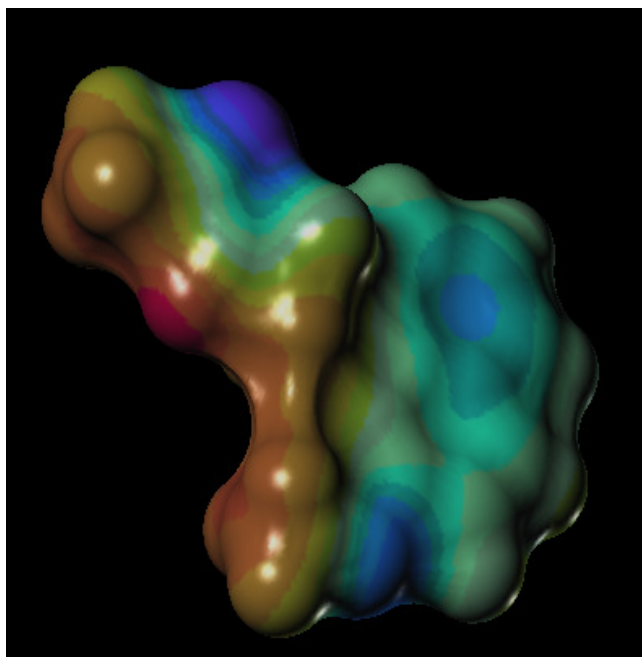
Figure 91. A picture showing the vdW calculations of the chemical structure of agomelatine with 183.9 mol/$Å^2$ density of dots.

25.1.2. Selection of initial parameters

Before creating a map, different parameters have to be configured. In this case, the first map has been created using MLT molecule as both, reference and test file (table 40). Next, these parameters have been changed for the purpose of obtaining a map with a smaller error. First, the changes have been made one by one (table 41) and in the next step, more than one parameter has been changed each time (table 42). The map with the smallest error has been selected and referred to as **MAP_0** (MLT_41192_map01_50L2.cod; table 4q); its parameters are referred to as **P0**.

Table 40. First trial to parameters selection in assay MOL_11_DI_50

Number of trials	100
Input data file name	MLT_41192.dat
Input test file name	MLT_41192.dat
Output map name	MLT_41192_map01.cod
X dimensión	50
Y dimensión	50
Topology type	Rect
Neighborhood	Bubble
Training length of 1st part	6
Training rate of 1st part	0.05
Radius of 1st part	35
Training length of 2nd part	15
Training rate of 2nd part	0.01
Radius of 2nd part	15
ERROR	**1.566830**

Table 41. First change of parameters in assay MOL_11_DIM_50.

From MLT_41192_map01.cod	to…	MAP NAME	ERROR
Topology type	Hexa	MLT_41192_map01_hexa.cod	1.564530
Neighborhood	gaussian	MLT_41192_map01_gaus.cod	1.590763
Training length of 1st part	20	MLT_41192_map01_20L1.cod	1.567794
Training length of 1st part	20	MLT_41192_map01_2L1.cod	1.575290
Training rate of 1st part	0.1	MLT_41192_map01_01Rt1.cod	1.559479
Training rate of 1st part	0.01	MLT_41192_map01_001Rt1.cod	1.580813
Radius of 1st part	50	MLT_41192_map01_50Rd1.cod	1.564245
Radius of 1st part	15	MLT_41192_map01_15Rd1.cod	1.571196
Training length of 2nd part	**50**	**MLT_41192_map01_50L2.cod (MAP_0)**	**1.547186**
Training length of 2nd part	5	MLT_41192_map01_5L2.cod	1.570553
Training rate of 2nd part	0.1	MLT_41192_map01_01Rt2.cod	1.554943
Training rate of 2nd part	0.001	MLT_41192_map01_0001Rt2.cod	1.572482
Radius of 2nd part	35	MLT_41192_map01_35Rd2.cod	1.566047
Radius of 2nd part	5	MLT_41192_map01_5Rd2.cod	1.572050

Table 42. Second change of parameters in assay MOL_11_DIM_50.

From MLT_41192_map01.cod	to…	MAP NAME	ERROR
Topology type	hexa	MLT_41192_map02.cod	1.828723
Training rate of 1st part	0.1		
Radius of 1st part	50		
Training length of 2nd part	50		
Training rate of 2nd part	0.1		
Radius of 2nd part	35		
From MLT_41192_map01.cod	to…	MAP NAME	ERROR
Topology type	hexa	MLT_41192_map03.cod	1.573001
Training rate of 1st part	0.075		
Radius of 1st part	42.5		
Training length of 2nd part	32.5		
Training rate of 2nd part	0.05		
Radius of 2nd part	25		
From MLT_41192_map02.cod	to…	MAP NAME	ERROR
Topology type	rect	MLT_41192_map02_rect,cod	1.751403
From MLT_41192_map03.cod	to…	MAP NAME	ERROR
Topology type	rect	MLT_41192_map03_rect,cod	1.752541

25.1.3.   MAP_1 construction and input file generation

Next, a real map has been constructed starting with the previously selected parameters, **P0**, and using MLT molecule as reference and the KNN_11_MOL file as test file (table 43). Different maps have been created in this step, changing the initial parameters. No improvement has been achieved so this map has been termed **MAP_1** and its parameters **P1**.

Table 43. First trial for MAP_1 construction in assay MOL_11_DIM_50

Number of trials	100
Input data file name	MLT_41192.dat
Input test file name	KNN_11_MOL.dat
Output map name	**MLT_11_map01.cod (MAP_1)**
X dimensión	50
Y dimensión	50
Topology type	Rect
Neighborhood	bubble
Training length of 1st part	6
Training rate of 1st part	0.05
Radius of 1st part	35
Training length of 2nd part	50
Training rate of 2nd part	0.01
Radius of 2nd part	15
RANDOM	68
ERROR	**9.126225**

At this point, the aim is to obtain a single file in a matrix format where each line represents all the information of one molecule. Therefore, this matrix will contain as many lines as the number of molecules in the training set. In this way, the first approach has been to make the visualization of each molecule of the training set in **MAP_1** in order to obtain one potential map per molecule. Next, all the information has been placed in a single line and the **MATRIX_11** has been created.

25.1.4.   MAP_2 generation and molecule classification

Finally, another map has been created with the generated input file as both reference file and test file, using the parameters **P1** (table 44). After several changes (only the best one has been reported in table 45), the map containing the smallest error has been selected and termed **MAP_2**, with the aim of visualizing the molecules on it. The visualization of molecules has not been possible because there is too much data and the software was not able to handle such a large quantity. New trials have been carried out using reduced dimensions. These assays have been interesting when attempting to learn how the manipulation and data analysis of the application should be carried out. However, since the results were not very useful, they have not been reported in this work.

Table 44. First trial for MAP_2 construction in assay MOL_11_DIM_50.

Number of trials	100
Input data file name	MATRIX_11dat
Input test file name	MATRIX_11dat
Output map name	Final_map.cod
X dimensión	50
Y dimensión	50
Topology type	Rect
Neighborhood	Bubble
Training length of 1st part	6
Training rate of 1st part	0.05
Radius of 1st part	35
Training length of 2nd part	50
Training rate of 2nd part	0.01
Radius of 2nd part	15
RANDOM	83
ERROR	**678.995911**

Table 45. First trial for MAP_2 construction in assay MOL_11_DIM_50.

From Final_map.cod	to…	MAP NAME	ERROR
Training length of 1st part	300	**Final_map_ch5.cod (MAP_2)**	3.382098
Training rate of 1st part	0.5		
Training length of 2nd part	500		
Training rate of 2nd part	0.1		

25.2. Assay 2: MOL_110_DIM_25

25.2.1. Database generation

First, an extensive revision has been carried out on the database of the *"Institute de Recherches Servier"* in order to select several molecules with great structural variation and different MT_1 binding affinities. Next, their configuration has been corrected, taking the MLT molecule as reference (figure 86) and they have been superimposed in order to create the following sets:

• KNN_110_MOL: 110 molecules have been selected as first training set of the KNN study.

• KNN_19_SET: 19 molecules have been selected to validate the KNN map as well as to fill it up.

The vdW surface of each molecule has been calculated with 183.9 mol/Å2 density of dots. Finally, the input file has been generated placing all the information (coordination and dots) of each molecule in it.

25.2.2. MAP_1 construction and input file generation

As explained previously, the first step before constructing a map is to set up the different parameters to be used to create a KNN map. This step has not been carried out due to the fact that the parameters have already been selected from the previous assays.

Next, a real map has been constructed starting with the selected parameters, using MLT molecule as reference and the KNN_110_MOL file as test file. In this step, different maps have been created by changing the initial parameters a little. The best map has been selected as **MAP_1** and its parameters as **P1** (tables 46 and 47).

Table 46. First trial for **MAP_1** generation in assay MOL_11_DIM_25.

Number of trials	100
Input data file name	MLT,dat
Input test file name	KNN_110.dat
Output map name	MLT_110_map1_dim25.cod
X dimension	25
Y dimension	25
Topology type	Rect
Neighborhood	Bubble
Training length of 1st part	6
Training rate of 1st part	0.05
Radius of 1st part	35
Training length of 2nd part	15
Training rate of 2nd part	0.01
Radius of 2nd part	15
ERROR	**4.431020**

Table 47. Change of parameters for **MAP_1** generation in assay MOL_110_DIM_25.

From MLT_110_map1_dim25.cod	to…	MAP NAME	ERROR
Radius of 1st part	20	MLT_110_map1_dim25_ch1.cod	4.099879
Radius of 2nd part	10		
From MLT_110_map1_dim25.cod	to…	MAP NAME	ERROR
Radius of 1st part	10	**MLT_110_map1_dim25_ch2.cod (MAP_1)**	4.074740
Radius of 2nd part	5		

The MATRIX_110 has then been created as explained previously.

25.2.3.   *MAP_2 generation and molecule classification*

Finally, another map has been created with the MATRIX_110 as both, reference file and test file, using the parameters **P1**.

Table 48. First trial for **MAP_2** generation in assay MOL_11_DIM_25

Number of trials	100
Input data file name	MATRIX_110.dat
Input test file name	MATRIX_110.dat
Output map name	MAP_FINAL_110_dim25.cod
X dimension	25
Y dimension	25
Topology type	Rect
Neighborhood	Bubble
Training length of 1st part	3000
Training rate of 1st part	0.05
Radius of 1st part	10
Training length of 2nd part	1500
Training rate of 2nd part	0.01
Radius of 2nd part	5
ERROR	**179**

After changing the parameters, several maps have been created.

209

Table 49. Change of parameters for **MAP_2** generation in assay MOL_110_DIM_25.

From FMAP_110_map1_dim25.cod	to…	MAP NAME	ERROR
Training rate of 1st part	0.1	MAP_FINAL_110_map1_dim25_ch1.cod	158
Training rate of 2nd part	0.05		
From FMAP_110_map1_dim25.cod	to…	MAP NAME	ERROR
Training rate of 1st part	0.01	MAP_FINAL_110_map1_dim25_ch2.cod	200
Training rate of 2nd part	0.005		
From FMAP_110_map1_dim25.cod	to…	MAP NAME	ERROR
Radius of 1st part	20	MAP_FINAL _110_map1_dim25_ch3.cod	217
Radius of 2nd part	10		
From FMAP_110_map1_dim25.cod	to…	MAP NAME	ERROR
Training length of 1st part	1000	MAP_FINAL _110_map1_dim25_ch4.cod	207
Training length of 2nd part	700		
From FMAP_110_map1_dim25.cod	to…	MAP NAME	ERROR
Training length of 1st part	5000	MAP_FINAL _110_map1_dim25_ch5.cod	170
Training length of 2nd part	2500		

The maps containing the smallest error (highlighted in grey in tables 48 and 49) have been selected in order to visualize the molecules of the input file on it. Looking at the maps below (figure 92) it can be said that the most promising one is the first one (MAP_FINAL_110_dim25.cod) because the repartition of the molecules is the clearest for visualizing separating the active and inactive molecules. It should be pointed out that since there are some uncertain cells on it, this map should be developed in order to optimize it. On KNN maps, several compounds can be placed in the same node but normally in a correct map the pki of these compounds should be similar. If the pki is not similar an uncertainty is generated.

MAP_FINAL_dim25_pki

MAP_FINAL_dim25_ch1_pki

MAP_FINAL_dim25_ch2_pki

MAP_FINAL_dim25_ch5_pki
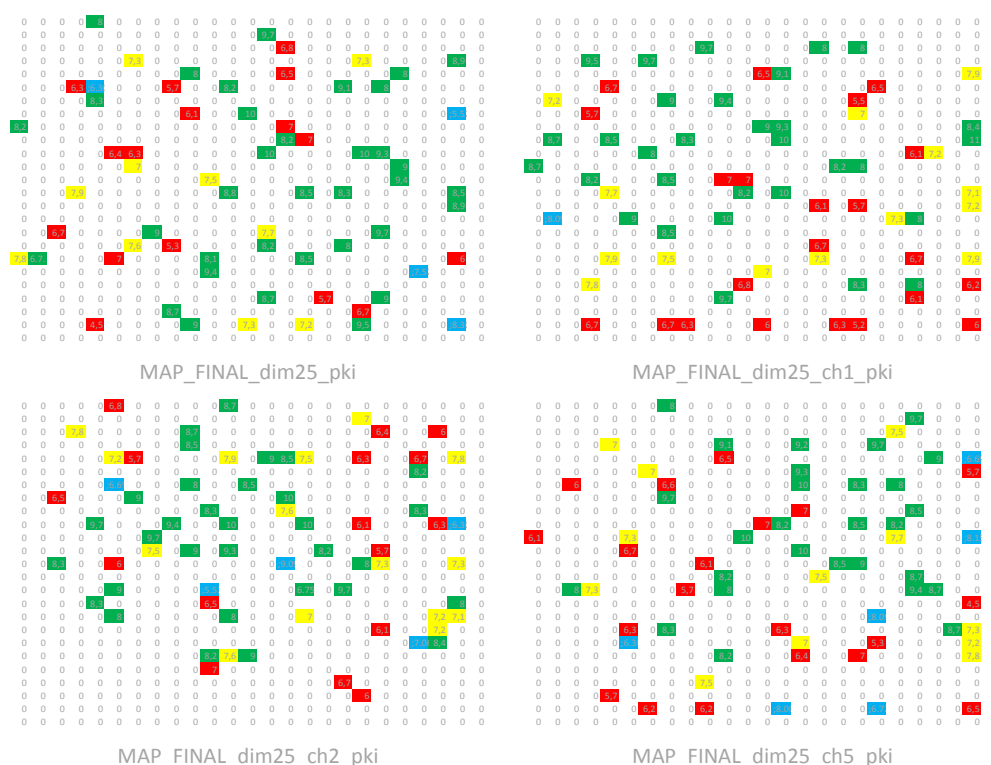
Figure 92. Representation of matrix_110 on the best maps of tables 49 and 50. pKi>8 (red) / 7<pki<8 (yellow) / pki<7(red) / Uncertainty (blue)

On the maps showed in figure 92 there are some uncertain cells but as can be observed, these areas are mainly placed near the border of the maps. This fact suggests that the map is not big enough so as to be able to correctly place all the data. Therefore, the program has comprised and superimposed the data. In order to probe this hypothesis, the map has been re-dimensioned; all the parameters have been maintained when creating the new map but the dimensions have been changed as follows (tables 50 and 51, figures 93 and 94).

25.2.3.1. REDIMENSION_2530

Table 50. Redimension of map MAP_FINAL_110_dim25.cod from 25 to 30

Number of trials	100
Input data file name	MATRIX_110.dat
Input test file name	MATRIX_110.dat
Output map name	MAP_FINAL_110_redim2530.cod
X dimension	30
Y dimension	30
Topology type	Rect
Neighborhood	Bubble
Training length of 1st part	3000
Training rate of 1st part	0.05
Radius of 1st part	10
Training length of 2nd part	1500
Training rate of 2nd part	0.01
Radius of 2nd part	5
ERROR	**169**

MAP_FINAL_redim2530_pki

Figure 93. Representation of matrix_110 on MAP_FINAL_110_redim2530.cod
pKi>8 (red) / 7<pki<8 (yellow) / pki<7(red) / Uncertainty (blue)

Although the new map shows some uncertain cells, as can be observed in figure 93, when carrying out re-dimension of the map, the uncertainty decreased.

25.2.3.2. REDIMENSION_2535

Table 51. Redimension of map MAP_FINAL_110_dim25.cod from 25 to 35.

Number of trials	100
Input data file name	MATRIX_110.dat
Input test file name	MATRIX_110.dat
Output map name	MAP_FINAL_110_redim2535.cod
X dimension	35
Y dimension	35
Topology type	Rect
Neighborhood	Bubble
Training length of 1st part	3000
Training rate of 1st part	0.05
Radius of 1st part	10
Training length of 2nd part	1500
Training rate of 2nd part	0.01
Radius of 2nd part	5
ERROR	**161**

MAP_FINAL_110_dim_35_pki   MAP_FINAL_110_dim_35_pki

Figure 94. Representation of matrix_110 on MAP_FINAL_110_redim2535.cod
pKi>8 (red) / 7<pki<8 (yellow) / pki<7(red)

As can be observed looking at the map in figure 94, the uncertainty has disappeared and the distribution of molecules for visually separating the active ones from the inactive ones is quite good. This map has been set as **MAP_2**. In graph 7 the virtual **MAP_2** can be observed with the molecules represented as active (X); inactive (o) and intermediate (I).
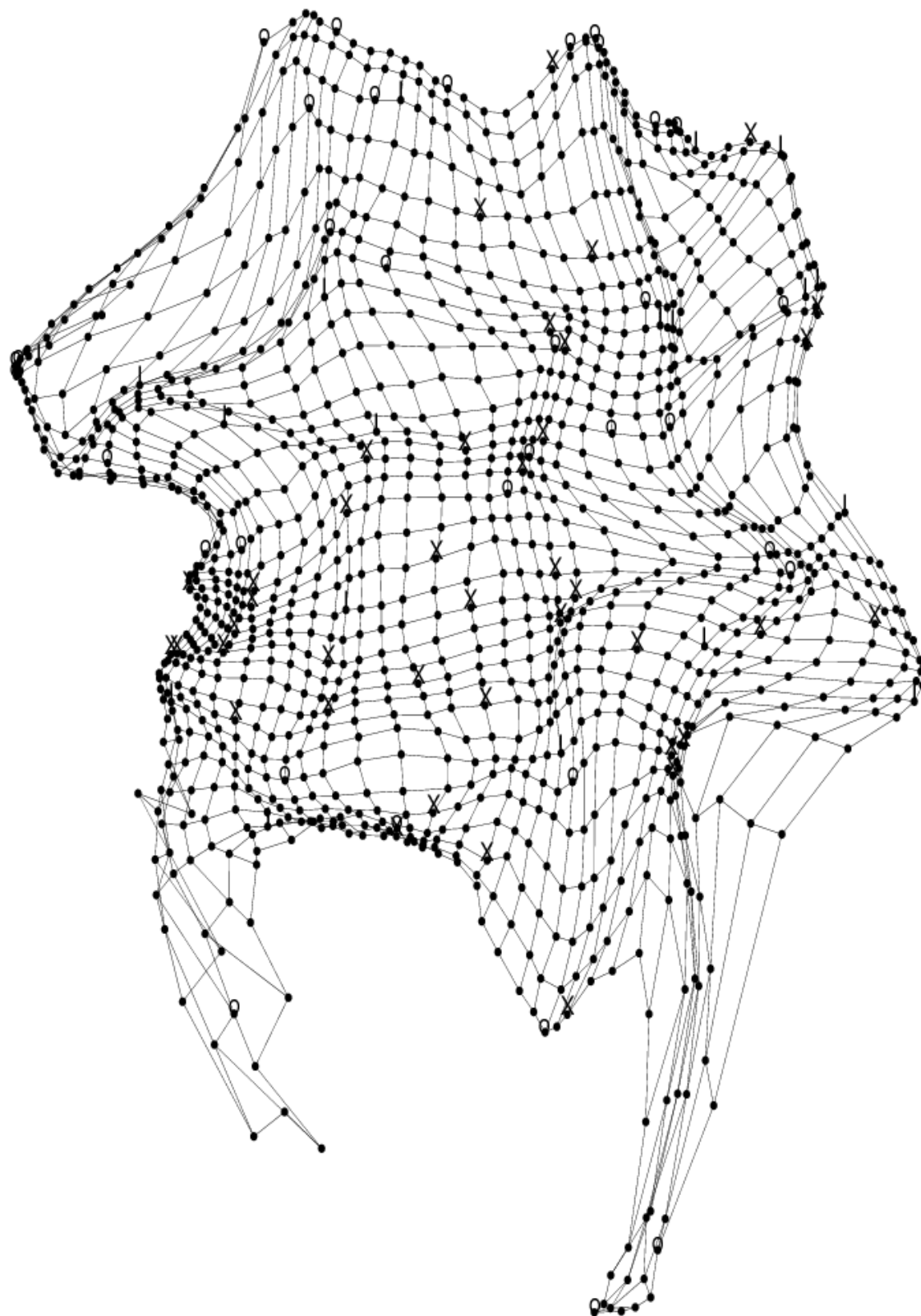
Graph 7. Graphical representation of MAP_2 with 110 molecules on it.
Active compounds (X) / Inactive compounds (o) / Intermediate compounds (I)

25.2.4.   Validation of MAP_2

At this point, 19 new molecules have been used to validate and fill up the map. After carrying out modelization of the molecules and calculating their vdW surface, their potential map has been created through their visualization in **MAP_1**. Then the matrix_19 has been made containing these potential maps.

Both of the matrixes, matrix_19 and matrix_110, have been put together to form a new input file named matrix_129. Finally, visualization of the molecules on **MAP_2** has been carried out representing the 19 new molecules in purple in order to carry out visual classification (figure 95, table 52).

MAP_FINAL_129_dim_35_name_pki
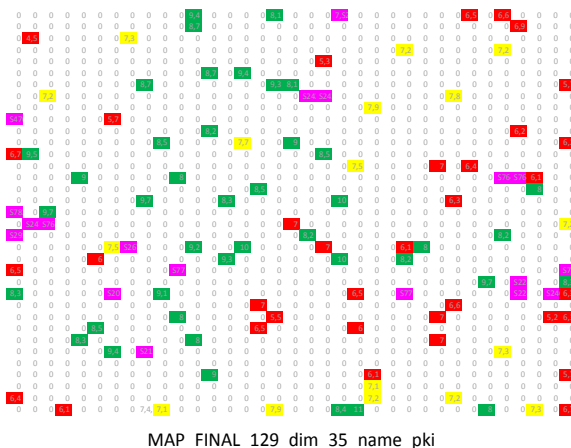
Figure 95. Representation of the molecules placed in matrix_129 on MAP_2. pKi>8 (red) / 7<pki<8 (yellow) / pki<7(red) / 19 new molecules (purple)

Table 52. Visual prediction of the 19 new compounds

Molecules	Predicted class	Real class (pKi)	Precisions
S20760	Act	Act (9.0)	☑
S21302	Act	Interm (7.0)	?
S22415	Inact	Inact (4.0)	☑
S22416	Act	Act (9.0)	☑
S24225	Act	Interm (7.7)	?
S24226	Act	Act (9.0)	☑
S24639	Inact	Act (9.3)	☒
S24788	Inact	Act (11.0)	☑
S26700	Act	Act (8.3)	☑
S29104	Inact	Interm (7.7)	?
S47607	Inact	Inact (6.0)	☑
S76227	Act	Interm (7.3)	?
S76418	Inact	Act (9.2)	☒
S76424	Inact	Act (8.0)	☒
S77359	Act	Inact (6.5)	☑
S77361	Inact	Inact (5.7)	☑
S78329	Inact	Act (8.7)	☒
S20933	Act	Interm (7.7)	?
S76425	Inact	Inact (5.4)	☑

As can be observed, several molecules have been placed at the border of the active/inactive area. As their classification is not clear enough, the molecules have been used to fill up the map in order to aid in future predictions (figure 96).

MAP_FINAL_129_dim_35_pki   MAP_FINAL_129_dim_35_pki

Figure 96. Representation of the molecules placed in matrix_129 on MAP_2.
pKi>8 (red) / 7<pki<8 (yellow) / pki<7(red).

In graph 8, the virtual **MAP_2** can be observed with all the molecules on it, represented as active (X); inactive (o) and intermediate (I).

Graph 8. Graphical representation of **MAP_2** with 129 molecules on it.
Active compounds (X) / Inactive compounds (o) / Intermediate compounds (I)

215

25.2.5.   Prediction of new designed molecules

Finally, the molecules placed in the file "MT1predictions35" (*see Ch3 – 24.1*) have been predicted. In reality, only 28 of the 35 indoles have been predicted because it was impossible to calculate the vdW surface of the other 7 compounds. After the modelization and their vdW surface calculation, the potential map of the selected compounds was created through their visualization in **MAP_1**. Then the MATRIX_28 was made containing these potential maps.

Next, MATRIX_28 and MATRIX_129 were put together to form a new input file named MATRIX_157. Finally, visualization of the molecules on **MAP_2** was carried out representing the 28 new molecules in purple in order to carry out visual classification (figure 97, graph 9, table 53).
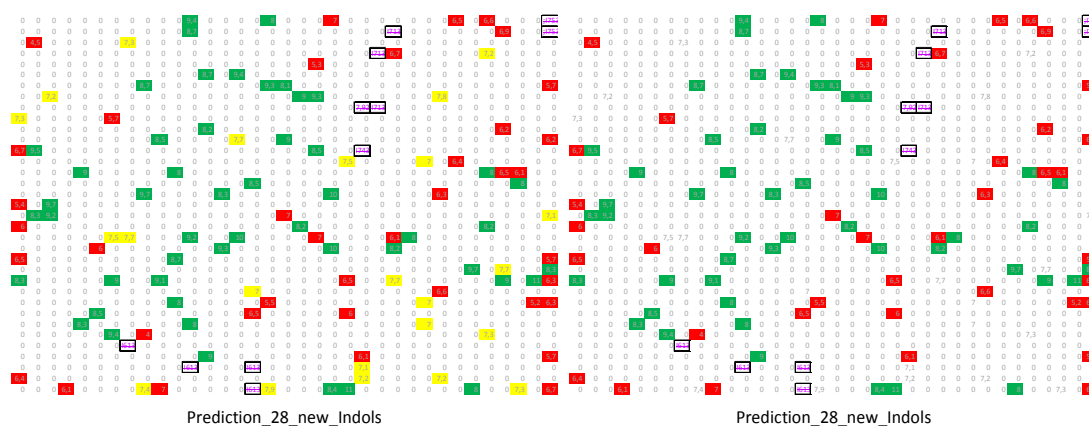
Prediction_28_new_Indols   Prediction_28_new_Indols
Figure 97. Visualization of the 28 newly designed indoles in **MAP_2**

pKi>8 (red) / 7<pki<8 (yellow) / pki<7(red) / 19 new molecules (purple)

The 28 molecules were placed on 11 different nodes, and after visualization in both figure 97 and graph 9, they were predicted as follows (table 53).

Table 53. Prediction of the selected indoles

| Node	Indoles	Prediction
1	**In1-3.XII; In1-3.COH; In1-3.XI; In1-3.IX**	I/X (act/interm)
2	**In1-3.CSCF3**	X (act)
3	**In1-3.COCF3**	X (act)
4	**In1-3.CSH**	I (interm)
5	**In2-3.COCF3; In2-3.CSCF3**	I (interm)
6	**I743CF3**	I (interm)
7	**In2-3.COH**	o (inact)
8	**In2-3.CSH**	I (interm)
9	**In2-3.XII; In2-3.XI**	o (inact)
10	**I753CSA; I753COcP; I753CSE; I753COH; I753COM; I753COP; I753CSA; I753CScP; I753CSE; I753CSH; I753CSM; I753CSP**	o (inact)
11	**I753COCF3; I753CSCF3**	o (inact)
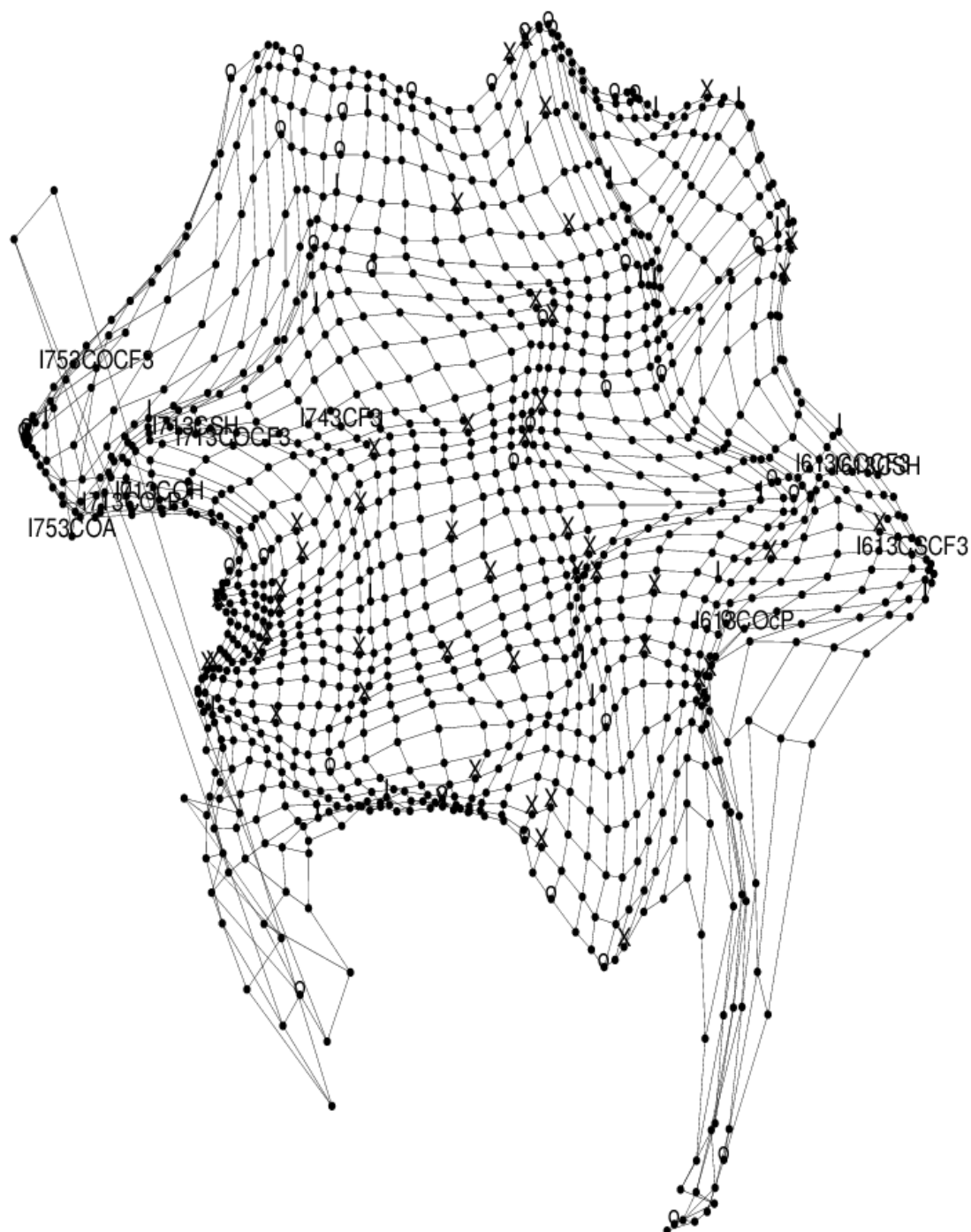
Graph 9. Graphical representation of MAP_2 with 157 molecules on it.
Active compounds (X) / Inactive compounds (o) / Intermediate compounds (I)

The obtained results are really interesting because they coincide with the observations made after the 3D-QSAR study (*see Ch3 - 23.3*).

1. The "In1" series is more active than its analogue series "In2", most likely due to the position of the methoxy group.

2. In the file "MT1predictions35", there are some molecules where the length of the linker is modified and therefore, carbonyl group is displaced in comparison with the carbonyl group of MLT. The predictions suggest that these compounds would present a smaller affinity for MT_1/MT_2 receptors.

26. COMPARISON BETWEEN DIFFERENT PREDICTIONS

As can be observed in table 54, all compounds set up as active in KNN maps have been already predicted as active in Bayesian models. In contrast, there are some compounds predicted as active by the Bayesian model but have not been set up as active in KNN. As previously stated, since the credibility of KNN map is higher, only the compounds predicted as active by KNN map have been taken into account.

At the time of creating this model, it has been decided that the biological data of the compounds already synthesized (**In1-4.XII, In1-4.XI, In1-3.X, In2-3.IX, In2-4.XII, In2-4.XI**) would be used when obtained in order to determine the validity of this method. If good results are obtained, the compounds not yet synthesized and predicted as active by both models would be the priority molecules so as to be able to synthesize them first, if possible (**In1-3.CSCF3, In1-3.COCF3, In1-3.COH**).

Table 54. Comparison of the obtained predictions through Bayesian model and KNN maps

Indoles	Prediction Bayesian	Indoles	Prediction KNN
In1-3.COCF3	act	In1-3.COCF3	act
In1-4.XII	act	In1-4.XII	interm/act
In1-3.COH	act	In1-3.COH	interm/act
In1-4.XI	act	In1-4.XI	interm/act
In1-3.CSCF3	act	In1-3.CSCF3	act
In1-3.CSH	inact	In1-3.CSH	interm
In1-3.IX	act	In1-3.IX	interm/act
In2-3.IX	inact	In2-3.IX	interm
In2-4.XII	inact	In2-4.XII	inact
In2-3.COH	inact	In2-3.COH	inact
In2-4.XI	act	In2-4.XI	inact
In2-3.CSCF3	inact	In2-3.CSCF3	interm
In2-3.CSH	inact	In2-3.CSH	interm
I743CF3	act	I743CF3	interm
In3-3.I	inact	I753COA	inact
In3-3.IX	inact	I753COCF3	inact
In3-4.XII	inact	I754COM	inact
In3-3.II	inact	I753COE	inact
In3-3.COH	inact	I753COH	inact
In3-4.XI	act	I754COM	inact
In3-3.III	inact	I753COP	inact
In3-3.IV	inact	I753CSA	inact
In3-3.CSCF3	inact	I753CSCF3	inact
In3-3.VIII	act	I753CScP	inact
In3-3.V	inact	I753CSE	inact
In3-3.CSH	inact	I753CSH	inact
In3-3.VII	inact	I753CSM	inact
In3-3.VI	act	I753CSP	inact
I623	inact		
I624N	inact		
I624O	Inact		
I624S	Inact		
I625N	Inact		
I625O	Inact		
I625S	Inact		

27.COMPARISON BETWEEN PREDICTIONS AND REAL DATA

When the biological data of the compounds already synthesized **(In1-4.XII, In1-4.XI, In1-3.X, In2-3.IX, In2-4.XII, In2-4.XI**) was obtained, it was used for determining the validity of the KNN model.

The activity was established as before; pKi>8 (active) / 7<pki<8 (interm) / pki<7(inactive). As can be observed in table 55, the KNN model provided a fairly good prediction, making few errors, mostly in the frontier between two different classes as in the case of compound **In2-4.XI**, which was predicted as inactive, and since its pki is 7.1, it should be marked as intermediate (8>pki>7).

Table 55. Comparison of predicted activity and real activity of some compounds from file MT1prediction35.

Indoles	Prediction KNN	Real Ki(nM)	Real pKi	Real Activity	Precisions
In1-3.IX	Interm/Act	12	7.9	Interm	☑
In1-4.XI	Interm/Act	4	8.4	Act	☑
In1-4.XII	Interm/Act	50	7.3	Interm	☑
In2-3.IX	Interm	12	7.9	Interm	☑
In2-4.XI	Inact	75	7.1	Interm	?
In2-4.XII	Inact	I	I	Inact	☑

Although this model should be filled up, it can be assumed that it is already a valid model to provide a prediction regarding the activity of compounds such as MT$_1$ agonists.

Once the validity of the model has been confirmed, compounds **In1-3.CSCF3, In1-3.COCF3** and **In1-3.COH** have been chosen as priority in order to synthesize them first. Only compound **In1-3.COH** (marked then as **In1-3.XIII**) has been achieved. The high complexity that the synthesis of compounds **In1-3.CSCF3** and **In1-3.COCF3** presents, in addition to the lack of time at the end of the project, has led us to consider these compounds as momentarily inaccessible.

XIII. FINAL DISCUSSION AND FUTURE PERSPECTIVES

Based on the information above, it can be concluded that the MT_1/MT_2 receptor agonists that present a space of 6 atoms as the distance between the methoxy group and the first nitrogen atom of the aliphatic chain obtained by a two-methylene linker, display the best binding affinities for both receptors. Thus, the pharmacophore described as our initial hypothesis can be redefined as follows (figure 98).
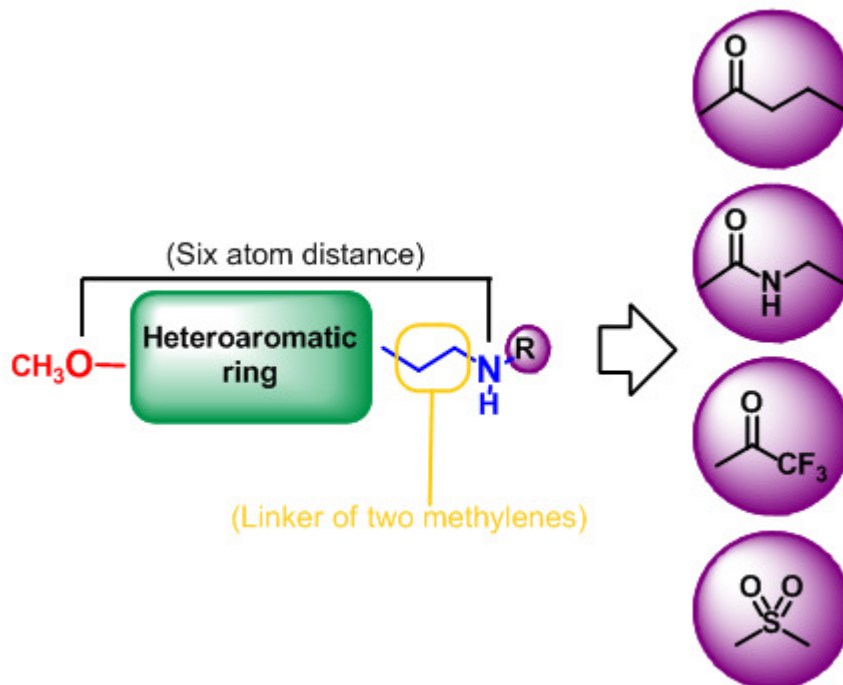
Figure 98. Redefinition of initial hypothesis

Among the variations introduced at the end of the aliphatic linker, it has been suggested that urea derivatives are probably more active than the amide derivatives (*see Ch3 - 22.1.1*) as well as more active than the thiourea derivatives (*see Ch3 - 22.1.1*). In addition, the urea derivatives usually present better metabolic stability than amide derivatives and therefore, the insertion of urea groups is preferred, if possible.

The best melatoninergic agonists have been obtained upon introducing a *N*-butyramide chain attached to the aliphatic linker in the case of series of quinoxalines (Q1, Q2 and Q3) (*see Ch3 - 22.1.1*) and a *N*-ethylurea chain, a trifluoromethyl group and a *N*-methanesulfonamide group in the case of series of indoles (In1 and In2) (*see Ch3 - 22.1.1*). Therefore, these groups can also be introduced as a structural requirement in the new pharmacophore structure (figure 98).

It is also remarkable that the length of the *N*-butyramide chain and *N*-ethylurea chain is almost the same, a three-atom space from the carbonyl group. This fact is logical, since it has been stated in the introduction that a synthetic MLT analog, which contains a C_3H_7 group attached to the carbonyl group instead of the methyl group found in MLT, displays the greatest activity but any larger degree of increase or branching leads to decreased binding affinity (*see Ch1 - 1.2.4*). These findings suggested the existence of a small hydrophobic receptor pocket in this area of the binding site, which has been corroborated by these new findings.[73]

As modeling assays have demonstrated, the vdW surface of the derivatives is the main characteristic to take into account when designing compounds for the purpose of obtaining MT_1 receptor agonists. Therefore, although any molecule which fulfills the pharmacophore described above would probably display good MT_1 affinity, it is also probable that other, different molecules can be found which, in spite of not fitting into the pharmacophore, may still generate MT_1 activation if they present a similar vdW surface.

The future work plan of the project can be centered on designing new structures taking into account the new pharmacophore described. In addition, the KNN map would be filled out and all the new derivatives would be predicted before their synthesis in order to reject the potentially inactive ones, saving time and effort.

As most neural networks, the architecture of Kohonen self-organising map is a "black box" which requires further processing to provide a meaningful interpretation of the results.[167] As the modeling assays include in this thesis have been performed on a short-term stay in which the main purpose has been to obtain a predicting model, no further investigations have been performed in the field of results interpretation and thus, this area remains as an open door for future development. In this sense it has been propose that a 3D-QSAR model could be generated using the physico-chemical parameters on molecular surface of each molecule by the use of KNN in order to overcome the technical problems associated to the high dimensionality of this molecular property.[98] As it has been explained, before building the KNN maps the vdW surface of different molecules has been calculated and then the 3D coordinates of all dots on the van der Waals surface have been projected into the 2D map by KNN obtaining this way a molecular electrostatic potential (MEP) map per molecule. Then, the correlation between the MEP values of all 2D maps and the biological activity can be analyzed for example by the use of 3-way PLS analysis.

As stated before, sleep disturbance is closely linked to major depression (*see Ch1 - 2.2.5*). Agomelatine, a potent melatonin receptor MT_1/MT_2 agonist and partial 5-HT_{2c} antagonist, is an effective antidepressant, which also has a notably positive impact on the synchronization of wake-sleep rhythms and effectively reduces sleep complaints in depressed patients (*see Ch1 - 2.2.6.2.5*). Therefore, agomelatine-like profile is a much desired characteristic for drugs focusing on the treatment of sleep disorders. A proposal has been made to test the 5-HT_{2C} affinity of the compounds that present the best MT_1/MT_2 activity in order to determine whether or not they also contain an agomelatine-like profile. If non-serotoninergic affinity is detected, modification of these compounds will be proposed in order to achieve 5-HT_{2C} affinity, but always maintaining MT_1/MT_2 affinity. A rapid bibliographic overview has revealed that the introduction of chloro and bromo atoms in some structures increases their 5-HT_{2C} affinity.[168, 169] Therefore, a future series of analogues to In1 can be designed by introducing some chloro or bromo atoms on the structure in order to determine if some 5-HT_{2C} activity can be obtained.

Chapter 4

Conclusions

This work includes the synthesis and structural characterization of twenty-three new quinoxaline derivatives and twenty-five new indole derivatives, as well as their biological evaluation as MT_1 and MT_2 receptor agonists carried out by the *"Institute de Recherches Servier"* in France. Molecular modeling studies performed in "le Laboratoire de Chimie Thérapeutique 2" at the University Lille 2 have also been included. The research work has led to the conclusions listed below.

From a synthetic point of view, the following contributions have been proposed:

1) The substitution of one of the two chlorine atoms of 2,3-dichloroquinoxaline by a methoxy group has been carried out selectively by using sodium methoxide in THF.

2) The removal of the primary arylamine in series Q3 has been achieved by means of diazotization and subsequent deamination performed in one-pot reaction, leading to improved reaction efficacy.

3) For the synthesis of urea derivatives of series Q1 (**Q1-3.VI**, **Q1-3.VII**, **Q1-3.VIII**, **Q1-3.IX**), one-pot reaction has been carried out due to the instability of the amine **Q1-2**, diminishing the synthetic steps and thereby saving time and effort.

4) Due to the lack of commercial availability of the reagents methyl isocyanate and cyclopropyl isocyanate, an alternative synthetic route using 1,1'-carbonyldiimidazole has been designed for obtaining compounds **.XI** and **.XII** of series In1 and In2.

As a result of the molecular modeling studies, it has been established that:

5) For an optimum superimposition of the molecules with the conformation of MLT when it is bound to the receptor MT_1, the following requirements are necessary:

   a) Ligands should contain a two-methylene linker between the central scaffold and the first nitrogen atom of the side chain.

   b) The methoxy group present in most of the MT_1/MT_2 receptor agonists should be placed in a MLT-like position in order to avoid detriment in the affinity of the ligands for MT_1/MT_2 receptors.

6) The 3D-QSAR CoMFA and CoMSIA models created with the program Sybyl, with PLS analysis as well as with SIMCA analysis, have not generated the desired results for quantitative or qualitative prediction.

7) Several Bayesian models have been obtained based on different fingerprint types and different pki ranges. Although a good predicting model (BM_act8_inact6_SEFC20) has been obtained, its use is not suitable because this model is based on "Sybyl atom type".

8) The KNN maps created, based on the vdW surface of the molecules, have provided a good separation between active and inactive compounds. The obtained model has been suggested as a good tool for predicting MT_1 affinity of new ligands.

9) Twenty-eight molecules have been represented on KNN map, obtaining significantly promising results. Refilling the initial map has been proposed as future work in order to permit interpretation of the activity that any molecule placed on the map might have.

With regard to biological activity and structure-activity relationship, the following conclusions have been suggested:

10) The biological evaluation of quinoxaline derivatives such as MT_1 and MT_2 receptor ligands has demonstrated that:

a) The distance of six atoms between the methoxy group and the first nitrogen of the side chain is essential.

b) It has not been possible to establish whether or not quinoxaline ring is suitable as central scaffold of MT_1/MT_2 receptor ligands due to the lack of a series of quinoxalines which completely fulfills the structural requirements of the pharmacophore described in this work.

c) With regard to the amide/urea variation, it cannot be concluded which group leads to the best affinity values for MT_1/MT_2 receptors.

11) The biological evaluation of indole derivatives, which shows their affinity for MT_1 and MT_2 receptors in the order of nanomolar, has evidenced that:

a) The distance of six atoms between the methoxy group and the first nitrogen of the side chain is essential.

b) With regard to the urea/thiourea variation, it has been demonstrated that urea derivatives generally show better affinity values for MT_1/MT_2 receptors.

12) After the SAR study performed, a new pharmacophore is proposed in which the following structural requirements are included:

a) A central core constituted by an aromatic ring.

b) A methoxy group substituted on the central scaffold.

c) A two-methylene linker substituted on the central ring and united to different functionalized chains such as *N*-butyramide, *N*-ethylurea, 2,2,2-trifluoro-*N*-acetamide and *N*-methylsulfonamide.

d) A space corresponding to a six-atom distance between the methoxy group and the first nitrogen atom of the aliphatic chain.

13) Compounds **In1-3.II**, **In1-3.IX** and **In1-3.X** have been selected as hits in this project due to their affinity values as well as their full agonist profile. These promising results open a new line of research for the development of future MT_1/MT_2 receptor agonists.

Capítulo 4

Conclusiones

El presente trabajo incluye la síntesis y caracterización estructural de veintitrés nuevos derivados de quinoxalina y veinticinco nuevos derivados de indol, así como su evaluación biológica como agonistas de los receptores MT$_1$ y MT$_2$ llevada a cabo en el *"Institute de Recherches Servier"* en Francia. Además, se incluyen también los estudios de modelización molecular realizados en el *"Laboratoire de Chimie Thérapeutique 2"* de la Universidad Lille 2. El trabajo de investigación realizado ha dado como resultado las siguientes conclusiones.

Desde el punto de vista sintético se proponen las siguientes aportaciones:

1) La sustitución de uno de los dos átomos de cloro del compuesto 2,3-dicloroquinoxalina por un grupo metoxilo, ha sido llevada a cabo selectivamente mediante el uso de metóxido sódico en THF.

2) La eliminación de la arilamina primaria de la serie Q3, ha sido llevada a cabo mediante una diazotización y su posterior deaminación en un solo paso de reacción, mejorando de este modo la eficacia del proceso sintético.

3) Para la síntesis de los derivados de urea de la serie Q1 (**Q1-3.VI**, **Q1-3.VII**, **Q1-3.VIII**, **Q1-3.IX**) y debido a la inestabilidad de la amina **Q1-2**, se ha puesto a punto una reacción en un solo paso, disminuyendo de este modo el número de pasos sintéticos empleados y logrando así un ahorro de tiempo y esfuerzo.

4) Debido a que los reactivos metil isocianato y ciclopropil isocianato no son comercialmente accesibles, se ha diseñado una ruta sintética alternativa para la obtención de los compuestos **XI** y **XII** de las series In1 e In2, empleando 1,1'-carbonildiimidazol.

Como consecuencia de los estudios de modelización molecular se ha establecido que:

5) Para lograr una óptima superposición de las moléculas con la conformación que presenta la MLT cuando está unida con el receptor MT$_1$, se necesitan los siguientes requisitos:

   a) Los ligandos deben contener un espaciador de dos metilenos entre el anillo central y el primer átomo de nitrógeno de la cadena lateral.

   b) El grupo metoxilo presente en la mayoría de los agonistas de los receptores MT$_1$ y MT$_2$, debe estar colocado en una posición análoga al de la MLT para evitar que se dé un detrimento en la afinidad de dichos ligandos.

6) Los modelos 3D-QSAR CoMFA y CoMSIA creados con el programa Sybyl tanto con los análisis PLS como SIMCA, no han generado los resultados esperados ni para una predicción cuantitativa ni para una predicción cualitativa.

7) Se han obtenido varios modelos Bayesianos basados en diferentes tipos de "fingerprints" y con diferentes rangos de pki. Pese a haberse obtenido un buen modelo predictivo (BM_act8_inact6_SEFC20), su uso no es conveniente ya que este modelo está construido con el "tipo de átomo de Sybyl".

8) Los mapas KNN creados basados en la superficie de vdW de las moléculas ha proporcionado una buena separación entre compuestos activos e inactivos. El modelo obtenido parece ser una buena herramienta para predecir la afinidad MT_1 de nuevos ligandos.

9) Se han representado veintiocho moléculas en el mapa KNN, obteniéndose resultados significativamente prometedores. Como futuro trabajo se propone rellenar el mapa inicial para permitir la interpretación de la actividad de cualquier molécula representada en dicho mapa.

Respecto a la actividad biológica y la relación estructura-actividad, se sugieren las siguientes conclusiones:

10) La evaluación biológica de los derivados de quinoxalinas como ligandos de los receptores de melatonina MT_1 y MT_2 ha demostrado que:

   a) La distancia de seis átomos entre el grupo metoxilo y el primer átomo de nitrógeno de la cadena lateral es esencial.

   b) No ha sido posible establecer si el anillo de quinoxalina es adecuado como núcleo central de los ligandos de los receptores MT_1/MT_2 debido a la falta de series de quinoxalina que cumplan completamente los requerimientos estructurales del farmacóforo descrito en este trabajo.

   c) Con respecto a la variación amida/urea, no ha podido concluirse que grupo conduce a una mejor afinidad por los receptores MT_1/MT_2.

11) La evaluación biológica de los derivados de indol, los cuales presentan una afinidad con los receptores MT_1/MT_2 en el orden de nanomolar, evidencia que:

   a) La distancia de seis átomos entre el grupo metoxilo y el primer átomo de nitrógeno de la cadena lateral es esencial.

   b) Con respecto a la variación urea/tiourea, se ha demostrado que los derivados de urea muestran en general unos mejores valores de afinidad con los receptores MT_1/MT_2.

12) Tras el estudio SAR se propone un nuevo farmacóforo en el que se incluyen los siguientes requisitos estructurales:

    a) Un núcleo central constituido por un anillo aromático.

    b) Un grupo metoxilo sustituido en el núcleo central.

    c) Un espaciador de dos metilenos sustituido sobre el anillo central y unido a diferentes cadenas funcionalizadas como *N*-butiramida, *N*-etilurea, 2,2,2-trifluoro-*N*-acetamida y *N*-metanosulfonamida

    d) Un espacio de seis átomos de distancia entre el grupo metoxilo y el primer átomo de nitrógeno de la cadena alifática.

13) Los compuestos **In1-3.II**, **In1-3.IX** e **In1-3.X** han sido seleccionados como cabezas de serie del proyecto debido a sus valores de afinidad y a su perfil de agonistas completos. Estos prometedores resultados abren una nueva línea de investigación para el desarrollo de futuros agonistas de los receptores MT_1/MT_2.

Bibliography

1. Velayos, J. L.; Moleres, F. J.; Irujo, A. M.; Yllanes, D.; Paternain, B., Bases anatómicas del sueño. *Anales del Sistema Sanitario de Navarra* **2007,** 30, (Supl. 1), 7-17.

2. Erro, M. E., Presentación. *Anales del Sistema Sanitario de Navarra* **2007,** 30, (Supl. 1), 1-2.

3. Siegel, J. M., Clues to the functions of mammalian sleep. *Nature* **2005,** 437, (7063), 1264-1271.

4. Cirelli, C.; Tononi, G., Is sleep essential? *Public Library of Science. Biology* **2008,** 6, (8), e216.

5. Walker, M. P., Sleep, memory and emotion. *Progress in Brain Research* **2010,** Volume 185, 49-68.

6. Reimund, E., The free radical flux theory of sleep. *Medical Hypotheses* **1994,** 43, (4), 231-233.

7. Siegel, J. M.; Rogawski, M. A., A function for REM sleep: Regulation of noradrenergic receptor sensitivity. *Brain Research Reviews* **1988,** 13, (3), 213-233.

8. Cuba-MINSAP, *Síntomas y signos frecuentes en la adolescencia (Capitulo VII)*. Ministerio de Salud Pública (MINSAP): La Habana (Cuba), 1999.

9. Aguirre-Navarrete, R. I., Bases Anatómicas y Fisiológicas del Sueño *Revista Ecuatoriana de Neurología* **2006,** 15, (2-3).

10. Hudetz, A.; Pearce, R., *Suppresing the mind: Anesthesic modulation of memory and consciousness*. 1st ed.; Humana Press: USA, 2010.

11. Haas, L. F., Hans Berger (1873–1941), Richard Caton (1842–1926), and electroencephalography. *Journal of Neurology, Neurosurgery & Psychiatry* **2003,** 74, (1), 9.

12. Deak, M.; Epstein, L. J., The history of polysomnography. *Sleep Medicine Clinics* **2009,** 4, (3), 313-321.

13. Woo, J., A short history of the development of ultrasound in obstetrics and gynecology. Alfred Loomis.

14. California-University Electroencephalogram lab. http://www.csulb.edu/~cwallis/482/eeg/eeg.html

15. Pace-Schott, E. F.; Hobson, J. A., The neurobiology of sleep: Genetics, cellular physiology and subcortical networks. *Nature Reviews Neuroscience* **2002,** 3, (8), 591-605.

16. Smith, H. R.; Comella, C.; Högl, B., *Sleep medicine*. 1st ed.; Cambridge (UK), 2008; p 11.

17. Rosenthal, M. S., Physiology and neurochemistry of sleep (Galley) *American Journal of Pharmaceutical Education* **1998,** 62, (2), 204-208.

18. Kirschfeld, K., The physical basis of alpha waves in the electroencephalogram and the origin of the "Berger effect". *Biological Cybernetics* **2005,** 92, (3), 177-185.

19. Colten, H. R.; Altevogt, B. M., *Sleep disorders and sleep deprivation: An unmet public health problem*. National Academies Press: Washington (USA), 2006; p 34-37.

20. Colten, H. R.; Altevogt, B. M., *Sleep disorders and sleep deprivation: An unmet public health problem*. National Academies Press: Washington (USA), 2006; p 339.

21. Colten, H. R.; Altevogt, B. M., *Sleep disorders and sleep deprivation: An unmet public health problem*. National Academies Press: Washington (USA), 2006; p 43-49.

22. Carskadon, M. A.; Dement, W. C., *Chapter 2 - Normal human sleep: An overview*. W.B. Saunders: Philadelphia, 2011; p 16-26.

23. Colten, H. R.; Altevogt, B. M., *Sleep disorders and sleep deprivation: An unmet public health problem*. National Academies Press: Washington (USA), 2006; p 39-41.

24. Tononi, G.; Cirelli, C., Sleep function and synaptic homeostasis. *Sleep Medicine Reviews* **2006,** 10, (1), 49-62.

25. Pandi-Perumal, S. R.; Srinivasan, V.; Spence, D. W.; Cardinali, D. P., Role of the melatonin system in the control of sleep: Therapeutic implications. *CNS Drugs* **2007,** 21, (12), 995-1018.

26. Konturek, S. J.; Konturek, P. C.; Brzozowska, I.; Pawlik, M.; Sliwowski, Z.; Cześnikiewicz-Guzik, M.; Kwiecień, S.; Brzozowski, T.; Bubenik, G. A.; Pawlik, W. W., Localization and biological activities of melatonin in intact and diseased gastrointestinal tract (GIT). *Journal of Physiology and Pharmacology* **2007,** 58, (3), 381-405.

27. Reiter, R. J.; Tan, D. X.; Fuentes-Broto, L., *Chapter 8 - Melatonin: A multitasking molecule*. Elsevier: 2010; Vol. 181, p 127-151.

28. Colten, H. R.; Altevogt, B. M., *Sleep disorders and sleep deprivation: An unmet public health problem*. National Academies Press: Washington (USA), 2006; p 41-43.

29. Pevet, P.; Challet, E., Melatonin: Both master clock output and internal time-giver in the circadian clocks network. *Journal of Physiology-Paris*, (0).

30. Reppert, S. M.; Weaver, D. R., Coordination of circadian timing in mammals. *Nature* **2002,** 418, (6901), 935-941.

31. Sukumaran, S.; Almon, R. R.; DuBois, D. C.; Jusko, W. J., Circadian rhythms in gene expression: Relationship to physiology, disease, drug disposition and drug action. *Advanced Drug Delivery Reviews* **2010,** 62, (9-10), 904-917.

32. Agez, L.; Laurent, V.; Pévet, P.; Masson-Pévet, M.; Gauer, F., Melatonin affects nuclear orphan receptors mRNA in the rat suprachiasmatic nuclei. *Neuroscience* **2007,** 144, (2), 522-530.

33. Maronde, E.; Stehle, J. H., The mammalian pineal gland: Known facts, unknown facets. *Trends in Endocrinology & Metabolism* **2007,** 18, (4), 142-149.

34. Monteleone, P.; Martiadis, V.; Maj, M., Circadian rhythms and treatment implications in depression. *Progress in Neuro-Psychopharmacology and Biological Psychiatry* **2011,** 35, (7), 1569-1574.

35. Hardeland, R.; Tan, D.-X.; Reiter, R. J., Kynuramines, metabolites of melatonin and other indoles: the resurrection of an almost forgotten class of biogenic amines. *Journal of Pineal Research* **2009,** 47, (2), 109-126.

36. Brainard, G. C.; Kavet, R.; Kheifets, L. I., The relationship between electromagnetic field and light exposures to melatonin and breast cancer risk: A review of the relevant literature. *Journal of Pineal Research* **1999,** 26, (2), 65-100.

37. Boutin, J. A.; Audinot, V.; Ferry, G.; Delagrange, P., Molecular tools to study melatonin pathways and actions. *Trends in Pharmacological Sciences* **2005,** 26, (8), 412-419.

38. Jockers, R.; Maurice, P.; Boutin, J. A.; Delagrange, P., Melatonin receptors, heterodimerization, signal transduction and binding sites: What's new? *British Journal of Pharmacology* **2008,** 154, (6), 1182-1195.

39. Witt-Enderby, P. A.; Bennett, J.; Jarzynka, M. J.; Firestine, S.; Melan, M. A., Melatonin receptors and their regulation: Biochemical and structural mechanisms. *Life Sciences* **2003,** 72, (20), 2183-2198.

40. Farce, A.; Chugunov, A. O.; Logé, C.; Sabaouni, A.; Yous, S.; Dilly, S.; Renault, N.; Vergoten, G.; Efremov, R. G.; Lesieur, D.; Chavatte, P., Homology modeling of MT_1 and MT_2 receptors. *European Journal of Medicinal Chemistry* **2008,** 43, (9), 1926-1944.

41. von Gall, C.; Stehle, J.; Weaver, D., Mammalian melatonin receptors: Molecular biology and signal transduction. *Cell and Tissue Research* **2002,** 309, (1), 151-162.

42. Dubocovich, M. L.; Rivera-Bermudez, M. A.; Gerdin, M. J.; Masana, M. I., Molecular pharmacology, regulation and function of mammalian melatonin receptors. *Frontiers in Bioscience* **2003,** 8, d1093-1108.

43. Dubocovich, M. L., Melatonin receptors: Role on sleep and circadian rhythm regulation. *Sleep Medicine* **2007,** 8, Supplement 3, (0), S34-S42.

44. Hardeland, R.; Cardinali, D. P.; Srinivasan, V.; Spence, D. W.; Brown, G. M.; Pandi-Perumal, S. R., Melatonin—A pleiotropic, orchestrating regulator molecule. *Progress in Neurobiology* **2011,** 93, (3), 350-384.

45. Monti, J. M.; Pandi-Perumal, S. R.; Möhler, H., *GABA and sleep: Molecular, functional and clinical aspects* Springer Basel: 2010; p 286-287.

46. Pandi-Perumal, S. R.; Trakht, I.; Srinivasan, V.; Spence, D. W.; Maestroni, G. J. M.; Zisapel, N.; Cardinali, D. P., Physiological effects of melatonin: Role of melatonin receptors and signal transduction pathways. *Progress in Neurobiology* **2008,** 85, (3), 335-353.

47. Masana, M. I.; Dubocovich, M. L., Melatonin receptor signaling: Finding the path through the dark. . *Science Signaling: Signal Transduction Knowledge Environment* **2001,** 2001, (107), pe. 39.

48. Gállego, J.; Toledo, J. B.; Urrestarazu, E.; Iriarte, J., Clasificación de los trastornos del sueño. *Anales del Sistema Sanitario de Navarra* **2007,** 30, (Supl. 1), 19-36.

49. Colten, H. R.; Altevogt, B. M., *Sleep disorders and sleep deprivation: An unmet public health problem*. National Academies Press: Washington (USA), 2006; p 56.

50. Iriarte, J.; Artieda, J., Trastornos del sueño. *Revista de Medicina de la Universidad de Navarra* **2005,** 49, (1), 6-9.

51. Brain basics: Understanding sleep. NIH publication No.06-3440-c (2007). http://www.ninds.nih.gov/disorders/brain_basics/understanding_sleep.htm#sleep_disorders

52. AASM The international classification of sleep disorders, revised: Diagnostic and coding manual.

53. Summers, M. O.; Crisostomo, M. I.; Stepanski, E. J., Recent developments in the classification, evaluation, and treatment of insomnia. *Chest* **2006,** 130, (1), 276-286.

54. Lee-Chiong, T. L., *Sleep: A comprehensive handbook*. Wiley-Liss, a John Wiley & Sons Inc. publication: New Yersey (USA), 2006.

55. Mayer, G.; Jennum, P.; Riemann, D.; Dauvilliers, Y., Insomnia in central neurologic diseases – Occurrence and management. *Sleep Medicine Reviews* **2011,** 15, (6), 369-378.

56. Bjorvatn, B.; Pallesen, S., A practical approach to circadian rhythm sleep disorders. *Sleep Medicine Reviews* **2009,** 13, (1), 47-60.

57. Thorpy, M. J., Which clinical conditions are responsible for impaired alertness? *Sleep Medicine* **2005,** 6, Supplement 1, (0), S13-S20.

58. Monti, J. M.; Pandi-Perumal, S. R.; Möhler, H., *GABA and sleep: Molecular, functional and clinical aspects* Springer Basel: 2010; p 344.

59. Colten, H. R.; Altevogt, B. M., *Sleep disorders and sleep deprivation: An unmet public health problem*. National Academies Press: Washington (USA), 2006; p 88.

60. Iber, C., Sleep-related breathing disorders. *Neurologic Clinics* **2005,** 23, (4), 1045-1057.

61. Colten, H. R.; Altevogt, B. M., *Sleep disorders and sleep deprivation: An unmet public health problem*. National Academies Press: Washington (USA), 2006; p 82.

62. Freedom, T., Classification of sleep disorders. *Disease-a-Month* **2011,** 57, (7), 323-327.

63. Neubauer, D.; Erman, M. K.; Zee, P., New perspectives in the diagnosis and management of insomnia. *CNS Spectrums. The International Journal of Neuopsychiatic Medicine* **2008,** 13, (12 (Suppl 17)).

64. Léger, D.; Bayon, V., Societal costs of insomnia. *Sleep Medicine Reviews* **2010,** 14, (6), 379-389.

65. *Grupo de trabajo de la guía de práctica clínica para el manejo de pacientes con insomnio en atención primaria. Guía de práctica clínica para el manejo de pacientes con insomnio en atención primaria*. Plan de calidad para el sistema nacional de salud del ministerio de sanidad y política social. Unidad de evaluación de tecnologías sanitarias. Agencia Laín Entralgo. Guías de práctica clínica en el SNS: UETS Nº 2007/5-1: Comunidad de Madrid (Spain), 2009; p 23-25.

66. Roth, T., Insomnia: Definition, prevalence, etiology, and consequences *Journal of Clinical Sleep Medicine* **2007,** 3, (Suppl 5), S7–S10.

67. Colten, H. R.; Altevogt, B. M., *Sleep disorders and sleep deprivation: An unmet public health problem*. National Academies Press: Washington (USA), 2006; p 75-78.

68. Colten, H. R.; Altevogt, B. M., *Sleep disorders and sleep deprivation: An unmet public health problem*. National Academies Press: Washington (USA), 2006; p 138-140.

69. Staner, L., Comorbidity of insomnia and depression. *Sleep Medicine Reviews* **2010,** 14, (1), 35-46.

70. Ebert, B.; Wafford, K. A.; Deacon, S., Treating insomnia: Current and investigational pharmacological approaches. *Pharmacology & Therapeutics* **2006,** 112, (3), 612-629.

71. Alarma-Estrany, P.; Pintor, J., Melatonin receptors in the eye: Location, second messengers and role in ocular physiology. *Pharmacology & Therapeutics* **2007,** 113, (3), 507-522.

72. Spadoni, G.; Bedini, A.; Rivara, S.; Mor, M., Melatonin receptor agonists: New options for insomnia and depression treatment. *CNS Neuroscience & Therapeutics* **2011,** 17, (6), 733-741.

73. Zlotos, D. P., Recent advances in melatonin receptor ligands. *Archiv der Pharmazie* **2005,** 338, (5-6), 229-247.

74. Buscemi, N.; Vandermeer, B.; Pandya, R.; Hooton, N.; Tjosvold, L.; Hartling, L.; Baker, G.; Vohra, S.; Klassen, T. *Melatonin for treatment of sleep disorders.* ; AHRQ (Agency for Healthcare Research and Quality): Rockville (USA), 2004.

75. *Withdrawal assessment report for Ramelteon Takeda global research and development centre (Europe) LTD*; EMEA (European Medicines Agency): London (UK), 2008.

76. Srinivasan, V.; Brzezinski, A.; Pandi-Perumal, S. R.; Spence, D. W.; Cardinali, D. P.; Brown, G. M., Melatonin agonists in primary insomnia and depression-associated insomnia: Are they superior to sedative-hypnotics? *Progress in Neuro-Psychopharmacology and Biological Psychiatry* **2011,** 35, (4), 913-923.

77. Hickie, I. B.; Rogers, N. L., Novel melatonin-based therapies: potential advances in the treatment of major depression. *The Lancet* **2011,** 378, (9791), 621-631.

78. ADIR, 284017. *Drug Data Report* **2000,** 22, (2), 123.

79. Servier, 284685. *Drug Data Report* **2000,** 22, (3), 220.

80. ADIR, 294907. *Drug Data Report* **2001,** 23, (2), 117.

81. Bristol-Myers-Squibb., 299825. *Drug Data Report* **2001,** 23, (6), 535.

82. Bristol-Myers-Squibb., 341109. *Drug Data Report* **2003,** 25, (7), 595.

83. Bristol-Myers-Squibb., 338239. *Drug Data Report* **2003,** 25, (5), 405.

84. Servier, 372483. *Drug Data Report* **2004,** 26, (7), 611-612.

85. Servier, 396644. *Drug Data Report* **2005,** 27, (5), 413.

86. Takeda, Rozerem®. Ramelteon. 255673. *Drug Data Report* **2005,** 27, (27), 798.

87. Bristol-Myers-Squibb., VEC-162. 31846. *Drug Data Report* **2007,** 29, (11), 984-985.

88. Takeda, 468450. *Drug Data Report* **2008,** 30, (2), 109.

89. Takeda, 630597. *Drug Data Report* **2008,** 30, (5), 390.

90. Takeda, 634689. *Drug Data Report* **2008,** 30, (7), 576.

91. Li, P.-K.; Chu, G.-H.; Gillen, M. L.; Witt-Enderby, P. A., Synthesis and receptor binding studies of quinolinic derivatives as melatonin receptor ligands. *Bioorganic & Medicinal Chemistry* **1997,** 7, (17), 2177-2180.

92. Cohen, N. C., *Guidebook on molecular modeling in drug design*. London (UK), 1996; p 1-5.

93. *HyperChem® computational chemistry. Practical guide.*; Canada, 1996; p 7.

94. Cohen, N. C., *Guidebook on molecular modeling in drug design*. London (UK), 1996; p 64-76.

95. SYBYL®-X 1.1 Force field manual. **2010**.

96. *Chemistry collection: Basic chemistry. User guide.*; San Diego (USA), 2008

97. *Data modeling collection: Basic data modeling. User guide.*; San Diego (USA), 2008.

98. Hasegawa, K.; Matsuoka, S.; Arakawa, M.; Funatsu, K., New molecular surface-based 3D-QSAR method using Kohonen Neural Network and 3-way PLS. . *Computers & Chemistry* **2002,** 26, 583-589.

99. Kohonen, T.; Hynninen, J.; Kangas, J.; Laaksonen, J., SOM PAK: The Self-Organizing Map Program Package. *Helsinki University of Technology. Faculty of Information Technology. Laboratory of Computer and Information Science.* **1996**, (Report A31).

100. Waring, M.; Ben-Hadda, T.; Kotchevar, A.; Ramdani, A.; Touzani, R.; Elkadiri, S.; Hakkou, A.; Bouakka, M.; Ellis, T., 2,3-Bifunctionalized quinoxalines: Synthesis, DNA interactions and evaluation of anticancer, anti-tuberculosis and antifungal activity. . *Molecules* **2002,** 7, 641-656.

101. Burguete, A.; Pontiki, E.; Hadjipavlou-Litina, D.; Ancizu, S.; Villar, R.; Solano, B.; Moreno, E.; Torres, E.; Pérez, S.; Aldana, I.; Monge, A., Synthesis and Biological Evaluation of New Quinoxaline Derivatives as Antioxidant and Anti-Inflammatory Agents. *Chemical Biology & Drug Design* **2011,** 77, (4), 255-267.

102.   Moreno, E.; Ancizu, S.; Pérez-Silanes, S.; Torres, E.; Aldana, I.; Monge, A., Synthesis and antimycobacterial activity of new quinoxaline-2-carboxamide 1,4-di-*N*-oxide derivatives. *European Journal of Medicinal Chemistry* **2010,** 45, (10), 4418-4426.

103.   Ancizu, S.; Moreno, E.; Solano, B.; Villar, R.; Burguete, A.; Torres, E.; Pérez-Silanes, S.; Aldana, I.; Monge, A., New 3-methylquinoxaline-2-carboxamide 1,4-di-*N*-oxide derivatives as anti-mycobacterium tuberculosis agents. *Bioorganic & Medicinal Chemistry* **2010,** 18, 2713–2719.

104.   Ancizu, S.; Moreno, E.; Torres, E.; Burguete, A.; Pérez-Silanes, S.; Benítez, D.; Villar, R.; Solano, B.; Marín, A.; Aldana, I.; Cerecetto, H.; González, M.; Monge, A., Heterocyclic-2-carboxylic acid (3-cyano-1,4-di-*N*-oxidequinoxalin-2-yl)amide derivatives as hits for the development of neglected disease drugs. *Molecules* **2009,** 14, (6), 2256-2272.

105.   Olayiwola, G.; Obafemi, C. A.; Taiwo, F. O., Synthesis and neuropharmacological activity of some quinoxalinone derivatives. *African Journal of Biotechnology* **2007,** 6, (6), 777-786.

106.   Smith, M. B.; March, J., *March's advanced organic chemistry. Reactions, mechanisms, and structure*. 5th ed.; John Wiley & Sons, Inc.: New York (USA), 2001; p 862.

107.   Beavers, M. P.; Dudash, J.; Zhang, Y., WO/2005/067932. **2005**.

108.   Smith, M. B.; March, J., *March's advanced organic chemistry. Reactions, mechanisms, and structure*. 5th ed.; John Wiley & Sons, Inc.: New York (USA), 2001; p 856.

109.   Bunnett, J. F., Aromatic-substitution by Srn1 mechanism. *Accounts of Chemical Research* **1978,** 11, (11), 413-420.

110.   Smith, M. B.; March, J., *March's advanced organic chemistry. Reactions, mechanisms, and structure*. 5th ed.; John Wiley & Sons, Inc.: New York (USA), 2001; p 867.

111.   Smith, M. B.; March, J., *March's advanced organic chemistry. Reactions, mechanisms, and structure*. 5th ed.; John Wiley & Sons, Inc.: New York (USA), 2001; p 454-455.

112.   Ellis, G. P.; Romney-Alexander, T. M., Cyanation of aromatic halides. *Chemical Reviews* **1987,** 87, (4), 779-794.

113.   Hermann, K.; Simchen, G., Synthese cyan-substituierter heterocyclen mit tetraethyl-ammoniumcyanid. *Liebigs Annalen der Chemie* **1981,** 1981, (2), 333-341.

114.   Smith, M. B.; March, J., *March's advanced organic chemistry. Reactions, mechanisms, and structure*. 5th ed.; John Wiley & Sons, Inc.: New York (USA), 2001; p 857.

115.   Smith, M. B.; March, J., *March's advanced organic chemistry. Reactions, mechanisms, and structure*. 5th ed.; John Wiley & Sons, Inc.: New York (USA), 2001; p 850.

116.   Thomas, J. M.; Thomas, W. J., *Principles and practice of heterogeneous catalysis*. 3th ed.; VCH Verlagsgesellschaft mbH and VCH Publishers Inc.: Weinheim (Germany) and New York (USA), 2005; p 36.

117.   Smith, M. B.; March, J., *March's advanced organic chemistry. Reactions, mechanisms, and structure*. 5th ed.; John Wiley & Sons, Inc.: New York (USA), 2001; p 1006.
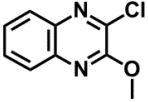
118.   Garcia Alonso, F. J.; Garcia Sanz, M.; Riera, V.; Anillo Abril, A.; Tiripicchio, A.; Ugozzoli, F., Reactions of unsaturated dihydrido carbonyl complexes of manganese(I) with nitriles and isonitriles. Preparation and characterization of the first binuclear μ, η^1, η^2-NCR derivatives *Organometallics* **1992,** 11, (2), 801-808.

119.   Chojecki, A. Selective hydrogenation of butyronitrile over Raney-metals. Universität München zur Erlangung des akademischen Grades eines, 2004.

120.   Smith, M. B.; March, J., *March's advanced organic chemistry. Reactions, mechanisms, and structure*. 5th ed.; John Wiley & Sons, Inc.: New York (USA), 2001; p 1204.
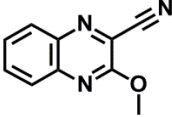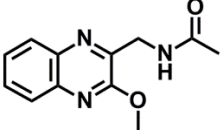
121.   Barrault, J.; Pouilloux, Y., Synthesis of fatty amines. Selectivity control in presence of multifunctional catalysts. *Catalysis Today* **1997,** 37, (2), 137-153.

122.   Wallez, V.; Durieux-Poissonnier, S.; Chavatte, P.; Boutin, J. A.; Audinot, V.; Nicolas, J.-P.; Bennejean, C.; Delagrange, P.; Renard, P.; Lesieur, D., Synthesis and structure−affinity−activity relationships of novel benzofuran derivatives as MT_2 melatonin receptor selective ligands. *Journal of Medicinal Chemistry* **2002,** 45, (13), 2788-2800.

123.   Carey, F. A., *Organic chemistry*. 3rd ed.; Department of Chemistry. University of Virginia. A Division of Mc-Gram-Hill companies.: New York (USA), 1996; p 833.

124.   Clayden, J.; Greeves, N.; Warren, S.; Wothers, P., *Organic chemistry*. 1st ed.; Oxford University Press Inc.: New York (USA), 2001; p 284.

125.   Galiano, S.; Ceras, J.; Cirauqui, N.; Pérez, S.; Juanenea, L.; Rivera, G.; Aldana, I.; Monge, A., Novel series of substituted biphenylmethyl urea derivatives as MCH-R1 antagonists for the treatment of obesity. *Bioorganic & Medicinal Chemistry* **2007,** 15, (11), 3896-3911.

126.   Mann, F. G.; Saunders, B. C., *Practical organic chemistry*. 4th ed.; Logman Inc.: New York (USA), 1978; p 375.

127.   Arnold, R. G.; Nelson, J. A.; Verbanc, J. J., Recent advances in isocyanate chemistry. *Chemical Reviews* **1957,** 57, (1), 47-76.

128.   Yuste, F.; Saldaña, M.; Walls, F., Selective reduction of aromatic nitro compounds containing *O*- and *N*-benzyl groups with hydrazine and Raney nickel. *Tetrahedron Letters* **1982,** 23, (2), 147-148.

129.   Smith, M. B.; March, J., *March's advanced organic chemistry. Reactions, mechanisms, and structure*. 5th ed.; John Wiley & Sons, Inc.: New York (USA), 2001; p 1553.

130.   Li, J. J., *Beirut reaction. Name reactions*. Springer Berlin Heidelberg: New Yersey (USA), 2006; p 43-44.

131.   Haddadin, M. J.; Taha, M. U.; Jarrar, A. A.; Issidorides, C. H., Reaction of benzofurazan oxide with unsymmetrical 1,3-diketones; Steric and polar effects. *Tetrahedron* **1976,** 32, (6), 719-724.

132.   Abushanab, E.; Alteri, N. D., Quinoxaline 1,4-dioxides. Substituent effects on the reaction of benzofurazan 1-oxides with carbonyl compounds. *The Journal of Organic Chemistry* **1975,** 40, (2), 157-160.

133.   Vicente, E.; Pérez-Silanes, S.; Lima, L. M.; Ancizu, S.; Burguete, A.; Solano, B.; Villar, R.; Aldana, I.; Monge, A., Selective activity against Mycobacterium tuberculosis of new quinoxaline 1,4-di-*N*-oxides. *Bioorganic & Medicinal Chemistry* **2009,** 17, (1), 385-389.

134.   Carey, F. A., *Organic chemistry*. 3rd ed.; Department of Chemistry. University of Virginia. A Division of Mc-Gram-Hill companies.: New York (USA), 1996; p 927-930.

135.   Carey, F. A., *Organic chemistry*. 3rd ed.; Department of Chemistry. University of Virginia. A Division of Mc-Gram-Hill companies.: New York (USA), 1996; p 933.

136.   Smith, M. B.; March, J., *March's advanced organic chemistry. Reactions, mechanisms, and structure*. 5th ed.; John Wiley & Sons, Inc.: New York (USA), 2001; p 934-935.
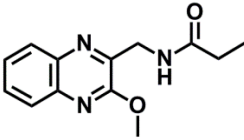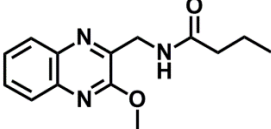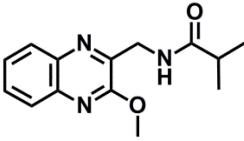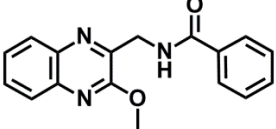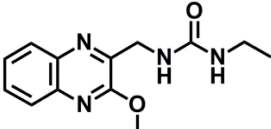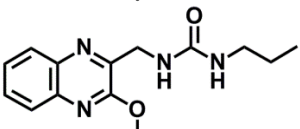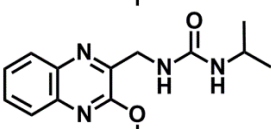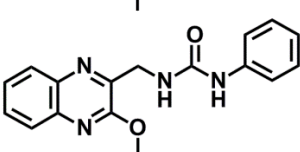
137.   Doyle, M. P.; Dellaria, J. F.; Siegfried, B.; Bishop, S. W., Reductive deamination of arylamines by alkyl nitrites in *N,N*-dimethylformamide. A direct conversion of arylamines to aromatic hydrocarbons. *The Journal of Organic Chemistry* **1977,** 42, (22), 3494-3498.

138.   Monge, A.; Martinez-Crespo, F. J.; Lopez de Cerain, A.; Palop, J. A.; Narro, S.; Senador, V.; Marin, A.; Sainz, Y.; Gonzalez, M., Hypoxia-selective agents derived from 2-quinoxalinecarbonitrile 1,4-di-*N*-oxides. 2. *Journal of Medicinal Chemistry* **1995,** 38, (22), 4488-4494.

139.   Smith, M. B.; March, J., *March's advanced organic chemistry. Reactions, mechanisms, and structure*. 5th ed.; John Wiley & Sons, Inc.: New York (USA), 2001; p 928.

140.   Ek, F.; Axelsson, O.; Wistrand, L.-G.; Frejd, T., Aromatic allylation via diazotization: Metal-free C−C bond formation. *The Journal of Organic Chemistry* **2002,** 67, (18), 6376-6381.

141.   Solano, B.; Junnotula, V.; Marín, A.; Villar, R.; Burguete, A.; Vicente, E.; Pérez-Silanes, S.; Aldana, I.; Monge, A.; Dutta, S.; Sarkar, U.; Gates, K. S., Synthesis and biological evaluation of new 2-arylcarbonyl-3-trifluoromethylquinoxaline 1,4-di-*N*-oxide derivatives and their reduced analogues. *Journal of Medicinal Chemistry* **2007,** 50, (22), 5485-5492.

142.   Haddadin, M. J.; Zahr, G. E.; Rawdah, T. N.; Chelhot, N. C.; Issidorides, C. H., Deoxygenation of quinoxaline *N*-oxides and related compounds. *Tetrahedron* **1974,** 30, (5), 659-666.

143.   Blankenhorn, G.; Moore, E. G., Sulfoxylate Ion (HSO^{2-}), the hydride donor in dithionite-dependent reduction of NAD^+ analogs. *Journal of the American Chemical Society* **1980,** 102, (3), 1092-1098.

144.   Lough, S. M.; McDonald, J. W., Synthesis of tetraethylammonium dithionite and its dissociation to the sulfur dioxide radical anion in organic solvents. *Inorganic Chemistry* **1987,** 26, (13), 2024-2027.

145.   Gassman, P. G.; Rasmy, O. M.; Murdock, T.; Saito, K., Mechanism of sodium dithionite reduction of aldehydes and ketones. *The Journal of Organic Chemistry* **1981,** 46, (26), 5457-5458.

146.   Joule, J. A.; Mills, K., Heterocyclic chemistry. **2010**, 386-387.

147.   Guida, W. C.; Mathre, D. J., Phase-transfer alkylation of heterocycles in the presence of 18-crown-6 and potassium *tert*-butoxide. *The Journal of Organic Chemistry* **1980,** 45, (16), 3172-3176.

148.   Smith, M. B.; March, J., *March's advanced organic chemistry. Reactions, mechanisms, and structure*. 5th ed.; John Wiley & Sons, Inc.: New York (USA), 2001; p 574.

149.   Clayden, J.; Greeves, N.; Warren, S.; Wothers, P., *Organic chemistry*. 1st ed.; Oxford University Press Inc.: New York (USA), 2001; p 1166.

150.   Gilchrist, A., *GPCR molecular pharmacology and drug targeting. Shifting paradigms and new directions*. John Wiley & Sons, Inc.: Hoboken, New Jersey (USA), 2010; p 215.

151.   Pleuvry, B. J., Receptors, agonists and antagonists. **2004,** 5, (10), 350-352.

152.   Gilchrist, A., *GPCR molecular pharmacology and drug targeting. Shifting paradigms and new directions*. John Wiley & Sons, Inc.: Hoboken, New Jersey (USA), 2010; p 209-211.

153.   Harrison, C.; Traynor, J. R., The [^{35}S]GTPγS binding assay: Approaches and applications in pharmacology. *Life Sciences* **2003,** 74, (4), 489-508.
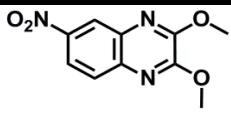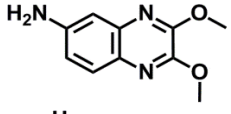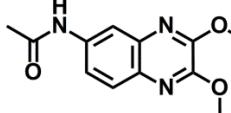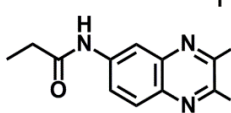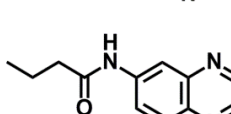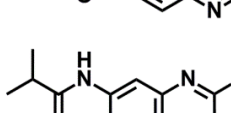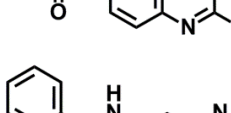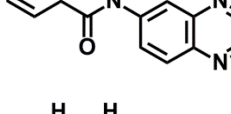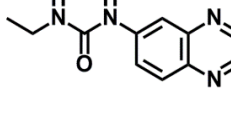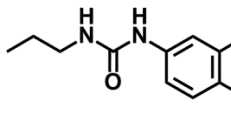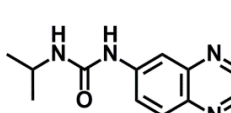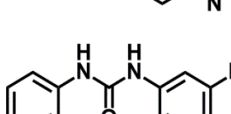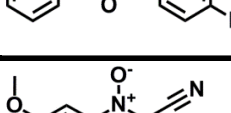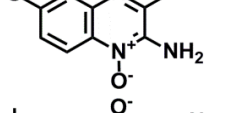
154.   De Vries, L.; Zheng, B.; Fischer, T.; Elenko, E.; Farquhar, M. G., The regulator of G protein signaling family. *Annual Review of Pharmacology and Toxicology* **2000,** 40, (1), 235-271.

155.   Conway, S.; Drew, J. E.; Canning, S. J.; Barrett, P.; Jockers, R.; Strosberg, A. D.; Guardiola-Lemaitre, B.; Delagrange, P.; Morgan, P. J., Identification of Mel$_{1a}$ melatonin receptors in the human embryonic kidney cell line HEK293: Evidence of G protein-coupled melatonin receptors which do not mediate the inhibition of stimulated cyclic AMP levels. *FEBS Letters* **1997,** 407, (1), 121-126.

156.   Stanford-University. School of medicin. High-throughput bioscience center (HTBC). ISIS/DRAW. http://htbc.stanford.edu/ISISDRAW.pdf

157.   Accelrys®. Pipeline Pilot overview.

158.   Accelrys®. Discovery Studio.

159.   *Spartan'08 for Windows, Macintosh and Linux. Tutorial and user's guide*. Wavefunction, Inc. in collaboration with Q-Chem, Inc: USA.

160.   Zarranz, B.; Jaso, A.; Aldana, I.; Monge, A., Synthesis and antimycobacterial nctivity of New quinoxaline-2-carboxamide 1,4-di-*N*-oxide derivatives. *Bioorganic & Medicinal Chemistry* **2003,** 11, 2149–2156.

161.   Pretsch, E.; Bühlmann, P.; Affolter, C.; Herrera, A.; Martínez, R., *Determinación estructural de compuestos orgánicos*. Elservier Masson: Barcelona (Spain), 2002; p 193.

162.   Pretsch, E.; Bühlmann, P.; Affolter, C.; Herrera, A.; Martínez, R., *Determinación estructural de compuestos orgánicos*. Elservier Masson: Barcelona (Spain), 2002; p 196.

163.   Greenberg, A.; Breneman, C. M.; Liebman, J. F., *The amide linkage: Structural significance in chemistry, biochemistry and materials science*. John wiley &Sons, Inc. : New Jersey (USA), 2003.

164.   Pretsch, E.; Bühlmann, P.; Affolter, C.; Herrera, A.; Martínez, R., *Determinación estructural de compuestos orgánicos*. Elservier Masson: Barcelona (Spain), 2002; p 180-183.

165.   Pretsch, E.; Bühlmann, P.; Affolter, C.; Herrera, A.; Martínez, R., *Determinación estructural de compuestos orgánicos*. Elservier Masson: Barcelona (Spain), 2002; p 168.

166.   Pretsch, E.; Bühlmann, P.; Affolter, C.; Herrera, A.; Martínez, R., *Determinación estructural de compuestos orgánicos*. Elservier Masson: Barcelona (Spain), 2002; p 176-178.

167.   Malone, J.; McGarry, K.; Wermter, S.; Bowerman, C., Data mining using rule extraction from Kohonen self-organising maps. *Neural Comput. Appl.* **2006,** 15, (1), 9-17.

168.   Cummings, D. F.; Canseco, D. C.; Sheth, P.; Johnson, J. E.; Schetz, J. A., Synthesis and structure–affinity relationships of novel small molecule natural product derivatives capable of discriminating between serotonin 5-HT$_{1A}$, 5-HT$_{2A}$, 5-HT$_{2C}$ receptor subtypes. *Bioorganic & Medicinal Chemistry* **2010,** 18, (13), 4783-4792.

169.   Hamprecht, D.; Micheli, F.; Tedesco, G.; Checchia, A.; Donati, D.; Petrone, M.; Terreni, S.; Wood, M., Isoindolone derivatives, a new class of 5-HT$_{2C}$ antagonists: Synthesis and biological evaluation. *Bioorganic & Medicinal Chemistry Letters* **2007,** 17, (2), 428-433.
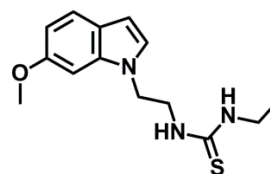
Relationship of synthesized compounds

Quinoxalines	Chemical Name	Structures
[Q1-1]	2-Chloro-3-methoxyquinoxaline	
[Q1-2]	3-methoxyquinoxaline-2-carbonitrile	
[Q1-3.I]	*N*-[(3-methoxyquinoxalin-2-yl)methyl]acetamide	
[Q1-3.II]	*N*-[(3-methoxyquinoxalin-2-yl)methyl]propionamide	
[Q1-3.III]	*N*-[(3-methoxyquinoxalin-2-yl)methyl]butyramide	
[Q1-3.IV]	*N*-[(3-methoxyquinoxalin-2-yl)methyl]isobutyramide	
[Q1-3.V]	*N*-[(3-methoxyquinoxalin-2-yl)methyl]benzamide	
[Q1-3.VI]	1-ethyl-3-[(3-methoxyquinoxalin-2-yl)methyl]urea	
[Q1-3.VII]	1-[(3-methoxyquinoxalin-2-yl)methyl]-3-propylurea	
[Q1-3.VIII]	1-Isopropyl-3-[(3-methoxyquinoxalin-2-yl)methyl]urea	
[Q1-3.IX]	1-[(3-methoxyquinoxalin-2-yl)methyl]-3-phenylurea	

[Q2-1]	2,3-dimethoxy-6-nitroquinoxaline	
[Q2-2]	2,3-dimethoxyquinoxaline-6-amine	
[Q2-3.I]	N-(2,3-dimethoxyquinoxalin-6-yl)acetamide	
[Q2-3.II]	N-(2,3-dimethoxyquinoxalin-6-yl)propionamide	
[Q2-3.III]	N-(2,3-dimethoxyquinoxalin-6-yl)butyramide	
[Q2-3.IV]	N-(2,3-dimethoxyquinoxalin-6-yl)isobutyramide	
[Q2-3.V]	N-(2,3-dimethoxyquinoxalin-6-yl)benzamide	
[Q2-3.VI]	1-(2,3-dimethoxyquinoxalin-6-yl)-3-ethylurea	
[Q2-3.VII]	1-(2,3-dimethoxyquinoxalin-6-yl)-3-propylurea	
[Q2-3.VIII]	1-(2,3-dimethoxyquinoxalin-6-yl)-3-isopropylurea	
[Q2-3.IX]	1-(2,3-dimethoxyquinoxalin-6-yl)-3-benzamide	
[Q3-1]	2-Amino-3-cyano-1.4-di-N-oxido-6-methoxyquinoxaline	
[Q3-2]	2-Cyano-1.4-di-N-oxido-7-methoxyquinoxaline	
[Q3-3]	7-methoxyquinoxaline-2-carbonitrile	

[Q3-4.I]	*N*-[(7-methoxyquinoxalin-2-yl)methyl]acetamide	
[Q3-4.II]	*N*-[(7-methoxyquinoxalin-2-yl)methyl]propionamide	
[Q3-4.III]	*N*-[(7-methoxyquinoxalin-2-yl)methyl]butyramide	
[Q3-4.IV]	*N*-[(3-methoxyquinoxalin-2-yl)methyl]isobutyramide	
[Q3-4.V]	*N*-[(7-methoxyquinoxalin-2-yl)methyl]benzamide	

Indoles	Chemical *N*ame	Structures
[In1-1]	2-(6-methoxyindolyl)acetonitrile	
[In1-2]	2-(6-methoxyindolyl)ethanamine	
[In1-3.CDI]	*N*-(2-(6-methoxyindolyl)ethyl)-imidazole-1-carboxamide	
[In1-3.I]	1-allyl-3-(2-(6-methoxyindolyl)ethyl)urea	
[In1-3.II]	1-ethyl-3-(2-(6-methoxyindolyl)ethyl)urea	
[In1-3.III]	1-(2-(6-methoxyindolyl)ethyl)-3-propylurea	
[In1-3.IV]	1-allyl-3-(2-(6-methoxyindolyl)ethyl)thiourea	

249

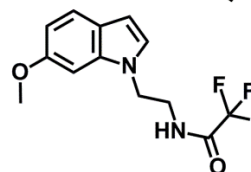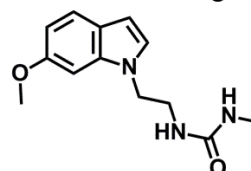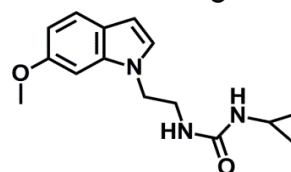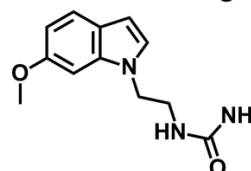|		
[In1-3.V]	1-ethyl-3-(2-(6-methoxyindolyl)ethyl)thiourea	
[In1-3.VI]	1-(2-(6-methoxyindolyl)ethyl)-3-propylthiourea	
[In1-3.VII]	1-(2-(6-methoxyindolyl)ethyl)-3-methylthiourea	
[In1-3.VIII]	1-cyclopropyl-3-(2-(6-methoxyindolyl)ethyl)thiourea	
[In1-3.IX]	*N*-(2-(6-methoxyindolyl)ethyl)methanesulfonamide	
[In1-3.X]	2,2,2-trifluoro-*N*-(2-(6-methoxyindolyl)ethyl)acetamide	
[In1-4.XI]	1-(2-(6-methoxyindolyl)ethyl)-3-methylurea	
[In1-4.XII]	1-cyclopropyl-3-(2-(6-methoxyindolyl)ethyl)urea	
[In1-3.XIII]	1-(2-(6-methoxyindolyl)ethyl)urea	
[In2-1]	2-(7-methoxyindolyl)acetonitrile	
[In2-2]	2-(7-methoxyindolyl)ethanamine	
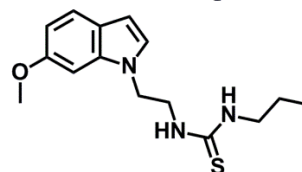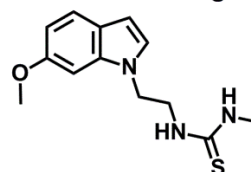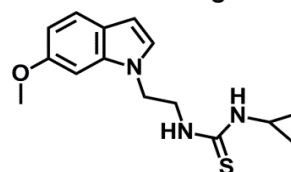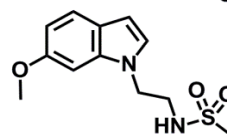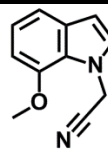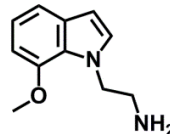
[In2-3.CDI]   *N*-(2-(7-methoxyindolyl)ethyl)-imidazole-1-carboxamide

[In2-3.I]   1-allyl-3-(2-(7-methoxyindolyl)ethyl)urea

[In2-3.II]   1-ethyl-3-(2-(7-methoxyindolyl)ethyl)urea

[In2-3.III]   1-(2-(7-methoxyindolyl)ethyl)-3-propylurea

[In2-3.IV]   1-allyl-3-(2-(7-methoxyindolyl)ethyl)thiourea

[In2-3.V]   1-ethyl-3-(2-(7-methoxyindolyl)ethyl)thiourea

[In2-3.VI]   1-(2-(7-methoxyindolyl)ethyl)-3-propylthiourea

[In2-3.VII]   1-(2-(7-methoxyindolyl)ethyl)-3-methylthiourea

[In2-3.VIII]   1-cyclopropyl-3-(2-(7-methoxyindolyl)ethyl)thiourea

[In2-3.IX]   *N*-(2-(7-methoxyindolyl)ethyl)methanesulfonamide

[In2-3.X]   2,2,2-trifluoro-*N*-(2-(7-methoxyindolyl)ethyl)acetamide
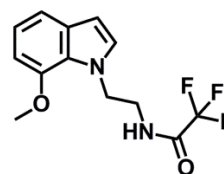
[In2-4.XI]   1-(2-(7-methoxyindolyl)ethyl)-3-methylurea

[In2-4.XII]   1-cyclopropyl-3-(2-(7-methoxyindolyl)ethyl)urea