# Drug Discovery Targeted to Transthyretin Related Amyloidosis

Daniel Blasi Pérez

**ADVERTIMENT**. La consulta d’aquesta tesi queda condicionada a l’acceptació de les següents condicions d'ús: La difusió d’aquesta tesi per mitjà del servei TDX (**www.tdx.cat**) ha estat autoritzada pels titulars dels drets de propietat intel·lectual únicament per a usos privats emmarcats en activitats d’investigació i docència. No s’autoritza la seva reproducció amb finalitats de lucre ni la seva difusió i posada a disposició des d’un lloc aliè al servei TDX. No s’autoritza la presentació del seu contingut en una finestra o marc aliè a TDX (framing). Aquesta reserva de drets afecta tant al resum de presentació de la tesi com als seus continguts. En la utilització o cita de parts de la tesi és obligat indicar el nom de la persona autora.

**ADVERTENCIA**. La consulta de esta tesis queda condicionada a la aceptación de las siguientes condiciones de uso: La difusión de esta tesis por medio del servicio TDR (**www.tdx.cat**) ha sido autorizada por los titulares de los derechos de propiedad intelectual únicamente para usos privados enmarcados en actividades de investigación y docencia. No se autoriza su reproducción con finalidades de lucro ni su difusión y puesta a disposición desde un sitio ajeno al servicio TDR. No se autoriza la presentación de su contenido en una ventana o marco ajeno a TDR (framing). Esta reserva de derechos afecta tanto al resumen de presentación de la tesis como a sus contenidos. En la utilización o cita de partes de la tesis es obligado indicar el nombre de la persona autora.

**WARNING**. On having consulted this thesis you’re accepting the following use conditions: Spreading this thesis by the TDX (**www.tdx.cat**) service has been authorized by the titular of the intellectual property rights only for private uses placed in investigation and teaching activities. Reproduction with lucrative aims is not authorized neither its spreading and availability from a site foreign to the TDX service. Introducing its content in a window or frame foreign to the TDX service is not authorized (framing). This rights affect to the presentation summary of the thesis as well as to its contents. In the using or citation of parts of the thesis it’s obliged to indicate the name of the author.

Programa de Química Orgánica

Tesis Doctoral

# Drug Discovery Targeted to Transthyretin Related Amyloidosis

Daniel Blasi Pérez

Dirigida y revisada por:

| Dr. Jordi Ramon Quintana Ruiz<br>(Parc Científic de Barcelona) | Dra Nuria Boada Centeno<br>(Universitat Pompeu Fabra) |
|---|---|

Tutelada por:

Dr. Fernando Albericio Palomera
(Universitat de Barcelona)

Barcelona, 2012

Memoria presentada por

Daniel Blasi Pérez

para optar al grado de Doctor por la Universidad de Barcelona

Dirigida y revisada por:

Dr. Jordi Ramon Quintana Ruiz

Dra. Nuria B. Centeno

Tutelada por:

Dr. Fernando Albericio Palomera

Programa de Química Orgánica

Barcelona, 2012

*A mi niña…*

"Life is what happens to you while you're busy making other plans"

John Lennon

## Acknowledgements

Sonará tópico, pero estas palabras son las más difíciles, con diferencia, de todas las escritas en esta tesis.

A lo largo de estos 4 años he tenido la suerte de compartir muchos momentos (buenos, regulares y malos) con personas muy especiales que me han hecho madurar, como persona y como científico. Por ello quiero dedicarle unas palabras a cada uno, ahora que finaliza esta etapa de mi vida en la que todos habéis aportado un granito de arena y en uno u otro momento me habéis servido de apoyo para llegar al final de esta aventura.

En primer lugar, me gustaría dar las gracias a mis dos directores, el **Dr. Jordi Quintana** y la **Dra. Nuria B. Centeno**. Me siento muy afortunado de haber tenido la oportunidad de compartir tantas horas hablando, debatiendo y planeando los pasos a seguir con ambos a lo largo de toda la tesis. Gracias **Jordi** por haber dirigido esta tesis y enseñarme tantas cosas a lo largo de estos años. Con tu paciencia, tu sentido del humor y tu optimismo has sabido contagiarme ese entusiasmo sano y sincero por la ciencia y por la investigación, motivándome en los momentos difíciles y haciendo que todo fuera siempre mucho más fácil. Eres un ejemplo a seguir, tanto en lo profesional como en lo personal. Gracias también a ti **Nuria**, a tus sabios consejos en esos momentos de duda y a tus preguntas clave en los instantes adecuados que consiguieron encender la chispa necesaria para seguir adelante con el trabajo. Sin vosotros no habría sido posible llegar a este momento.

También quiero dedicar una líneas a todos los integrantes del Consorcio del proyecto; un grupo de personas increíbles con las que he tenido la suerte de colaborar y compartir muy buenos ratos debatiendo, filosofando y discutiendo sobre los misteriosos caminos de nuestra a veces amada y a veces odiada TTR; a la **Dra. Gemma Arsequell** y al **Dr. Gregori Valencia** por su inagotable energía y su motivador entusiasmo ante alguno de los estudios realizados; al **Dr. Toni Planas** por sus siempre acertados e incisivos comentarios y preguntas y por toda la ayuda para conocer en detalle la biofísica de la proteína; a **Lluis** y **Mónica** (unos cracks capaces de sintetizar cualquier cosa), y a **Marta** y **Ellen** (las cracks de los ensayos in vitro). Ha sido una gozada el poder colaborar con un grupo de gente así. Así mismo quiero aprovechar para agradecer la ayuda, entusiasmo y energías que me ha transmitido una persona a la que tuve la casualidad de conocer en los cursos del master como profesor, luego como colaborador y por último como persona: al **Dr. Celerino Abad-Zapatero**. Muchas gracias por todo **Cele**, lo que empezó como una explicación por encima de algo llamado LEI ha acabado creando prácticamente una secta de la que nos has cedido el cetro. Ha sido un verdadero placer y un honor poder compartir ideas, inquietudes y horas de trabajo contigo.

Otra persona de la que tampoco quiero olvidarme: **Marta Pinto**. Gracias a ti y a tu inestimable ayuda en los primeros meses del proyecto, he conseguido llegar al final de este viaje.

Por último, aunque no por ello menos importante, quiero dirigir mi más sincero agradecimiento al **Dr. Fernando Albericio**. Él fue la persona que tuvo la paciencia y la disposición de escuchar a un atormentado chico que no sabía muy bien por donde seguir su carrera y que, tras ayudarme de forma totalmente desinteresada, me puso en contacto con toda la gente a la que he citado anteriormente. Por todo ello, y por haber accedido no sólo a ayudarme, sino a ser mi tutor en la tesis, muchísimas gracias **Fernando**.

En el día a día y a lo largo de estos 4 años, he tenido la suerte de compartir el laboratorio con gente que, en su momento llegaron y dejaron un gran hueco en el mismo y que me gustaría citar: gracias por las horas compartidas en el laboratorio **Mercè**; seguro que, a pesar de los vientos y tormentas que acechan entre todos conseguimos sacar adelante la Plataforma. Gracias también a todas las Passa l'Estiu al Parc que tuvieron la suerte (o desgracia) de compartir unos meses con nosotros en el lab: creo que aprendí yo más de vosotras que al revés. Las cuatro (**Laura**, **Diana**, **Morena** y **Marta**) habéis contribuido a que este tesis sea hoy una realidad. Hemos compartido momentos muy divertidos y os deseo a todas lo mejor.

Pero no sólo puedo sentirme afortunado por trabajar en un lab con la gente que he mencionado; desde el momento en el que entré a trabajar en el Parc (hace ya la friolera de 6 años) he tenido la inmensa suerte de sentirme "adoptado" en otro laboratorio del que casi casi me considero integrante: **Pharmamar**.

A lo largo de todo este tiempo he tenido la suerte de compartir momentos increíbles, dentro y fuera del entorno Parc con la gente que ha formado parte de Pharmamar y alrededores (léase 100 y 300). Tanto las viejas glorias como los nuevos fichajes han aportado su granito de arena en los mejores (y algunos en los peores) momentos que he vivido durante estos años y por eso me gustaría, aun a riesgo de olvidarme a alguien, dedicarles unas palabras.

**Lorena**!! (aka la cazadora de bolets estonios, jejeje!) Se te echa un poco de menos (tampoco mucho no te vayas a creer) pero me alegro que la aventura por tierras de Mordor (no porque haya volcanos, orcos o Saurones, sino porque hablan un idioma que nadie más sabe o se atreve a pronunciar :P) esté yendo tan bien. Y junto a Lorenaki, aunque técnicamente era un infiltrado del 100 (pero pharmaboy de adopción)… **Pau** (aka mi co)… otro al que se echa muchísimo de menos (momentos raje, Wii, tapeos varios, viajes en familia, puf!!) aunque también me alegra que su periplo por el norte (Winter is Coming!!) le esté yendo tan bien; te lo mereces co, y la verdad es que me siento muy afortunado de poder llamarte co y de haber compartido tantas cosas (todas buenas) los 4. Reservad hueco que en breve toca visitilla, ¿eh? Siguiendo con el repasón a Pharmacalculta… digoooo… Pharmamar (en que estaría yo pensando). **Carles** (aka **Xarly**): Montoya tio!! Jejeje! Si me pongo a hacer un resumen de todos los "efectos Mas-Blasi" que hemos compartido salen unos agradecimientos más largos que la propia tesis, y tampoco es plan, ¿no? Mejor me los guardo y seguimos ampliando la colección de efectos… ¡**Marta Paradís**! (aka Paradise)… sobreviviendo a la locura diaria del lab ya te has ganado el cielo, jejejeje! **Horten**… la que se va y vuelve siempre: habrá que poner orden a son cubano e ir preparando la visita a Cuba, ¿no? Ni que sea unos con **Ernesto** y otras contigo o todos revueltos o como sea, pero en mi lista está anotado como "pendiente & urgente", dicho queda.

¿Y el cártel Colombiano? **Xime** & **Vida** (con los respectivos "peor es nada" como se comenta ahora por allí, "**JF**" y "**Maikol**"), no me olvido de vosotras; cada comida, desayuno o fiesta es un espectáculo y un mar de risas.

De los nuevos fichajes / incorporaciones más recientes: **Lidia** (aka la intérprete… ta bueeeeeeno!!) otra que se va a ir de aventura por tierras del norte, junto a **Ramón** (aka Tituuu!). Me alegro que al final el plan esté saliendo, más o menos como lo habíais planeado y estéis ya preparando los primeros pasos del viaje. Habeis sido también miembros importante de mi "piña" particular del entorno Parc (tanto dentro como fuera) y aunque se os vaya a echar de menos, no perdemos dos compis… ganamos un rinconcito de Europa al que ya tenemos una excusa (aunque no nos suelen hacer falta muchas) para ir, jejejeje!! Y qué decir del duo dinámico **Pau** (aka Pau Pollo) e **Iván** (aka… iba a hacer alguna coña con el polvillo blanco famoso pero me la reservo para hacértela en directo, jejeje!): los pipiolillos del grupo pero que ya se han hecho un huequito y que, quizás sin saberlo, ya están en la lista de los imprescindibles. Ánimo con la tesis, que como veis, no se acaba tan mal… ¿o si? Jejejeje!

Bufff!! Es imposible no dejar sin mencionar a nadie (**Cami**, **Kate**, **Thanos**, **Iván**, **Diver**, **Bárbara**, **Ivonne** y una larguísima lista de gente que, de un modo u otro ha dejado su huella en mi camino)… de verdad, gracias a tod@s.

Cruzando el pasillo, nos adentramos en las profundidades del 300, ese lab donde, si te despistas 3 minutos, ya hay alguien nuevo que no conoces, jejeje!! Si me pongo a recordar momentos vividos con algunos de los integrantes del grupo (tanto viejas glorias como nuevos fichajes). Tantas casas rurales, cenas, fiestas, fiestas con finales de noche surrealistas, excursiones, charlas, cafés cervezas, tapas, cenas en casa, debates filosóficos sobre la ciencia, la vida, y un larguísimo etcétera. Como me pasa en Pharmamar, la lista de gente a la que podría agradecerle algo es tan larga que seguro que me voy a dejar a más de la mitad de las personas y lo siento: escribiendo estas palabras me vienen a la mente tantos momentos que me cuesta ordenar mis ideas.

Del bando de Giralt; **Nessim**, que desde el primer momento me demostró que es un tío de **** ***** (censurado) y con el que cualquier recuerdo que me venga a la cabeza en el que él esté implicado hace que sonría sin remedio (Amigos Invisibles rules!!!). Y a **Rodri**, que aunque sea un poco bohemio y "malcarado"… es nuestro bohemio y "malcarado" favorito. Gracias también a **Roger** (el hombre misterioso según algunos) el nuevo businessman del lab, jejeje! **Laura** (aka la mamma del lab :P) como dijeron no hace mucho de ti a cierta italiana hace poco "¡qué peligro! Cuando quieras fundamos la empresa de edición/producción de vídeos si vemos que esto de la química no es lo nuestro. ¡**Sol**! Guarda ese reposapiés del que tienes ahora el lujo de disfrutar, ¿eh? Y si necesitas los servicios de música en vivo/performance varias, ya sabes que el DJ Ofrécete siempre hace precio amigos. ¡**Miguelón**! Menudo crack estás hecho. Fuera bromas, fue un honor que nos dejaras compartir contigo (y con **Rubí**) el festival que significó estar… ¡¡¡3 semanas seguidas con graves efectos secundarios!!! Gracias por todo a los dos.

Y precisamente con **Rubí** (el terremoto mexicano) enlazo con el bando de Albericio para agradecer a **Albert** (tituuuuuuu!!!), **Peter** (ofrécete, no te escondas Peter), **Pelay** (si, lo siento Marta, para mí siempre serás "la Pelay" :P), a **Juan** (¡¡¡¡ponte traje Juan!!!! ¡Que sino no te respetamos!) todos los momentos vividos con todos ellos (y en general con toda la gente que me dejo y que ha estado en algún momento (**Fabio**, **Pep**, **Mendi**, y un larguísimo etcétera)

Tampoco me gustaría olvidarme de la gente que me acogió en Perugia durante 3 meses hace un año, con los que compartí una etapa, corta, pero muy intensa: gracias a tod@s, empezando por el propio profesor **Cruciani** y pasando por **Jean Pierre** (my dear flatmate), **Marco** (futittinni!), **Francesco**, **Lidia** (now in Barcelona… really glad for you), **Fabio**, **Paolo**, **Manu**, **Laura** y **Daniele**. Grazzie tanti a tutti! Ci vediamo presto :)

Un agradecimiento general también al resto de gente del PCB/cursos/doctorado con la que hemos compartido penurias, cerveza, piti a deshoras con una copa en la otra mano o momentos filosóficos (**Melga**, **Carol**, **Vilas**, los **Bernis**, **Xavi**, etc etc).

And of course, a mention to the symposium/congress networking team made through several hard beer-drinking sessions in Rome or even more beers sessions in Groningen: nice time there with incredible people (**Nelson**, **Nadia**, **Paolo**, **Adam**, **Ana**, **Javi**, **Santi**, **Alberto**, **Manué** and all the guys that I've meet during this years). Thanks guys!

Tampoco quiero olvidarme de la gente de la uni: **Kiko** (si al final haces el desfalco… acuérdate de nosotros, eh?), **Edu** (la noche es joven), **Raúl** (Raulitu chicu!), **Tito** & **Dulce**, **Ali** & **Eloy**… gracias a todos por estar siempre ahí aunque ya no nos veamos tan a menudo. Es genial vernos de tanto en tanto y tener la sensación que nada ha cambiado, que las conversaciones y las risas surgen solas, como si no hubiera pasado el tiempo.

Y para acabar, me he reservado dos de las cosas más importantes en mi vida sin las que, este día nunca habría llegado.

He tenido la suerte de encontrar una segunda familia en la que desde el primer momento fui aceptado como uno más y cada día que pasa me siento más afortunado de que me dejen formar parte de ella. Para mi sois mi familia de sangre y si he llegado a este punto el camino es gracias a vuestros ánimos, vuestros abrazos en los momentos duros y las risas que hemos compartido en los momentos alegres. **Puri**, **Paquito**, **Sergio**, **Miguel**, **Domi**, **Antonio**, **Abuela** y a **toda la familia** en general… muchas gracias. Os quiero.

Y estas últimas palabras son para la persona más importante de mi vida, a la que hace 11 años tuve la valentía de expresarle mis sentimientos y desde entonces ha sido mi luz y mi guía en todo momento: **Myriam**. Gracias por estar siempre ahí apoyándome, ayudándome a levantarme cada vez que me he caído y acompañándome cada día en todo momento. A lo largo de todos estos años has conseguido arrancarme siempre una sonrisa y entender el significado de muchas palabras que sin ti ya no tendrían sentido. Eres la persona más especial de mi vida mi niña y esta tesis es tan mía como tuya. Te quiero. Miut.

# INDEX

Introduction ........................................................................ 21

Amyloid disease ........................................................................ 23

Transthyretin ........................................................................ 23

TTR-related amyloid diseases ........................................................................ 27

Therapeutical strategies for TTR-related amyloidoses ........................................ 28

Bibliography ........................................................................ 32

Objectives ........................................................................ 37

Part I: Repurposing approach ........................................................ 39

The drug discovery process ........................................................................ 41

The repurposing approach ........................................................................ 43

Building up a repurposing tool ........................................................................ 43

Fingerprint modeling: Similarity search studies ........................................ 45

In vitro assays: kinetic turbidimetric assay ........................................ 48

Bibliography ........................................................................ 51

Part II: Pharmacophore modeling ................................................ 53

Pharmacophore modeling ........................................................................ 55

Flavones as TTR stabilizers ........................................................................ 55

Rat vs. human TTR ........................................................................ 57

Biological activity of a set of selected flavonoids ........................................ 62

Bibliography ........................................................................ 65

Part III: Ligand efficiency indices ................................................ 67

Ligand Efficiency Indices (LEI) for TTR Drug Discovery ........................................ 69

TTR binding pocket properties & Docking studies ........................................ 81

Experimental validation of the LEI-docking computational analysis of TTR ligands ........................................................................ 97

Bibliography ........................................................................ 104

Part IV: Metabolic stability studies .......... 105

Introduction .......... 107

In silico predictions: MetaSite and CYP's Consortium Model Software .......... 109

In vitro assays: predicted vs. experimental metabolites .......... 119

Tafamidis .......... 120

Meclofenamic .......... 121

Tolfenamic .......... 122

Lumiracoxib .......... 123

Vedaprofen .......... 123

Diflunisal & Iododiflunisal .......... 123

Further work .......... 126

Bibliography .......... 128

Conclusions .......... 133

Resumen en castellano .......... 135

Introducción .......... 137

Amiloidosis .......... 137

Transtirretina .......... 137

Objetivos de la presente tesis .......... 138

*Repurposing* basado en Fingerprints .......... 139

Modelización basada en farmacóforo .......... 140

Ligand Eficiency Indices (LEIs) .......... 140

Estudios *in silico* de Estabilidad Metabólica .......... 141

Estudio *in silico* .......... 142

Estudio *in vitro* .......... 143

Conclusiones .......... 144

Bibliografía .......... 146

**Annexes** .......................................................................................... 149

**List of publications** .......................................................................................... 151

**List of PDB crystallographic structures** ........................................................... 153

# INTRODUCTION

## Amyloid diseases

The folding process is a crucial step for the correct function of mature native proteins. For this reason, cells have developed complex mechanisms in order to protect folding and avoid any related malfunction[1,2]. More than 40 human proteins are known to misfold driving the formation of toxic aggregates, such as oligomers, protofibrils and amyloid fibrils[2,3,4]. Protein misfolding and aggregation is the cornerstone of a significant number of human diseases such as Alzheimer's disease, Huntington's disease, or type II diabetes. Many of these diseases have their molecular origin in misfolding mechanisms related to certain proteins. A relevant group in this typology is amyloidosis. Amyloid diseases are linked to alteration or damage of an organ function produced by accumulation of extracellular deposits, known as amyloids[5].

All amyloids deposits have very similar properties in terms of morphology and structural features:

1. They form fluorescent complexes with thioflavines S and T.
2. In presence of Congo red, using a polarized light microscopy, they show a green birefringence.
3. Most of the deposits are formed by fibrils with a diameter from 7 to 10 nm[6].
4. These fibrils are mainly constituted by polypeptidic chains arranged in a β-sheet extended shape and containing the β-strands perpendicular to the axis of the fibril.

Nowadays, about 20 human proteins have been identified as amyloid precursors in a wide range of pathologies. Transthyretin protein (TTR) is one of them. Independently of the protein involved in the amyloidogenic process, it is commonly accepted that the formation of amyloid deposits follows the same pathway in all cases, with a conformational change as the starting point. These deposits are formed by small protein-containing polymers with a low molecular weight, also called amyloid fibrils. This type of deposits is highly insoluble due to their β-sheet structural folding.

## Transthyretin

Transthyretin (TTR, also called prealbumin) is a tetrameric protein produced in the liver hepatocytes, plexus choroideus[7] and retina. TTR is involved in the extracellular transport of thyroid hormones and vitamin A, through a complex with serum retinol-binding protein. TTR functions as backup transporter for these kinds of hormones (such as thyroxine, also called T4) in plasma, and as a main transporter across the blood brain barrier.

Each TTR monomer contains 127 aminoacids, and has a β-sandwich conformation, consisting of eight β strands (designated A to H) arranged in two β sheets (labelled DAGH and CBEF), and a short α helix between the E and F strands[8] (**Figure 1**). Both β-sheets are separated by about 10 Å.

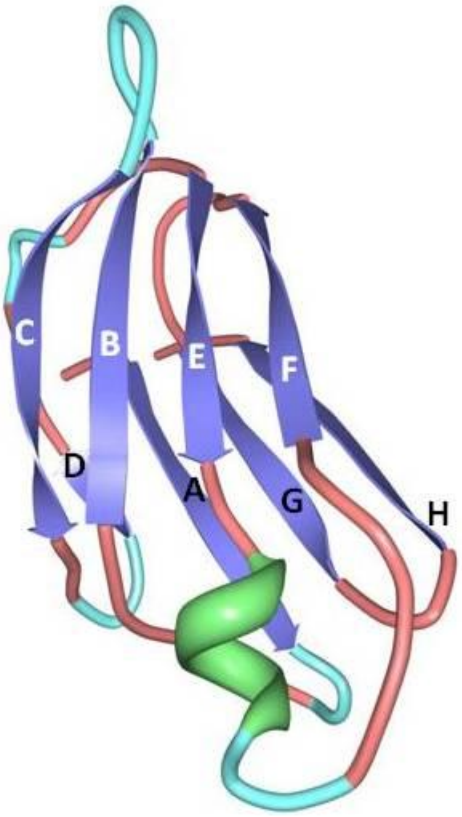


**Figure 1.** TTR monomer structure.

The TTR dimer (AB) is stabilized by an antiparallel hydrogen bonding network between the two adjacent F and H strands of both monomers. The functional TTR homotetramer is assembled by two TTR dimers (AB-A'B'). This TTR tetramer presents a central channel in the dimer-dimer interface where the two equivalent binding sites are located, which can accommodate simultaneosly two molecules of the natural ligand thyroxine (T4), as well as other small molecules (**Figure 2**).

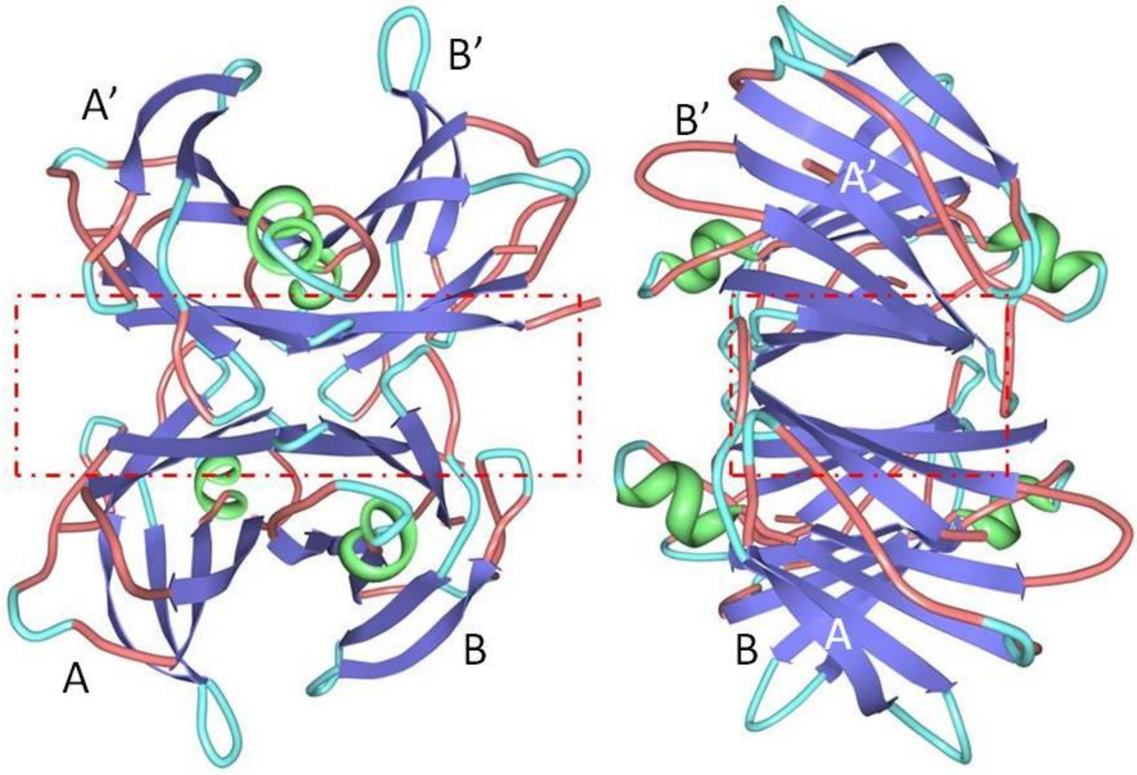


**Figure 2.** TTR tetramer structure representation, highlighting the hormone binding channel.

Presently, around 200 crystallographic structures of TTR protein have been solved and reported in the bibliography (the atomic coordinates of most of these structures are deposited in the Protein Data Bank, www.rcsb.org). This extremely comprehensive work has allowed a full characterization of the TTR binding sites, which has been key in the development of this thesis. The main features of the TTR binding site are:

a) The existence of three regions with different chemical features (**Figure 3**):

  - A charged region on the binding pocket entrance
  - A hydrophobic region, located in the central part
  - A polar region at the bottom of the binding pocket, near the center of the TTR homotetramer structure.

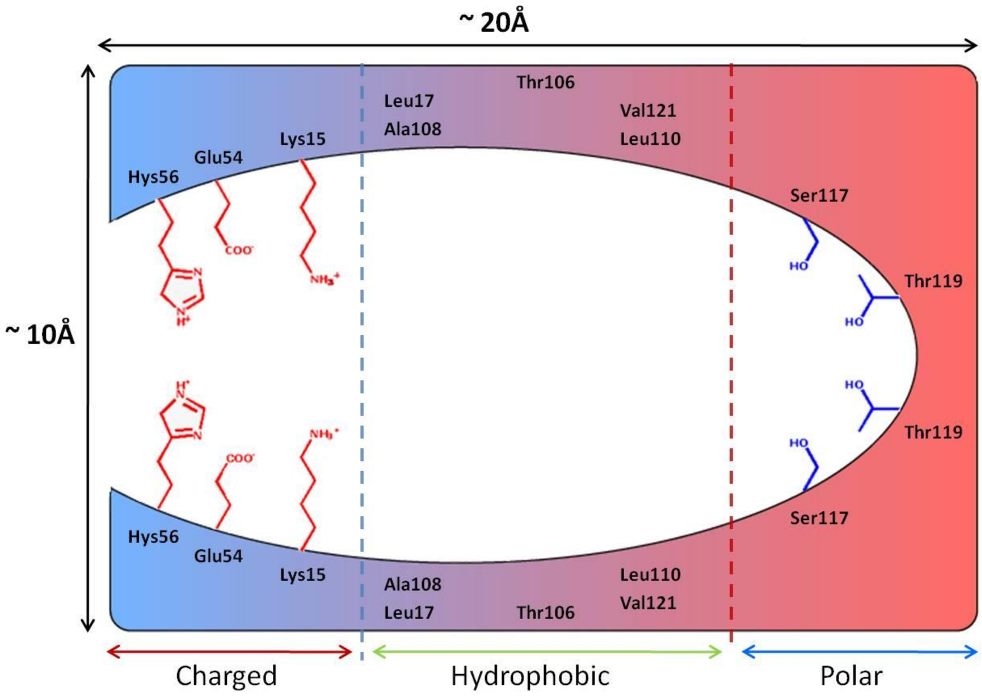

~ 20Å
~ 10Å
Hys56
Glu54
Lys15
COO⁻
NH₃⁺
H⁺
Leu17
Ala108
Thr106
Val121
Leu110
Ser117
Thr119
HO
Charged
Hydrophobic
Polar

**Figure 3.** Chemical nature of the ligand binding site of TTR. TTR binding site can be divided in 3 regions, depending on the chemical functionality of the aminoacid residues.

b) The presence in each binding site of six (three symetry-related pairs) small hydrophobic depressions – termed halogen binding pockets (HBP) – as revealed by the crystal structure of TTR in complex with tyroxine hormone (T4)[9,10] (**Figure 4**).

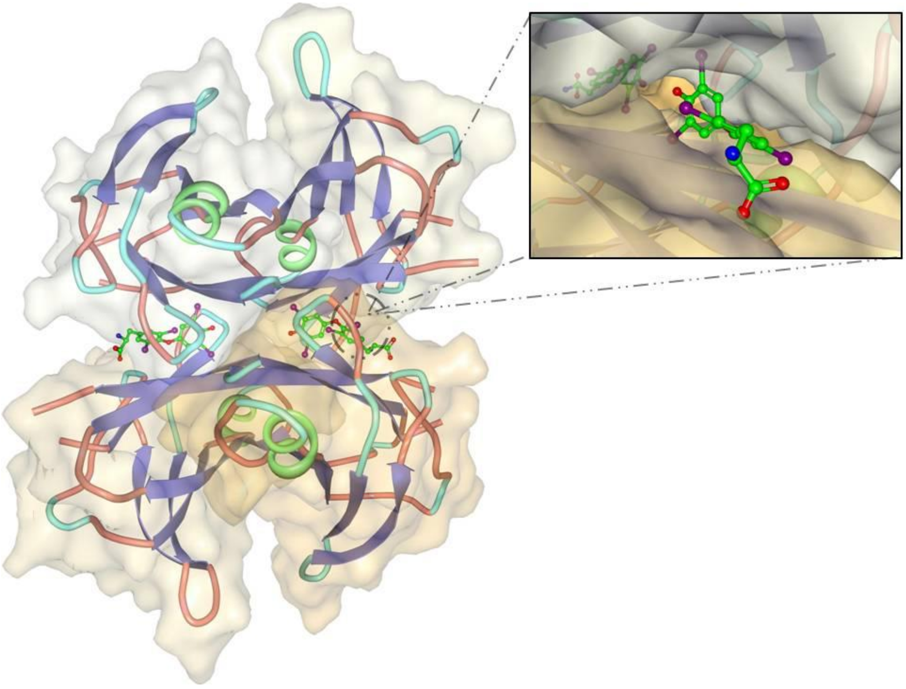


**Figure 4.** The crystallographic structure of the complex of TTR and thyroxine T4, reported on the Protein Data Bank (PDB ID: 1ICT)[10].

The innermost HBP pockets (HBP3 and HBP3'), located at 6 Å from the centre of the tetramer, are formed by the side chains and backbone atoms of Ala108, Ala109, Leu110, Ser117 and Thr119 of the monomers A and A' respectively. HBP2 and HBP2' are defined by residues Ala108, Ala109 and Leu110 of one of the two neighbouring β strands and Lys15 and Leu17 of the other one. Finally, the outermost pockets, termed HBP1 and HBP1' and located about 13 Å of the tetramer centre, are composed by the side chains of Thr106, Ala108, Met 13 and Lys15 (**Figure 5**).

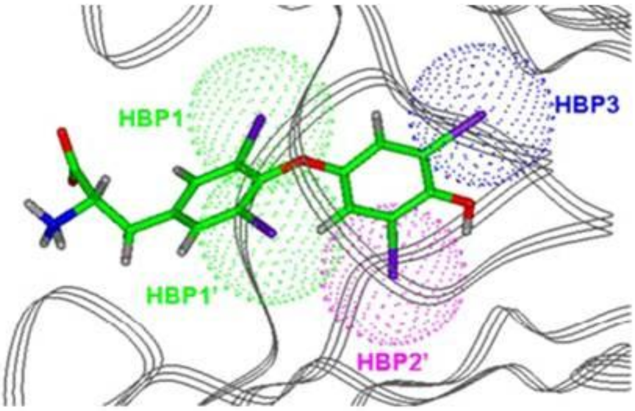


**Figure 5.** HBP pair pockets located on the TTR binding site, from the TTR-T4 crystal structure[10].

### TTR-related amyloid diseases

The malfunction of TTR is the cause of a series of rare amyloid diseases: familial amyloid polyneuropathy (**FAP**), familial amyloid cardiomyopathy (**FAC**), central nervous system selective amyloidosis (**CNSA**) and senile systemic amyloidoses (**SSA**).

For SSA, the deposit (mainly constituted by wild-type TTR protein) can be found on the brain, heart, pancreas and spleen. This disease affects around 25% of the population over 80 years[11].

Certain mutations cause CNS amyloidosis[12]. Patients with this involvement are rare to find and usually, they present deposits in the leptomeningeal vessels and in the subarachnoid membrane. In addition, it is common to observe the presence of non-CNS amyloids also in the kidney, heart and skin.

FAP is a hereditary amyloidosis caused by single aminoacid mutations across the peptidic sequence of the protein. The disease affects pain and temperature perception, and causes gastrointestinal, cardiac or even renal disfunctions.

In the case of FAC, fibrils are accumulated on the patient's heart. Both the wild-type TTR protein and the TTR mutant V122I can produce this disease. This single-point mutation is present in 4% of the Afro-American population[13,14].

More than 100 mutations have been reported for the TTR protein, and only 13 are non-amyloidogenic[15]. The mutations are distributed along all the protein sequence, but the most amyloidogenically aggressive mutations are related to residue 55[16,17].

TTR amyloid deposits are produced by a still poorly characterized mechanism that involves several steps such as: dissociation of the tetramer, monomer conformational changes, aggregation of modified monomers into non-fibrillar oligomers, that later form protofibrils and further elongate into mature fibrils[18] (**Figure 6**).

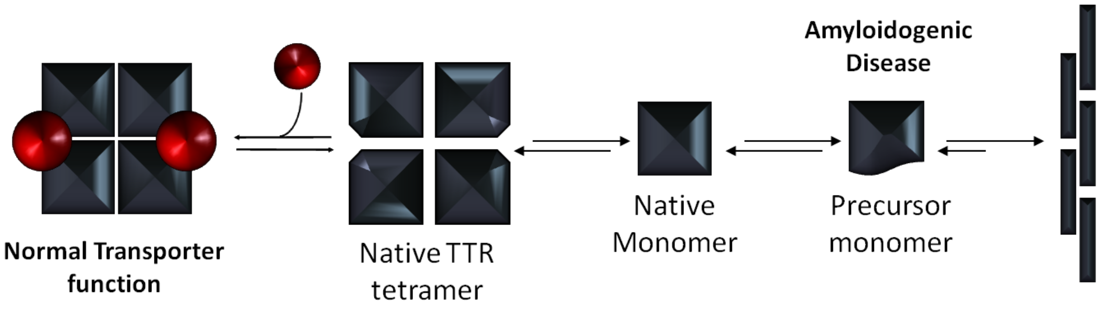


**Figure 6**. Proposed mechanism of amyloid fibrils formation.

The generally accepted hypothesis[19] supports that the first step of this mechanism is the tetramer dissociation into non-native monomers, with conformational changes that drive the system into a decrease of their thermodynamic stability. The modified monomers form soluble amyloid aggregates that are able to evolve into insoluble fibrils through autoassembling processes. *In vivo* conditions for the formation of amyloid fibrils are still unknown, but many studies have revealed that the crucial point of this pathway is the low stability of the TTR tetramer[20]. *In vitro*, the formation of fibrils may be triggered by decreasing the pH of the media. In addition, it is possible to correlate the pH value at which amyloidogenicity is induced and the pathogenicity of a given TTR mutant[21].

## Therapeutical strategies for TTR-related amyloidoses

Liver transplantation[22,23] is currently the only therapy for FAP and FAC amyloidosis, with limited clinical success, and not applicable to mutants that develop amyloidosis in the central nervous system, where TTR is also synthetized. In addition, liver transplantation has some disadvantages:

- The disease is only slowed down, because TTR mutants can be also synthesized on the plexus choroideus and retina.
- Symptomatology is not improved due to the already existing amyloid deposits.
- High difficulty of finding organ donors and the concomitant risk of surgery.

- Expensive costs
- Lifelong administration of immunosupressants
- Sometimes, development of cardiac and ocular dissorders

Other TTR-related amyloidosis therapies have been studied, such as:

1. Gene therapy to increase stability of the protein by specific mutations in order to delete the amyloidogenic TTR variants[24].
2. Amyloid deposits disolution[25].
3. Avoiding aggregation mechanism disrupting fibrils interactions[26].
4. TTR tetramer stabilization by binding of small molecules on the hormone site channel.

The design of new TTR ligands to stabilize the tetramer is the framework of this thesis. It is based on the hypothesis, suggested by Kelly et al[22], that the binding of small molecules to the T4 binding site, produces the kinetic stabilization of the tetramer, and will prevent the initial tetramer dissociation required for the fibril formation[21].

It has been demonstrated that many small molecules, with a wide range of chemical scaffolds, are able to mimic the hormone capability to bind TTR, thereby stabilizing the whole tetramer[27,28,29] (**Figure 7**).

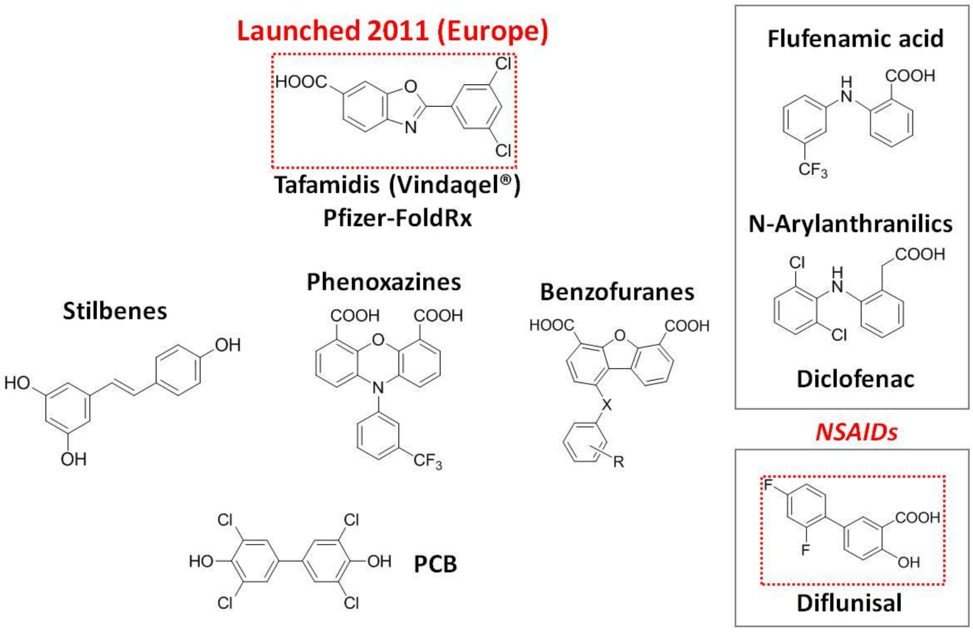

Launched 2011 (Europe)
HOOC
O
N
Cl
Cl
Tafamidis (Vindaqel®)
Pfizer-FoldRx
Flufenamic acid
H
N
COOH
CF3
N-Arylanthranilics
Cl
H
N
COOH
Cl
Diclofenac
NSAIDs
F
F
COOH
OH
Diflunisal
Stilbenes
HO
OH
OH
Phenoxazines
COOH
COOH
O
N
CF3
Benzofuranes
HOOC
O
COOH
X
R
Cl
Cl
HO
OH
Cl
Cl
PCB

**Figure 7**. Some of the TTR stabilizers reported in the literature.

One of most promising TTR amyloidogenic inhibitors, which has recently been launched on the market as an orphan drug for FAP disease, is Tafamidis (also called Fx-1006A or Vyndaqel®)[30]. This drug has been approved by the European Medicines Agency (EMA) for the FAP treatment as a meglumine salt. It has been possible to demonstrate that Tafamidis slows the disease progression in adult patients with an early stage of FAP. It was discovered by Jeffery W. Kelly et al.[31] through a structure-based drug design strategy, and it was developed at FoldRx Pharmaceuticals, a company that was acquired in 2010 by Pfizer.

This orphan drug interacts with the TTR tetramer by a kinetic stabilization pathway, providing pharmacological evidences that the process of amyloid fibril formation causes this disease.

Treatment with Tafamidis dramatically slows the process of amyloid fibril formation and the degeneration of post-mitotic tissue. Tafamidis is currently being considered for approval by the United States Food and Drugs Administration (FDA).

In this context, a multidisciplinary Consortium was created in Barcelona in 2000 with the aim of discovering new TTR ligands, with expertise in three different areas:

- Compounds synthesis: carried out by the team leaded by Dr. Gemma Arsequell and Dr. Gregorio Valencia, at the Institut de Química Avançada de Catalunya (IQAC-CSIC, Barcelona)
- Biochemistry and protein engineering: performed under the supervision of Dr. Antoni Planas from Institut Químic de Sarriá (Universitat Ramon Llull, Barcelona).
- Computational studies: lead by Dr. Nuria B. Centeno from Research Programme on Biomedical Informatics (Universitat Pompeu Fabra, Barcelona) and Dr. Jordi Quintana from Drug Discovery Platform (Parc Científic de Barcelona, Barcelona).

In addition, this Consortium has cooperated closely with the Neurobiology Unit of Molecular and Cellular Biology Institute of the University of Oporto, which carries out biomedical research related to amyloid diseases.

Several rational drug design workflows have been developed by our Consortium through the last years working on TTR fibril inhibition, and more than 800 compounds have been synthesized and tested against this protein. A database compiling all this chemico-biological information was built and it is used nowadays within our Consortium to define new and faster ways to find new and potent TTR fibrillogenesis inhibitors.

Due to the large amount of work previously performed, it has been possible to establish a chemico-biological data and knowledge base, from which all the computational work in this thesis project has been carried out.

The following chapters on this tesis include a specific introduction to the particular topics covered in each chapter. The thesis contains several published scientific articles, which are listed as bibliographic references in Annex 1 (List of Publications). In addition, Annex 2 (List of PDB crystallographic structures), includes the bibliographic references for several crystal structures of TTR-ligand complexes, and other proteins (Cytochromes CYP450), which have been used throughout the research described in the thesis chapters.

## Bibliography

(1) Radford, S. E.; Dobson, C. M. From Computer Simulations to Human Disease: Emerging Themes in Protein Folding. *Biopolymers* **1999**, *97*, 291–298.

(2) Dobson, C. M. The structural basis of protein folding and its links with human disease. *Philosophical Transactions of the Royal Society of London. Series B, Biological Xciences* **2001**, *356*, 133–145.

(3) Chiti, F.; Dobson, C. M. Protein misfolding, functional amyloid, and human disease. *Annual Review of Biochemistry* **2006**, *75*, 333–366.

(4) Forman, M. S.; Trojanowski, J. Q.; Lee, V. M. Neurodegenerative diseases: a decade of discoveries paves the way for therapeutic breakthroughs. *Nature Medicine* **2004**, *10*, 1055–1063.

(5) Pepys, M. B. Amyloidosis. *Annual review of medicine* **2006**, *57*, 223–241.

(6) Buxbaum, J. Diseases of protein conformation: what do in vitro experiments tell us about in vivo diseases? *Trends in Biochemical Sciences* **2003**, *28*, 585–592.

(7) Aleshire, S. L.; Bradley, C. a.; Richardson, L. D.; Parl, F. F. Localization of human prealbumin in choroid plexus epithelium. *Journal of Histochemistry & Cytochemistry* **1983**, *31*, 608–612.

(8) Pinto, M.; Blasi, D.; Nieto, J.; Arsequell, G.; Valencia, G.; Planas, A.; Quintana, J.; Centeno, N. B. Ligand-binding properties of human transthyretin. *Amyloid* **2011**, *18*, 51–54.

(9) Blake, C. C. F.; Geisow, M. J.; Oatley, S. J.; Rérat, B.; Rérat, C. Structure of prealbumin: Secondary, tertiary and quaternary interactions determined by Fourier refinement at 1.8 Å. *Journal of Molecular Biology* **1978**, *121*, 339–356.

(10) Wojtczak, A.; Cody, V.; Luft, J. R.; Pangborn, W. Structures of human transthyretin complexed with thyroxine at 2.0 A resolution and 3’,5'-dinitro-N-acetyl-L-thyronine at 2.2 A resolution. *Acta Crystallographica. Section D, Biological Crystallography* **1996**, *52*, 758–765.

(11) Cornwell, G. G.; Westermark, P. Senile amyloidosis: a protean manifestation of the aging process. *Journal of Clinical Pathology* **1980**, *33*, 1146–1152.

(12) Hammarström, P.; Sekijima, Y.; White, J. T.; Wiseman, R. L.; Lim, A.; Costello, C. E.; Altland, K.; Garzuly, F.; Budka, H.; Kelly, J. W. D18G transthyretin is monomeric, aggregation prone, and not detectable in plasma and cerebrospinal fluid: a prescription for central nervous system amyloidosis? *Biochemistry* **2003**, *42*, 6656–6663.

(13) Buxbaum, J. N.; Tagoe, C. E. The genetics of the amyloidoses. *Annual Review of Medicine* **2000**, *51*, 543–569.

(14) Sunda, M.; Blake, C. C. F. From the globular to the figrous state: protein structure and structural conversion in amyloid formation. *Quarterly Reviews of Biophysics* **1998**, *31*, 1–39.

(15) Connors, L. H.; Lim, A.; Prokaeva, T.; Roskens, V. a; Costello, C. E. Tabulation of human transthyretin (TTR) variants, 2003. *Amyloid* **2003**, *10*, 160–184.

(16) Jacobson, D. R.; Mcfarlin, D. E.; Buxbaum, J. N. Transthyretin Pro 55, a variant associated with early-onset, aggressive, diffuse amyloidosis with cardiac and neurologic involvement. *Human Genetics* **1992**, *89*, 353–356.

(17) Yamamoto, K.; Hsu, S. P.; Yoshida, K.; Ikeda, S.; Nakazato, M.; Shiomi, K.; Cheng, S. Y.; Furihata, K.; Ueno, I.; Yanagisawa, N. Familial amyloid polyneuropathy in Taiwan: identification of transthyretin variant (Leu55-->Pro). *Muscle & Nerve* **1994**, *17*, 637–641.

(18) Foss, T. R.; Wiseman, R. L.; Kelly, J. W. The pathway by which the tetrameric protein transthyretin dissociates. *Biochemistry* **2005**, *44*, 15525–15533.

(19) Quintas, a; Vaz, D. C.; Cardoso, I.; Saraiva, M. J.; Brito, R. M. Tetramer dissociation and monomer partial unfolding precedes protofibril formation in amyloidogenic transthyretin variants. *The Journal of Biological Chemistry* **2001**, *276*, 27207–27213.

(20) Brito, R. M. M.; Damas, A. M.; Saraiva, M. J. Amyloid Formation by Transthyretin: From Protein Stability to Protein Aggregation. *Current Medicinal Chemistry - Immunology, Endocrine & Metabolic Agents* **2003**, *3*, 349–360.

(21) Kelly, J. K.; Lansbury, P. T. A chemical approach to elucidate the mechanism of transthyretin and P-protein amyloid fibril formation. *Amyloid: the international journal of experimental and clinical investigation* **1994**, *1*, 186–205.

(22) Ando, Y. Liver transplantation and new therapeutic approaches for familial amyloidotic polyneuropathy (FAP). *Medical Molecular Morphology* **2005**, *38*, 142–154.

(23) Ando, Y.; Ueda, M. Diagnosis and therapeutic approaches to transthyretin amyloidosis. *Current Medicinal Chemistry* **2012**, *19*, 2312–2323.

(24) Nakamura, M.; Ando, Y. Applications of gene therapy for familial amyloidotic polyneuropathy. *Expert Opinion on Biological Therapy* **2004**, *4*, 1621–1627.

(25) Merlini, G.; Joa, M.; Sebastia, M. P. The molecular interaction of 4’-iodo-4'-deoxydoxorubicin with Leu-55Pro transthyretin “amyloid-like” oligomer leading to disaggregation. *Biochemical Journal* **2000**, *351*, 273–279.

(26) O’Nuallain, B.; Wetzel, R. Conformational Abs recognizing a generic amyloid fibril epitope. *Proceedings of the National Academy of Sciences of the United States of America* **2002**, *99*, 1485–1490.

(27) Munro, S. L.; Lim, C.-F.; Hall, J. G.; Barow, J. W.; Craik, D. J.; Topliss, D. J.; Stockigt, J. R. Drug Competition for Thyroxine Binding to Transthyretin (Prealbumin): Comparison with Effects on Thyroxine-Binding Globulin. *Journal of Clinical Endocrinology & Metabolism* **1989**, *68*, 1141–1147.

(28) Peterson, S. a; Klabunde, T.; Lashuel, H. a; Purkey, H.; Sacchettini, J. C.; Kelly, J. W. Inhibiting transthyretin conformational changes that lead to amyloid fibril formation. *Proceedings of the National Academy of Sciences of the United States of America* **1998**, *95*, 12956–12960.

(29) Almeida, M. R.; Gales, L.; Damas, A. M.; Cardoso, I.; Saraiva, M. J. Small transthyretin (TTR) ligands as possible therapeutic agents in TTR amyloidoses. *Current Drug Targets. CNS and Neurological Disorders* **2005**, *4*, 587–596.

(30) Razavi, H.; Palaninathan, S. K.; Powers, E. T.; Wiseman, R. L.; Purkey, H. E.; Mohamedmohaideen, N. N.; Deechongkit, S.; Chiang, K. P.; Dendle, M. T. a; Sacchettini, J. C.; Kelly, J. W. Benzoxazoles as transthyretin amyloid fibril inhibitors: synthesis, evaluation, and mechanism of action. *Angewandte Chemie (International ed. in English)* **2003**, *42*, 2758–2761.

(31) Connelly, S.; Choi, S.; Johnson, S. M.; Kelly, J. W.; Wilson, I. a Structure-based design of kinetic stabilizers that ameliorate the transthyretin amyloidoses. *Current Opinion in Structural Biology* **2010**, *20*, 54–62.

# OBJECTIVES

The general goal of this thesis is the contribution to the design and discovery of new compounds able to interact with high affinity with the hormone binding site of the homotetrameric protein transthyretin (TTR), and stabilize this tetramer, becoming drug candidates to treat several rare amyloid diseases associated with TTR.

This general goal has been approached through the development of several computational workflows and chemico-biological databases, and in collaboration with two experimental research laboratories of our TTR Consortium (one contributing with the chemical synthesis or acquisition of the designed compounds, and the other contributing with the biological activity assay results for the synthesized or acquired compounds).

The specific objectives of this thesis are:

a) The generation of a chemico-biological database containing the historical and newly generated results of the TTR Consortium, containing the chemical structures and biological activities of the TTR ligands.
b) Explore the possibility of using repurposing techniques applied to the discovery of new TTR inhibitors among the existing drugs, with particular focus on anti-inflammatory drugs, which are known to be good TTR ligands.
c) Design of new flavonoid compounds as TTR ligands by means of structure-based drug design.
d) Incorporate the Ligand Efficiency Indices analysis (both retrospective and prospective) as a new tool for designing new compounds with increased efficiency as TTR ligands.
e) The computational development of a combined predictive/experimental workflow for the analysis of the metabolic stability of TTR ligands, as a tool for improving the prioritized compounds in our in-house database to obtain new compounds with better metabolic and pharmacokinetic properties.

# PART I

# REPURPOSING APPROACH

## The drug discovery process

The main focus of a drug discovery and development process is to help patients with a given disease to improve their quality of life, through the administration of the developed compounds that need to be approved by regulatory agencies.

This process is designed in order to ensure that a medicine will be effective, safe and available for the patients in the shortest possible time.

The first step in drug development is the discovery of the best targets for treating or preventing a given disease. Targets are usually proteins which are associated with the disease, or proteins in a microorganism causing the disease. The challenge in this step is the identification of the relevant protein or proteins, and the confirmation of their role in the key pathways of the disease. It is also important to take into account that, sometimes, a possible target can be involved in secondary pathways which contain other proteins that can be more determinant for the treatment of the disease.

Several methods like high-throughput-screening (HTS) or computer-aided ligand or drug design are used to find new molecular entities or biologics that bind to the identified targets linked to the disease. The aim of this step in the drug discovery process is to alter in an expected manner the protein function, obtaining active molecules, also known as “hits”, that interact with the protein targets. These hits will eventually be optimized, considering effectiveness and safety, becoming drug candidates, able to go into advanced clinical phases, and finally reaching the market to treat the disease.

Discovering and bringing a new drug entity to the market typically takes around 12 - 15 years and costs around 1 billion dollars. The clinical and other challenges to reach the market are enormous and usually, only 1 compound per 10000 studied reaches the final step on this path[1] (**Figure 1**).

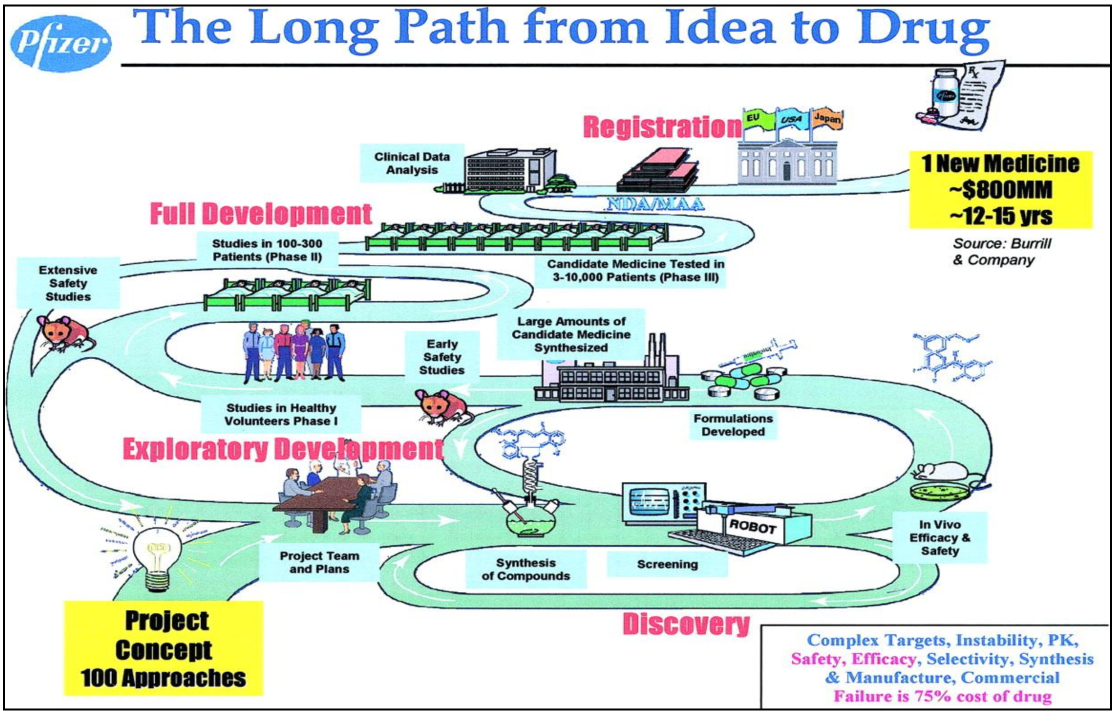

Pfizer
The Long Path from Idea to Drug
Registration
EU
USA
Japan
Clinical Data Analysis
NDA/MAA
1 New Medicine
~$800MM
~12-15 yrs
Source: Burrill & Company
Full Development
Studies in 100-300 Patients (Phase II)
Candidate Medicine Tested in 3-10,000 Patients (Phase III)
Extensive Safety Studies
Large Amounts of Candidate Medicine Synthesized
Early Safety Studies
Studies in Healthy Volunteers Phase I
Formulations Developed
Exploratory Development
ROBOT
In Vivo Efficacy & Safety
Project Team and Plans
Synthesis of Compounds
Screening
Discovery
Project Concept
100 Approaches
Complex Targets, Instability, PK, Safety, Efficacy, Selectivity, Synthesis & Manufacture, Commercial
Failure is 75% cost of drug

**Figure 1.** The “market way” is hard, costly and slow[1].

It is widely accepted that this process is not efficient. This issue occurs mainly in clinical trials due to liabilities related to poor pharmacokinetics, poor efficacy and high toxicity and it means an enormous cost against a high failure rate[2].

Despite the enormous increase in Research and Development (R+D) spending, the number of new drugs launched to the market every year is still quite low (**Figure 2**)

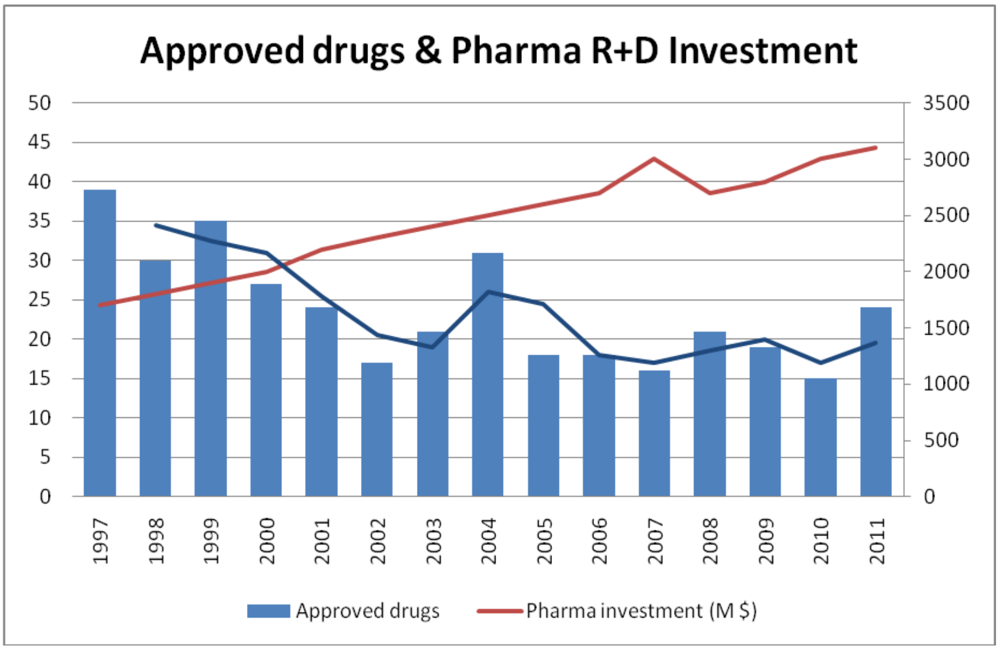


**Figure 2.** FDA approved drugs since the last 14 years compared with the investment done by the pharma industries[3].

## The repurposing approach

Since many years, companies have increased significantly the investment in the drug discovery pipeline optimization, using more advanced techniques such as structure-based drug design, combinatorial chemistry, high-throughput screening (HTS) and a huge range of "omics" technologies. However, the impact of these improvements has not been significant on increasing the efficiency of the drug discovery process.

For this reason, in recent years, new approaches have been consolidated as possible alternatives to reach an efficient and safety treatment for a given disease more rapidly. One of these approaches is the drug repositioning or drug repurposing[4].

This approach is based on finding new medical indications for existing drugs that have reached the market (or drugs that have been discontinued in clinical trials for other reasons than safety concerns). For a launched drug, the safety profile and the whole pharmacokinetic properties are known; as a consequence of this, the clinical development for the new indication can be performed with less investment, potentially faster than the *de novo* development, and with a lower risk in terms of toxicity and safety.

Some successful cases of repurposed drugs have been reported such as sildenafil as therapeutical agent against erectile dysfunction and pulmonary arterial hypertesion[5] or thalidomide for leprosy[6] and multiple myeloma[7].

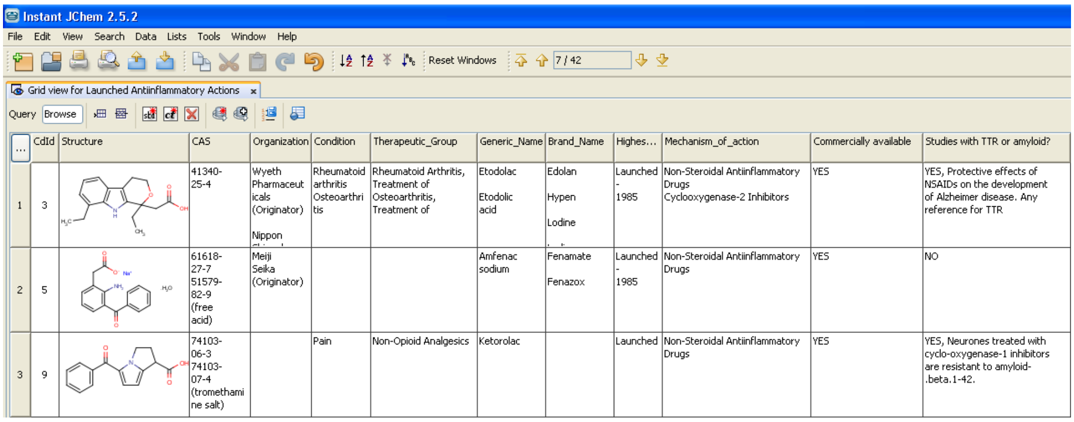
For these reasons, the repurposing approach seems to be quite reasonable and feasible, to explore the possibility to find a good drug candidate for a given disease, especially when an orphan or rare disease is the target of the drug discovery process. Because of the reduced potential market for these orphan or rare diseases, the repurposing approach may contribute significantly to finding efficient treatments for these diseases.

## Building up a repurposing tool

Taking into account the potential advantage that can be obtained from a repurposing strategy (in terms of time and costs), an *in silico* approach has been established and performed to find new therapeutical applications of known drugs for TTR protein stabilization.

First of all, a database containing drugs that have reached the market has been built. For this, 3248 launched drug were extracted from Integrity Prous database (http://integrity.thomson-pharma.com/integrity/) and compiled in a local database where the compounds chemical structures were refined (adding hydrogens, standardizing the structures, etc), using the software MOE, from Chemical Computing

Group[8], and other useful information such as commercial availability or reported amyloidogenesis inhibition activity was also included. (**Figure 3**)


Instant JChem 2.5.2
File Edit View Search Data Lists Tools Window Help
Reset Windows
7 / 42
Grid view for Launched Antiinflammatory Actions
Query Browse

| | CdId | Structure | CAS | Organization | Condition | Therapeutic_Group | Generic_Name | Brand_Name | Highes... | Mechanism_of_action | Commercially available | Studies with TTR or amyloid? |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 3 | | 41340-25-4 | Wyeth Pharmaceuticals (Originator)<br>Nippon | Rheumatoid arthritis Osteoarthritis | Rheumatoid Arthritis, Treatment of Osteoarthritis, Treatment of | Etodolac<br>Etodolic acid | Edolan<br>Hypen<br>Lodine | Launched - 1985 | Non-Steroidal Antiinflammatory Drugs<br>Cyclooxygenase-2 Inhibitors | YES | YES, Protective effects of NSAIDs on the development of Alzheimer disease. Any reference for TTR |
| 2 | 5 | | 61618-27-7<br>51579-82-9 (free acid) | Meiji Seika (Originator) | | | Amfenac sodium | Fenamate<br>Fenazox | Launched - 1985 | Non-Steroidal Antiinflammatory Drugs | YES | NO |
| 3 | 9 | | 74103-06-3<br>74103-07-4 (tromethamine salt) | | Pain | Non-Opioid Analgesics | Ketorolac | | Launched | Non-Steroidal Antiinflammatory Drugs | YES | YES, Neurones treated with cyclo-oxygenase-1 inhibitors are resistant to amyloid-.beta.1-42. |

**Figure 3**. An overview of the database constructed from Integrity Prous.

The whole database was analyzed and several filters for the selection of compounds were applied sequentially with the workflow shown on **Figure 4**:

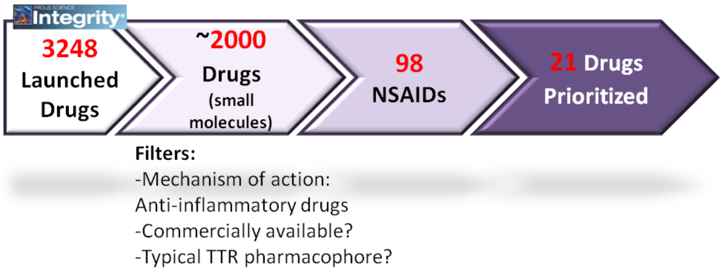

Integrity
3248
Launched
Drugs
~2000
Drugs
(small
molecules)
98
NSAIDs
21 Drugs
Prioritized
Filters:
-Mechanism of action:
Anti-inflammatory drugs
-Commercially available?
-Typical TTR pharmacophore?

**Figure 4**. Designed workflow for the TTR-focused repurposing strategy.

1) The dataset was reduced to 2000 molecules selecting compounds with molecular weight below 700 Daltons.
2) Only anti-inflammatory drugs were selected, since it is known that many non-steroidal anti-inflammatory drugs (also called NSAIDs) show high stabilization and binding activity against TTR[9,10].
3) Commercial availability was set as a mandatory condition for a given compound to be included in the training dataset.
4) Only compounds with a certain substructure shape, common in the better TTR inhibitors known to date[11], were selected (**Figure 5**).

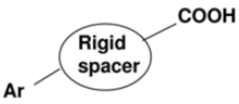

COOH
Rigid
spacer
Ar

**Figure 5**. Well-known TTR stabilizer scaffold[11].

After this first filtering phase, 98 launched drugs, commercially available and with a privileged substructure, were ready to be studied using fingerprint modeling.

### Fingerprint modeling: Similarity search studies

Similarity searching using molecular fingerprints is a common virtual screening technique used in data mining of compound libraries[12,13,14]. It is a simple method quite efficient because it is computationally inexpensive, and it needs a reference structure (usually an active compound) and a similarity metric system able to determine how many descriptors are shared by the reference and the compound under study. This approach has been used for many years to find new drug-like compounds based on known active molecules, exploring a given region of the chemical space metrically via the *Tanimoto coefficient* $S(i, j)$[15], which, for two compounds $i$ and $j$, with fingerprints of length $a$ and $b$, respectively, is given by the mathematical expression:

$$S(i, j) = c / [a + b - c],$$

where $a$ is the number of bits in molecule $i$, $b$ is the number of bits in molecule $j$, and $c$ is the number of bits in common between $a$ and $b$ (**Figure 6**).

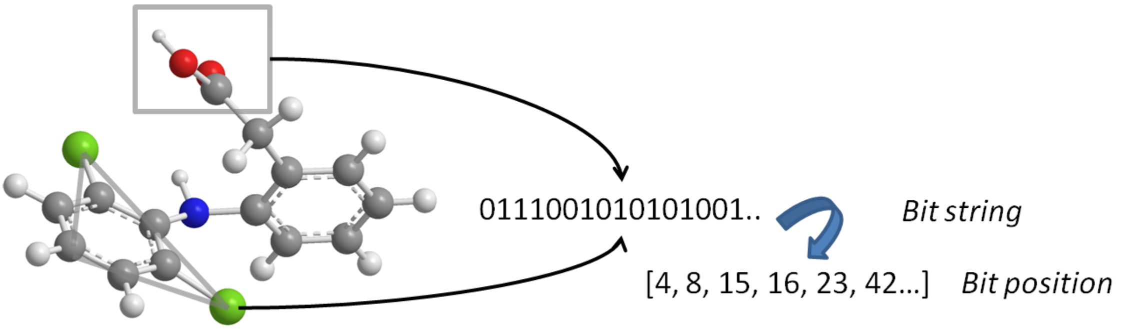


**Figure 6**. Molecular fingerprints example; each bit indicates the presence/absence (1/0) of a structural motif. The fingerprint may be a *bit string* (indicating the on/off state of all the bits) or a *bit position vector* (indicating the positions of the bits which are turned on).

The similarity between the reference compound used and each compound is computed and the whole library is sorted based on decreasing similarity rank. This similarity relies on the similarity property principle[16] which determines that similar molecules exhibit similar biological behavior.

This fingerprint modeling technique has some advantages:

1) It is computationally inexpensive.
2) It allows the user to mine datasets composed of thousands of molecules quite fast.
3) The experimental crystal structure of the system under study is not required.

Traditionally, similarity searching was performed using a single active compound as reference, but it has been demonstrated that merging multiple similarity scores through the setting of more reference compounds improves the results and performance of this methodology.

For this reason, for the TTR-focused repurposing approach, using fingerprint modeling, 3 references were used: Mefenamic acid[17], Tafamidis[18] and Iododiflunisal[19] (**Figure 7**). All of them have been reported as high potent TTR stabilizers and amyloidogenic inhibitors. In fact, in 2011 and 2012, Tafamidis has been recently accepted to be marketed in Europe and pre-registered in USA, under the orphan drug designation for the Familial Amyloid Polyneuropathy (FAP) treatment. This will be the first-in-class disease-modifying compound that has been designed to inhibit the formation of amyloid deposits by preventing the TTR protein misfolding and deposition.

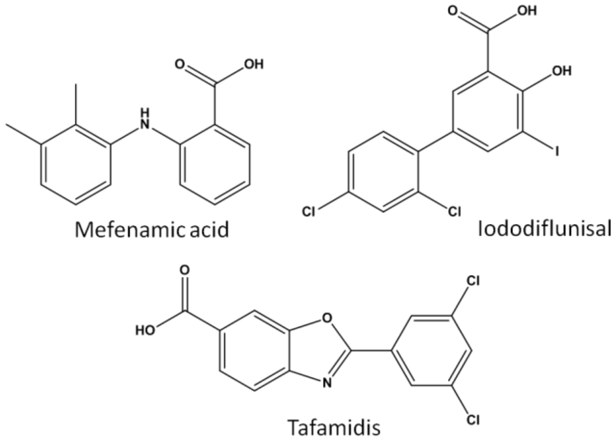


**Figure 7.** Known TTR inhibitors set as references in the fingerprint modeling similarity search performed.

Apart from using 3 different reference molecules to expand the chemical space exploration in the dataset selected, 3 different types of fingerprints were used in the similarity searches, all of them available at the Molecular Operating Environment (MOE)[8]:

1) **MACCS keys**: each bit of this kind of keys (or each feature) indicates the presence or absence of one of the 166 public MDL structural keys (a set of pre-selected functional groups)
2) **Typed Atom Fingerprint (TAF)**: each atom of the molecule is assigned to an atom type from the set (donor, acceptor, polar, anion, cation and hydrophobe). Under this category, two TAF were chosen: TAD (2-point pharmacophore based fingerprint where all pairs are encoded as features, using the inter-atomic distance of each pair) and TAT (3 point pharmacophore based fingerprint where all atom type triplets are coded as features using the three inter-atomic distances).

3) **GpiDAPH3 (Pharmacophore-based fingerprints):** this fingerprint is 3-point pharmacophore based. All triplets of atoms are coded as features using the three graph distances and associated atoms types of each triangle.

Using this combination of fingerprint types, the dataset of 98 launched drugs was compared with the three selected reference compounds setting the Tanimoto coefficient of similarity to 75% (**Figure 8**).

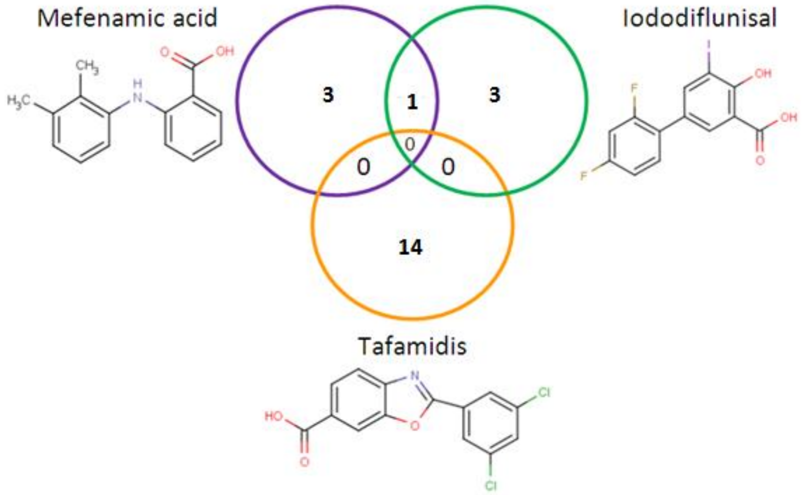
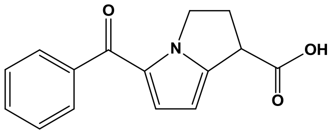
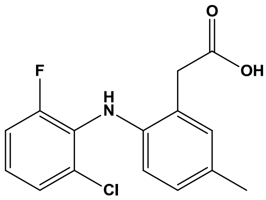
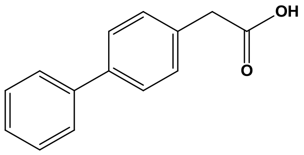


**Figure 8.** Results from the similarity search performed show that some compounds are structurally similar to the reference compounds.

21 compounds were prioritized based on their similarity against the 3 reference molecules. From this dataset prioritized, 9 drugs (**Table 1**) were acquired and tested in *in vitro* assays in order to confirm the predictions as possible new amyloid inhibitors for TTR.

| Compound number | Structure | Name | Therapeutic Group | Reference matching |
|---|---|---|---|---|
| **1** |  | Ketorolac | Non-Opioid Analgesic | Tafamidis |
| **2** |  | Lumiracoxib | Non-Opioid Analgesic | Mefenamic acid |
| **3** |  | Felbinac | Antiarthritic Drug | Iododiflunisal |

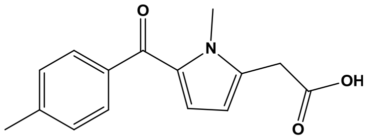
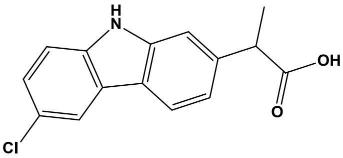
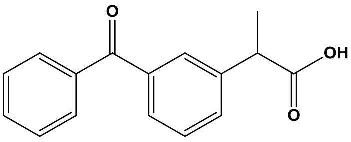
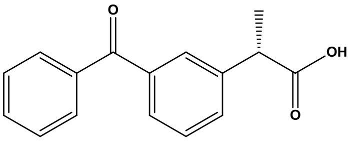
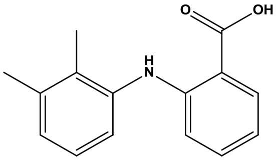
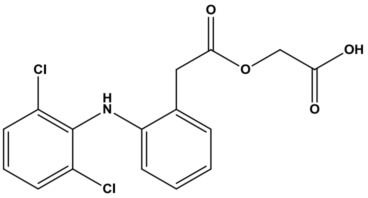
| Compound number | Structure | Name | Therapeutic Group | Reference matching |
|---|---|---|---|---|
| **4** |  | Tolmetin | Treatment of Osteoarthritis | Mefenamic acid |
| **5** |  | Carprofen | Antiarthritic Drug | Mefenamic acid |
| **6** |  | Ketoprofen | Antiarthritic Drug | Mefenamic acid |
| **7** |  | Dexketoprofen | Non-Opioid Analgesic | Mefenamic acid |
| **8** |  | Tolfenamic acid | Antiarthritic Drug | Mefenamic acid |
| **9** |  | Aceclofenac | Analgesic Drug | Mefenamic acid |

**Table 1.** Compounds acquired to validate the similarity results obtained.

## In vitro assays: kinetic turbidimetric assay

In order to evaluate the amyloid inhibition activity of the selected candidates against TTR, an *in vitro* kinetic turbidimetric assay was perfomed[20].

In this kind of experiment, two different activity values are obtained: the fibrillogenesis inhibition capability is determined using $IC_{50}$ (concentration of inhibitor in which the initial rate of fibril formation is reduced to 50%) and RA (% reduction of fibril formation). The experimental protocol is summarized on **Figure 9**.

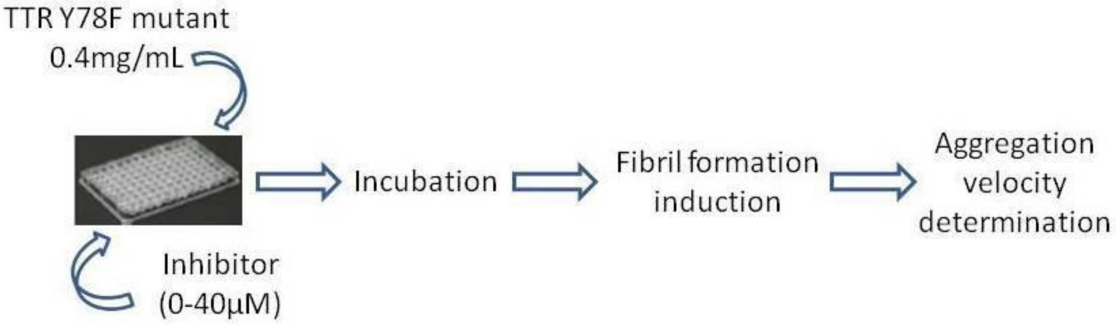

TTR Y78F mutant
0.4mg/mL
Inhibitor
(0-40µM)
Incubation
Fibril formation induction
Aggregation velocity determination

**Figure 9.** Schematic representation of the workflow used for the *in vitro* testing of the selected compounds[20].

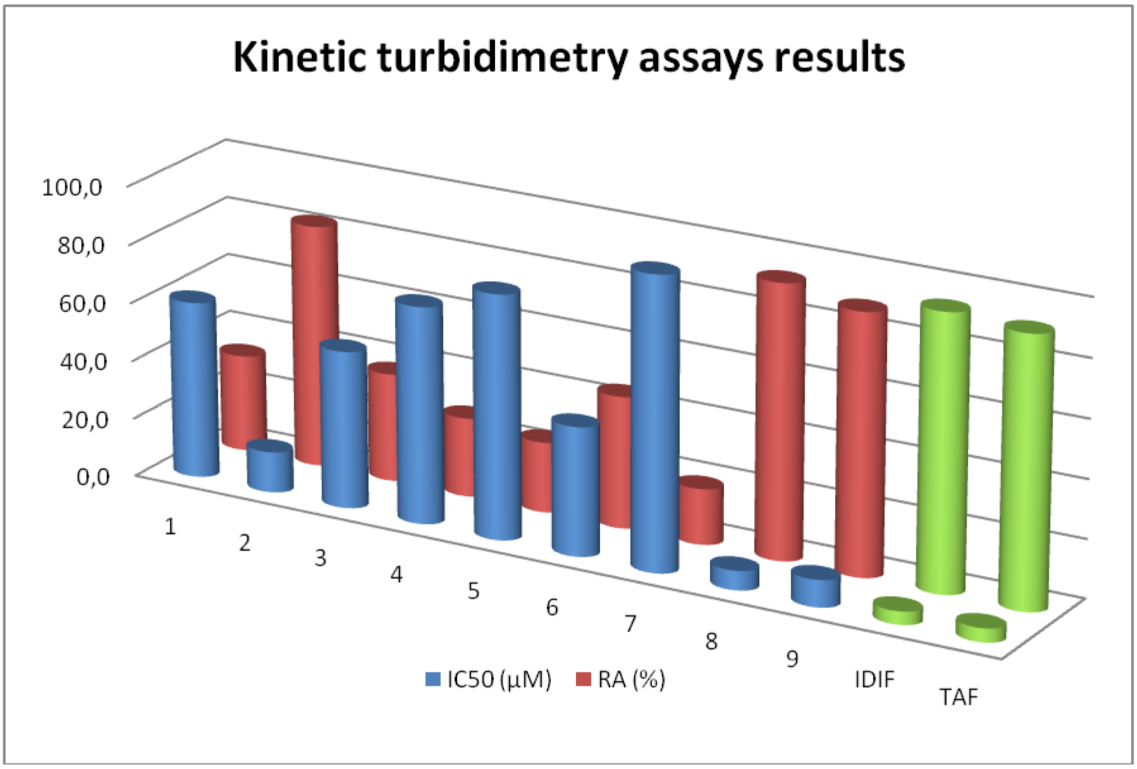


**Figure 10.** In vitro results for the 9 tested compounds against TTR Y78F mutant. Green bars refer to the Iododiflunisal and Tafamidis activities that have been used as references. Tested compounds are labeled as in **Table 1**.

From the results (**Figure 10**), it is possible to identify 3 compounds (**2**: Lumiracoxib, **8**: Tolfenamic acid and **9**: Aceclofenac) that are able to inhibit the TTR amyloid formation and show a good activity profile in terms of $IC_{50}$ and RA, compared to Iododiflunisal and Tafamidis (both well-known potent TTR stabilizers[18,19]). These results validate the computational protocol established and demonstrate that anti-inflammatory drugs are a good starting point for a repurposing approach centered on the TTR protein.

For years, development of therapeutical agents for orphan diseases has been largely overlooked by pharmaceutical industries due to small market thereof. Repurposing techniques, combined with *in silico* approaches, such as fingerprint modeling applied to similarity searching, are currently emerging to identify secondary or “off-target” drug actions for new relevant indications.

The results reported in this chapter (paper under preparation) have enabled us to identify a set of promiscuous drugs able to bind and stabilize the TTR tetramer *in vitro*, thereby becoming candidates that could be advanced to further preclinical and clinical studies related to TTR-mediated amyloidoses.

## Bibliography

(1) Kennedy, S. P.; Bormann, B. J.; Ormann, B. J. B. Effective Partnering of Academic and Physician Scientists with the Pharmaceutical Drug Development Industry. *Experimental Biology and Medicine* **2006**, *23*, 1690–1694.

(2) Pammolli, F.; Magazzini, L.; Riccaboni, M. The productivity crisis in pharmaceutical R&D. *Nature Reviews. Drug Discovery* **2011**, *10*, 428–438.

(3) Hudes, M. Fresh from the biotech pipeline— 2011. *Nature Biotechnology* **2012**, *30*, 128–131.

(4) Ekins, S.; Williams, A. J.; Krasowski, M. D.; Freundlich, J. S. In silico repositioning of approved drugs for rare and neglected diseases. *Drug Discovery today* **2011**, *16*, 298–310.

(5) Kuschner, W. G. Sildenafil citrate therapy for pulmonary arterial hypertension. *The New England Journal of Medicine* **2006**, *354*, 1091–1093.

(6) Teo, S. K.; Resztak, K. E.; Scheffler, M. a; Kook, K. a; Zeldis, J. B.; Stirling, D. I.; Thomas, S. D. Thalidomide in the treatment of leprosy. *Microbes and Infection / Institut Pasteur* **2002**, *4*, 1193–202.

(7) Program, L.; Cancer, S. C.; Rock, L. Correction: Antitumor Activity of Thalidomide in Refractory Multiple Myeloma. *The New England Journal of Medicine* **2000**, *342*, 364.

(8) Group, C. C. Molecular Operating Environment (MOE), Chemical Computing Group Inc.: 1010 Sherbooke St. West, Suite #910, Montreal, QC, Canada, H3A 2R7, 2011.

(9) Adamski-Werner, S. L.; Palaninathan, S. K.; Sacchettini, J. C.; Kelly, J. W. Diflunisal analogues stabilize the native state of transthyretin. Potent inhibition of amyloidogenesis. *Journal of Medicinal Chemistry* **2004**, *47*, 355–374.

(10) Munro, S. L.; Lim, C.-F.; Hall, J. G.; Barow, J. W.; Craik, D. J.; Topliss, D. J.; Stockigt, J. R. Drug Competition for Thyroxine Binding to Transthyretin (Prealbumin): Comparison with Effects on Thyroxine-Binding Globulin. *Journal of Clinical Endocrinology & Metabolism* **1989**, *68*, 1141–1147.

(11) Johnson, S. M.; Connelly, S.; Fearns, C.; Powers, E. T.; Kelly, J. W. The transthyretin amyloidoses: from delineating the molecular mechanism of aggregation linked to pathology to a regulatory-agency-approved drug. *Journal of Molecular Biology* **2012**, *421*, 185–203.

(12) Sheridan, R. P.; Kearsley, S. K. Why do we need so many chemical similarity search methods? *Drug Discovery Today* **2002**, *7*, 903–911.

(13) Miller, M. a Chemical database techniques in drug discovery. *Nature Reviews. Drug Discovery* **2002**, *1*, 220–227.

(14) Oprea, T. I.; Informatics, E. S. T. L.; Astrazeneca, R.; Mölndal, D.; Mölndal, S.- Virtual Screening in Lead Discovery: A Viewpoint. *Molecules*. **2002**, 51–62.

(15) Willett, P.; Barnard, J. M.; Downs, G. M.; Information, B. C.; Road, U.; Sheffield, S. Chemical Similarity Searching. **1998**, 983–996.

(16) Klopmand, G. Concepts and applications of molecular similarity, by Mark A. Johnson and Gerald M. Maggiora, eds., John Wiley & Sons, New York, **1990**, 393 pp.

(17) Klabunde, T.; Petrassi, H. M.; Oza, V. B.; Raman, P.; Kelly, J. W.; Sacchettini, J. C. Rational design of potent human transthyretin amyloid disease inhibitors. *Nature Structural Biology* **2000**, *7*, 312–321.

(18) Bulawa, C. E.; Connelly, S.; Devit, M.; Wang, L.; Weigel, C.; Fleming, J. a; Packman, J.; Powers, E. T.; Wiseman, R. L.; Foss, T. R.; Wilson, I. a; Kelly, J. W.; Labaudinière, R. Tafamidis, a potent and selective transthyretin kinetic stabilizer that inhibits the amyloid cascade. *Proceedings of the National Academy of Sciences of the United States of America* **2012**, *109*, 9629–9634.

(19) Mairal, T.; Nieto, J.; Pinto, M.; Almeida, M. R.; Gales, L.; Ballesteros, A.; Barluenga, J.; Pérez, J. J.; Vázquez, J. T.; Centeno, N. B.; Saraiva, M. J.; Damas, A. M.; Planas, A.; Arsequell, G.; Valencia, G. Iodine atoms: a new molecular feature for the design of potent transthyretin fibrillogenesis inhibitors. *PloS One* **2009**, *4*, e4124.

(20) Dolado, I.; Nieto, J.; Saraiva, M. J. M.; Arsequell, G.; Valencia, G.; Planas, A. Kinetic assay for high-throughput screening of in vitro transthyretin amyloid fibrillogenesis inhibitors. *Journal of Combinatorial Chemistry* **2005**, *7*, 246–252.

# PART II

# PHARMACOPHORE MODELING

## Pharmacophore modeling

A pharmacophore is the ensemble of steric and electronic features that is necessary to ensure the optimal intermolecular interactions with a specific biological target structure, and to block (or trigger) its biological response. A pharmacophore can be also defined as the largest common denominator shared by a set of active molecules that defines the interaction with a set of complementary sites in the target due to its intrinsic molecular recognition mechanism.

The pharmacophore modeling has evolved successfully as a drug design methodology in the last decades[1]. The introduction of the tridimensionality in the nature of these interactions with the target and the spatial relationships between the pharmacophoric features has been a key factor for the impact of contemporary pharmacophore research in the drug discovery process.

A pharmacophore model can be developed following two main strategies:

1. **Structure-based**: extracting the relevant chemical features of a ligand or set of ligands from a known X-ray of NMR structure for a given ligand-receptor complex.
2. **Ligand-based**: extracting a maximum common set of relevant features from a set of active compounds against a given receptor.

Taking into account the large amount of information available for the TTR protein concerning both crystallographic structures of ligand-TTR complexes, and information about active molecules able to stabilize the TTR tetramer, avoiding amyloid formation, a TTR-based pharmacophore model was developed in this thesis.

## Flavones as TTR stabilizers

It is well-known that several plant derived flavones and xanthones[2,3,4] have shown a high affinity towards TTR, and a good profile of amyloid formation inhibition *in vitro*. This class of flavone compounds, like other small organic molecules, binds preferentially in the hormone-binding site of the TTR protein.

T4 displacement studies carried out by Kohrle et al[5] revealed that a synthetic plant flavonoid EMD21388 (3-methyl-4’,6-dihydroxy-3’,5’-dibromoflavone) (**Figure 1**) was a strong thyroxine (T4) binding competitor, and it was an antagonist able to alter the amount of total hormone circulating, and the percentage of free thyroid hormones and of serum thyrotropin concentrations.

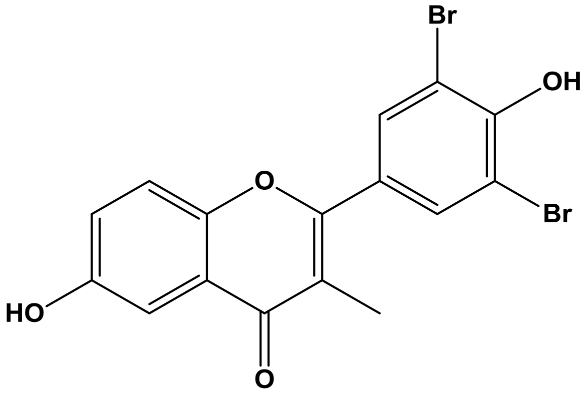


**Figure 1**. EMD21388 (3-methyl-4’,6-dihydroxy-3’,5’-dibromoflavone) structure.

The structure of a rat TTR protein, co-crystallized with the flavone EMD21388, was reported by Cody et al in 2002[6] (**Figure 2**).

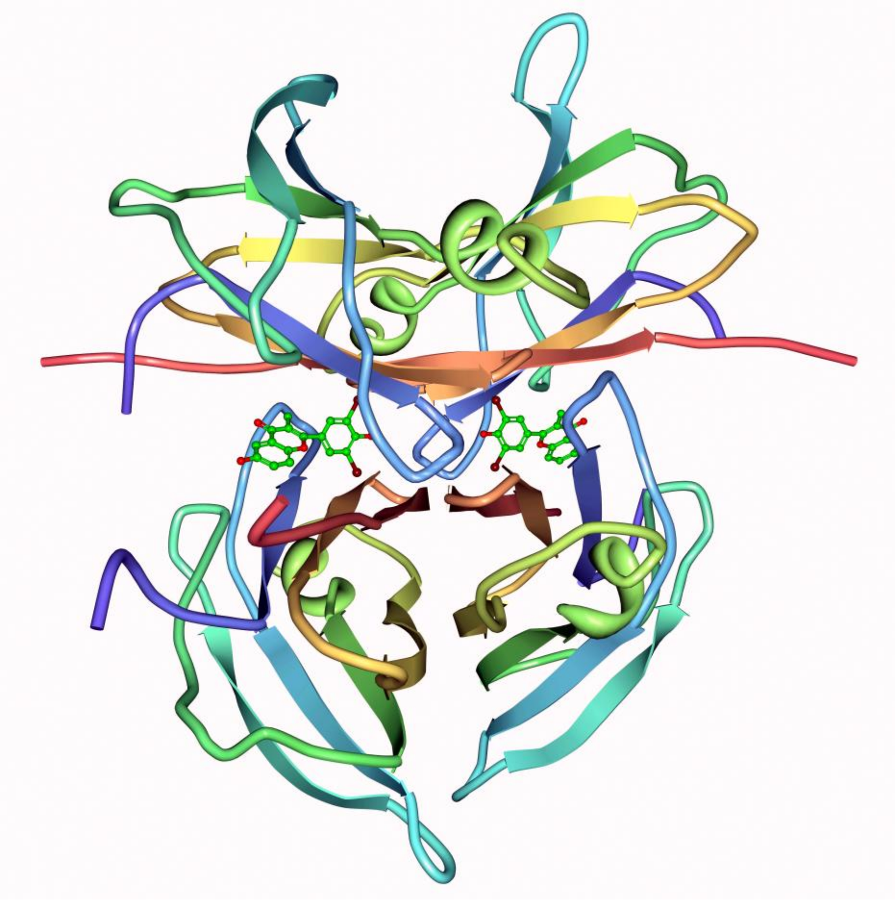


**Figure 1**. The crystallopraphic structure of the complex ratTTR and flavones EMD21388, reported in the Protein Data Bank (PDB ID: 1KGJ)

This structure has been used in this thesis as a starting point to develop a local pharmacophore to find interesting flavonoid compounds as possible TTR tetramer stabilizers.

## Rat vs. human TTR

Previous to the study itself, a comparison between the rat and the human TTR proteins was carried out in order to establish whether the rat TTR crystal structure could be used as the basis for the pharmacophore modeling study.

With this objective, the TTR binding pocket (**Figure 3**) was analyzed in order to determine which are the main parts and the residues involved in the interactions with the molecules reported as stabilizers.

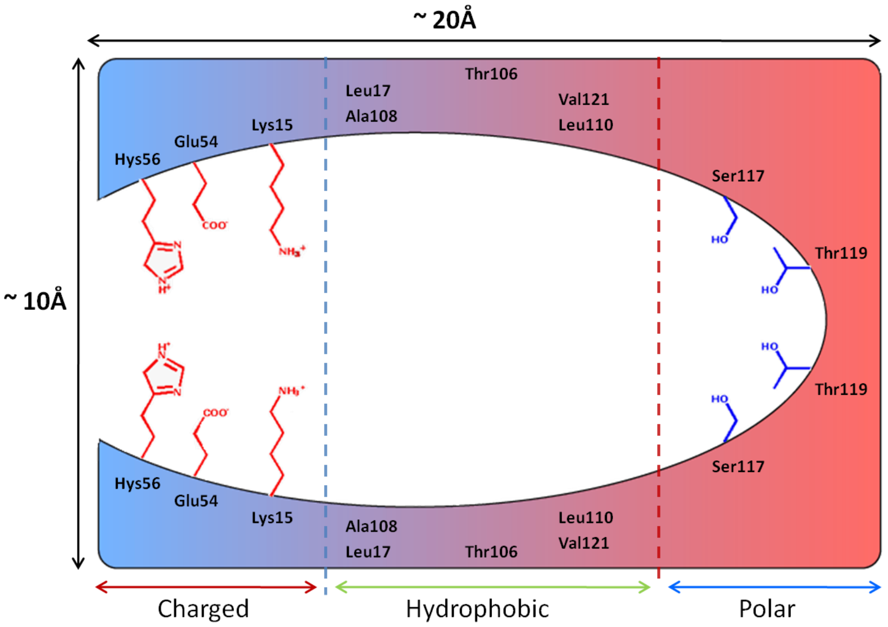
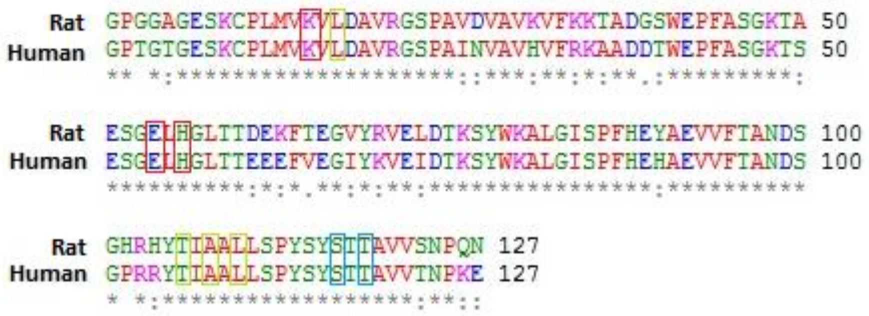

~ 20Å
~ 10Å
Hys56
Glu54
Lys15
Leu17
Ala108
Thr106
Val121
Leu110
Ser117
Thr119
Thr119
Ser117
Hys56
Glu54
Lys15
Ala108
Leu17
Thr106
Leu110
Val121
COO⁻
NH₃⁺
Charged
Hydrophobic
Polar

**Figure 2.** 2D representation of the main residues responsible of the protein-ligand interactions in most of the reported TTR crystal structures.

The second step was to determine if these key residues in the TTR binding pocket were conserved in the rat and in the human TTR protein sequences, and to analyze if both rat and human structures have any significant structural differences, in order to perform the proposed pharmacophore modeling study from an optimal starting point.



```
  Rat  GPGGAGESKCPLMVKVLDAVRGSPAVDVAVKVFKKTADGSWEPFASGKTA 50
Human  GPTGTGESKCPLMVKVLDAVRGSPAINVAVHVFRKAADDTWEPFASGKTS 50
       ** *:********************::***:**:*:**.:*********:

  Rat  ESGELHGLTTDEKFTEGVYRVELDTKSYWKALGISPFHEYAEVVFTANDS 100
Human  ESGELHGLTTEEEFVEGIYKVEIDTKSYWKALGISPFHEHAEVVFTANDS 100
       **********:*:*.**:*:**:****************:**********

  Rat  GHRHYTIAALLSPYSYSTTAVVSNPQN 127
Human  GPRRYTIAALLSPYSYSTTAVVTNPKE 127
       * *:******************:**::
```

**Figure 3**. Rat and Human TTR sequence aligned. Residues present in the human binding site are highlighted along the rat sequence.

As can be seen from **Figure 3** all the human wtTTR residues from the binding site are conserved in the rat TTR protein. For this reason, the rat TTR was considered as a good template for the structure-based pharmacophore modelling.

The proposed methodology and the preliminary results obtained in this pharmacophore modeling analysis of the TTR-ligands interactions were reported in an article published in the Amyloid Journal in 2011 which is shown below[7].

Amyloid Downloaded from informahealthcare.com by Joaquin Ibanez Esteb on 07/05/11
For personal use only.

## Drug discovery targeted at transthyretin cardiac amyloidosis: rational design, synthesis, and biological activity of new transthyretin amyloid inhibitors

D. Blasi[1], M. Pinto[2], J. Nieto[3], G. Arsequell[4], G. Valencia[4], A. Planas[3], N. B. Centeno[2], & J. Quintana[1]

[1]*Drug Discovery Platform (PDD), Parc Cientific Barcelona, Barcelona, Spain,* [2]*Computer-Assisted Drug Design Laboratory, Research Group on Biomedical Informatics (GRIB) IMIM-Universitat Pompeu Fabra, Barcelona, Spain,* [3]*Laboratory of Biochemistry, Institut Quimic de Sarrià, Universitat Ramon Llull, Barcelona, Spain, and* [4]*Unit of Glycoconjugate Chemistry, Institut de Química Avançada de Catalunya, I.Q.A.C.-C.S.I.C., Barcelona, Spain*

**Abstract:** We have previously designed and synthesized ligands that stabilize the transthyretin (TTR) tetramer, in order to obtain therapeutically active compounds for familial amyloid polyneuropathy. We are hereby reporting a drug design strategy to optimize these ligands to target familial amyloid cardiomyopathy, through the following steps: (a) Structure Activity Relationship (SAR) analyses of the ligands described previously for the TTR tetramer, classified in structurally similar families; (b) drug design/optimization of TTR ligands through docking in the TTR tetramer three-dimensional structure and through optimization of physicochemical/pharmacokinetic/selectivity properties; (c) comparative structural analyses of selected amyloidogenic and non-amyloidogenic TTR mutants and native TTR structures; and (d) virtual screening of commercially available ligands and therapeutically active compounds (repurposing) towards wild-type and mutant TTR tetramer structures. First results in steps (a) and (d) of this strategy will be reported.

**Introduction:** Transthyretin (TTR) is a homotetrameric protein that functions as the backup transporter for thyroxine hormone (T4) in plasma, and it is the main transporter across blood brain barrier. Also, TTR is the main carrier of vitamin A through the formation of a TTR–retinol-binding protein complex.

TTR exhibits an amyloidogenic behavior, with more than 100 TTR mutants associated with hereditary amyloidosis. Although details of the TTR aggregation mechanism are still unknown, it involves several steps such as: dissociation of the tetramer; conformational changes in the monomer; aggregation of these modified monomers into non-fibrillar oligomers, that later form protofibrils, and further elongate into mature fibrils [1], TTR is in the cause of several rare amyloid diseases such as familial amyloid polyneuropathy (FAP), familial amyloid cardiomyopathy, central nervous system selective amyloidosis and senile systemic amyloidosis.

The crystal structure of TTR in complex with T4 reveals that iodine atoms ocuppy four out of six (three symmetry-related pairs) small hydrophobic depressions – termed halogen binding pockets (HBP) – present in each binding site [2]. Considering that several non-steroidal antiinflammatory drugs (NSAIDs) are able to bind to TTR [3,4], some iodinated NSAIDs derivatives that could present high binding affinities to TTR have been reported as therapeutical active compounds for FAP disease [5]. Taking advantage of this knowledge, a drug repurposing approach has been carried out in order to find new amyloid inhibitors, based on *in silico* identification of secondary or ‘off-target’ drug actions, followed by the development of the compound, already in the market, in a new relevant indication targeting TTR protein.

**Methods**: ***Hardware specifications:*** Molecular modeling studies were carried out on an Intel Pentium Core 2 Duo 3.00 GHz running Windows XP.

***Software specifications:*** The following software programs were used for this study: MOE 2009.10 [6] for complex protein-ligand preparation, conformational Low MD calculation and plotting alignment results; and LigandScout 2.03 [7] for the pharmacophore elucidation and visualization of compounds alignments based on the ligand-based pharmacophore model obtained.

***Preparation of crystallographic complex:*** Crystallographic structure 1KJG [8] was extracted from Protein Data Bank (http://www.pdb.org) and hydrogen atoms were added using LigX in MOE 2009.10 package. The minimization of the system was carried out also in MOE 2009.10, using the MMF95 ForceField.

***Data set:*** A set of 3000 Flavonoids was extracted from MMsINC Database [9]. The conformational sampling of the MMsINC subset selected was performed in MOE 2009.10 with the Conformational Search approach, using a Low MD search protocol.

**Results and discussion:** It is well known that drugs interact with multiple targets, and drug discovery repurposing efforts are based upon a compound’s potentially ‘desirable promiscuity’, following the ‘known compound-new target’ approach. Within our program to obtain new drugs towards TTR-related amyloidosis, through repurposing, we have generated a structure-based pharmacophore model that we applied to study a database of commercially available flavonoids, extracted from the MMsINC DataBase [9]. The compounds compiled in this database are non-redundant and biomedically relevant chemical structures, richly annotated with crucial chemical

RIGHTSLINK Copyright Clearance Center

properties. Moreover, MMsINC is integrated with other primary data collectors, such as PubChem, Protein Data Bank, the Food and Drug Administration database of approved drugs, ZINC and also commercial catalogs from several companies.

The methodology presented here involves four steps: (a) identification and pharmacophore extraction from the selected crystallographic complex, 1KGJ; (b) conformational search calculation using Low MD approach for the 3000 flavonoids selected; (c) alignment of the conformers with the pharmacophore extracted from the TTR-ligand complex and (d) visualization and plotting of the results obtained.

The starting point of this study was the crystal structure of a wild-type rat TTR complexed with the 3′,5′-dibromoflavone (EMD21388). This structure was selected since it is known that several flavones, including this one, exhibit a potent inhibition activity against TTR amyloid formation [10], and therefore the pharmacophore extracted from such a complex could be reliable for the repurposing protocol. In addition, all the human residues present on the wild type TTR binding sites are identical, and therefore confirming that the rat TTR structure could be a good template for structure-based pharmacophore modeling. The pharmacophore analysis showed the presence of five features that represent the interaction with the TTR binding site: two hydrophobic volumes located in the halogen atoms and another one over one of the aromatic rings; one hydrogen bond interaction between the carbonyl group and a water molecule; and another hydrogen bond interaction with the hydroxyl group in the phenolic ring (Figure 1). In order to validate this model, 11 flavones synthesized in house were tested *in silico*; a plot of the pharmacophoric scores of the conformational sets of each molecule showed a correlation with the *in vitro* activity of these compounds in an amyloid fibril formation inhibition assay, thus demonstrating that the model is able to discriminate between active and inactive compounds for this small data set.

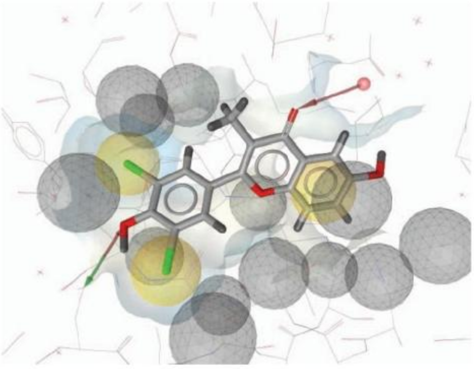


Figure 1. 3D representation of the pharmacophore extracted from crystallographic rat TTR-EMD21388 complex.

A conformational search, using LowMD calculation approach, of a 3000 commercially available flavonoids subset, extracted from the MMsINC database, was next performed, and an alignment protocol was carried out based on the pharmacophore model obtained from the structure of the rat wtTTR - EMD21388 complex. Results from

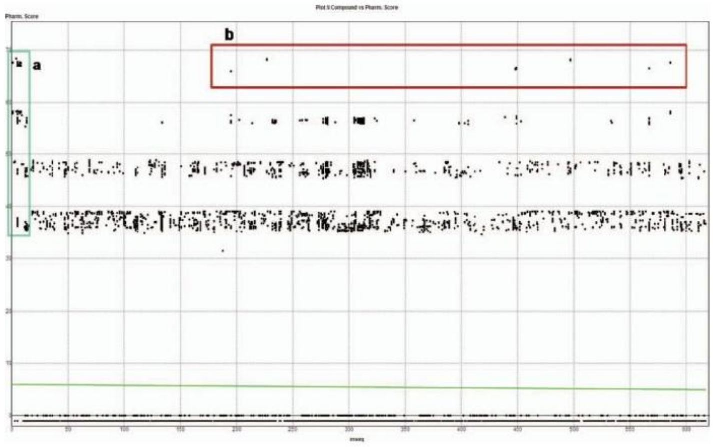
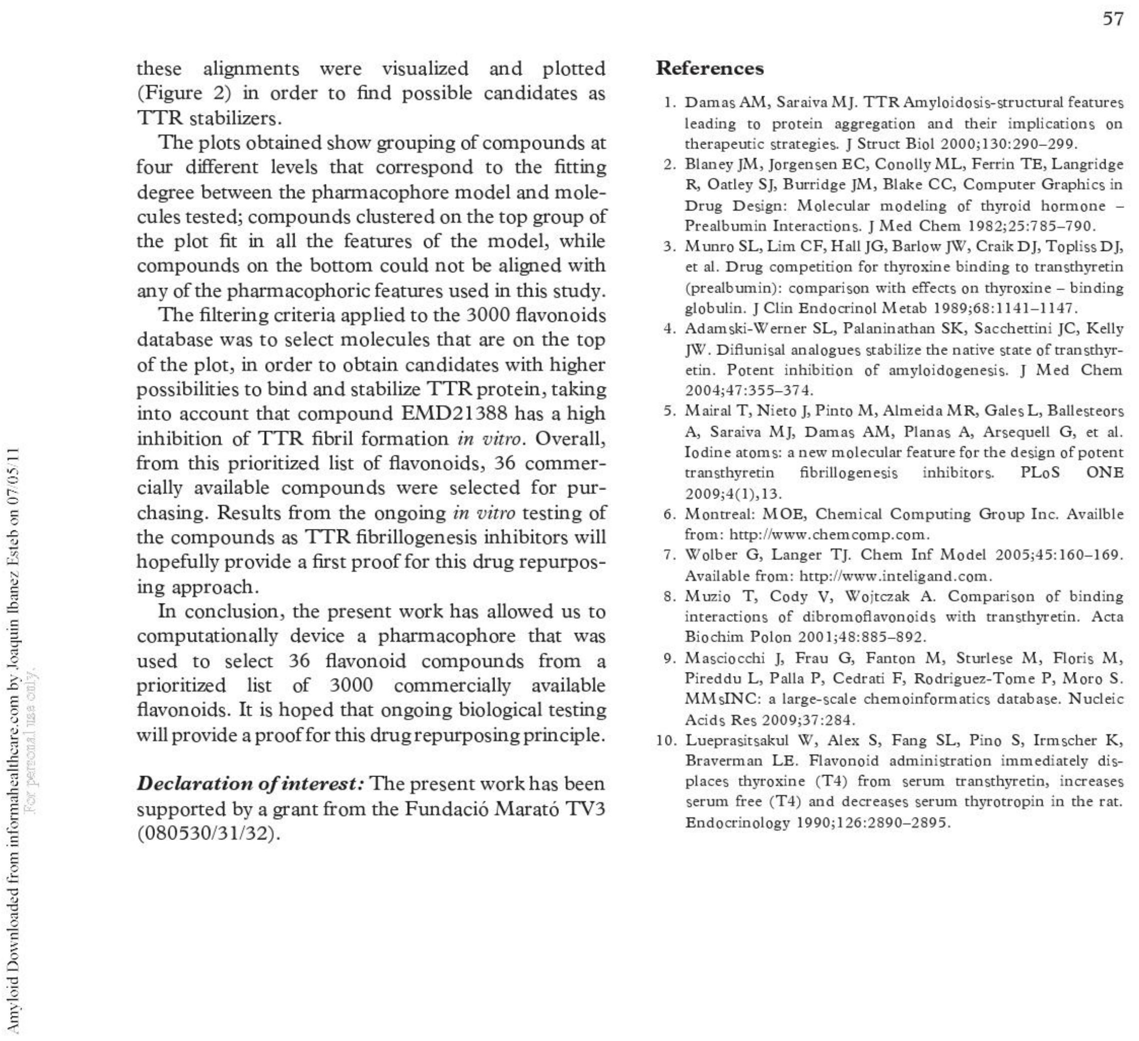


Figure 2. Section of the whole visualization plot obtained after alignment calculations between pharmacophore model and flavonoids conformers. (a) Validation subset from 11 flavonoids synthesized in house; (b) region considered as highly promising to find new potential TTR stabilizers.

Amyloid Downloaded from informahealthcare.com by Joaquin Ibanez Esteb on 07/05/11
For personal use only.



these alignments were visualized and plotted (Figure 2) in order to find possible candidates as TTR stabilizers.

The plots obtained show grouping of compounds at four different levels that correspond to the fitting degree between the pharmacophore model and molecules tested; compounds clustered on the top group of the plot fit in all the features of the model, while compounds on the bottom could not be aligned with any of the pharmacophoric features used in this study.

The filtering criteria applied to the 3000 flavonoids database was to select molecules that are on the top of the plot, in order to obtain candidates with higher possibilities to bind and stabilize TTR protein, taking into account that compound EMD21388 has a high inhibition of TTR fibril formation *in vitro*. Overall, from this prioritized list of flavonoids, 36 commercially available compounds were selected for purchasing. Results from the ongoing *in vitro* testing of the compounds as TTR fibrillogenesis inhibitors will hopefully provide a first proof for this drug repurposing approach.

In conclusion, the present work has allowed us to computationally device a pharmacophore that was used to select 36 flavonoid compounds from a prioritized list of 3000 commercially available flavonoids. It is hoped that ongoing biological testing will provide a proof for this drug repurposing principle.

***Declaration of interest:*** The present work has been supported by a grant from the Fundació Marató TV3 (080530/31/32).

Amyloid Downloaded from informahealthcare.com by Joaquin Ibanez Esteb on 07/05/11
For personal use only.

## References

1. Damas AM, Saraiva MJ. TTR Amyloidosis-structural features leading to protein aggregation and their implications on therapeutic strategies. J Struct Biol 2000;130:290–299.
2. Blaney JM, Jorgensen EC, Conolly ML, Ferrin TE, Langridge R, Oatley SJ, Burridge JM, Blake CC, Computer Graphics in Drug Design: Molecular modeling of thyroid hormone – Prealbumin Interactions. J Med Chem 1982;25:785–790.
3. Munro SL, Lim CF, Hall JG, Barlow JW, Craik DJ, Topliss DJ, et al. Drug competition for thyroxine binding to transthyretin (prealbumin): comparison with effects on thyroxine – binding globulin. J Clin Endocrinol Metab 1989;68:1141–1147.
4. Adamski-Werner SL, Palaninathan SK, Sacchettini JC, Kelly JW. Diflunisal analogues stabilize the native state of transthyretin. Potent inhibition of amyloidogenesis. J Med Chem 2004;47:355–374.
5. Mairal T, Nieto J, Pinto M, Almeida MR, Gales L, Ballesteors A, Saraiva MJ, Damas AM, Planas A, Arsequell G, et al. Iodine atoms: a new molecular feature for the design of potent transthyretin fibrillogenesis inhibitors. PLoS ONE 2009;4(1),13.
6. Montreal: MOE, Chemical Computing Group Inc. Availble from: http://www.chemcomp.com.
7. Wolber G, Langer TJ. Chem Inf Model 2005;45:160–169. Available from: http://www.inteligand.com.
8. Muzio T, Cody V, Wojtczak A. Comparison of binding interactions of dibromoflavonoids with transthyretin. Acta Biochim Polon 2001;48:885–892.
9. Masciocchi J, Frau G, Fanton M, Sturlese M, Floris M, Pireddu L, Palla P, Cedrati F, Rodriguez-Tome P, Moro S. MMsINC: a large-scale chemoinformatics database. Nucleic Acids Res 2009;37:284.
10. Lueprasitsakul W, Alex S, Fang SL, Pino S, Irmscher K, Braverman LE. Flavonoid administration immediately displaces thyroxine (T4) from serum transthyretin, increases serum free (T4) and decreases serum thyrotropin in the rat. Endocrinology 1990;126:2890–2895.

## Biological activity of a set of selected flavonoids

Among the 36 flavonoids described in the paper as promising TTR inhibitors, 4 compounds (**Table 1**) were selected due to their commercial availability. Their activity was evaluated by means of *in vitro* kinetic turbidimetric assay[8] explained in chapter 1 . The results are graphically depicted in **Figure 4**.

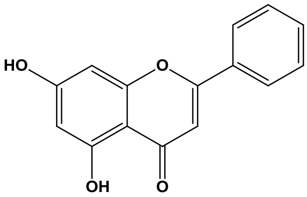
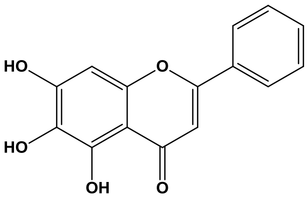
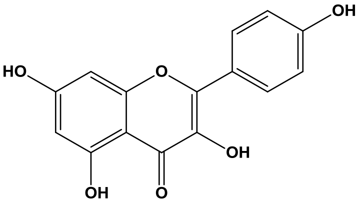
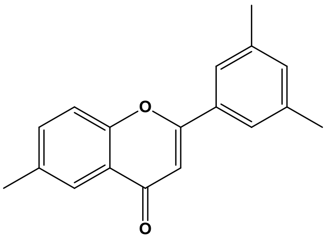
| Compound name | Structure | Name | Pharm. Score* |
|---|---|---|---|
| **1** |  | Chrysin | 46,5 |
| **2** |  | Baicalein | 46,4 |
| **3** |  | Kaempferol | 47,5 |
| **4** |  | Axp224 | 48,1 |

**Table 1**. List of compounds tested against TTR protein. (*) Pharmacophoric Score is a mesure of the alignment between molecules and pharmacophore extracted from the rat TTR crystal structure, using LigandScout 3.0 software[9].

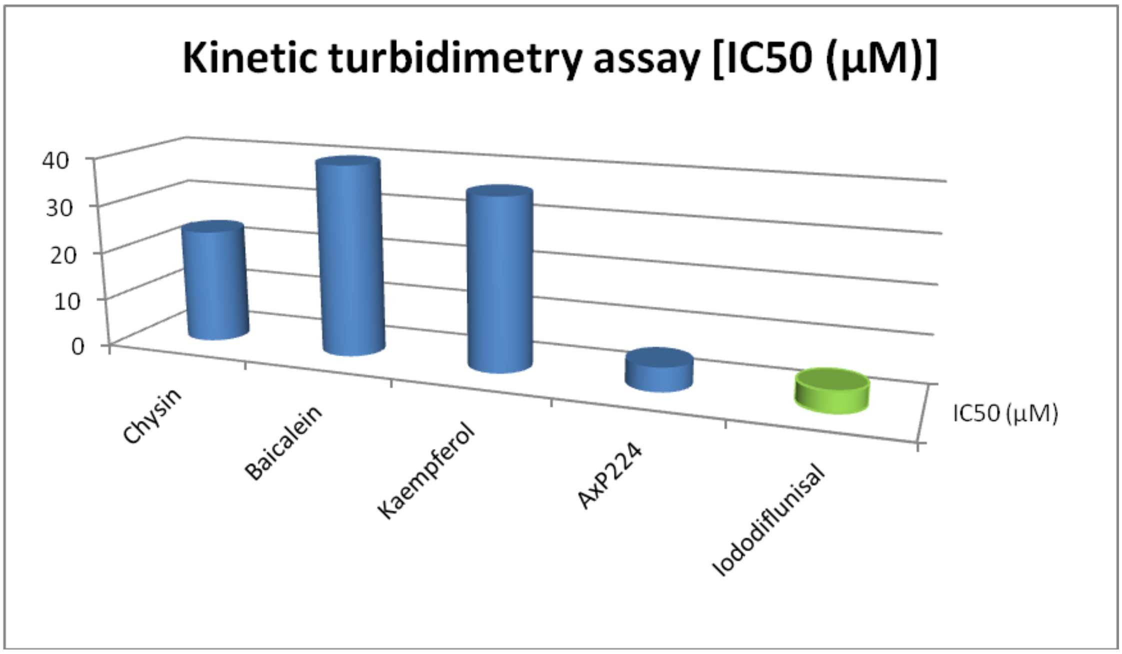


**Figure 4.** In vitro results for the 4 tested compounds against TTR Y78F mutant. Green bars refers to the Iododiflunisal activity that has been used as reference.

Interestingly, the AxP224 compound shows the best pharmacophore score due to the good alignment pattern obtained for this compound after comparison with the interaction map obtained from the TTR:EMD2138 crystal structure, and it can be also deduced from the analysis of the fitting of AxP224 against the interaction map used in the study (**Figure 5**),

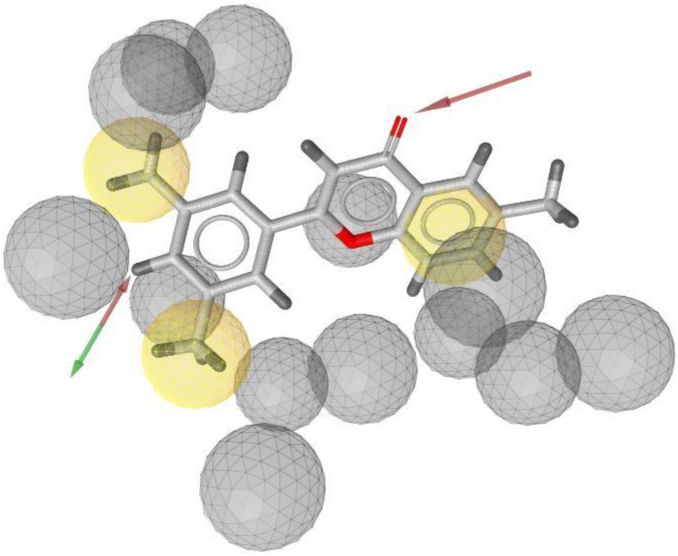


**Figure 5.** AxP224 fitting with the extracted pharmacophore used in this study. The molecule shows a good alignment with 4 of the 5 features.

Therefore, the good biological activity of AxP224 can be seen as a proof of this repurposing methodology, since using this workflow, it has been possible to find an active molecule with a stabilizing behavior similar to one of the most potent TTR amyloid inhibitors known (Iododiflunisal).

The methodology developed in this thesis chapter, structure-based drug design focused on pharmacophoric research, would also be interesting to be applied to an expanded set of drugs launched into the market, in a repurposing drug finding study for TTR inhibitors.

## Bibliography

(1) Drews, J. Drug Discovery: A Historical Perspective. *Science* **2000**, *287*, 1960–1964.

(2) Baures, P. W.; Peterson, S. a; Kelly, J. W. Discovering transthyretin amyloid fibril inhibitors by limited screening. *Bioorganic & Medicinal Chemistry* **1998**, *6*, 1389–1401.

(3) Green, N. S.; Foss, T. R.; Kelly, J. W. Genistein, a natural product from soy, is a potent inhibitor of transthyretin amyloidosis. *Proceedings of the National Academy of Sciences of the United States of America* **2005**, *102*, 14545–14550.

(4) Maia, F.; Rosa, M.; Kijjoa, A.; Pinto, M. M. M. The binding of xanthone derivatives to transthyretin. *Biochemical Pharmacology* **2005**, *70*, 1861–1869.

(5) Kohrle, J.; Fang, S. L.; Yang, Y.; Irmscher, K.; Hesch, R. D.; Pino, S.; Alex, S.; Braverman, L. E. Rapid Effects of the Flavonoid EMD 21388 on Serum Thyroid Hormone Binding and Thyrotropin Regulation in the Rat. *Endocrinology* **1989**, *125*, 532–537.

(6) Cody, V. Mechanisms of molecular recognition: crystal structure analysis of human and rat transthyretin inhibitor complexes. *Clinical Chemistry and Laboratory Medicine* **2002**, *40*, 1237–1243.

(7) D. Blasi, M. Pinto, J. Nieto, G.Arsequell, G. Valencia, A. Planas, N.B. Centeno, J. Q. Drug discovery targeted at transthyretin cardiac amyloidosis: rational design, synthesis, and biological activity of new transthyretin amyloid inhibitors. *Amyloid* **2011**, *18*, 55–57.

(8) Dolado, I.; Nieto, J.; Saraiva, M. J. M.; Arsequell, G.; Valencia, G.; Planas, A. Kinetic assay for high-throughput screening of in vitro transthyretin amyloid fibrillogenesis inhibitors. *Journal of Combinatorial Chemistry* **2005**, *7*, 246–252.

(9) Wolber, G.; Langer, T. LigandScout: 3-D pharmacophores derived from protein-bound ligands and their use as virtual screening filters. *Journal of Chemical Information and Modeling* **2005**, *45*, 160–169.

# PART III

# LIGAND EFFICIENCY INDICES

## Ligand Efficiency Indices (LEI) for TTR Drug Discovery

The concept of Ligand Efficiency (LE)[1] used as a tool to guide the fragment and lead optimization in the drug discovery process has recently emerged. The use of a LE relating the affinity ($K_i$, $IC_{50}$ or similar) to the molecular size (molecular weight, number of non-hydrogen atoms, number of heavy atoms and others) is now well accepted. The combination of two Ligand Efficiency Indices (LEI) such as binding efficiency index (BEI; relating the potency to the molecular weight of the Ligand) and surface efficiency index (SEI; relating the potency to the polar surface area), in comparing retrospectively chemical series in an optimization plane BEI-SEI, has been published recently[2,3].

In order to find new effective ways to guide the TTR-focused drug discovery strategy to optimize amyloid inhibitors, the LEIs based approach has been applied to map different chemical series in the TTR chemical-biology space (CBS).

The use of this LEI methodology has been illustrated by a 2D graphical representation (nBEI-NSEI plane), to establish a prospective way to guide a lead optimization process based on theoretical potency calculation (using docking protocols) of possible new TTR amyloid inhibitors, combined with the LEIs approach.

For this study, Iododiflunisal[4] was chosen as the molecular template for this *in silico* design of new and more efficient compounds able to stabilize TTR tetramer. We have published two scientific papers included in this thesis chapter: one presenting a review of the LEI methodology and applications[5], and the other[3] about the LEI methodology applied to the discovery of new TTR inhibitors. We are also including a third published scientific article[6] in this thesis chapter, to describe the computational methodology followed to use the LEI approach in a prospective manner, so as to design new more efficient compounds on the basis of a combined LEI / molecular docking approach.

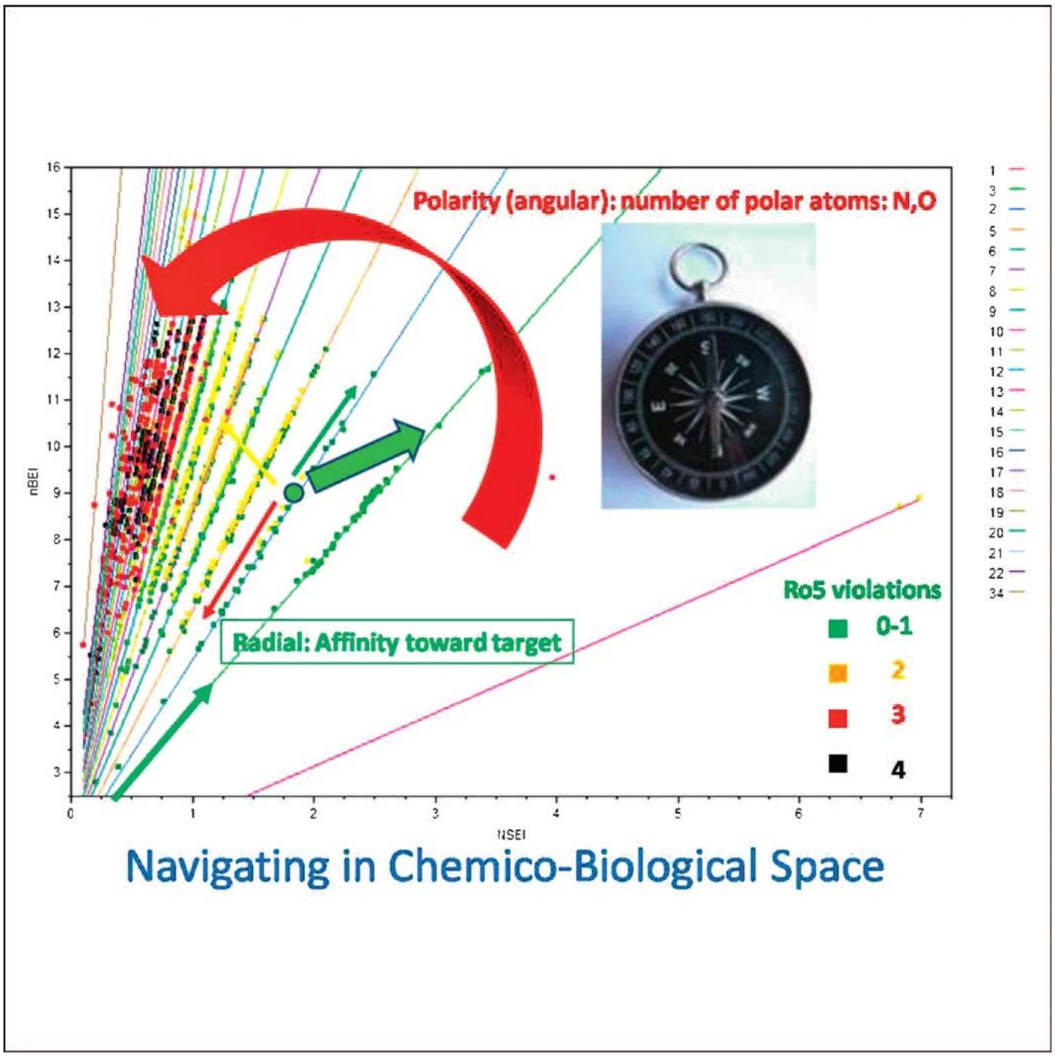
**Review**

DOI: 10.1002/minf.201000161

# Ligand Efficiency Indices (LEIs): More than a Simple Efficiency Yardstick

Cele Abad-Zapatero*[a, b] and Daniel Blasi[b]

*Presented at the 18th European Symposium on Quantitative Structure Activity Relationships, EuroQSAR 2010, Rhodes, Greece*
To the younger generations of drug discoverers: may they find the way



© 2011 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

molecular
informatics

**Abstract**: The concept of ligand efficiency and the usage of ligand efficiency values to assess the quality of fragments and compounds is becoming more accepted in the practice of medicinal chemistry. This is particularly true as it refers to the efficiency of ligands per unit size (i.e., binding affinity/number of non-hydrogen atoms or binding affinity/MW). The use of the Ligand Efficiency Indices (LEIs) as variables for a Cartesian mapping of chemico-biological space, the concept of AtlasCBS, has been presented in a recent publication with some initial drug-discovery applications. In this communication, we present additional applications of the concept in three domains of drug discovery: i) analyze and compare the content of databases: inhibitors vs. drugs; ii) polypharmacology; and iii) applications to Fragment-Based strategies. We suggest that the combined use of LEIs in a Cartesian representation of Chemico-Biological Space (AtlasCBS) could be a useful tool in various aspects of drug-discovery in the future.

**Keywords**: Ligand efficiency indices · Chemico-biological space · AtlasCBS · LEIs

## 1 Introduction

Through history, graphical representation of data has played a critical role in providing insight and understanding of complex processes and phenomena. Within the biomedical sciences, the challenge has been finding ways to represent adequately and connect the complex universe of bioactive chemical compounds with the corresponding biological targets. The vastness of Chemico-Biological space as it relates to biology and medicine has been discussed by Oprea and collaborators[1–3] and more recently by Lipinski and Hopkins[4] and others.

Recently, the concept of an atlas-like representation of chemico-biological space (AtlasCBS) has been proposed and its usage in various domains of drug discovery has been reported.[5] This concept allows the representation of the content of the databases containing affinity data (PDBBind,[6,7] BindingDB,[8] WOMBAT[9] and ChEMBL[10] (www.ebi.ac.uk/chembldb) in various Cartesian planes based on pairs of variables derived from the definition of Ligand Efficiency Indices (LEIs).[2] Applications of this concept to describe trajectories in chemico-biological space and to guide the drug-discovery process have been presented and illustrated. This concept would suggest that LEIs can be considered to be more than just a simple 'efficiency yardstick' to compare fragments or compounds. Rather, they can be used in combination to provide an effective mapping of chemico-biological space that could be useful to guide the drug-discovery efforts.[11]

In this review, we present additional applications of the AtlasCBS concept in three domains of drug discovery: i) analyze and compare the content of databases, in particular as it relates to comparing the mapping of inhibitors vs. drugs; ii) polypharmacology, as it relates to the multiple activity of approved drugs in different targets; and iii) three applications to Fragment-Based strategies. We suggest that a combination of an effective graphical representation of chemico-biological space and a robust numerical framework could provide a more efficient drug-discovery strategy in the near future. We recognize the limitations of this retrospective analysis but hope that by illustrating the power of this graphical and numerical framework based on LEIs, future prospective usage of this concept and ideas would help to test their validity and expedite and optimize drug discovery.

## 2 Concepts and Applications

### 2.1 The AtlasCBS Concept

The idea of AtlasCBS has been conceived as a series of physical or electronic pages representing on the different Cartesian planes each target:ligand pair as a point. The Cartesian coordinates of each ligand:target are given by pairs of values derived from the affinity of the ligand toward the target ($K_i$ or equivalent) in relation to its size (MW or number of non-hydrogen atoms NHEA), and its polarity (expressed as Polar Surface Area: PSA; or number of polar atoms, $NPOL = \text{count of } N + O$). The first one represents a size-related Ligand Efficiency Index (LEI) and the second a polarity-related LEI. It is the combination of these three variables (affinity, size and polarity) in a Cartesian plane using the two Ligand Efficiency Indices (LEIs) that permit the representation of CBS in a series of atlas-like pages.[5]

The character and appearance of each of these pages depends on the choice of the complementary pairs of variables. An initial pair relating affinity to PSA and MW (SEI-BEI: *x,y*; respectively) has been introduced and illustrated in previous publications.[12,13] A novel definition of LEIs relating the affinity of the ligand to its atomic composition provides a more distinct separation of the target:ligand pairs in the corresponding Cartesian planes (see Table 1 for the definition of the different LEIs pairs). These Cartesian planes have been referred to as 'efficiency planes' since they are derived from variables related to the LEIs of each target:ligand pair.[5]

---

[a] *C. Abad-Zapatero*
*Center for Pharmaceutical Biotechnology, University of Illinois at Chicago*
*900 So. Ashland St, MBRB Building, Room 3020; (M/C870), Chicago, IL 60607-7173, USA*
**e-mail: caz@uic.edu*

[b] *C. Abad-Zapatero, D. Blasi*
*Drug Discovery Platform, Parc Científic Barcelona (PCB)*
*Baldiri Reixac 4–6, E-08028 Barcelona, Spain*

© 2011 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

**Table 1.** Definitions of Ligand Efficiency Indices (LEIs) pairs used in each Efficiency Plane.

| Name | Definition | Example value[a] | Equation |
|---|---|---|---|
| *BEI* | p(*K*), p($K_d$), or p($IC_{50}$)/MW(kiloDa) | 27 | 1 |
| *SEI* | p(*K*), p($K_d$), or p($IC_{50}$)/(PSA/100 Å$^2$) | 18 | 2 |
| Slope of lines | 10(PSA/MW) | | |
| Description | Efficiency plane based on physico-chemical properties of the ligand | | |
| *NSEI* | $-\log_{10} K_i/(NPOL) = pK_i/NPOL(N,O)$ | 1.5 | 3 |
| *NBEI* | $-\log_{10} K_i/(NHEA) = pK_i/(NHEA)$ | 0.36 | 4 |
| Slope of lines | *NPOL/NHEA* | | |
| Description | Efficiency plane based on atomic properties of the ligand | | |
| *nBEI* | $-\log_{10}[(K_i/NHEA)]$ | 10.25 | 5 |
| *NSEI* | $-\log_{10} K_i/(NPOL) = pK_i/NPOL(N,O)$ | 1.5 | 3 |
| *mBEI* | $-\log_{10}[(K_i/MW)]$ | 10.5 | 6 |
| Slope of lines | *NPOL* | | |
| Intercept: | $\log_{10}(NHEA)$ or $\log_{10}(MW)$[b] | | |
| Description | Efficiency plane based on atomic properties of the ligand with a special 'fan-like' appearance | | |

[a] Examples and reference values for the LEIs are calculated for each index using the following idealized values (units have been omitted in the table): $K_i$ or $IC_{50} = 1.0$ nM; $p(K_i) = -\log K_i = 9.00$; $MW = 0.333$ kDa; *NHEA* = Number of heavy atoms (non-Hydrogen in the compound) = 25; *NPOL* = Number of polar atoms (N,O) = 6; $NBEI/NSEI = NPOL(N,O)/NHEA(\text{non-H}) = 0.36/1.5 = 6/25 = 0.24$; $PSA = 50$ Å$^2$. [b] Supplementary Material[2]

Although the location of each target:ligand pair in each efficiency plane is given by its two efficiency components (polarity and size, *x* and *y*), it is important to note that the choice of variables permits a different interpretation of the appearance of the maps that facilitates its understanding.

Cele Abad-Zapatero received a Licenciado degree (Physics) from the University of Valladolid in Spain. He obtained his Ph.D. in Biophysics (Protein Crystallography) from the University of Texas at Austin, US. He did postdoctoral work in virus/macromolecular crystallography at Purdue Univ. mentored by M. G. Rossmann. He worked for more than twenty years at Abbott Laboratories doing SBDD. He is currently Adjunct Professor that the Center for Pharmaceutical Biotechnology of the University of Illinois at Chicago. His research interests are focused in developing more robust and predictive methods for structure, and fragment-based drug discovery.

Daniel Blasi received a Master's Degree in Organic Chemistry from the University of Barcelona in 2006, with a specialization in solid phase cyclic peptide synthesis. He spent two years at Enantia working as assistant researcher junior. He now is a Ph.D. student in the Drug Discovery Platform at the Scientific Park of Barcelona headed by Dr. Jordi Quintana. The focus of his thesis is the design of new amyloid inhibitors of Transthyretin protein using computational tools.

In the SEI-BEI plane, the slope of the lines is given by 10(PSA/MW), a quantity that depends only on the physico-chemical properties of the ligand. Thus the angular coordinate of any target:ligand pair in this representation depends only on the chemical composition of the ligand. On the other hand, the radial coordinate depends on the relative affinity toward the target. This conceptualization permits a 'decoupling' of chemical and biological variables.[5]

Interestingly, the *NSEI-nBEI* (*x*,*y*) pair allows an analogous representation whereby the slope of the lines in the Cartesian plane is given by the *NPOL* count (N + O) of the ligand and the intersect by $\log_{10}$(NHEA). The statistical analysis shows that most of the variance (> 98 %) can be explained by the slope of the lines (see in the literature,[5] Supplementary Material, Box S2). Therefore, the plots show a series of distinct lines characterized by a slope given by *NPOL*, with a striking 'fan-like' (or even 'harp-like') appearance. All these concepts have been illustrated in Figure1a where the content of PDBBind[6,7] has been represented. The content of any other database containing affinity data ($K_i$, IC50 or equivalent) can be plotted in similar efficiency planes and compared. By a suitable choice of the range of variables and scale, a certain subspace of the chemico-biological universe can be represented with the same appearance and characteristics. Figure 1b illustrates the mapping of the 2353 compounds that are active against the HIV protease within ChEMBL[14] and relate the angular coordinate of the chemical entities to the well known guidelines of Lipinski's 'Rule of 5' (Ro5). Compounds violating an increasing number of the Ro5 guidelines are colored from green to red to black (Figure 1b).

The limitations of the proposed variables to map CBS should be emphasized as well as extensions that are currently being investigated to address those issues. When the chemical space under investigation has similar (or the

© 2011 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

same) *NPOL* or *NPOL*/NHEA ratios (for instance peptides with the same composition but different sequences), then the choice of variables to map CBS should be based on PSA/MW ratios using the three-dimensional structure of the ligands to accurately calculate the PSA variable. It can be shown that in those cases the planes defined by mBEI vs. SEI (*y*, *x*) permit a separation in the angular coordinate with variations in the polar surface area as fine as PSA/100. This strategy permits the separation of peptides having the same *NPOL* number in a very fine sampling of chemical variables (i.e., polar surface area) of the ligand.

Another clarification is needed in regard to the definition of *NPOL* as the sum of N, O only. The initial choice of this definition was purely on the basis of simplicity. It is clearly a convenient choice to relate the maps to the Lipinski Ro5 guidelines (Figure 1b). For chemical series involving hypervalent P or compounds containing polar sulfurs an 'extended' definition could be used to analyze the series of terms of PSA/MW ratios, where the PSA can be accurately calculated as indicated above. It is the possibility of selecting different efficiency planes, defined by different variables and using widely different scales and regions, that gives versatility to the AtlasCBS concept.

### 2.2 Analyze and Compare the Content of Databases: Inhibitors vs. Drugs

To illustrate the use of these concepts we chose to study and compare three different databases. The public domain database of protein:ligand complexes (PDBBind 2007 release, extracted from the PDB), a small subset of marketed drugs from the ChEMBL database and the extensive database of bioactive compounds WOMBAT.[9] As illustrated before (Figure 1), the results for the overall representation of the content of PDBind in LEI space (*nBEI* vs. *NSEI*) is a uniformly distributed set of target:ligand complexes with a normally distributed set of affinity constants ($K_i$'s).[6,7] In contrast, the results for a limited subset of the drug from ChEMBL are shown in Figure 2a using the same pair of LEIs. This pair of variables was found most suitable before to highlight the separation in terms of the number of polar atoms of the drugs (*NPOL*).

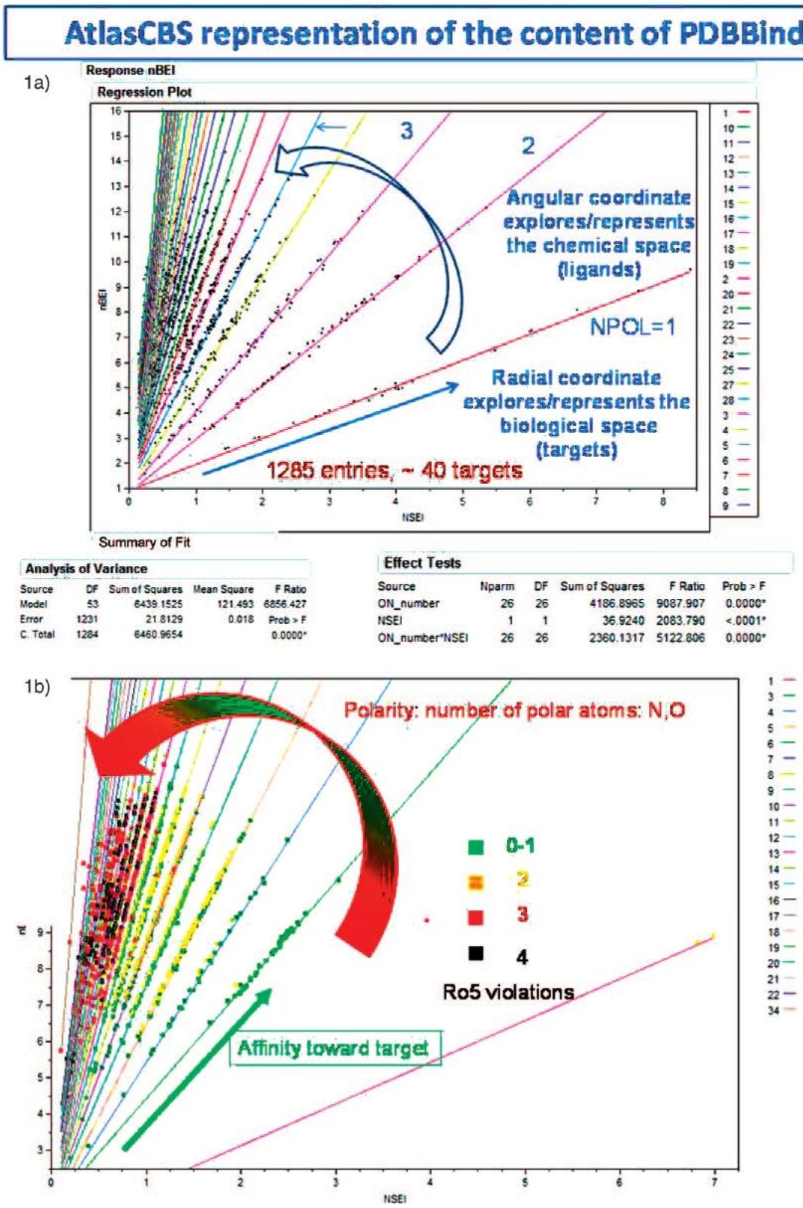
This 'fan-like' plot contrasts with the one in Figure 1 shown previously, obtained from PDBBind, which is a well-curated subset of the PDB with an approximately normal distribution of p$K_i$ values.[6,7] PDBBind has a much heavier density of compounds in the polar regions ($NPOL > 6$) of chemical space attesting to the higher abundance of polar compounds in the PDB. Naturally, this is related to the fact that the vast majority of the experimentally available protein:ligand complexes are soluble in aqueous media. In contrast, Figure 2a provides a discrete breakdown in terms of the number of polar atoms (*NPOL*) of a small subset of drugs in the market, with very polar compounds in the red lines and compounds with only one polar atom ($NPOL = 1$) in the blue line. Non-oral drugs are also highlighted (indicated with a + symbol). Similar plots for the content of different databases (and libraries, see below) provide a visual overview of the different contents in terms of enzyme classes and distribution of compounds in chemical space and the range of affinity values. This is illustrated below using the WOMBAT database where the content is not limited by the availability of three-dimensional structures.

Figure 2a is also remarkable in that it shows a dearth of compounds in the lower quadrant of the diagram. This is related to the fact that in order to be a drug, a compound has to achieve a certain potency against the target of interest. The initial analysis suggested that there might be a soft boundary ($nBEI \sim 5.5$, $NSEI \sim 5$), indicated with a soft-dark gradient of yellow color, below which the probability of finding drugs drops dramatically. These trends are also confirmed when the content of the more extensive WOMBAT drug database (containing over 7000 entries) is examined[9,16] (Figure 2b). Our efforts to analyze quantitatively WOMBAT and other databases (BindingDB, ChemEBL) have just began and it is likely that probability estimates of 'drug-likeness' can be extracted as a function of the position in the optimization plane.

### 2.3 LEIs Applications to Poly-Pharmacology

The appearance of the efficiency planes defined by the different LEIs discussed above is also useful in exploring the ligand efficiency profile of an active pharmacological entity towards a multiplicity of targets. To illustrate this application we chose to map retrospectively the set of antipsychotic drugs studied recently by Mestres and Gregori-Puigjané.[17] The use of the combined LEIs indices in the efficiency planes defined above complements the analysis previously reported whereby only the size efficiency index BEI (see Table 1) was considered.

As indicated earlier, since each antipsychotic drug considered in the sample (Clozapine, Haloperidol and Risperidone) has different affinities for the different targets, the corresponding points in the Cartesian efficiency planes (i.e., SEI, BEI; or *NSEI*, *nBEI*; *x*,*y*) would map along straight lines defined by slopes given by 10(PSA/MW) and *NPOL*, respectively, for each chemical entity. The unique position of each target:drug pair along the lines would depend on the efficiency values of each ligand toward the specific target. Figure 3a illustrates the relative 'multiple efficiency' ranking of the three compounds toward the different targets in the SEI-BEI plane (slopes: Clozapine: 30.87; Haloperidol: 40.54; and Risperidone: 61.94). Similarly, Figure 3b illustrates the corresponding 'multiple-efficiency' ranking of the three compounds toward the various targets in the *NSEI-nBEI* plane (slopes equal to *NPOL* values of 3, 4 and 6 for Haloperidol, Clozapine and Risperidone, respectively). These plots permit a quantitative assessment of the multiple affinities of a various ligands toward the ten selected targets in two different directions. On one hand, the relative 'ligand specificity index' (LSI) of each compound towards the dif-



Analysis of Variance

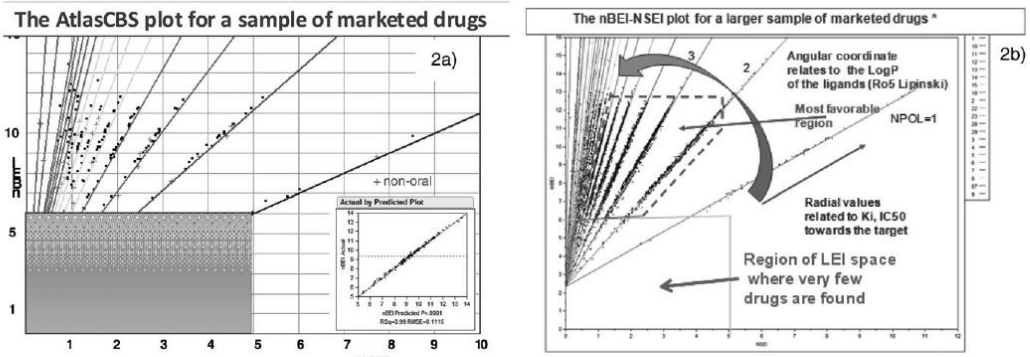
| Source | DF | Sum of Squares | Mean Square | F Ratio |
|---|---|---|---|---|
| Model | 53 | 6439.1525 | 121.493 | 6856.427 |
| Error | 1231 | 21.8129 | 0.018 | Prob > F |
| C. Total | 1284 | 6460.9654 | | 0.0000* |

Effect Tests

| Source | Nparm | DF | Sum of Squares | F Ratio | Prob > F |
|---|---|---|---|---|---|
| ON_number | 26 | 26 | 4186.8965 | 9087.907 | 0.0000* |
| NSEI | 1 | 1 | 36.9240 | 2083.790 | <.0001* |
| ON_number*NSEI | 26 | 26 | 2360.1317 | 5122.806 | 0.0000* |

**Figure 1.** Summary of the concepts of AtlasCBS. a) The content of PDBBind represented in the efficiency plane *NSEI-nBEI*. The content of the refined set (PDBBind2007) has been plotted using the LEIs variables indicated as defined in Table 1. The sample (1285) target:ligand pairs represents an evenly distributed set of complexes with a normal distribution of $K_i$ values and comprises data for approximately 40 targets. The panels below provide details of the statistical analysis as described in the literature,[5] Supplementary Material, Box S2; 1b) mapping of the set of compounds with activity towards HIV protease (2523 entries) from the ChEMBL database. An example of how to use the AtlasCBS concept to map compounds with activity towards a certain target. This representation illustrates the relationship between the angular coordinate in the plane to the physico-chemical properties of the ligands, in a continuous scale and to relate these values to the Lipinski Ro5 guidelines.[15]

© 2011 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim



*WOMBAT database (~7,000 entries) courtesy of T. Oprea

**Figure 2.** Comparing the content of different databases using AtlasCBS. a) Mapping of a small subset (198) of marketed drugs from ChEMBL in the *nBEI-NSEI* plane. The marketed drugs having *NPOL* (number of polar atoms) equal to 1 and above (up to 32) have been indicated. The + signs mark the position of non-orally administered drugs. Grey indicates a region of the plane where essentially no marketed drugs are found. See in the literature,[5] Supplementary Material, Box B for statistical analysis and further details. Inset shows the predicted vs. actual values of *nBEI* in the sample, as given by the model *NSEI*, *NPOL* and *NSEI*NPOL* cross term (highly significant $P < 0.0001$). The data were kindly provided by J. Overington and Bissan Al-Lazikani at the initial stages of the formation of the ChEMBL database. 2b) Mapping of the content of the WOMBAT (~7000 entries) drug database in the *nBEI-NSEI* plane. The grey box outlines the boundaries of the same box found in the much smaller subset obtained from HIV-I ligands (a). The marketed drugs having *NPOL* (number of polar atoms) equal to 1–3 have been indicated with *NPOL* ranging from 1 to 29 for the entire sample. The grey box indicates a region of the plane where much fewer marketed drugs are found. The dashed trapezoid outlines the boundaries within the plane where the majority of the marketed drugs are found. As indicated above, the arrow-headed ribbon indicates the increase of the polar coordinate (*NPOL*) related to the polarity of the ligands, and the blue arrowed line indicates the increasing direction of affinity towards the target. The WOMBAT data set contains approximately seven thousand entries and is the courtesy of T.Oprea.

ferent targets can be numerically estimated by the corresponding Cartesian distances of the points in the plane. On the other hand, the more efficient the compounds towards other targets, beside the one for which they were designed, the more likely that the compound would have a poly-pharmacological activity for the additional target. This representation and the underlying numerical values could provide a complementary quantitative estimate of conciliation between poly-pharmacological activity of the drugs.[17]

### 2.4 LEIs Applications in Fragment-Based Strategies

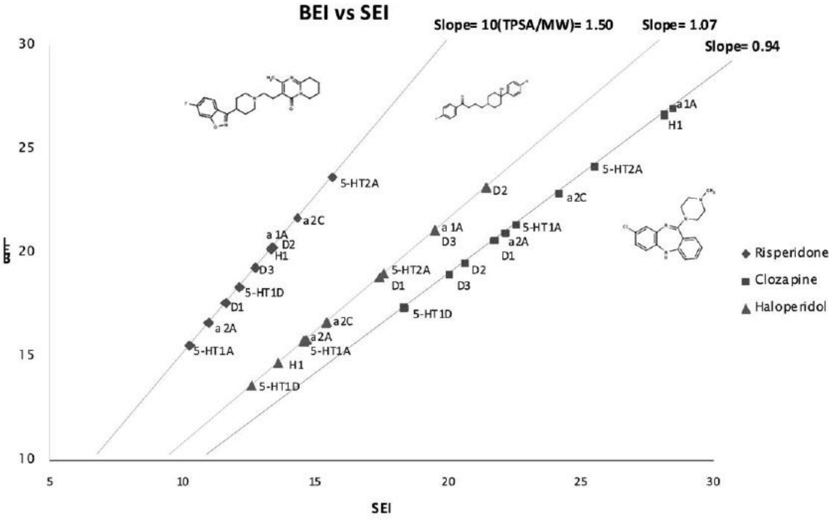
The initial introduction of the concept of efficiency by Hopkins and co-workers was for the purposes of lead selection and its subsequent optimization.[18] The extension of the concept for fragment optimization, fragment library design and fragment-based strategies soon followed.[19,20] Monitoring the size of fragments, fragment size and extracting optimal paths using concepts derived from efficiency criteria is widely accepted.[21,22] However, these analyses only consider the efficiency as it relates to size. We present here three examples where the addition of LEIs related to polarity (as in SEI and related indices) and the subsequent graphical representation of the data expands the analyses and provides additional insights into the use of LEIs to optimize Fragment-Based strategies.

The initial library design is an essential part of any Fragment-Based drug discovery effort. The limitations that the MW of the initial fragments puts on what is achievable has been discussed and illustrated recently.[19,20] However, it is necessary also to consider the polarity of the initial compounds and how they affect the course of the drug-discovery effort. We have explored this issue with a small (384 compounds) commercially available fragment collection (ActiveSight). As explained before, the position of the chemical fragments along lines on the LEI planes (BEI-SEI, *NBEI-NSEI* or *nBEI-NSEI*) is independent of the $K_i$'s toward specific targets. Thus, we used uniformly distributed values between 1 mM–1 μM as it might be obtained in any screening of a fragment library to graphically represent the chemical compounds contained in the library (Figure 4). The corresponding p$K_i$ values have been represented only to provide a crude scale of affinity. The image dramatically illustrates where the chemical entities that compose the library map and consequently what is possible in fragment-based drug discovery, given the chemical composition (size and polarity), of the fragments. It is immediately apparent that changing the slope of the lines (number of polar atoms, *NPOL*) is more effective in reaching the 'drug-like' regions of the map than changing the size (intercept: $\log_{10}$(NHEA)) of the fragments.

On the other hand, the concept of AtlasCBS allows flexibility in the choice of efficiency planes used to scrutinize

Efficiency Analisys: BEI vs SEI

3a)



Ligand Efficiency Indices: nBEI vs NSEI

3b)

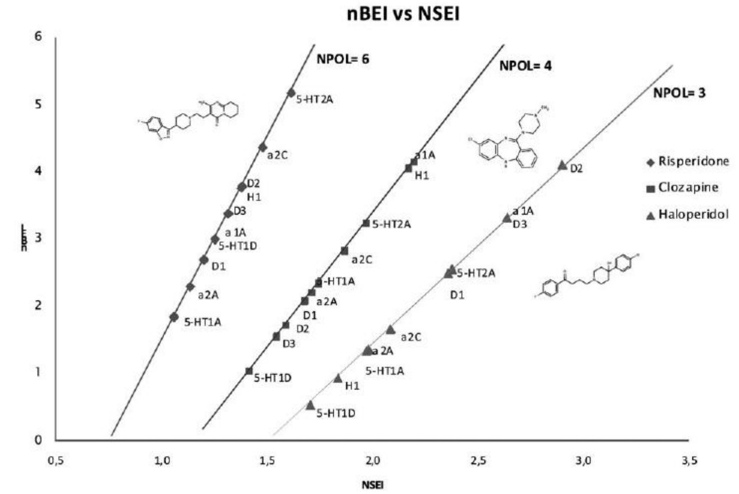
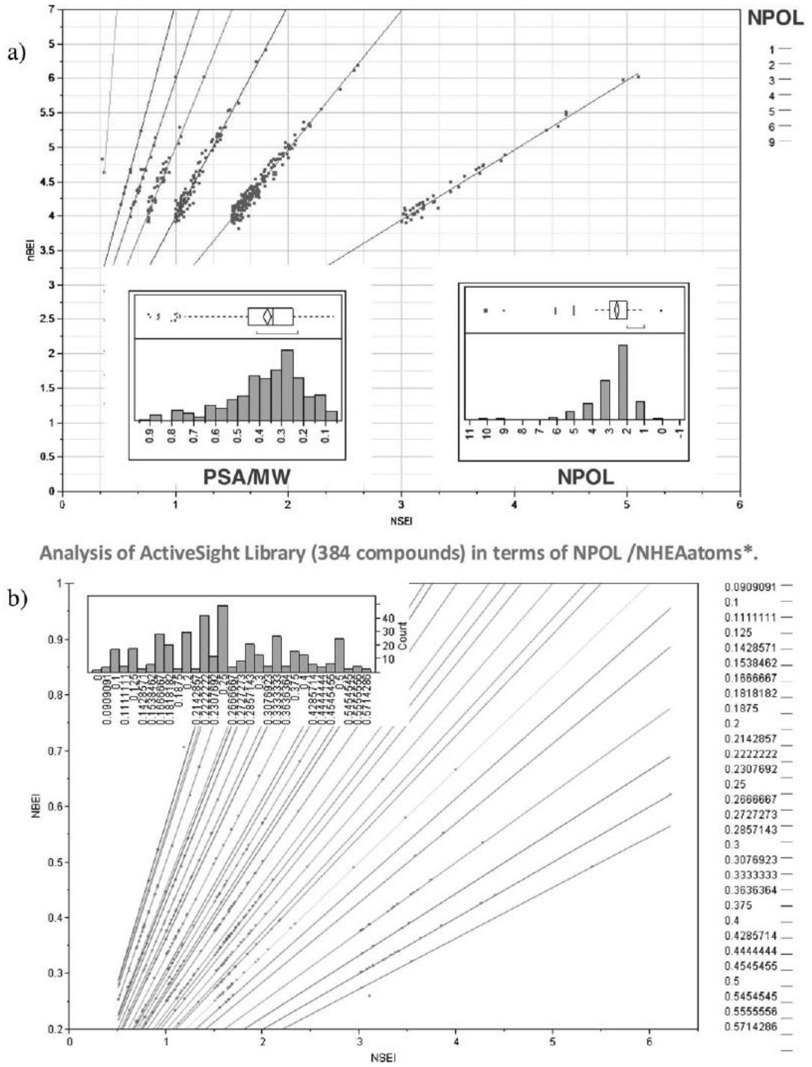


**Figure 3.** Using LEIs to analyze the concepts of poly-pharmacology. a) Multiple-efficiency analysis in the SEI, BEI plane. The efficiency plane ranks the compounds based on the efficiency towards the ten aminergic GPCRs target using size-related efficiency (BEI) and polarity-related efficiency (SEI). The slope of the lines corresponding to each compound is equal to 10(TPSA/MW) of the ligand. Each line corresponds to a compound as indicated in the right panel. The most efficiency compounds per size (*y*-axis) and per polarity (*x*-axis) for the different targets are furthest away from the origin. b) Multiple-efficiency analysis in the *NSEI*, *nBEI* plane. The efficiency plane ranks the compounds based on the efficiency towards the ten aminergic GPCRs target using size-related efficiency (*nBEI*) and polarity-related efficiency (*NSEI*) indices (See Table 1 for the variable definitions). The slope of the lines corresponds to the number of polar atoms (*NPOL* = count of N + O) of each compound, as indicated. Each line corresponds to a compound as indicated in the right panel. The most efficiency compounds per size (*y*-axis) and per polarity (*x*-axis) for the different targets are furthest away from the origin (slopes near one). Further explanations in the text. Note that the relative slopes of the three lines corresponding to the three compounds is not the same when based on *NPOL* vs. TPSA/MW.

© 2011 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim



Analysis of ActiveSight Library (384 compounds) in terms of NPOL /NHEAatoms*.

ActiSite Library in the NBEI vs. NSEI plane, Ki's random centered at 1 mM

**NOTE: slope of lines is always a rational number: NPOL/NHEA**

**Figure 4.** a. Fragment library representation in *nBEI-NSEI* plane. The content of the ActiveSight library (384 fragments) has been represented in the plane using uniform, randomly distributed $K_i$ values in the range between 1 m–1 μM for illustration purposes only. The chemical content of the fragment library is independent of the modeled $K_i$ values. The representation in the *nBEI-NSEI* plane shows the number of compounds containing a number of polar atoms (*NPOL*, slope of the lines) between 1 and 9 (see upper right panel for statistical modeling). This representation is equivalent to the right-hand histogram but adds the value of mapping the different affinities ($K_i$s) in the optimization plane. Thus the diagram illustrates what regions of the plane are accessible for the different affinities, efficiencies and chemical compositions (*NPOL*, NHEA). These regions can be compared with the areas of the map illustrated before for various targets and successful trajectories for specific targets. The insets show the distributions of PSA/MW and *NPOL* for the same library and have been included for comparison. b. The content of the same fragment library mapped into the efficiency plane defined by *NBEI-NSEI*. Notice the fine separation in the angular coordinate given by the ratio of *NPOL/NHEA*. The histogram illustrates the range, the size of the *NPOL/NHEA* bins and the frequency count. Although not illustrated in this diagram, it is possible to select a certain *NPOL* line from the library as illustrated in (a) (for example, *NPOL* = 2) and plot it in the variables of (b) for a detailed scrutiny of the library content.

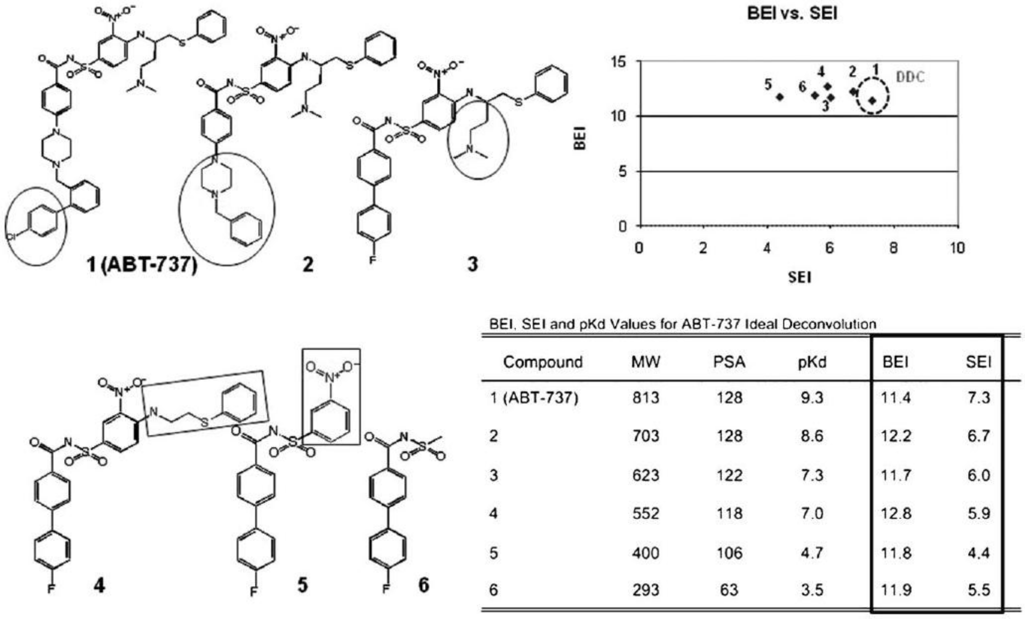
different issues in the drug-discovery process. Figure 4b presents the content of the same small library using the variables *NBEI*, *NSEI* (See Table 1). In this representation the chemical content of the library is finely sampled in terms of the *NPOL*/NHEA ratio and can be statistically modeled as shown. The histogram shows the range and sampling of the chemical matter present in the library in terms of polarity as it relates to size (i.e., *NPOL*/NHEA ranging from 1/11 to 4/7).

Selection of the efficiency of the initial fragments and their number of polar atoms, limits the regions of the efficiency plane that can be explored within chemical series. It should be mentioned that it in the near future, it is conceivable that theoretical estimates of $K_i$'s as obtained by computational methods[23] could provide more realistic estimation of the position of the fragments (ligands) in the plane and the possibilities that each fragment (ligand) offers for future optimization.

Extensive public domain data on successful fragment-based strategies and paths using LEIs criteria are not available. However, an intriguing concept has been put forward by Hajduk[20] suggesting that in an idealized (deconstructed *a posteriori*) path, the efficiency of the successive compounds should remain approximately constant as adding more atoms should increase the potency proportionally. Our analysis of the same data, combining the use of BEI and SEI, suggest that in the idealized pathway suggested by the retrospective trajectory indicated by Hajduk, while BEI remains indeed constant as suggested, it is SEI that is optimized in the process from 5.3 to 7.2 (Figure 5). Thus, the combination of LEIs may have more predictive value than either of them alone.

More recently, the concept of BEI has been used to dissect the potency of the natural product Argifin, a cyclic pentapeptide inhibitor of chitinase B1 from the fungus A. *Aspergillus fumigatus* (AfChiB1).[24] The authors show that the dimethylguanylurea fragment, representing only a quarter of the natural product in mass, contains all the critical interactions and binds to the target with very high efficiency.

Our analysis of the data (Figure 6) reveals that the smallest fragment optimizes both BEI and SEI. This dissection would suggest that the additional chemical structure of the intact cyclopentapeptide is necessary for specificity and possibly also for transport into the cell, since the guanylurea moiety is very polar.

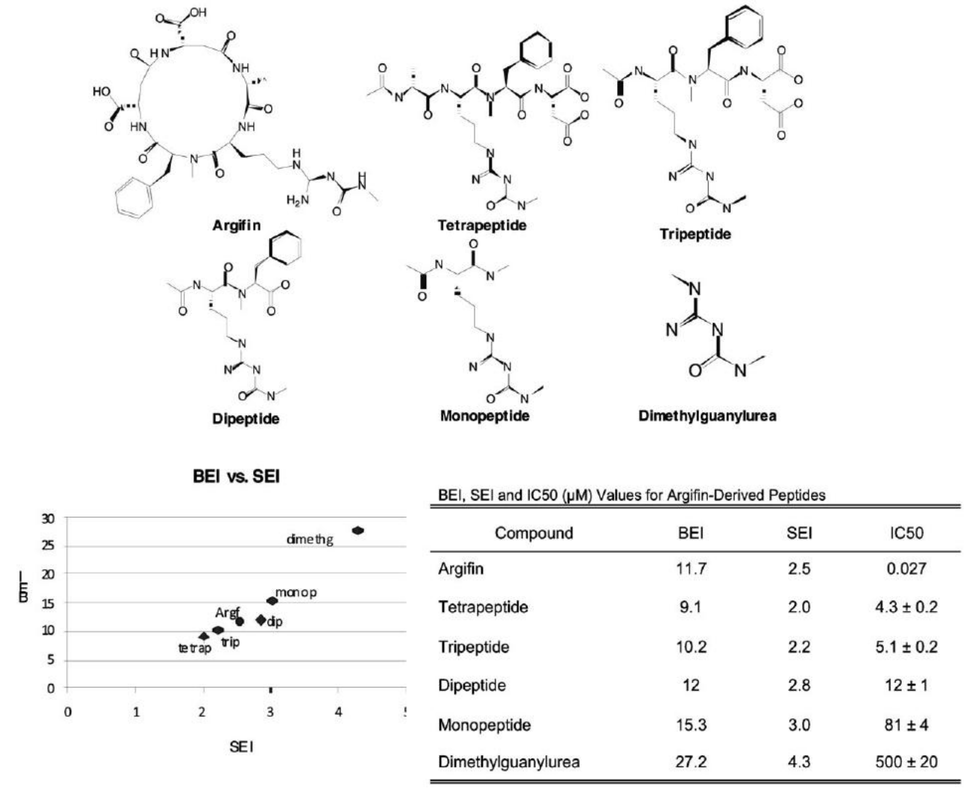


BEI, SEI and pKd Values for ABT-737 Ideal Deconvolution

| Compound | MW | PSA | pKd | BEI | SEI |
|---|---|---|---|---|---|
| 1 (ABT-737) | 813 | 128 | 9.3 | 11.4 | 7.3 |
| 2 | 703 | 128 | 8.6 | 12.2 | 6.7 |
| 3 | 623 | 122 | 7.3 | 11.7 | 6.0 |
| 4 | 552 | 118 | 7.0 | 12.8 | 5.9 |
| 5 | 400 | 106 | 4.7 | 11.8 | 4.4 |
| 6 | 293 | 63 | 3.5 | 11.9 | 5.5 |

*Adapted from Hajduk. P. How big is too big? J. Med. Chem. 2006. 49:6972-6976

**Figure 5.**
Ideal fragment reconstruction of a Drug Discovery Candidate (DCC). In the analysis presented by Hajduk, BEI remains approximately constant through the optimization process while SEI is optimized from *SEI* = 5.5 (compound 6) to *SEI* = 7.3 (1) for the Drug Discovery Candidate ABT-737. Adapted from Hajduk.[20]



BEI, SEI and IC50 (μM) Values for Argifin-Derived Peptides

| Compound | BEI | SEI | IC50 |
|---|---|---|---|
| Argifin | 11.7 | 2.5 | 0.027 |
| Tetrapeptide | 9.1 | 2.0 | 4.3 ± 0.2 |
| Tripeptide | 10.2 | 2.2 | 5.1 ± 0.2 |
| Dipeptide | 12 | 2.8 | 12 ± 1 |
| Monopeptide | 15.3 | 3.0 | 81 ± 4 |
| Dimethylguanylurea | 27.2 | 4.3 | 500 ± 20 |

**Figure 6.** Efficiency dissection of a natural pentapeptide. The *SEI–BEI* plot show how the smallest dimethylsanylurea fragment (last row of data table) of the natural product Argifin (first row) is the most efficient chemical entity of the series in both *BEI* and *SEI*. Adapted from Andersen et al.[24]

## 3 Outlook: More than an Efficiency Yardstick

Although proposed originally as a criterion to select lead compounds,[18] LEIs are becoming more accepted in the wider community of medicinal chemists. This is true of efficiency variables relating potency to the size of the ligand. Possibly, due to their less precise definition, efficiency indices relating potency to polarity (i.e., PSA or number of polar atoms: *NPOL*, or other descriptors of polarity) are still not widely used. In previous work and in this brief review, we suggest that the combination of size-related and polarity-related LEIs can be used as guides in different areas of drug-discovery and could play an important role for the medicinal chemistry community, as they can be regarded as much more than just "efficiency yardsticks". The concept of an atlas-like representation of the combined Chemico-Biological Space (AtlasCBS) has been introduced recently as a framework for the introduction of a graphical and algebraic framework that could put drug-discovery in a more numerically sound path. Only additional usage, exploring and testing by the drug-discovery community in prospective applications can provide the data and experience required to prove or disprove this suggestion.

## Acknowledgements

Access to the data in WOMBAT granted by Prof. *T. Oprea* is fully appreciated. The data for the initial dataset of marketed drugs from ChEMBL were kindly provided by *J. Overington* and *Bissan Al-Lazikani* at the initial stages of the formation of the ChEMBL database. Productive conversations and the assistance of Dr. *Patricia Anna Bento* in extracting the HIV activity data from the ChEMBL are greatly appreciated. The participation of Dr. *O. Perisic* (currently at USI, Lugano, Switzerland) in the early stages of this work is appreciated. The expertise and participation of *John Wass* in the statistical analysis of the initial data sets is greatly appreciated.

The financial support of the *AGAUR Foundation* to C. A.-Z. for a Visiting Professorship at the Platforms of Drug Discovery (Dr. *J. Quintana*) and Plataforma Automatitzada de Cristal·lografia (Dr. *A. Guasch*) is greatly appreciated as well as the hospitality of the *Parc Cientfic Barcelona*. The support of Prof. *M. Johnson* director of the Center for Pharmaceutical Biotechnology at the University of Illinois at Chicago (UIC) is greatly appreciated.

## References

[1] T. I. Oprea, J. Gottfries, *J. Comb. Chem.* **2001**, *3*, 157–166.
[2] T. Olsson, T. I. Oprea, *Curr. Opin. Drug. Discov. Devel.* **2001**, *4*, 308–313.
[3] T. I. Oprea, *Curr. Opin. Chem. Biol.* **2002**, *6*, 384–389.
[4] C. Lipinski, A. Hopkins, *Nature* **2004**, *432*, 855–861.
[5] C. Abad-Zapatero, O. Perisic, J. Wass, A. P. Bento, J. Overington, B. Al-Lazikani, M. E. Johnson, *Drug Discov. Today* **2010**, *15*, 804–811.
[6] R. Wang, X. Fang, Y. Lu, C. Y. Yang, S. Wang, *J. Med. Chem.* **2005**, *48*, 4111–4119.
[7] R. Wang, X. Fang, Y. Lu, S. Wang, *J. Med. Chem.* **2004**, *47*, 2977–2980.
[8] T. Liu, Y. Lin, X. Wen, R. N. Jorissen, M. Gilson, *Nucleic Acids Res.* **2006**, *35*, D198–D201.
[9] T. I. Oprea, T. K. Allu, D. C. Fara, R. F. Rad, L. Ostopovici, C. G. Bologa, *J. Comput. Aided Mol. Des.* **2007**, *21*, 113–119.
[10] J. P. Overington, B. Al-Lazikani, A. L. Hopkins, *Nat. Rev. Drug Discov.* **2006**, *5*, 993–996.
[11] E. Perola, *J. Med. Chem.* **2010**, *53*, 2986–2997.
[12] C. Abad-Zapatero, J. M. Metz, *Drug Discov. Today* **2005**, *10*, 464–469.
[13] C. Abad-Zapatero, *Expert Opin. Drug Discov.* **2007**, *2*, 469–488.
[14] J. Overington, *J. Comput. Aided Mol. Des.* **2009**, *23*, 195–198.
[15] C. A. Lipinski, *J. Pharmacol. Toxicol. Meth.* **2000**, *44*, 235–249.
[16] M. Olah, C. Bologa, T. I. Oprea, *J. Comput. Aided Mol. Des.* **2004**, *18*, 437–449.
[17] J. Mestres, E. Gregori-Puigjané, *Trends Pharmacol. Sci.* **2009**, *30*, 470–474.
[18] A. L. Hopkins, C. R. Groom, A. Alex, *Drug Discov. Today* **2004**, *9*, 430–431.
[19] P. J. Hajduk, J. Greer, *Nat. Rev. Drug Discov.* **2007**, *6*, 211–219.
[20] P. J. Hajduk, *J. Med. Chem.* **2006**, *49*, 6972–6976.
[21] M. Congreve, R. Carr, C. Murray, H. Jhoti, *Drug Discov. Today* **2003**, *8*, 876–877.
[22] M. Congreve, G. Chessari, D. Tisi, A. J. Woodhead, *J. Med. Chem.* **2008**, *51* (13), 3661–3680.
[23] M. Gilson, H. X. Zhou, *Ann. Rev. Biophysics Biolmol. Struct.* **2007**, *36*, 21–42.
[24] O. A. Andersen, A. Nathubhai, M. J. Dixon, I. M. Eggleston, D. M. van Aalten, *Chem. Biol.* **2008**, *15*, 295–301.

Received: November 7, 2010
Accepted: February 23, 2011
Published online: March 17, 2011

© 2011 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

## TTR binding pocket properties & Docking studies

With the aim to perform a prospective drug design study based on the LEIs methodology described above, a computational analysis of the TTR binding pocket was carried out, in order to find hints about the most favorable binding mode for TTR ligands, as well as which pocket regions in the TTR binding site are the key points for the interaction, and, if possible, to establish structure-activity relationships of structurally related series of TTR ligands.

In the following published scientific paper[6], 13 TTR crystal structures were analyzed in order to get useful information about: a) the flexibility of the TTR binding pocket; b) the binding mode preference (forward or reverse binding) for TTR ligands in different conditions and/or for structurally-related compounds families; and c) the relationship between the binding mode and the amyloidogenesis inhibition activity.

As a result of this study an important conclusion was obtained: the most active compounds in TTR amyloidogenesis inhibition activity, as measured in the inhibition of fibrils formation assay[7], are characterized by the presence of at least one halogen atom in one of the six (three symmetry-related pairs) small hydrophobic depressions known as halogen binding pockets (HBP) present in the TTR binding site[8].

Amyloid Downloaded from informahealthcare.com by Joaquin Ibanez Esteb on 07/05/11
For personal use only.

## Ligand-binding properties of human transthyretin

M. Pinto[1], D. Blasi[2], J. Nieto[3], G. Arsequell[4], G. Valencia[4], A. Planas[3], J. Quintana[2], & N. B. Centeno[1]

[1]*Computer-Assisted Drug Design Laboratory, Research Unit on Biomedical Informatics (GRIB), IMIM/UPF, Barcelona, Spain,* [2]*Drug Discovery Platform (PDD), Parc Científic de Barcelona, Barcelona, Spain,* [3]*Laboratory of Biochemistry, Institut Químic de Sarrià, Universitat Ramon Llull, Barcelona, Spain, and* [4]*Unit of Glycoconjugate Chemistry, Institut de Química Avançada de Catalunya, IQAC-CSIC, Barcelona, Spain*

**Abstract:** A computational analysis was performed on a selected group of 13 TTR-ligand crystallographic complexes in order to deduce information useful for drug design and discovery. The results obtained can be summarized as follows: (1) the binding site of TTR is a large and very flexible cavity, which is composed of three regions with different chemical features; (2) ligands bind to TTR in forward or reverse modes depending on the conformation adopted by the serine and threonine residues located at the end of the cavity; (3) no relationship could be found between the binding mode of the ligands and their TTR fibrillogenesis inhibitory activity; (4) regardless of the structure, chemical properties or binding mode of the ligand to TTR, there is always a contribution of residues Lys15, Leu17, Ala108, Leu110, Ser117 and Thr119 to ligand binding and finally, (5) the most active compounds are characterised by the presence of at least one halogen atom in the HBP1/HBP1′ or HBP3/HBP3′ pockets.

**Introduction:** Transthyretin (TTR) is one of the most studied amyloidogenic proteins. It is thought that the aggregation pathway of TTR into amyloid fibrils occurs by tetramer dissociation, a process which may be blocked by stabilization of the tetramer via ligand binding. In the last years, several ligands of structurally non-related families have been evaluated for their ability to bind and stabilize the native structure of TTR *in vitro*.

Today, more than 25 compounds have been co-crystallized with the protein, being their crystallographic structures available in the Protein Data Bank. However, in spite of this amount of structural data, no attempt to critically review the relevant information for ligand binding that is embedded on them has been published. To fill this gap, 13 crystallographic complexes were computationally analyzed in order to extract common features and correlate them with the fibrillogenesis inhibitory potency of the ligands. The choice of complexes was made in order to cover the maximum range of structural and functional diversity of the ligands (Figure 1).

**Methods**: ***Preparation of the crystallographic complexes:*** Crystal coordinates of selected TTR-ligand complexes were directly extracted from the Protein Data Bank. Hydrogen atoms were added to the X-ray structures by using the EDIT module from the AMBER v.6 package [1].

***Hydrogen atoms refinement:*** Added hydrogen atoms were energy-minimized by using the parm94 [2] force field implemented in AMBER. Ligand partial charges were obtained by computing the electrostatic-potentials around the optimized structures at the RH/6-31G* level using Gaussian 94 [3] and then fitting the charges using the RESP method [4]. Minimization was carried out by using a distance dependent dielectric constant of *1r* to screen for electrostatic interactions and a cutoff distance of 12 Å for van der Waals interactions. Hydrogen atoms refinement was accomplished using 5000 cycles of steepest descents followed by conjugate gradient until the maximum gradient of the AMBER energy was smaller than 0.001 kcal/mol Å [2].

***Energy decomposition:*** Molecular mechanics interaction energy between TTR and the ligand was decomposed on a residue basis by using the module ANAL of AMBER.

**Results and discussion:** It is well known that TTR is a homotetrameric protein. These four identical monomers assemble to form the functional tetramer, with a central channel at the dimer-dimer interface which contains the two chemically equivalent hormone binding sites, named A-A' and B-B' respectively.

Our modelling studies on these 13 structures have corroborated that each binding site is characterized by the presence of three symmetry-related depressions termed halogen binding pockets (HBPs) and located between two neighbouring $\beta$ strands. The innermost pockets (HBP3 and HBP3'), located at 6 Å from the centre of the tetramer, are formed by the side chains and backbone atoms of Ala108, Ala109, Leu110, Ser117, and Thr119 of the monomers A and A' respectively. HBP2 and HBP2' are defined by residues Ala108, Ala109 and Leu110 of one of the two neighbouring $\beta$ strands and Lys15 and Leu17 of the other one. The outermost pockets, termed HBP1 and HBP1', located about 13 Å from the tetramer centre, are formed by the side chains of Thr106, Ala108, Met 13 and Lys15 [5,6].

We have also confirmed that the side chains projected into each binding site define three regions with different chemical features. The first one, located at the entrance, is a highly charged region formed by the side chains of residues Lys15, Glu54, and His56 of both monomers A and A'. The second region, located at the centre of the binding site, is

Amyloid Downloaded from informahealthcare.com by Joaquin Ibanez Esteb on 07/05/11
For personal use only.

predominantly hydrophobic and is formed by the paired residues Leu17, Thr106, Ala108, Leu110, and Val121. Finally the last one, located near the centre of the tetramer, is a hydrophilic patch derived from the side chains of residues Ser112, Ser115, Ser117, and Thr119 and associated water molecules.

Our studies further support the fact that the binding site of transthyretin is also very flexible. This conformational flexibility is experimentally supported by the (1) presence in the crystallographic complexes of residues able to adopt multiple conformations in the same complex (for example, residues Leu17, Ser117 and Thr119 in the TTR-DIF complex); (2) the promiscuity in ligand binding (see below), and (3) the binding of ligands to different conformational states of TTR, each one of them characterized by different side-chain conformations. This conformational flexibility may provide an explanation for the ability of TTR to bind a broad diversity of chemical structures (Figure 1).

The analysis of these 13 crystallographic TTR-ligand complexes provides further evidences about the promiscuous ligand binding-properties of transthyretin. This TTR promiscuity is highlighted by (1) the different binding modes of structurally related ligands; (2) its ability to bind one single ligand molecule in more than one orientation and finally, by (3) the different conformations that can adopt a ligand in the TTR binding site.

Ligands bind to TTR in either forward or reverse modes. In the first one, the carboxylic group of the ligand is positioned at the entry of the cavity, forming electrostatic interactions with the residues Lys15 and Glu54 from one or two monomers. In the reverse binding mode, the carboxylic group of the ligand is positioned at the end of the binding site, favouring the formation of a hydrogen bond with the residues Ser117 and Thr119 of this region. Ligands like T4, FHI or REA bind to TTR in forward mode whereas others such as FBC, 3CA or DIF bind in reverse mode.

The existence of both binding modes may be attributed to the high flexibility of residues Ser117 and Thr119, which in turn depends on the water molecule positioned at HBP3/3'. In the apo-form, both pockets are occupied by a water molecule which acts as a bridge between the hydroxyl group of residues Ser117 and Thr119. The placement of a ligand substituent in these pockets displaces this initial water molecule, leading to a rotation of their side chains. The conformation adopted by residues Ser117 and Thr119 upon ligand binding depends on the nature of the ligand substituent positioned at this region and has two major consequences: on the one hand, it allows the formation of additional intermonomer interactions

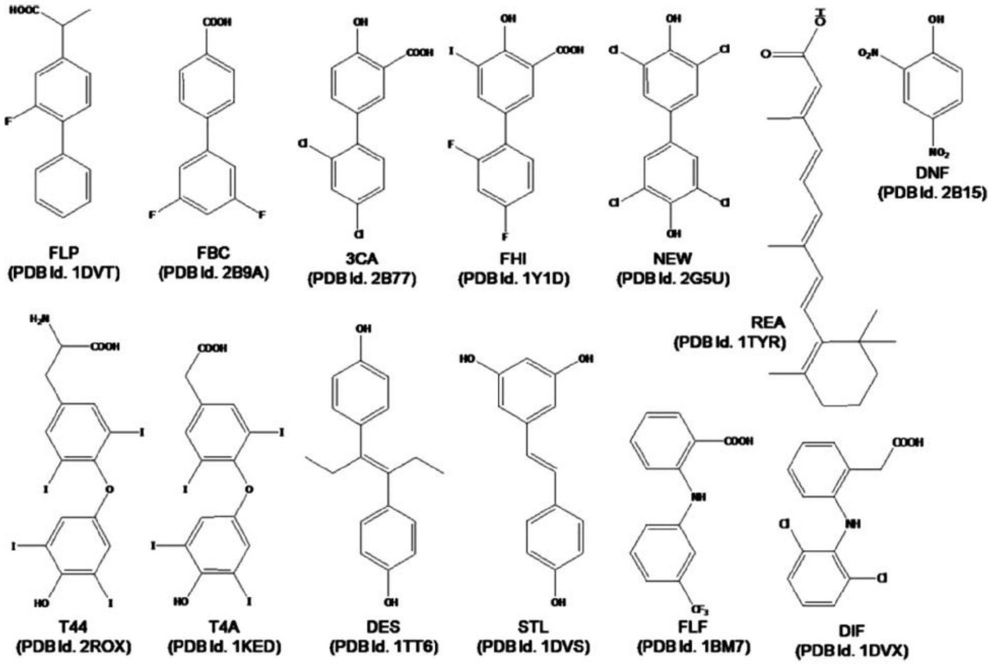
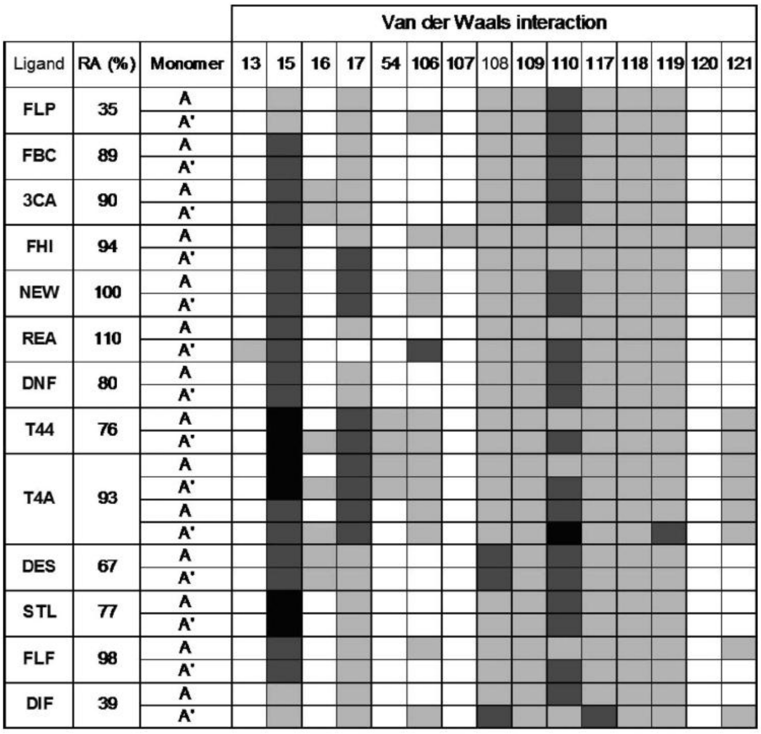


Figure 1. Structure of the crystallographic ligands studied in the present work.

Table 1. Hydrogen bond and van der Waals contributions of individual residues of the monomers A and A′ of TTR to ligand binding. Protein residues that make van der Waals contacts with the ligand are represented in different shades of grey, with the colour intensity directly related to the strength of the interaction.



| | | | Van der Waals interaction | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Ligand | RA (%) | Monomer | 13 | 15 | 16 | 17 | 54 | 106 | 107 | 108 | 109 | 110 | 117 | 118 | 119 | 120 | 121 |
| FLP | 35 | A | | ░ | | ░ | | | | ░ | ░ | ▓ | ░ | ░ | ░ | | |
| | | A' | | ░ | | ░ | | ░ | | ░ | ░ | ▓ | ░ | ░ | ░ | | |
| FBC | 89 | A | | ▓ | | ░ | | | | ░ | ░ | ▓ | ░ | ░ | ░ | | |
| | | A' | | ▓ | | ░ | | | | ░ | ░ | ▓ | ░ | ░ | ░ | | |
| 3CA | 90 | A | | ▓ | ░ | ░ | | | | ░ | ░ | ▓ | ░ | ░ | ░ | | |
| | | A' | | ▓ | ░ | ░ | | | | ░ | ░ | ▓ | ░ | ░ | ░ | | |
| FHI | 94 | A | | ▓ | | ░ | | ░ | ░ | ░ | ░ | ░ | ░ | ░ | ░ | ░ | ░ |
| | | A' | | ▓ | | ▓ | | | | ░ | ░ | ░ | ░ | ░ | ░ | | |
| NEW | 100 | A | | ▓ | | ▓ | | ░ | | ░ | ░ | ▓ | ░ | ░ | ░ | | ░ |
| | | A' | | ▓ | | ▓ | | ░ | | ░ | ░ | ▓ | ░ | ░ | ░ | | ░ |
| REA | 110 | A | | ▓ | | ░ | | | | ░ | ░ | ░ | ░ | ░ | ░ | | |
| | | A' | ░ | ▓ | | | | ▓ | | ░ | ░ | ▓ | ░ | ░ | ░ | | |
| DNF | 80 | A | | ▓ | | ░ | | | | ░ | ░ | ▓ | ░ | ░ | ░ | | |
| | | A' | | ▓ | | ░ | | | | ░ | ░ | ▓ | ░ | ░ | ░ | | |
| T44 | 76 | A | | █ | | ▓ | ░ | ░ | | ░ | ░ | ░ | ░ | ░ | ░ | | ░ |
| | | A' | | █ | ░ | ▓ | ░ | ░ | | ░ | ░ | ▓ | ░ | ░ | ░ | | ░ |
| T4A | 93 | A | | █ | | ▓ | ░ | ░ | | ░ | ░ | ░ | ░ | ░ | ░ | | ░ |
| | | A' | | █ | ░ | ▓ | ░ | ░ | | ░ | ░ | ▓ | ░ | ░ | ░ | | ░ |
| | | A | | ▓ | | ▓ | | ░ | | ░ | ░ | ▓ | ░ | ░ | ░ | | ░ |
| | | A' | | ▓ | ░ | ▓ | | ░ | | ░ | ░ | █ | ░ | ░ | ▓ | | ░ |
| DES | 67 | A | | ▓ | ░ | ░ | | | | ▓ | ░ | ▓ | ░ | ░ | ░ | | |
| | | A' | | ▓ | ░ | ░ | | | | ▓ | ░ | ▓ | ░ | ░ | ░ | | |
| STL | 77 | A | | █ | | ░ | | | | ░ | ░ | ▓ | ░ | ░ | ░ | | |
| | | A' | | █ | | ░ | | | | ░ | ░ | ▓ | ░ | ░ | ░ | | |
| FLF | 98 | A | | ▓ | | ░ | | ░ | | ░ | ░ | ░ | ░ | ░ | ░ | | ░ |
| | | A' | | ▓ | | ░ | | | | ░ | ░ | ▓ | ░ | ░ | ░ | | |
| DIF | 39 | A | | ░ | | ░ | | | | ░ | ░ | ▓ | ░ | ░ | ░ | | |
| | | A' | | ░ | | ░ | | ░ | | ▓ | ░ | ░ | ▓ | ░ | ░ | | ░ |

Amyloid Downloaded from informahealthcare.com by Joaquin Ibanez Esteb on 07/05/11
For personal use only.

which may contribute to stabilize the TTR tetramer; on the other hand, the ligand-induced conformational changes of residues Ser117 and Thr119 modify the polarity surface of the HBP3/3' pockets leading to the forward or reverse binding modes.

Our study of residue contribution has shown that regardless of the structure, chemical properties or binding mode of the ligand to TTR, there is always a contribution of residues Lys15, Leu17, Ala108, Leu110, Ser117 and Thr119 of monomers A and A' to the van der Waals interaction energy between TTR and the ligand (Table I). The estimated interaction values for Ala108, Ser117 and Thr119 are generally lower and remain constant across the series. However, the involvement of the latter residues in the hydrogen bond network of TTR-ligand complexes may enhance their contribution to the complex stability. In contrast to these residues, the contribution of residue Lys15 to ligand binding is highly variable, although for the most active ligands there is a large contribution of these residues to the binding (Table I). The conformation adopted by the Lys15 residue upon ligand binding allows to maximize the hydrophobic interaction of the side chain of Lys15 with the phenyl ring of the ligand positioned at the entry of the cavity.

In summary, our computational analysis of 13 TTR-ligand complexes could not establish a relationship between the binding mode of the ligands and their inhibitory potency. However, the most active compounds are characterized by the presence of halogen atoms interacting with the HBP1/HBP1' or HBP3/3' pockets [6].

***Declaration of interest:*** The present work has been supported by the Fundació Marató TV3 (2000/02).

## References

1. Case DA, Pearlman DA, Caldwell JW, Cheatham TE, III, Ross WS, Simmerling CL, Darden TA, Merz KM, Stanton RV, Cheng AL, et al. AMBER 6, University of California: San Francisco; 1999.

2. Cornell W, Cieplak P, Bayley CI, Gould I, Merz KM, Ferguson DM, Spellmeyer DC, Fox T, Caldwell JW, Kollman PA. A second generation force field for the simulation of proteins and nucleic acids. J Am Chem Soc 1995;117:5179–5197.
3. Frisch MJ, Trucks GW, Schlegel HB, Gill PMW, Johnson BG, Robb MA, Cheeseman JR, Keith T, Petersson GA, Montgomery JA, et al. Gaussian 94, revision E.1. Gaussian, Inc.: Pittsburgh, PA; 1996.
4. Bayly CI, Cieplak P, Cornell W, Kollman PA. A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model. J Phys Chem 1993;97:10269–10280.
5. Blake CCF, Geisow MJ, Oatley SJ. Structure of prealbumin: secondary, tertiary and quaternary interactions determined by Fourier refinement at 1.8 A. J Mol Biol 1978;121:339–356.
6. Mairal T, Nieto J, Pinto M, Almeida MR, Gales L, Ballesteros A, Barluenga J, Pérez JJ, Vázquez JT, Centeno NB, et al. Iodine atoms: a new molecular feature for the design of potent transthyretin fibrillogenesis inhibitors. PLoS One 2009;4:e4124.

Amyloid Downloaded from informahealthcare.com by Joaquin Ibanez Esteb on 07/05/11
For personal use only.

As a follow-up of the scientific paper published above, in order to check the feasibility of generating a LEI prospective map based on TTR ligands docking results, a computational study was performed in order to establish a methodology and determine its capability to reproduce crystal structure poses through TTR ligand docking.

5 crystallographic structures of complexes of ligands with wild-type (wt)-TTR were selected from the Protein Data Bank (PDB, www.rcsb.org) for this study (**Table 1**). They were selected taking into account two factors: high resolution of the crystal structure (below 1,9 Å), and pH between 5.0 and 7,0 (to cover a wide range of pH values because it is known that the ligand TTR binding mode (forward/reverse) is highly dependent on the pH value).

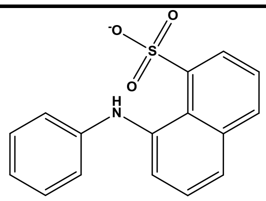
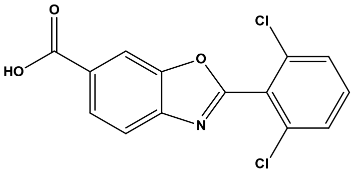
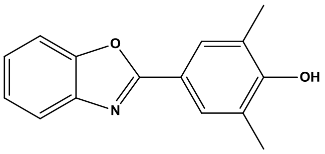
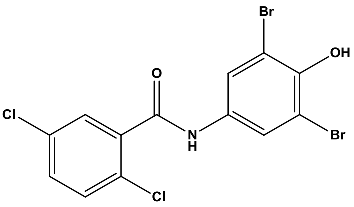
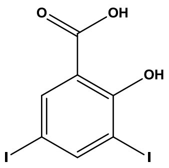
| PDB ID | LIGAND | pH | Resolution (Å) |
|---|---|---|---|
| **3CFN**[9] |  | 7,5 | 1,87 |
| **2F8I**[10] |  | 7,0 | 1,54 |
| **2QGC**[11] |  | 5,5 | 1,30 |
| **3ESO**[12] |  | 5,5 | 1,31 |
| **3B56**[13] |  | 5,0 | 1,55 |

**Table 1.** Crystal structures selected for the docking analysis.

Cross-docking experiments were carried out using these 5 TTR protein structural templates against 48 TTR ligands for which TTR-ligand complexes are available at the

PDB (www.rcsb.org). The aim of the study was to evaluate if the software package used (MOE 2009.10)[14] was able to reproduce the crystal poses for all the ligands selected.

The docking was performed using as ligand placement method the alpha triangle (AT) methodology (in the software MOE), in which the poses are generated by superposition of ligand atom triplets and triplets of alpha sphere centers in the receptor site. At each iteration, a random conformation is selected, and a random triplet of ligand atoms and a random triplet of alpha sphere centers are used to determine the pose. Several scoring functions in the MOE software were tested, and Alpha HB (AHB) gave the best results; in the Alpha HB scoring function, binding free energy is estimated by the sum of the geometric fit and a hydrogen bond term.

An overview of the cross-docking experiments is displayed on **Figure 1**.

The bar plots obtained represent the RMSD obtained from the comparison between the crystallographic and calculated poses for each of the 48 compounds analyzed.

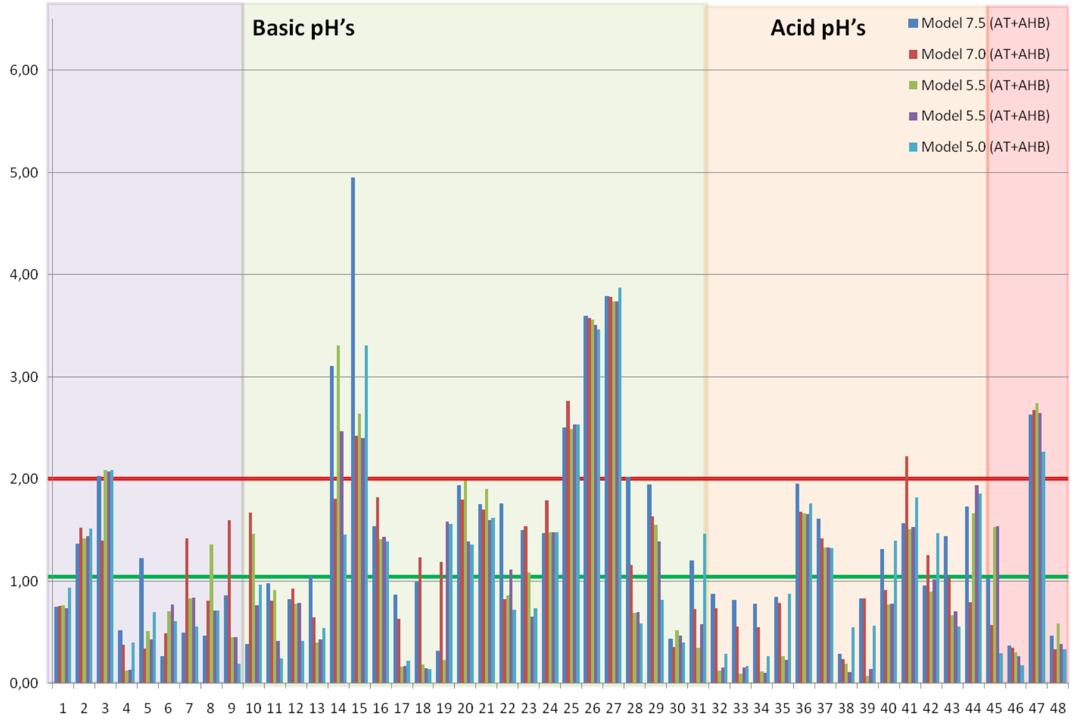
In terms of pH, the results show that the computational system for docking does not show any dependency with this parameter, since it is able to reproduce properly poses in a wide range of pH, with any of the template models available in the cross-docking experiments.

Flexible ligands, with long aliphatic chains, are the only compounds not correctly simulated in the cross-docking experiments, giving high RMSD values with respect to the corresponding crystallographic poses. In the docking protocol performed, the flexibility of the binding site residues was not considered, and this may be the reason why 6 out of the 48 ligands studied (the 6 larger and more flexible compounds) presented RMSD above 2 Å.

Therefore, the following conclusions were obtained from this cross-docking analysis:

1. The system is pH-independent.
2. The docking system is scaffold dependent; large ligands (bivalent, aliphatic chains or peptidic) are difficult to reproduce correctly.
3. 91% of the crystallographic poses are correctly reproduced using the MOE software package.

In summary, looking at these results, it is possible to perform reliable TTR ligand docking studies that give us good predictions useful for the planned LEI prospective strategy, which is described in the following published scientific paper[3], and in the subsequent text on this thesis chapter.



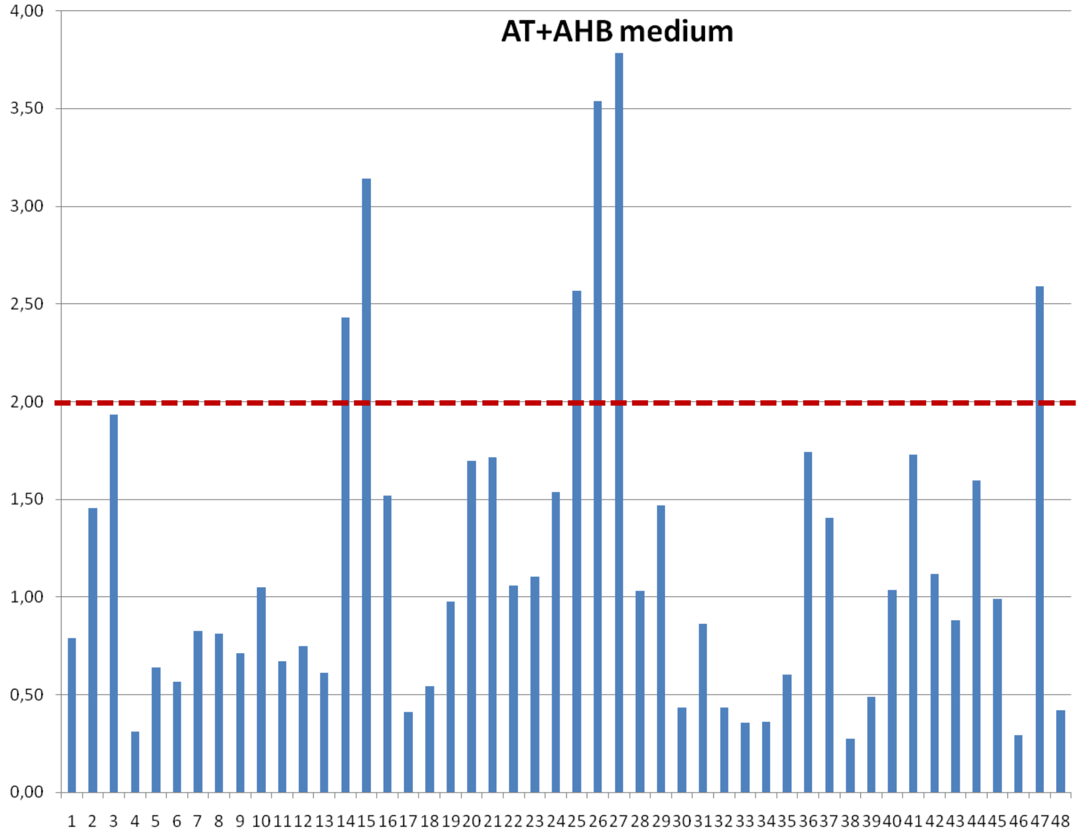


**Figure 1.** Bar plots show the RMSD (root mean square deviation) between the crystallographic and the calculated pose. Only six ligands are not correctly estimated. AT: alpha triangle; AHB: alpha HB, methodologies in MOE software.

molecular informatics

DOI: 10.1002/minf.201000157

# Retrospective Mapping of SAR Data for TTR Protein in Chemico-Biological Space Using Ligand Efficiency Indices as a Guide to Drug Discovery Strategies

Daniel Blasi,*[a] Gemma Arsequell,[b] Gregori Valencia,[b] Joan Nieto,[c] Antoni Planas,[c] Marta Pinto,[d] Nuria B. Centeno,[d] Cele Abad-Zapatero,[a, e] and Jordi Quintana[a]

*Presented at the 18th European Symposium on Quantitative Structure Activity Relationships, EuroQSAR 2010, Rhodes, Greece*

**Abstract**: We have previously reported the design and synthesis of ligands that stabilize Transthyretin protein (TTR) in order to obtain therapeutically active compounds for Familial Amyloid Polyneuropathy (FAP). We are hereby reporting a drug design strategy to optimize these ligands and map them in Chemico-Biological Space (CBS) using Ligand Efficiency Indices (*LEI*s). We use a binding efficiency index (*BEI*) based on the measured binding affinity related to the molecular weight (*MW*) of the compound combined with surface-binding efficiency index (SEI) based on Polar Surface Area (PSA). We will illustrate the use of these indices, combining three crucial variables (potency, MW and PSA) in a 2D graphical representation of chemical space, to perform a retrospective mapping of SAR data for a current TTR inhibitors database, and we propose prospective strategies to use these efficiency indices and chemico-biological space maps for optimization and drug design efforts for TTR ligands.

**Keywords:** TTR · Amyloid · Ligand efficiency indices · Chemico-biological space

## 1 Introduction

Transthyretin (TTR) is a homotetrameric protein (Figure 1) that functions as the backup transporter for thyroxine hormone (T4) in plasma and it is the main transporter across blood brain barrier. TTR is also the main carrier of vitamin A by forming a complex with retinol-binding protein.

TTR is the cause of a series of rare amyloid diseases: familial amyloid polyneuropathy (FAP), familial amyloid cardiomyopathy (FAC), central nervous system selective amyloidoses (CNSA) and senile systemic amyloidoses (SSA). TTR amyloid deposits are produced by a still poorly characterized mechanism that involves several steps such as: dissociation of the tetramer, monomer conformational changes, aggregation of modified monomers into non-fibrillar oligomers that later form protofibrils and further elongate into mature fibrils.[2]

The crystal structure of TTR in complex with T4 (Figure 2) reveals that iodine atoms occupy four out of six (three symmetry-related pairs) small hydrophobic depressions – termed halogen binding pockets (HBP) – present in each binding site.[3] Considering this and knowing that several non-steroidal anti-inflammatory drugs (NSAIDs) are able to bind to TTR,[4,5] we have reported some iodinated NSAIDS derivatives that could present high binding affinities to TTR.

One of these promising new fibrillogenesis inhibitors is based on the diflunisal core structure, a registered salicylate drug with NSAID activity now undergoing clinical trials for TTR amyloid diseases. Biochemical and biophysical evidences confirm that iodine atoms can be an important design feature in the search of candidate drugs for TTR related amyloidoses.[6]

In the present work, our objective was the implementation of the Ligand Efficiency (LE)[7] concept as a tool to guide our drug discovery process.

[a] D. Blasi, C. Abad-Zapatero, J. Quintana
Drug Discovery Platform, Parc Científic Barcelona (PCB)
Baldiri Reixac 4–6, E-08028 Barcelona, Spain
*e-mail: dblasip@pcb.ub.es

[b] G. Arsequell, G. Valencia
Unit of Glycoconjugate Chemistry, Institut de Química Avançada de Catalunya, I.Q.A.C.-C.S.I.C., Jordi Girona 18–26, E-08034 Barcelona, Spain

[c] J. Nieto, A. Planas
Laboratory of Biochemistry, Institut Químic de Sarrià, Universitat Ramon Llull
Via Augusta 390, E-08017 Barcelona, Spain

[d] M. Pinto, N. B. Centeno
Computer-Assisted Drug Design Laboratory, Research Group on Biomedical Informatics (GRIB), IMIM-Universitat Pompeu Fabra, Dr. Aiguader 88, E-08003 Barcelona, Spain

[e] C. Abad-Zapatero
Center for Pharmaceutical Biotechnology, University of Illinois at Chicago
900 So. Ashland St, MBRB Building, Room 3020; (M/C870), Chicago, IL 60607-7173, USA

© 2011 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

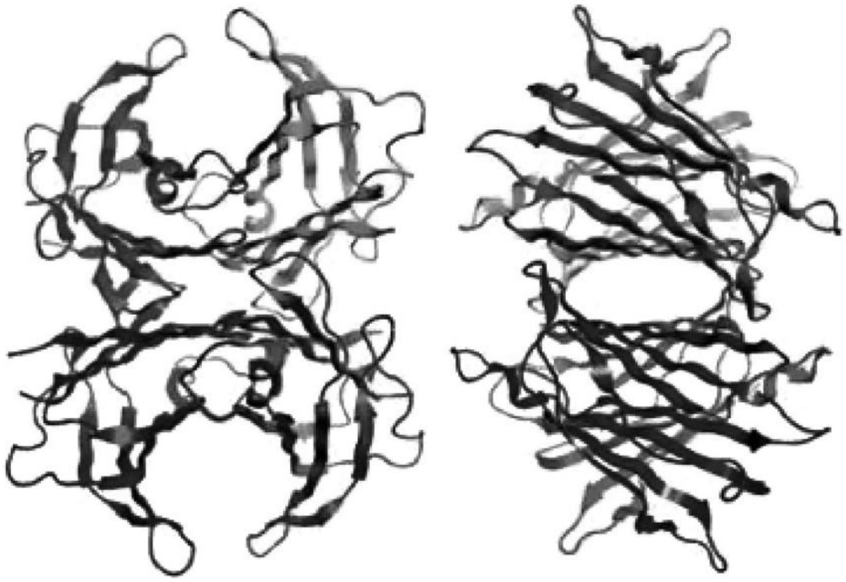


**Figure 1.** Frontal and lateral view of the quaternary structure of TTR (built from PDB ID: 2ROX[1] using MOE 2009.10 package).

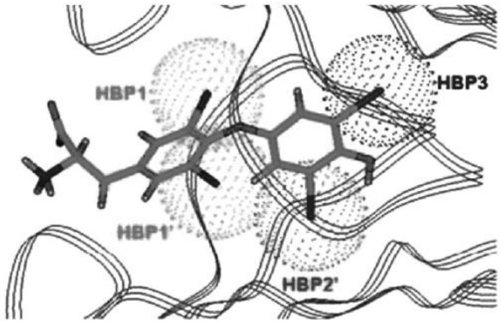


**Figure 2.** 3D. Binding site of TTR complexed with thyroxine (PDB ID: 2ROX[1]).

The LE definition relating the binding energy ($\Delta G$) of a ligand towards a target to a size parameter of the ligand (number of non-hydrogen atoms) was introduced by Hopkins et al. in 2005.[8] Modifications on this definition have been implemented as a useful way to improve lead optimization or HTS strategies,[9,10] among others applications. The use of a LE relating the affinity (Ki, IC50 or similar) to its molecular size (molecular weight (*MW*), number of non-hydrogen atoms, (*NHEA*) and others) is now well accepted. In 2005, two ligand efficiency indices (*LEI*s) were also introduced: binding efficiency index (*BEI*) and surface efficiency index (*SEI*). The first related the potency of the ligand to its *MW* (in kDa), and the second combined the potency with the Polar Surface Area (*PSA*) scaled to 100 Å$^2$. The combination of both indices in comparing retrospectively chemical series in an optimization plane *BEI-SEI* has been published.[11,12]

In order to find new effective ways to guide our drug discovery strategy to optimize TTR amyloid inhibitors, the *LEI*s based approach has been applied to map the diflunisal-related family of compounds reported previously[6] in our TTR chemical-biology space (CBS). We will hereby illustrate the use of these indices, combining three variables (potency, MW and PSA) in a 2D graphical representation (*nBEI-NSEI* plane) to perform a retrospective mapping of SAR data for a current TTR stabilizers database, and also, to establish a prospective way to guide a lead optimization process based on theoretical potency calculations (using docking protocols) of possible new TTR amyloidogenesis inhibitors, combined with the *LEI*s approach.

## 2 Methodology

### 2.1 Computational Methods

#### *2.1.1 Hardware Specifications*

Molecular modelling studies were carried out on an Intel Pentium Core 2 Duo 3.00 GHz running Windows XP.

### *2.1.2 Crystallographic Complex*

3D atomic coordinates of the TTR-iododiflunisal complex (PDB ID: 1Y1D)[13] used in the present work were obtained from the structural information available in the Protein Data Bank (PDB).[14]

Before the ligand-protein interactions mapping, a previous processing of the pdb file was needed. The asymmetric crystal unit of TTR complexes is formed by a dimer, two ligand molecules (one for each binding site) and water molecules; taking this into account, coordinates for the tetrameric form of TTR were obtained by applying the crystallographic symmetry transformations described in the pdb file. For residues with multiple conformations, we considered the one with the highest occupation factor.

### *2.1.3 Hydrogen Atoms Refinement*

Added hydrogen atoms were energy-minimized by using the Protonate 3D package implemented in MOE 2009.10.[15] Ligand partial charges were obtained by computing the electrostatic-potentials around the optimized structures using MOE 2009.10. Minimization was carried out using a distance dependent dielectric constant and a cutoff distance of 10 Å for Van der Waals interactions. Hydrogen atoms refinement was accomplished using 1000 cycles of steepest descents followed by conjugate gradient until the maximum gradient of the energy was smaller than 0.05 kcal/mol Å$^2$.

### *2.1.4 Data Set*

A set of 2300 biphenylic compounds was extracted from MMsINC Database.[16] LigX package (MOE 2009.10) was used for the hydrogen addition and ligand preparation.

### *2.1.5 Ligand-Protein Interaction Mapping*

LigandScout 2.03[17] was used to create and evaluate the most important interactions in the crystallographic complex (TTR-Iododiflunisal) used in this study, and to perform the pharmacophore-based search.

### *2.1.6 Docking Experiments*

MOE 2009.10 package was used to perform the docking studies of the data set selected with the crystallographic TTR complex. Alpha Triangle was used as placement method, Alpha HB as score function and MMF94 as forcefield in the refinement step of the docking solutions.

## 3 Results and Discussion

### 3.1 *LEI*s in Drug Discovery Strategies

As mentioned before, the concept of Ligand Efficiency (LE)[7] used as a tool to guide the fragment and lead selection optimization in the drug discovery process has emerged.

The use of this tool allows the navigation into the chemical-biology space in order to identify regions that contain biologically active compounds for particular biological targets.[12]

Several definitions of *LEI*s (Table 1) have been explored to validate this strategy. Initial formulations related potency ($K_i$, $IC_{50}$ or related measurements) to the *MW* (*BEI* = binding efficiency index) and the PSA (*SEI* = surface efficiency index). Other useful formulations, that refine and improve the navigation concept, use heavy atom count (*nBEI*) and the number of polar atoms (*NSEI*). These definitions, and maps obtained using them, have the advantage that medicinal chemists find them intuitive and easy to understand.

### 3.2 *LEI* Retrospective Map

The salicylate family looks very promising to generate TTR drug candidates due to their long therapeutic tradition and clinical applications. In order to test the effect of iodination over this type of compounds to obtain good TTR amyloid inhibitors, diflunisal was chosen as model template. Diflunisal was selected since it is an already registered drug, it has a biphenyl core like TTR inhibitors that have a good amyloid inhibitory profile, and it is under clinical trials for TTR-related amyloidosis.

A library of 40 salicylate derivatives (Table 2) was designed, prepared (without iodine atoms and with iodine at

**Table 1.** Definitions of ligand efficiency indices (*LEI*s) pairs used in each efficiency plane.

| Name | Definition | Example value[a] | Equation |
|---|---|---|---|
| *BEI* | p(*K*), p($K_d$), or p($IC_{50}$)/*MW*(kDa) | 27 | 1 |
| *SEI* | p(*K*), p($K_d$), or p($IC_{50}$)/(*PSA*/100 Å$^2$) | 18 | 2 |
| *NSEI* | −log10 $K_i$/(*NPOL*) = p$K_i$/*NPOL*(N,O) | 1.5 | 3 |
| *NBEI* | −log10 $K_i$/(*NHEA*) = p$K_i$/(*NHEA*) | 0.36 | 4 |
| *nBEI* | −log10[($K_i$/*NHEA*)] | 10.25 | 5 |

[a] Examples and reference values for the *LEI*s are calculated for each index using the following idealized values (units have been omitted in the table): $K_i$ or $IC_{50}$ = 1.0 nM; p(*K*) = −log $K_i$ = 9.00; *MW* = 0.333 kDa; *NHEA* = Number of heavy atoms (non-Hydrogen in the compound) = 25; *NPOL* = Number of polar atoms (N,O) = 6; *NBEI*/*NSEI* = *NPOL*(N,O)/*NHEA*(non-H) = 0.36/1.5 = 6/25 = 0.24; *PSA* = 50 Å$^2$

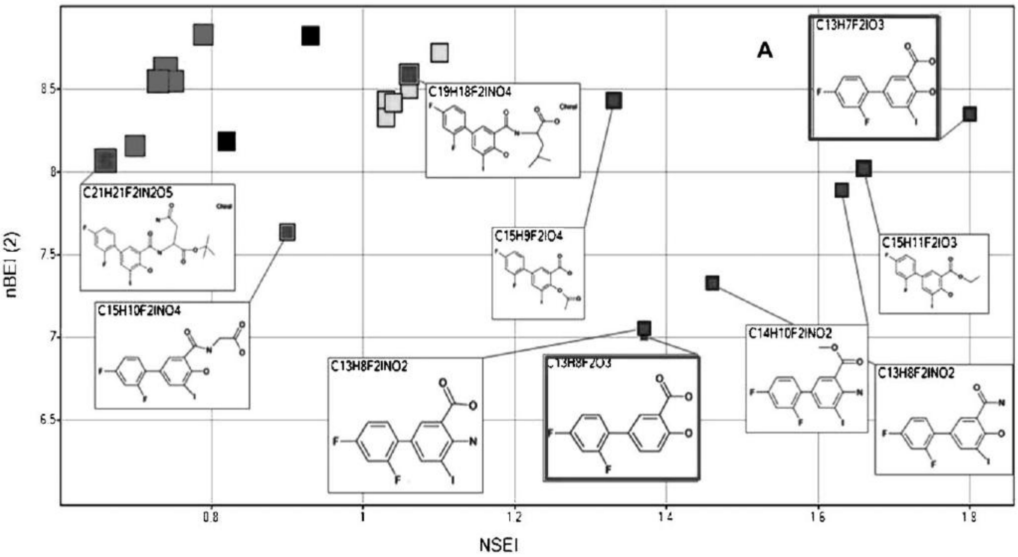


**Figure 3.** Retrospective map of diflunisal library: nBEI vs. NSEI plot shows that iododiflunisal (frame A) is the most effective TTR amyloid inhibitor for this chemical series, described in the literature.[5] Gray squares: some iododiflunisal derivatives mapped.

C5 position, in order to validate the hypothesis that iodinated compounds can be potent amyloid inhibitors) and tested in vitro.

The optimization plane of this new family of diflunisal-like compounds was analyzed using *nBEI* vs. *NSEI* plot (Figure 3).

The plot is a representation of this small CBS in polar coordinates where the angular component represents *NPOL* (number of polar atoms in chemical structure of compounds; increasing from right to left), and the radial component represents the potency value of each compound. The more effective compounds will be those that have been optimized, and maximized, in the two variables of the plane created by *nBEI-NSEI*.

For that reason, this *LEI* representation allows us to conclude that iododiflunisal is more effective for TTR protein than diflunisal, and this could be a good starting point to perform a new drug discovery strategy.

### 3.3 *LEI* Prospective Map

An iododiflunisal scaffold based search was carried out using MMsINC database.[16] The compounds compiled in this database are non-redundant and biomedically relevant chemical structures, richly annotated with crucial chemical properties. Moreover, MMsINC is integrated with other primary data collectors, such as PubChem,[18] Protein Data Bank,[14] the Food and Drug Administration database of approved drugs,[19] ZINC[20] and also commercial catalogs from several companies. Commercially available compounds extracted from this source were merged with 500 suggested molecules based on iododiflunisal, with different decorations in the aromatic rings. A final database of 2300 virtual and commercial biphenyl compounds was obtained.

The most important interactions between wild type TTR and iododiflunisal from 1Y1D crystallographic complex were analyzed using the software LigandScout (Figure 4).[17]

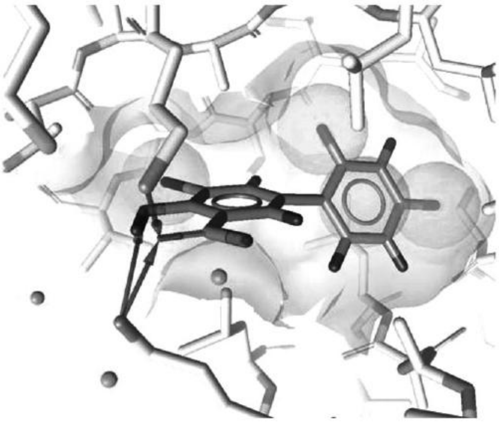
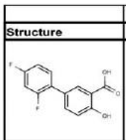
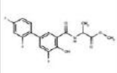
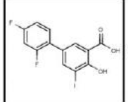
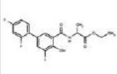
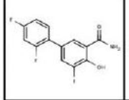
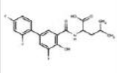
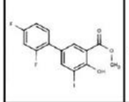
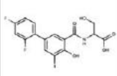
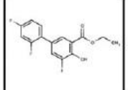
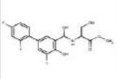
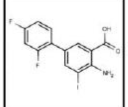
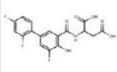
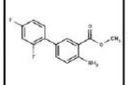
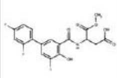
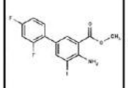
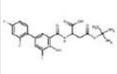
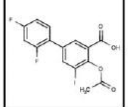
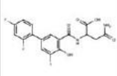
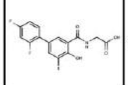
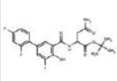
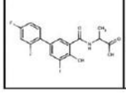
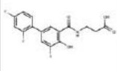


**Figure 4.** Crucial interactions in 1Y1D complex (TTR: iododiflunisal).

© 2011 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

**Table 2.** A sample of derivatives Diflunisal/Iododiflunisal derivatives used in the Retrospective map.

| | | Kinetic turbidity assay | | | | Kinetic turbidity assay | |
|---|---|---|---|---|---|---|---|
| **Structure** | | **RA (%)[a]** | **IC50 (μM)[b]** | **Structure** | | **RA (%)[a]** | **IC50 (μM)[b]** |
|  | **1** | 87 | 16,3 |  | **12** | 93 | 5,5 |
|  | **2** | 94 | 4,5 |  | **13** | 96 | 5,7 |
|  | **3** | 86 | 12,4 |  | **14** | 97 | 5,0 |
|  | **4** | 72 | 6,1 |  | **15** | 98 | 3,7 |
|  | **5** | 86 | 6,9 |  | **16** | 94 | 7,2 |
|  | **6** | 93 | 16,3 |  | **17** | 97 | 5,2 |
|  | **7** | 65 | 41 |  | **18** | 99 | 4,1 |
|  | **8** | 73 | 7,5 |  | **19** | 99 | 6 |
|  | **9** | 92 | 4,8 |  | **20** | 96 | 7,7 |
|  | **10** | 100 | 11,1 |  | **21** | 100 | 9,7 |
|  | **11** | 98 | 5,8 |  | **22** | 97 | 4,9 |

[a] Percent reduction of fibril formation rate at high inhibitor concentration relative to the rate at [I] = 0. [b] Concentration of inhibitor at which initial rate of fibril formation is one-half than at [I] = 0.

© 2011 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

This interaction map shows three hydrogen-bond acceptors located in the hydroxyl and carboxylic groups, one hydrophobic volume in the iodine atom and two hydrophobic volumes over each fluorine atom.

This map was used as our reference pharmacophore model, based on iododiflunisal, and a pharmacophore-based search was performed, using LigandScout software in order to filter the 2300 compounds database, obtaining a subset of 1200 compounds that share the iododiflunisal interaction pattern and fit with our model.

These molecules were docked against the wild type TTR protein model extracted and prepared, using the software MOE 2009.10,[15] from the iododiflunisal:TTR 1Y1D crystallographic complex. Alpha triangle was used as placement method, Alpha HB as score function and the

MMFF94 forcefield was applied in the refinement step for the docking solutions. In the placement step, using alpha triangle method, poses are generated by superposition of ligand atom triplets and triplets of receptor site points. The receptor site points are *alpha sphere* centres which represent locations of tight packing. At each iteration a random conformation is selected, a random triplet of ligand atoms and a random triplet of alpha sphere centres are used to determine the pose. After this placement process, with Alpha HB function, it is possible to obtain a binding free energy estimation considering the linear combination of a geometric fit of the ligand to the binding site term, and a hydrogen bonding effect term. The refinement of poses obtained was carried out using the conventional molecular mechanics forcefield (MMFF94). In order to speed up the calculation, residues over a cut-off distance of 6 Å away from all the pre-refined poses are ignored (during the refinement and the final energy evaluation). However, if after the refinement, any poses moved to within 80 % of the cutoff distance of any ignored atoms, the refinement stage was repeated with an increased cut-off distance. All receptor atoms (backbone and sidechains) are fixed. In addition, the final energy was evaluated using the generalized Born solvation model (GB/VI).[15] Finally, the energy obtained in this work-flow, includes only the interaction energy between the receptor and the ligand. The self-energies of both the receptor and the ligand are excluded.

After this docking protocol, 80 compounds were prioritized by the score value. Taking into account that the output of MOE docking is an estimation of the binding free energy, it is possible to establish a relationship between $K_i$ and $\Delta G$, through the equation: $\Delta G = -RT \ln K_i$.

Following this approximation, the theoretical binding energy was calculated relating the score obtained from each pose of the 80 selected molecules to the free binding energy. For the normalization of these values, the experimental $K_i$ of iododiflunisal related to its TTR binding, ob-

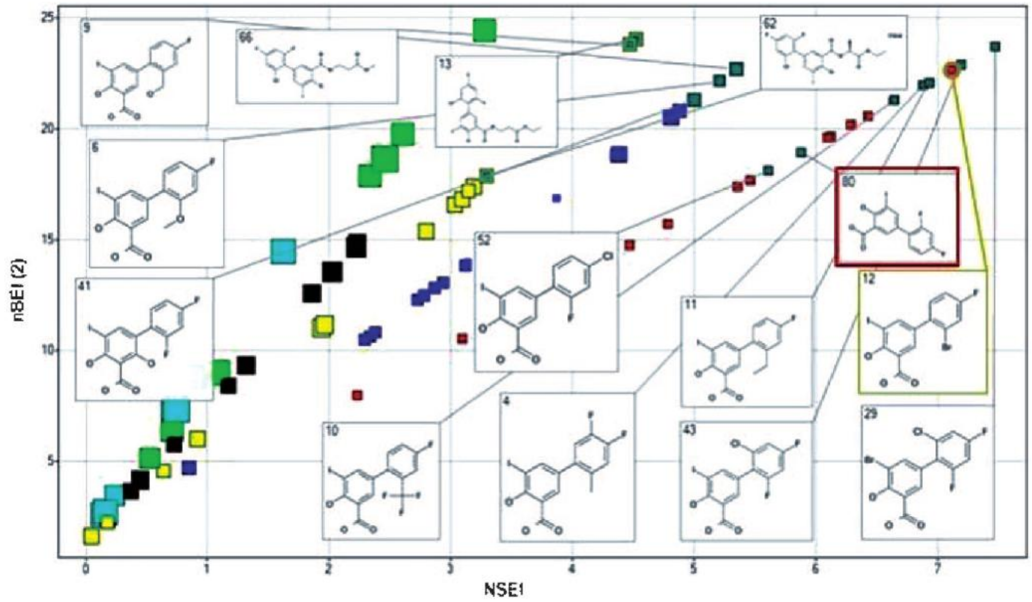


**Figure 5.** nBEI vs. NSEI plot of the 80 selected docked compounds (nBEI and NSEI as defined in Table 1). Red squares: *NPOL* = 3; blue squares: *NPOL* = 4; yellow squares: *NPOL* = 5; black squares: N*POL* = 6; green squares: *NPOL* = 7; cyan squares: *NPOL* = 8. Red frame: iododiflunisal, the reference compound for the retrospective mapping.

© 2011 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

tained by a calorimetric experiment (ITC),[21] was used. These theoretical $K_i$'s were needed to carry out the prospective *LEI* analysis in the *nBEI-SEI* optimization plane (Figure 5).

Using this methodology was possible to evaluate the efficiency of new biphenyl compounds, similar to diflunisal and iododiflunisal by adding or deleting polar groups to the scaffold. From these 80 compounds prioritized and classified by *NPOL*, after *nBEI-NSEI* plane visualization, we were able to select 12 promising molecules that are, theoretically, more effective than the reference compound (iododiflunisal) because they are optimized in the two directions of the plane (M*W* and *PSA* that means size and polarity). Another important point is that, these 12 compounds have the same binding pattern than iododiflunisal and all of them have the iodine atom, a possible key point to obtain selective and potent amyloid inhibitors for TTR. These molecules will be tested in fibrillogenesis inhibition studies[22] as possible new amyloidogenesis inhibitors for the TTR protein.

## 4 Conclusions

*LEI*s are a powerful tool in the drug discovery workflow. Recently, it has been demonstrated that using BEI-SEI optimization plane, it is possible to explore the CBS and find the crucial milestones to select the most promising candidates.

In TTR studies, we have mapped a local CBS centred on diflunisal-like compounds in order to evaluate the efficiency of this family, concluding that iododiflunisal is a good and efficient amyloid inhibitor.

A computational protocol, based on a target-ligand interaction mapping and docking, followed by a *LEI* analysis using theoretical binding energies, allowed us to prioritize 12 biphenylic compounds from a database of 2300 as possible new TTR stabilizers.

## Acknowledgements

This work was supported by a grant from the *Fundació Marató de TV3*, Barcelona, Spain (Project Number: 080530/31/32).

## References

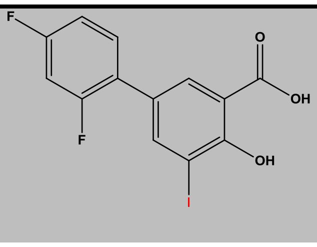
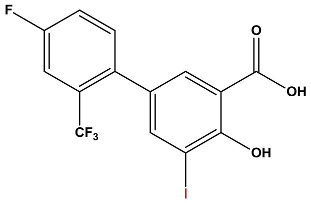
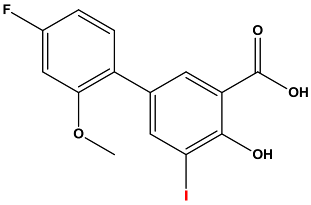
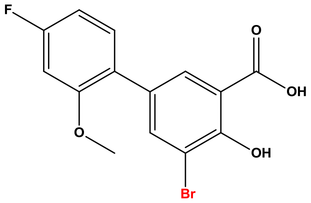
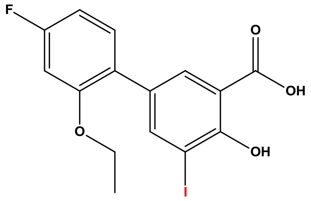
[1] *PDB ID: 2ROX*, A. Wojtczak, V. Cody, J. R. Luft, W. Pangborn, Structures of human transthyretin complexed with thyrosine at 2.0 A resolution and 3′,5′-dinitro-N-acetyl-L-thyrosine at 2.2 A resolution, **1996**.
[2] T. R. Foss, R. L. Wiseman, J. W. Kelly, *Biochemistry* **2005**, *44*, 15525 – 15533.
[3] J. M. Blaney, E. C. Jorgensen, M. L. Connolly, T. E. Ferrin, R. Langridge, S. J. Oatley, J. M. Burridge, C. C. Blake, *J. Med. Chem.***1982**, *25*, 785 – 90.
[4] S. L. Munro, C. F. Lim, J. G. Hall, J. W. Barlow, et al., *J. Clin. Endocrinol. Metab.* **1989**, *68*, 1141 – 1147.
[5] S. L. Adamski-Werner, S. K. Palaninathan, J. C. Sacchettini, J. W. Kelly, *J. Med. Chem.* **2004**, *4*, 355 – 374.
[6] T. Mairal, J. Nieto, M. Pinto, M. R. Almeida, L. Gales, A. Ballesteros, M. Saraiva, A. M. Damas, A. Planas, G. Arsequell, G. Valencia, et al., *PLoS ONE* **2009**, *4*(1).
[7] I. D. Kuntz, K. Chem, K. A. Sharp, P. A. Kollman, *Proc. Natl. Acad. Sci. USA,* **1999**, *9*(10), 430 – 431.
[8] A. L. Hopkins, C. R. Groom, A. Alex, *Drug Discov. Today* **2004**, *9*, 430 – 431.
[9] C. Abad-Zapatero, J. T. Metz, *Drug Discov. Today* **2005**, *10*(7), 464 – 469.
[10] D. Tanaka, Y. Tsuda, T. Shiyama, T. Nishimura, N. Chiyo, Y. Tominaga, N. Swada, T. Mimoto, N. Kusunose, *J. Med. Chem.* **2011**, *54*, 851 – 857.
[11] C. Abad-Zapatero, *Expert Opin. Drug Discov.* **2007**, *2*, 469 – 484.
[12] C. Abad-Zapatero, O. Perišić, J. Wass, A. P. Bento, J. Overington, B. Al-Lazikani, M. E. Johnson, *Drug Discov. Today,* **2010**, *15*(19 – 20), 804 – 811.
[13] *PDB ID: 1Y1D;* L. Gales, S. Macedo-Ribeiro, G.Arsequell, G. Valencia, M. J. Saraiva, A. M. Damas, Human transthyretin in complex with iododiflunisal: structural features associated with a potent amyloid inhibitor, **2005**.
[14] H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, P. E. Bourne, *Nucleic Acids Res.* **2000**, *28*(1), 235 – 242; RCSB Protein Data Bank, http://www.pdb.org
[15] *MOE*, Chemical Computing Group Inc., Montreal, http://www.chemcomp.com.
[16] J. Masciocchi, G. Frau, M. Fanton, M. Sturlese, M. Floris, M. L. Pireddu, P. Palla, F. Cedrati, P. Rodriguez-Tomé, S. Moro, *Nucleic Acids Res.* **2009**, *37*, D284.
[17] G. Wolber, T. Langer, *J. Chem. Inf. Model.* **2005**, *45*, 160 – 169; http://www.inteligand.com.
[18] *The PubChem Project*, http://pubchem.ncbi.nlm.nih.gov.
[19] *Drugs@FDA*, FDA Approved Drug Products, http://www.accessdata.fda.gov/scripts/cder/drugsatfda.
[20] J. J. Irwin, B. K. Shoicet, *J. Chem. Inf. Model.*, **2005**, *45*, 177 – 182.
[21] A. González, J. Quirante, J. Nieto, M. R. Almeida, M. J. Saraiva, A. Planas, G. Arsequell, G. Valencia, *Bioorg. Med. Chem. Lett.* **2009**, *19*, 5270 – 5273.
[22] I. Dolado, J. Nieto, M. J. Saraiva, G. Arsequell, G. Valencia, A. Planas, *J. Comb. Chem.* **2005**, *7*(2), 246 – 252.

Received: November 3, 2010
Accepted: February 24, 2011
Published online: March 18, 2011

© 2011 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

## Experimental validation of the LEI-docking computational analysis of TTR ligands

In order to validate the proposed strategy for prospective TTR drug design of more efficient inhibitors, combining the LEI retrospective analysis of efficiency of TTR ligands, with the TTR ligand docking studies, 10 Iododiflunisal analogs (**Table 2**), theoretically more efficient than the parent compound, were synthesized and their amyloid inhibition activity against TTR of these compounds was evaluated in the *in vitro* kinetic turbidimetric assay[7].

| Compound number | Structure | $IC_{50}$ (µM) | RA (%) |
|---|---|---|---|
| **IDIF** |  | 4,2 | 91 |
| **1** |  | 7,2 | 87 |
| **2** |  | 5,3 | 89 |
| **3** |  | 5,2 | 92 |
| **4** |  | 5,1 | 92 |

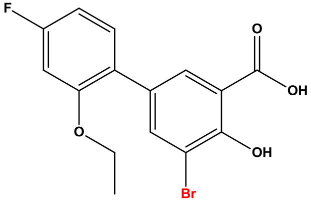
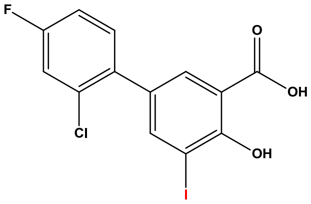
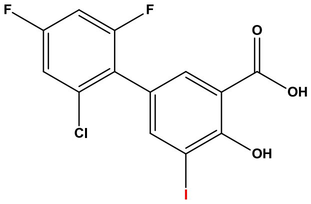
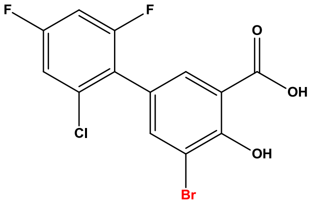
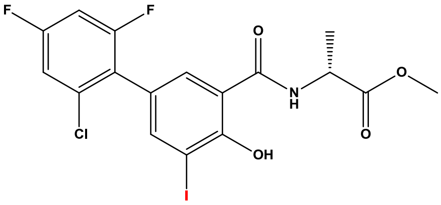
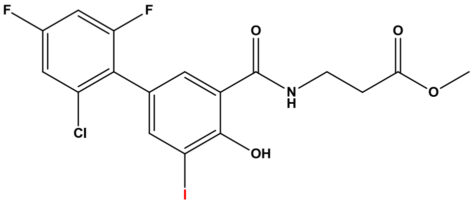
| Compound number | Structure | $IC_{50}$ (µM) | RA (%) |
|---|---|---|---|
| **5** |  | 4,3 | 89 |
| **6** |  | 3,9 | 88 |
| **7** |  | 4,6 | 91 |
| **8** |  | 3,6 | 89 |
| **9** |  | 6,0 | 89 |
| **10** |  | 3,7 | 90 |

**Table 2.** 10 Iododiflunisal analogs designed and selected as theoretically more efficient than Iododiflunisal in terms of ligand efficiency parameters, and their experimental amyloid inhibition activities.

The results are summarized in the **Figure 2**.

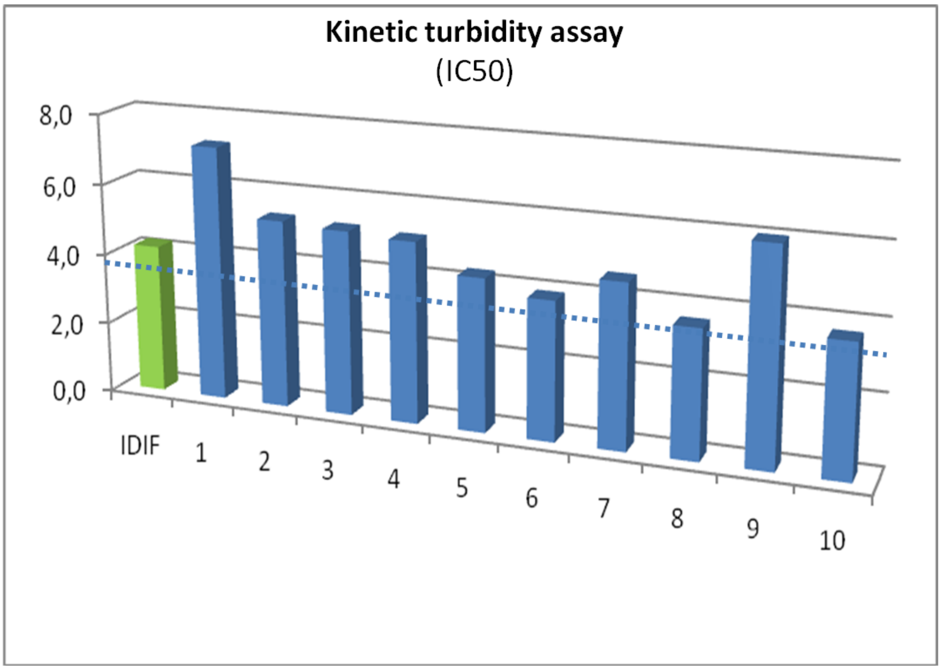


**Figure 2**. Histograms of $IC_{50}$ values for the amyloid inhibition experimental activities for the 10 compounds selected. In green, Iododiflunisal reference is highlighted in both graphics.

In terms of $IC_{50}$, 4 compounds (5, 6, 8 and 10) show an equal or higher fibrillogenesis inhibition activity compared with the reference (Iododiflunisal).

With these experimental amyloid inhibition activities results, it was possible to generate the following retrospective LEI map, using the experimental values to confirm the predictions previously reported[3].

**Figure 3** shows the final result of the experimental nBEI vs SEI map after the *in vitro* amyloid inhibition activity assays. In **Table 3**, the results are listed in detail.

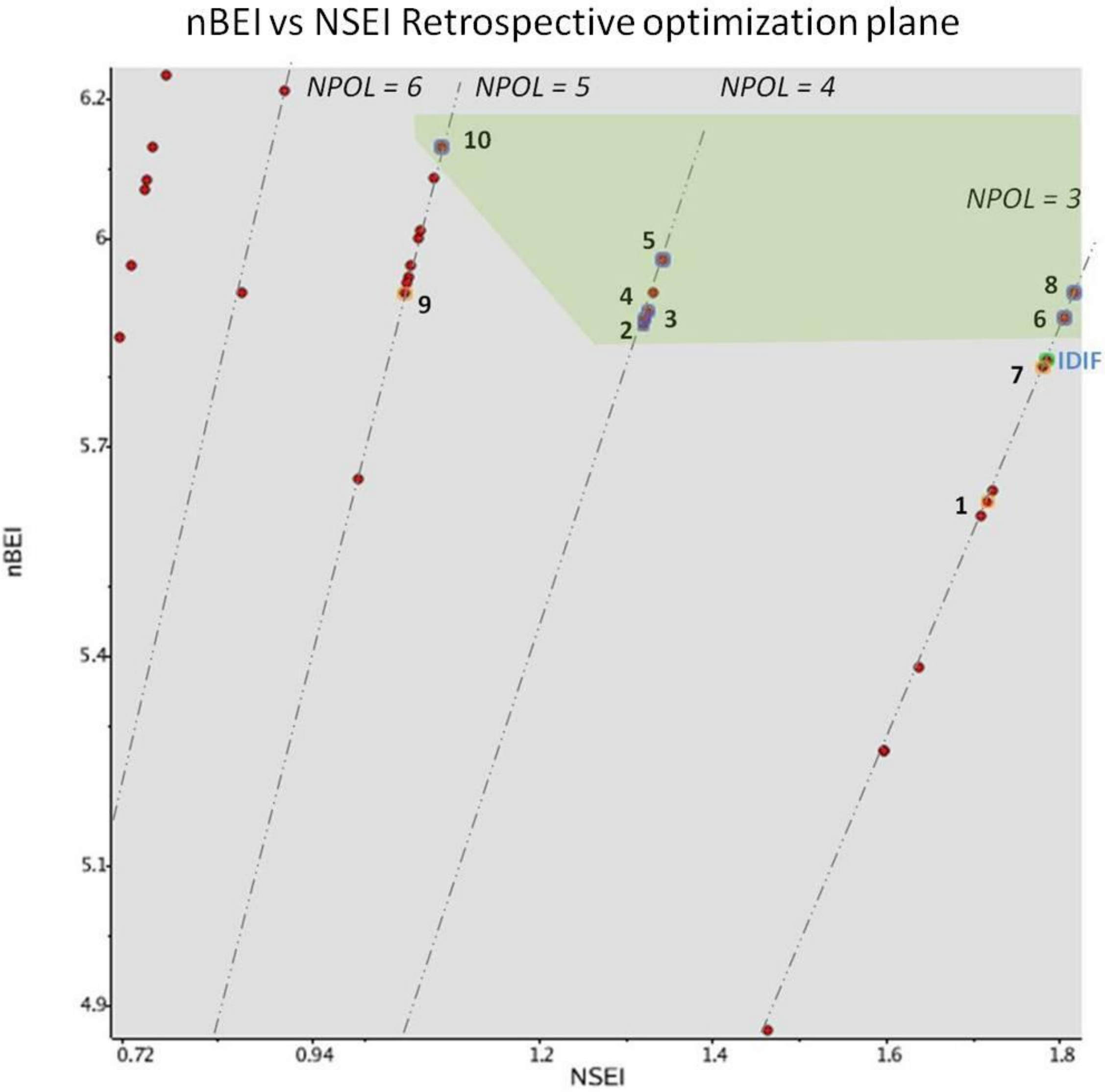
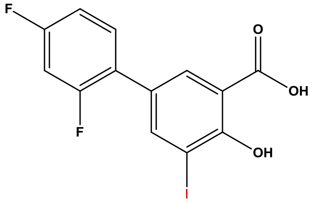
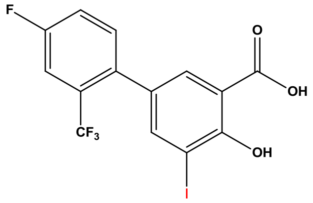
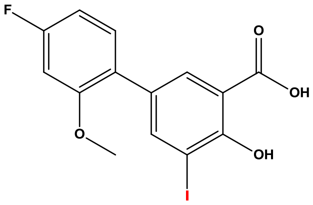
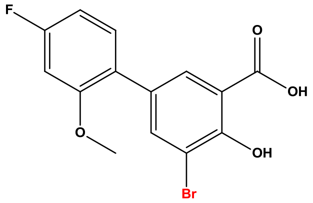
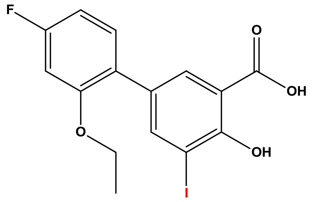
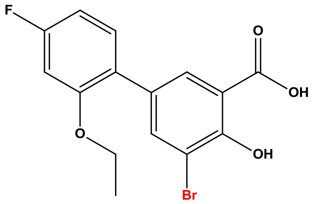
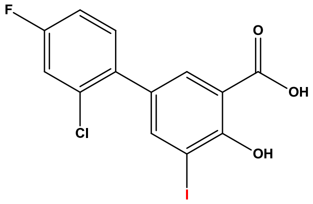


**Figure 3.** nBEI vs NSEI plot based on the experimental data ($IC_{50}$ values) obtained from the kinetic turbidimetry assay. Green spot corresponds to the reference compound (Iodiflunisal); blue spots are the compounds with a similar or higher efficiency compared with Iododiflunisal (in agreement with the predictions performed); orange spots highlight compounds that are less efficient than Iododiflunisal (where the prediction fails).

| Compound number | Structure | $IC_{50}$ (μM) | nBEI | NSEI | Prediction |
|---|---|---|---|---|---|
| **IDIF** |  | 4,2 | 14,2 | 9,3 | Reference |
| **1** |  | 7,2 | 12,1 | 8,9 | Mismatch |
| **2** |  | 5,3 | 13,6 | 7,9 | Match |
| **3** |  | 5,2 | 13,6 | 7,9 | Match |
| **4** |  | 5,1 | 13,1 | 7,9 | Match |
| **5** |  | 4,3 | 15,1 | 8,0 | Match |
| **6** |  | 3,9 | 13,9 | 9,4 | Match |

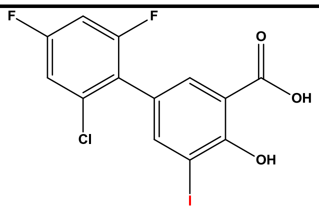
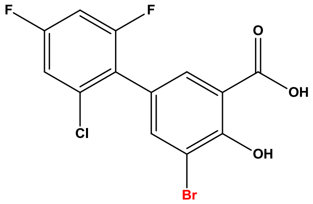
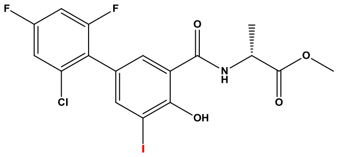
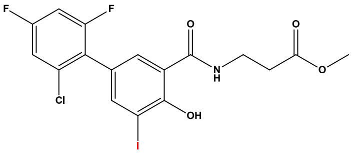
| Compound Number | Structure | $IC_{50}$ (µM) | nBEI | NSEI | Prediction |
|---|---|---|---|---|---|
| **7** |  | 4,6 | 13,0 | 9,2 | Mismatch |
| **8** |  | 3,6 | 14,9 | 9,4 | Match |
| **9** |  | 6,0 | 10,5 | 6,9 | Mismatch |
| **10** |  | 3,7 | 10,9 | 7,2 | Match |

**Table 3.** nBEI and NSEI values for Iododiflunisal (reference compound) and the 10 analogs designed. The field "Prediction" is refered to the match or mismatch between the efficiency predicted and the experimental results (*match*: the experimental value corresponds with the expected efficiency; *mismatch*: the efficiency of the compound was not correctly predicted).

From the 10 compound analogs designed by using this prospective LEI approach, only 3 were not correctly predicted; the rest of them had a good agreement with the predicted estimations of efficiency. In addition, the most promising compounds in the prospective LEI's map have also been the most efficient in the retrospective LEI's map.

Looking in detail to the nBEI-SEI optimization plane, it can be highlighted that compounds **5** and **8** in table 3 could be the most interesting of the series because both of them are optimizing the 3 crucial LEI variables at the same time, compared with the reference: potency, size and polarity.

All the information and results presented in this thesis chapter have helped us to answer the initial question: "is it possible to use the LEI methodology to guide the drug discovery process in a prospective manner?". The study we have presented here shows that the LEI analysis of the efficiency of a series of TTR ligands, combined with the docking analysis of these ligands in the TTR binding site, can be applied in such prospective approach to lead the drug design phase of TTR ligands optimization. We have also demonstrated that the LEI methodology (retrospective and prospective) may be combined and integrated efficiently with other computational workflows such us pharmacophore modelling or docking.

As a conclusion for this thesis chapter, it is possible to propose new more efficient TTR ligands in the drug discovery and development path, by Chemico-Biological Space (CBS) exploration, browsing the nBEI-NSEI optimization plane of the LEI methodology. Focusing on the TTR protein, it has been possible to optimize (in terms of size and polarity) a well-known active compound, to get more efficient TTR ligand analogs, through the analysis of the TTR CBS, and we have selected the most efficient ligand of our TTR nBEI-NSEI plane, as a possible candidate for the next steps of TTR ligand drug discovery optimization.

## Bibliography

(1) Kuntz, I. D.; Chen, K.; Sharp, K.; Kollman, P. The maximal affinity of ligands. *Proceedings of the National Academy of Sciences of the United States of America* **1999**, *96*, 9997–10002.

(2) Abad-Zapatero, C. Ligand efficiency indices for effective drug discovery. *Expert Opinion on Drug Discovery* **2007**, *2*, 469–488.

(3) Blasi, D.; Arsequell, G.; Valencia, G.; Nieto, J.; Planas, A.; Pinto, M.; Centeno, N. B.; Abad-Zapatero, C.; Quintana, J. Retrospective Mapping of SAR Data for TTR Protein in Chemico-Biological Space Using Ligand Efficiency Indices as a Guide to Drug Discovery Strategies. *Molecular Informatics* **2011**, *30*, 161–167.

(4) Mairal, T.; Nieto, J.; Pinto, M.; Almeida, M. R.; Gales, L.; Ballesteros, A.; Barluenga, J.; Pérez, J. J.; Vázquez, J. T.; Centeno, N. B.; Saraiva, M. J.; Damas, A. M.; Planas, A.; Arsequell, G.; Valencia, G. Iodine atoms: a new molecular feature for the design of potent transthyretin fibrillogenesis inhibitors. *PloS One* **2009**, *4*, e4124.

(5) Abad-Zapatero, C.; Blasi, D. Ligand Efficiency Indices (LEIs): More than a Simple Efficiency Yardstick. *Molecular Informatics* **2011**, *30*, 122–132.

(6) Pinto, M.; Blasi, D.; Nieto, J.; Arsequell, G.; Valencia, G.; Planas, A.; Quintana, J.; Centeno, N. B. Ligand-binding properties of human transthyretin. *Amyloid* **2011**, *18*, 51–54.

(7) Dolado, I.; Nieto, J.; Saraiva, M. J. M.; Arsequell, G.; Valencia, G.; Planas, A. Kinetic assay for high-throughput screening of in vitro transthyretin amyloid fibrillogenesis inhibitors. *Journal of Combinatorial Chemistry* **2005**, *7*, 246–252.

(8) Dobson, C. M. The structural basis of protein folding and its links with human disease. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* **2001**, *356*, 133–145.

(9) Lima, L. M. T. R.; Silva, V. D. A.; Palmieri, L. D. C.; Oliveira, M. C. B. R.; Foguel, D.; Polikarpov, I. Identification of a novel ligand binding motif in the transthyretin channel. *Bioorganic & Medicinal Chemistry* **2010**, *18*, 100–110.

(10) Razavi, H.; Palaninathan, S. K.; Powers, E. T.; Wiseman, R. L.; Purkey, H. E.; Mohamedmohaideen, N. N.; Deechongkit, S.; Chiang, K. P.; Dendle, M. T. a; Sacchettini, J. C.; Kelly, J. W. Benzoxazoles as transthyretin amyloid fibril inhibitors: synthesis, evaluation, and mechanism of action. *Angewandte Chemie (International ed. in English)* **2003**, *42*, 2758–2761.

(11) Johnson, S. M.; Connelly, S.; Wilson, I. a; Kelly, J. W. Biochemical and structural evaluation of highly selective 2-arylbenzoxazole-based transthyretin amyloidogenesis inhibitors. *Journal of medicinal chemistry* **2008**, *51*, 260–270.

(12) Johnson, S. M.; Connelly, S.; Wilson, I. a; Kelly, J. W. Toward optimization of the second aryl substructure common to transthyretin amyloidogenesis inhibitors using biochemical and structural studies. *Journal of medicinal chemistry* **2009**, *52*, 1115–1125.

(13) Gales, L.; Almeida, M. R.; Arsequell, G.; Valencia, G.; Saraiva, M. J.; Damas, A. M. Iodination of salicylic acid improves its binding to transthyretin. *Biochimica et biophysica acta* **2008**, *1784*, 512–517.

(14) Group, C. C. Molecular Operating Environment (MOE), Chemical Computing Group Inc.: 1010 Sherbooke St. West, Suite #910, Montreal, QC, Canada, H3A 2R7, 2011.

# PART IV

# METABOLIC STABILITY STUDIES

## Introduction

Early identification of metabolic biotransformations and pathways can be a crucial point in a drug discovery strategy. These biotransformations are the key elements that define bioavailability, activity/efficiency, toxicity, distribution and elimination of a given compound[1].

Cytochromes P450 (CYPs) constitute the most important family of biotransformation enzymes involved in drug metabolism[2]. This system is also responsible for many of the drug-drug interactions and metabolism-dependent toxicity effects reported in clinical phases. CYPs catalyze a variety of reactions, converting any xenobiotic to potentially reactive products as well as less toxic compounds. Most of these reactions are oxidations, but CYPs can also be involved in reductions, desaturations, ester cleavages, ring expansions and formations, dehydrations, coupling reactions, rearrangements and other reactions[3]. The purpose of most of these pathways is the bioactivation and detoxification of a xenobiotic chemical that enters the body.

All CYP structures reported contain a three-dimensional folding pattern quite similar and conserved, consisting, mainly, of a core region (where several alpha-helices are located), and a heme coordination region[4]. Other regions are more variable, depending on the CYP isoform (e.g., active site region, spanning helices region or β-strands). These variable regions are related to substrate access and molecular recognition processes. Some of the most relevant CYP structures reported for CYP2C9 and CYP3A4 are given on **Figure 1**, which shows CYP2C9 without ligand (PDB ID: 1OG2), with Flurbiprofen bound (PDB ID: 1R9O) and CYP3A4 without ligand (PDB ID: 1W0E). These CYP structures are extracted from the Protein Data Bank (www.rcsb.org).

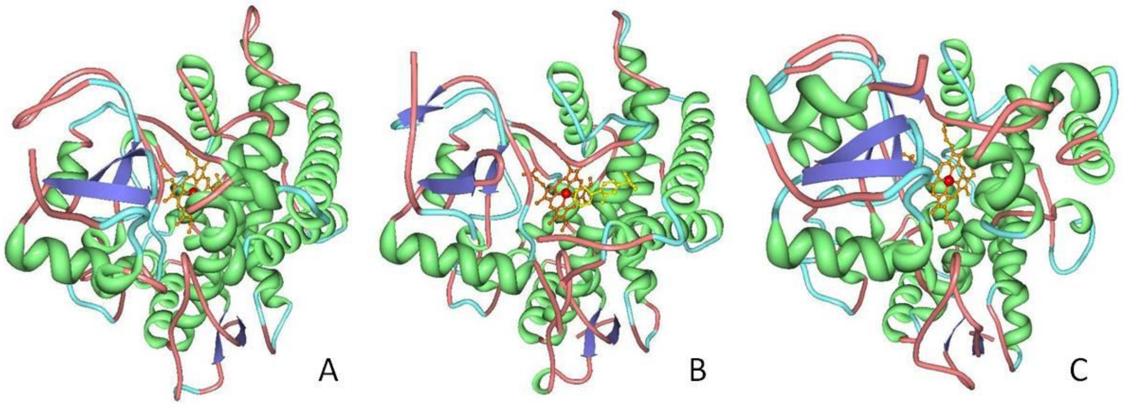


**Figure 1.** Most relevant CYP structures reported for CYP2C9 (A) (without ligand), CYP2C9 bound to Flurbiprofen (B) and CYP3A4 (C) (without ligand). Extracted from the Protein Data Bank (www.rcsb.org)

Any evidence about the metabolic profile of a drug candidate can be considered as crucial information for a drug discovery strategy. However, an experimental study on metabolic resistance and elucidation of the sites of metabolism (SOM), defined as the places in a molecule where the metabolic reactions take place, is a resource-demanding and time-consuming task[5]. It involves several experimental techniques and consumes significant amounts of each compound tested. However, any drug discovery process, in its preclinical and clinical phases, needs to characterize the SOMs to advance in the design of new compounds with an optimized pharmacokinetic profile. When a labile compound for a specific CYP isoform is found or when a toxic metabolite is detected, chemists are then able to: a) perform a synthetic strategy to add stabilizing groups at positions in the chemical structure where metabolic transformations may occur; or b) to shield some parts of the molecule using steric groups to prevent the interaction with a certain CYP enzyme; or c) to avoid any toxic effect by chemically protecting metabolically liable positions with suitable groups. For these reasons, an early identification of the SOM for a given molecule is crucial to enhance the speed and quality of the progress of a new drug entity to the market.

The increasing number of X-ray crystal structures of CYPs and the improvements in hardware and software, in terms of speed calculation through faster computational processors and architectures, as well as new algorithms, have led to more powerful and predictive techniques, such as quantum methods or docking protocols, for *in silico* SOM computations, during the last years. However, due to the flexibility of the CYP structures, the results obtained have not been quite accurate[6,7,8]. Some other methods based on *ab initio* calculations are slow and they don’t take into account the substrate-enzyme recognition and orientation processes[9]. In addition, attempts to develop rules-based methods have shown low quality predictions[10]. They show a high dependence on the training set and, generally, they over-predict the metabolic transformations, leading to thousands of possible metabolites and failing on the detection of crucial pathways.

Under the claim “The chemists need a new approach”, Prof. Gabriele Cruciani and collaborators at the University of Perugia, Italy, have worked intensely[5] on the development of new computational tools that may be able to answer the following three “most wanted” questions: a) which CYP isoform is involved in a given metabolic degradation?; b) which are the most probable sites of metabolism (SOM) in our molecules of interest?; and, c) are we able to rationally modify the compound to prevent this metabolic degradation? The computational tool developed from this research is MetaSite[5] and nowadays it is a well established methodology in the medicinal chemistry field. Using a molecular interaction field similarity method to compare the CYP enzymes and the substrates, as well as covering all the possible reactivity sites inside the CYP enzymes through a fast procedure, MetaSite is able to give metabolic pathway predictions for any molecule through the analysis of its interactions with the main CYP enzymes.

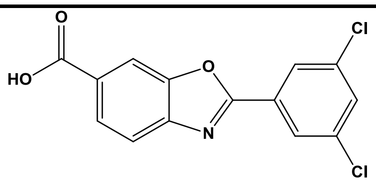
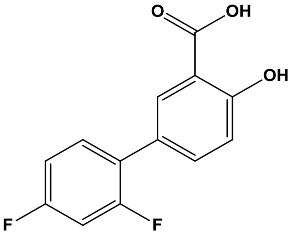
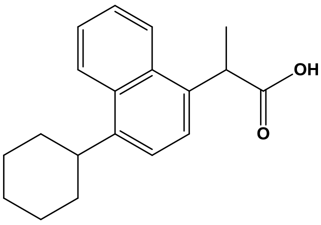
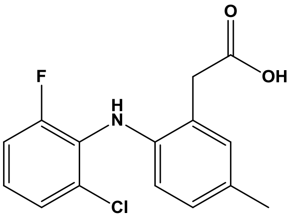
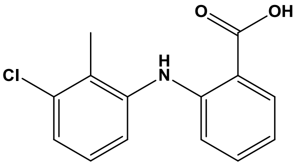
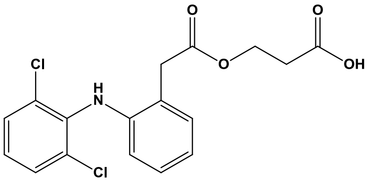
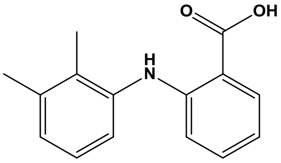
The Metasite software is presented in a user-friendly interface simple to use, whereas the system utilizes complex algorithms, already reported on the bibliography[5]. The methodology involves the calculation of two sets of molecular descriptors: one for CYP and one for the potential substrate (Chemical FingerPrint). The set of descriptors used for the CYP characterization is based on GRID-MIFs (GRID flexible molecular interaction field descriptors)[11]. GRID probes based on pharmacophore recognition (atoms categorization by hydrophobicity, hydrogen-bond donor, hydrogen-bond acceptor and charge capabilities) are computed in the case of substrate. Finally, both sets of descriptors are used to compare the fingerprints of the enzyme and substrate, providing to the user an estimation of the accessibility of each substrate toward the reactive heme group of the CYP isoform considered. Substrate reactivity is also computed and taken into account. SOMs are the result of the accessibility and reactivity calculations.

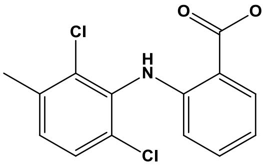
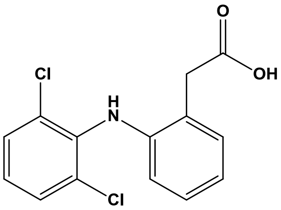
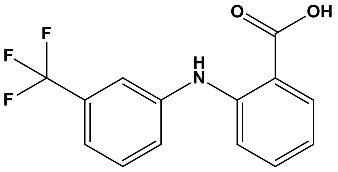
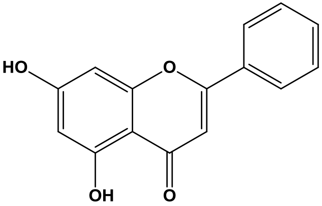
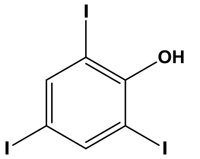
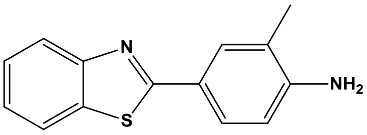
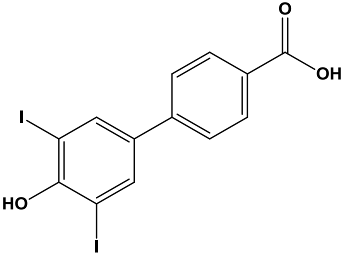
All these internal calculations require only a few seconds per molecule and they are automatically performed when a molecule or a set of molecules are provided by the user in any of the file formats available.

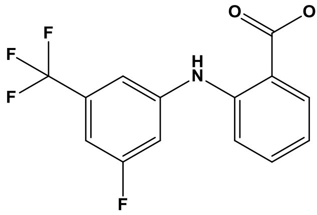
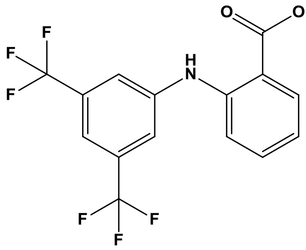
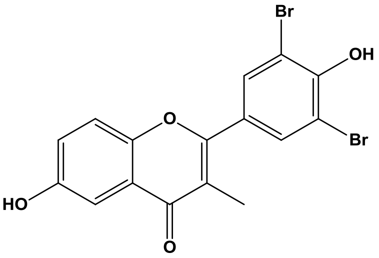
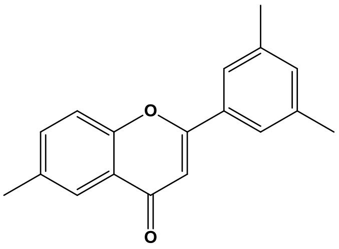
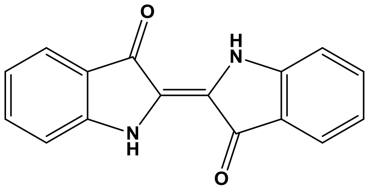
As it has been mentioned before in this thesis, more than 800 compounds have been tested against TTR amyloidosis inhibition activity assays in our TTR Consortium during the last 10 years, finding promising candidates and new interesting chemical scaffolds as TTR inhibition ligands. Among them, we have selected a small dataset (**Table 1**) of interesting "hits" (26 compounds with good $IC_{50}$ values in the inhibition activity assays from different chemical families) and we have performed *in silico* pharmacokinetic studies in collaboration with the modeling group of the University of Perugia, led by Prof. Gabriele Cruciani. The final goal of study is the analysis of the physicochemical, pharmacokinetic and metabolic stability properties of this group of TTR-inhibitors, in terms of absorption, interaction with different CYP450 proteins, possible sites of metabolism, and other pharmacokinetic and physicochemical parameters that could drive our TTR drug discovery workflow to the design of new molecules based on ADME predictions.

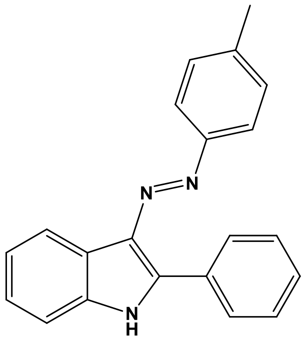
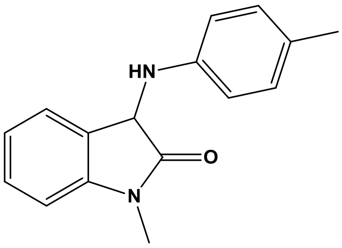
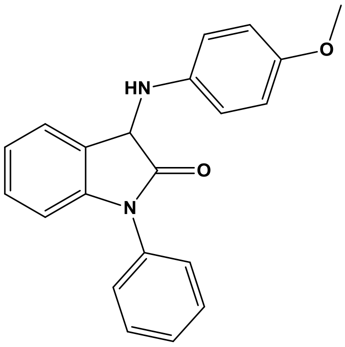
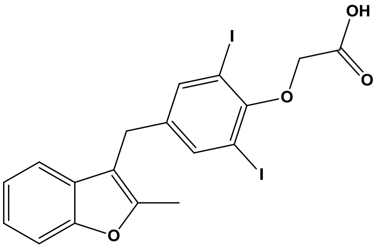
### In silico predictions: MetaSite and CYP’s Consortium Model Software

**Table 1** shows the dataset of studied compounds selected from our in-house database of TTR inhibitors, that have been studied in terms of metabolic resistance using predictive tools and experimental assays developed at Prof. Cruciani’s group at the Università degli Studi di Perugia These 26 compounds have all either reached the market or clinical phases of development (compounds **1** to **13**), or they represent compounds derived from the work of the TTR Consortium group (compounds **14** to **26**) that had shown high activity in the *in vitro* fibril formation inhibition assay.

| Compound Number | Structure | Name* | Compound status |
|---|---|---|---|
| **1** |  | Tafamidis | Drug |
| **2** |  | Diflunisal | Drug |
| **3** |  | Vedaprofen | Drug |
| **4** |  | Lumiracoxib | Drug |
| **5** |  | Tolfenamic | Drug |
| **6** |  | Aceclofenac | Drug |
| **7** |  | Mefenamic | Drug |

| Compound Number | Structure | Name* | Compound status |
|---|---|---|---|
| **8** |  | Meclofenamic | Drug |
| **9** |  | Diclofenac | Drug |
| **10** |  | Flufenamic | Drug |
| **11** |  | Chrysin | Clinical |
| **12** |  | TIP | Clinical |
| **13** |  | Benzothiazol | Preclinical |
| **14** |  | 261 | In vitro hit |

| Compound Number | Structure | Name* | Compound status |
|---|---|---|---|
| **15** |  | 14 | In vitro hit |
| **16** |  | 66 | In vitro hit |
| **17** |  | 480 | In vitro hit |
| **18** |  | AxP224 | In vitro hit |
| **19** |  | 44 | In vitro hit |

| Compound Number | Structure | Name* | Compound status |
|---|---|---|---|
| **20** |  | 30 | In vitro hit |
| **21** |  | 45 | In vitro hit |
| **22** |  | 146 | In vitro hit |
| **23** |  | KB13 | In vitro hit |

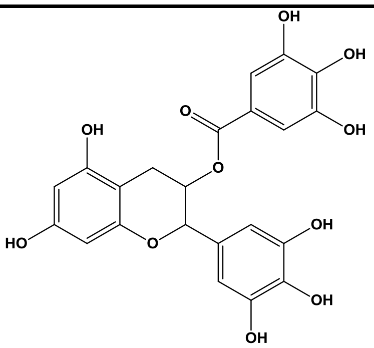
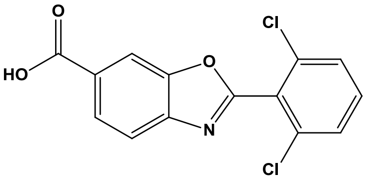
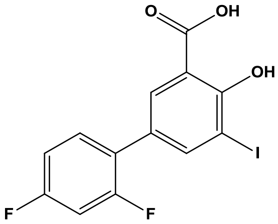
| Compound Number | Structure | Name* | Compound status |
|---|---|---|---|
| **24** |  | EGCG | In vitro hit |
| **25** |  | 477 | In vitro hit |
| **26** |  | IDIF | In vitro hit |

**Table 1.** Set of TTR amyloidogenic inhibitors considered in this study.

(*) For some compounds, the name is the internal numerical code used in our TTR Consortium.

The computational and experimental study presented in this chapter has the aim to answer two questions: a) which regions of our TTR stabilizers molecular skeleton would be targeted and degraded through metabolism?; and b) which CYP isoform will be mainly responsible for these CYP-derived metabolic modifications? The answers to these questions will be the key for the design of new compounds resistant to CYP metabolism and, in addition, it will add value to our drug discovery program on TTR amyloidogenic inhibitors, through the understanding of their metabolic pathways related to the CYPs considered.

The methodology used for this study involves the software package MetaSite 3.0 (site of metabolism (SOM) prediction; http://www.moldiscovery.com), developed by Prof. Gabriele Cruciani and collaborators[5]. As mentioned before, this software is simple to use and it presents a user-friendly interface, but the algorithms used are very complex. All the internal processes are fully automated and the system has pre-computed all the information from the molecular interaction fields calculated for the enzymes. Therefore, only a few seconds are needed to perform the complete calculations when the user introduces the compounds to be analyzed.

In Metasite, two sets of molecular descriptors are computed for the CYP enzymes (already calculated and stored on the system) and for the substrates. For the enzymes, GRID forcefield[12] is used in order to create the molecular interaction fields needed inside the CYP active site. The GRID uses an object (called “Probe”) to measure the electrostatic potential at each point in a given space (box) where the macromolecule (in this case the human CYP) is positioned. It is possible to use different probes that represent chemical groups in order to characterize different macromolecular properties. In addition, by setting the GRID on flexible mode, it is possible to obtain the molecular interaction fields (MIF) needed to construct the set of molecular descriptors of the CYP. This flexible mode allows some side chains of the enzyme to move based on attraction and repulsion effects due to the probe used. After this, all the 3D information compiled is referenced to the catalytic center of the enzyme in order to obtain the CYP fingerprints through the GRIND technology[11]. At the end of this process, each MIF descriptor computed from the probe interaction map is transformed in correlograms representing the distance between the reactive center of the CYP and all the chemical regions mapped inside the active site.

In the case of the molecules (or substrates) introduced on the system, they are modelled taking in consideration a population of different sets of low-energy minimum conformations. These conformations are selected by the interaction fields and the shapes of the CYPs active sites. To compute the molecular descriptors that will characterize the substrates, molecular orbital calculations are initially performed in order to obtain a 3D pharmacophore for the molecule. Next step is the classification of the atoms in the molecules through GRID probes, based on their hydrophobic, hydrogen-bond donor, or hydrogen-bond acceptor capabilities. Finally, all this information is clustered and converted to molecular descriptors that are used in the substrate-CYP comparison phase, where the system is able to define the SOM predicted for the considered compounds.

This comparison is driven by the complementarity between the properties computed for the CYPs and the substrates. For the SOM prediction, the hypothesis applied is the following: the correlation between the reactive center of the enzyme, the position of the atoms in the substrate and the distance between them, will define if a certain position in the chemical structure of the substrate will be a good SOM candidate.

Accessibility and reactivity are also considered during the internal calculations. Accessibility is represented by the recognition between the CYP and substrate when it exposes its atoms to the heme group, using a score function (proportional to this ligand exposure). The reactivity is taken into account thanks to the huge amount of information available of CYP biotransformations[13]. CYP catalyzes oxidative and reductive reactions (oxidative pathways are more frequent). This component is quantified by the system using fast *ab initio* methods.

In summary, through these internal (and automated) processes, the user only needs to introduce the structure of the molecules using a SMILES, 2D SDF or 3D mol2 file and a histogram is then obtained (**Figure 2**) where a ranking based on scoring (each score represents the probability for the sites of metabolism) shows the SOM predicted for the compounds considered.

The focus of this study was the CYP isoforms 2C9, 2D6 and 3A4. Crystal structures corresponding to human isoforms 2C9[14] and 3A4[15] used by the software were extracted from the Protein Data Bank (www.rcsb.org) and the 2D6 structure was provided by DeRienzo et al[16].

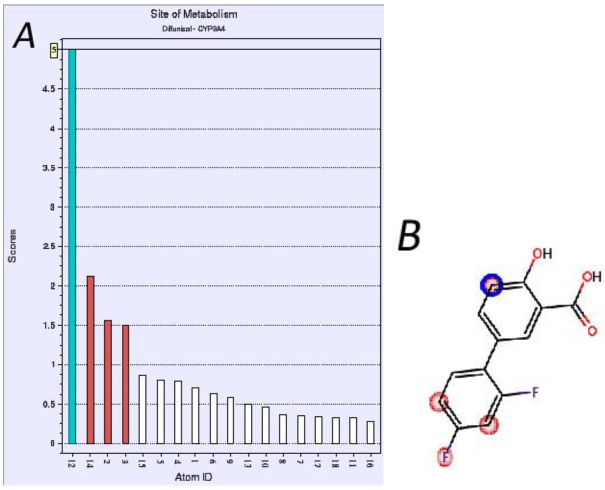
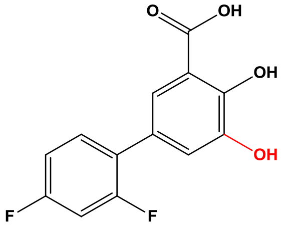
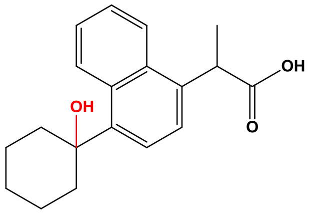
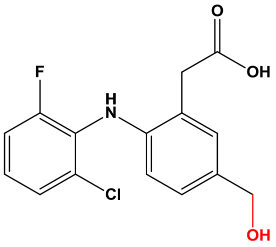
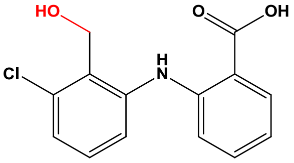
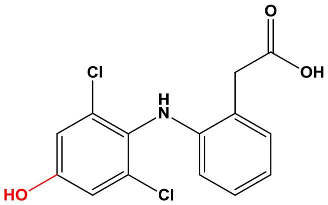
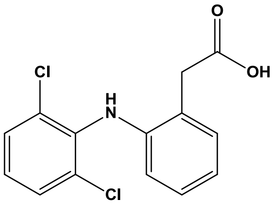
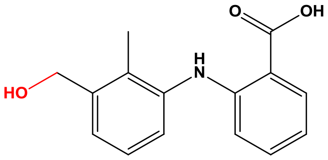


**Figure 2.** Example of MetaSite 3.0 output: A) score histogram representing the probability for the sites of metabolism in each substrate atom; B) SOM predicted for compound Diflunisal into the CYP3A4. Most important SOM's are highlighted in blue.

The 26 selected TTR stabilizing compounds were introduced on the Metasite 3.0 software system and, after running the calculations described, the predicted SOMs and the most likely metabolites for each compound, were obtained. As mentioned before, many of these TTR amyloidogenic inhibitors are drugs already in the market. All these drugs were used to perform a validation set of the prediction, comparing the results obtained from MetaSite with the information extracted from the bibliography about metabolites reported for each one. The results of this simple validation process demonstrate that MetaSite was able to predict, for every drug, the corresponding reported metabolite (**Table 2**)

| Metabolite Structure | Parent compound | MetaSite Score | Reported in |
|---|---|---|---|
|  | Diflunisal | 5,00 | (17) |
|  | Vedaprofen | 7,25 | (18) |
|  O<br>F<br>OH<br>H<br>N<br>Cl<br>OH | Lumiracoxib | 9,61 | (19) |
|  | Tolfenamic | 10,17 | (20) |
|  O<br>Cl<br>OH<br>H<br>N<br>HO<br>Cl | Diclofenac | 9,33 | (20) |
|  | Aceclofenac | 8,33 | (21) |
|  | Mefenamic | 10,44 | (22) |

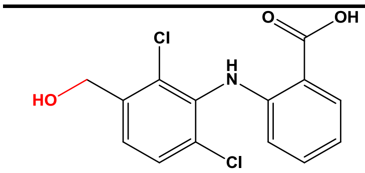
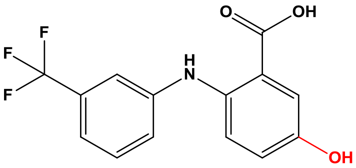
| Metabolite Structure | Parent compound | MetaSite Score | Reported in |
|---|---|---|---|
|  | Meclofenamic | 12,00 | (23) |
|  | Flufenamic | 7,00 | (24) |

**Table 2**. Table of metabolites predicted (using Metasite 3.0) for drugs included in the dataset under analysis, reported on the bibliography. In all the cases, the metabolite that shows the biggest score in Metasite fit in with the most prevalent metabolite experimentally found and reported.

Another method used to define the metabolic pathways of the TTR stabilizing compounds selected, was an internally developed software by Cruciani et al., called CYP Model Consortium. This system uses a technology quite related to MetaSite's in order to compute and predict CYP selectivity for a given substrate. The software has not yet been released and most of the details about the algorithm running inside have not been published. The procedure may be summarized as follows: molecular interaction fields obtained from the different CYPs are used to rank the substrate selectivity for human P450 enzymes based on global pharmacophoric similarity patterns. This system allows to create selectivity tables (**Table 3**), by means of comparing the activity of a given molecule in front of different CYPs.

| Mol-ID | Mol-name | Class-MetStab-HLM | Class-MetStab-2C9 | Class-MetStab-2D6 | Class-MetStab-3A4 | Class-IsofSel-overall |
|---|---|---|---|---|---|---|
| 1 | Tafamidis | HLM_STABLE | 2C9_MEDIUM_STABLE | 2D6_STABLE | 3A4_STABLE | STABLE_3CYPS |
| 2 | 477 | HLM_MEDIUM_STABLE | 2C9_STABLE | 2D6_STABLE | 3A4_STABLE | STABLE_3CYPS |
| 3 | Iododiflunisal | HLM_STABLE | 2C9_STABLE | 2D6_STABLE | 3A4_STABLE | STABLE_3CYPS |
| 4 | Diflunisal | HLM_STABLE | 2C9_STABLE | 2D6_STABLE | 3A4_STABLE | STABLE_3CYPS |
| 5 | 261 | HLM_STABLE | 2C9_STABLE | 2D6_STABLE | 3A4_STABLE | STABLE_3CYPS |
| 6 | 14 | HLM_STABLE | 2C9_STABLE | 2D6_STABLE | 3A4_STABLE | STABLE_3CYPS |
| 7 | 66 | HLM_MEDIUM_STABLE | 2C9_MEDIUM_STABLE | 2D6_STABLE | 3A4_MEDIUM_STABLE | STABLE_3CYPS |
| 8 | MERCK-480 | HLM_MEDIUM_STABLE | 2C9_MEDIUM_STABLE | 2D6_STABLE | 3A4_MEDIUM_STABLE | STABLE_3CYPS |
| 9 | Chrysin | HLM_MEDIUM_STABLE | 2C9_MEDIUM_STABLE | 2D6_STABLE | 3A4_STABLE | MAJOR_2C9 |
| 10 | AxP224 | HLM_MEDIUM_STABLE | 2C9_MEDIUM_STABLE | 2D6_STABLE | 3A4_MEDIUM_STABLE | UNSTABLE_2D6_3A4 |
| 11 | 44 | HLM_MEDIUM_STABLE | 2C9_STABLE | 2D6_MEDIUM_STABLE | 3A4_MEDIUM_STABLE | STABLE_3CYPS |
| 12 | 30 | HLM_MEDIUM_STABLE | 2C9_STABLE | 2D6_STABLE | 3A4_MEDIUM_STABLE | UNSTABLE_2D6_3A4 |
| 13 | 45 | HLM_MEDIUM_STABLE | 2C9_STABLE | 2D6_STABLE | 3A4_MEDIUM_STABLE | MAJOR_3A4 |
| 14 | 146 | HLM_MEDIUM_STABLE | 2C9_STABLE | 2D6_STABLE | 3A4_MEDIUM_STABLE | UNSTABLE_2D6_3A4 |
| 15 | Vedaprofen | HLM_MEDIUM_STABLE | 2C9_MEDIUM_STABLE | 2D6_STABLE | 3A4_STABLE | UNSTABLE_3CYPS |
| 16 | Lumiracoxib | HLM_STABLE | 2C9_MEDIUM_STABLE | 2D6_STABLE | 3A4_STABLE | STABLE_3CYPS |
| 17 | Tolfenamic | HLM_MEDIUM_STABLE | 2C9_MEDIUM_STABLE | 2D6_STABLE | 3A4_STABLE | STABLE_3CYPS |
| 18 | Aceclofenac | HLM_MEDIUM_STABLE | 2C9_STABLE | 2D6_STABLE | 3A4_STABLE | STABLE_3CYPS |
| 19 | KB13 | HLM_MEDIUM_STABLE | 2C9_MEDIUM_STABLE | 2D6_STABLE | 3A4_STABLE | MAJOR_3A4 |
| 20 | TIP | HLM_MEDIUM_STABLE | 2C9_STABLE | 2D6_MEDIUM_STABLE | 3A4_STABLE | STABLE_3CYPS |
| 21 | EGCG | HLM_STABLE | 2C9_STABLE | 2D6_STABLE | 3A4_STABLE | STABLE_3CYPS |
| 22 | Benzothiazol | HLM_MEDIUM_STABLE | 2C9_STABLE | 2D6_STABLE | 3A4_MEDIUM_STABLE | MAJOR_3A4 |
| 23 | Mefenamic | HLM_STABLE | 2C9_MEDIUM_STABLE | 2D6_STABLE | 3A4_STABLE | MAJOR_2C9 |
| 24 | Meclofenamic | HLM_MEDIUM_STABLE | 2C9_MEDIUM_STABLE | 2D6_STABLE | 3A4_STABLE | UNSTABLE_3CYPS |
| 25 | Diclofenac | HLM_MEDIUM_STABLE | 2C9_MEDIUM_STABLE | 2D6_STABLE | 3A4_STABLE | STABLE_3CYPS |
| 26 | Flufenamic | HLM_STABLE | 2C9_MEDIUM_STABLE | 2D6_STABLE | 3A4_STABLE | STABLE_3CYPS |

**Table 3.** Table of CYP selectivity for the 26 TTR inhibitors selected. Human Liver Microsome (HLM), CYP2C9, 2D6 and 3A4 have been checked in order to identify the selectivity of the molecules analyzed. STABLE = molecule is not metabolized by the enzyme; MEDIUM STABLE = the enzyme is able to metabolize the molecules; UNSTABLE = the given molecule is substrate of the enzyme; MAJOR = enzyme that will be responsible for the substrate metabolism.

In the results shown in the **Table 3**, for the 26 TTR stabilizing compounds, it is possible to see that most of them are predicted to be degraded by the CYP isoform 2C9 and, moreover, this sample of the chemical-biology space, focused on the TTR protein, is metabolically quite stable; most of these compounds are considered by the system as "stable", remarking that they are not only interesting for their amyloidosis antifibrillogenic activity, but they also present a promising pharmacokinetic profile.

**In vitro assays: predicted vs. experimental metabolites**

In order to carry out the experimental determination for the metabolites, Human Liver Microsomes (HLM, obtained from patient samples) and CYP isoforms 2C9, 2D6 and 3A4, were used.

The protocol perfomed was the following:

Compounds to be tested were diluted in 10 ml of a mixture $H_2O$:$CH_3CN$ (50:50) to obtain a solution 0.25 mM. In a 1.2 ml tube (from Micronic), 234 µL of buffer (phosphate 0.01 M pH=7.4) was added. After that, 6.3 µl of a solution of HLM UltraPool™ HLM 150 (cod. 452117 from BD Bioscience) **or** 25 µL of human CYP3A4 + Reductase + b5 Supersomes – 0.5 nmole (from BD Bioscience) **or** 25 µL of human CYP2C9 + Reductase + b5 Supersomes – 0.5 nmole (from BD Bioscience) **or** 25 µL of human CYP2D6 + Reductase + b5 Supersomes – 0.5 nmole (from BD Bioscience) were added.

5 µL of the 0.25 mM compound solution prepared previously were added. Finally, 5 µL of NADPH-GS solution (for 1 ml of solution 15 mg NADPH, 13.8 G-6-P, 6 µL of G-6-P dehydrogenase, 1000 µL of $MgCl_2$ 48 mM) were added and the incubation was started.

For each CYP experiment and compound, samples are taken at three different incubation times (0, 15 min and 30 min).

When the incubations were finished, the samples were centrifugated during 10 min, at 4 ºC (10000 r.p.m). Finally each sample was injected on a LC-MS-MS (Agilent, QToF detector): 2 µL of sample, 95:5 $H_2O$-Ammonium formiate/$CH_3CN$, 0.3 mL/min, 9 min.

The compounds analyzed in these experimental assays for the determination of metabolites, from the list of 26 TTR ligands analyzed computationally, were:

- Tafamidis
- Meclofenamic
- Tolfenamic
- Lumiracoxib

- Vedaprofen
- Diflunisal
- Iododiflunisal

The results obtained for each of these compounds, and the analysis of these results, are presented below.

**Tafamidis**

This registered TTR protein stabilizer for the FAP treatment presents the results displayed in **Figure 3**.

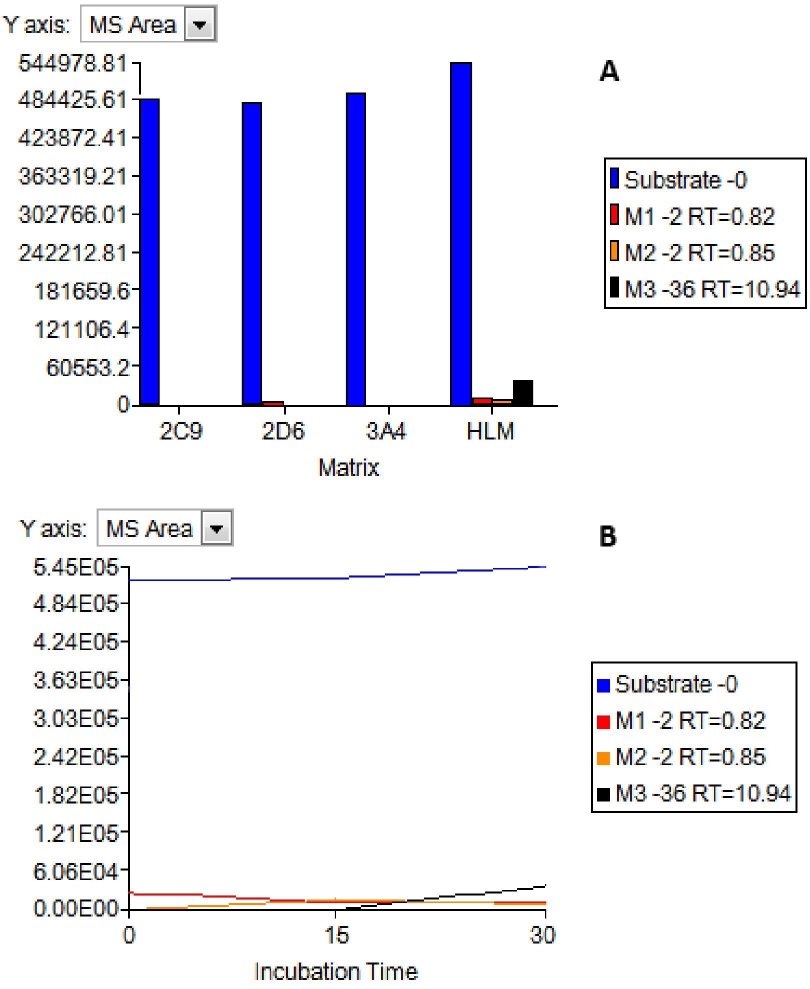


**Figure 3.** (A) Bar plot showing the different species detected among the experiment for all the cytochromes tested and the HLM for Tafamidis. (B) Parent (blue) and metabolites detected concentrations represented against the three times sampled for HLM (0, 15, 30 min).

No relevant metabolite was detected for this compound, Tafamidis, showing a high stability against the three cytochromes tested. Only a small amount of one metabolite (parent mass -36 units) was detected in the HLM experiment (not related with the predicted metabolic profile due to CYP450).

### Meclofenamic

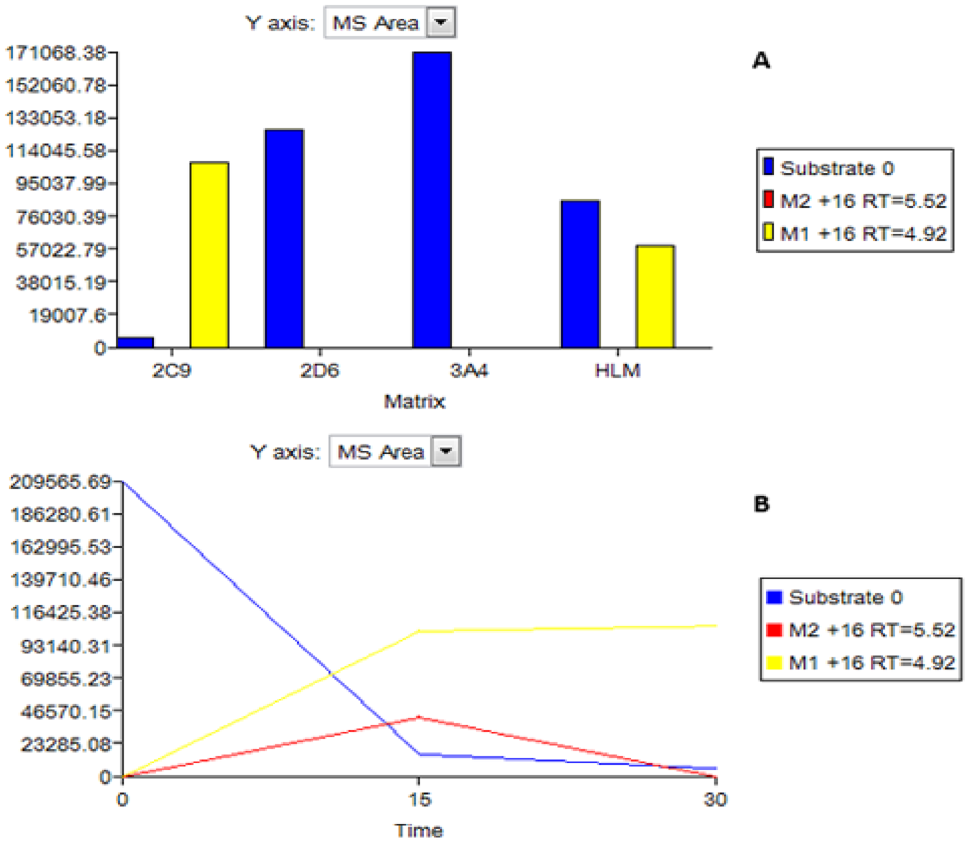


**Figure 4.** (A) Bar plot showing the different species detected among the experiment for all the cytochromes tested and the HLM for Meclofenamic. (B) Parent (blue) and metabolites detected concentrations represented against the three times sampled for 2C9 (0, 15, 30 min).

For this compound (**Figure 4**), Meclofenamic, one metabolite was detected (shown on **Table 2**, 8th entry). It was possible to confirm which Cytochrome P450 was responsible to degrade the parent compound (CYP2C9), and in addition, the *in silico* prediction was also confirmed.

## Tolfenamic

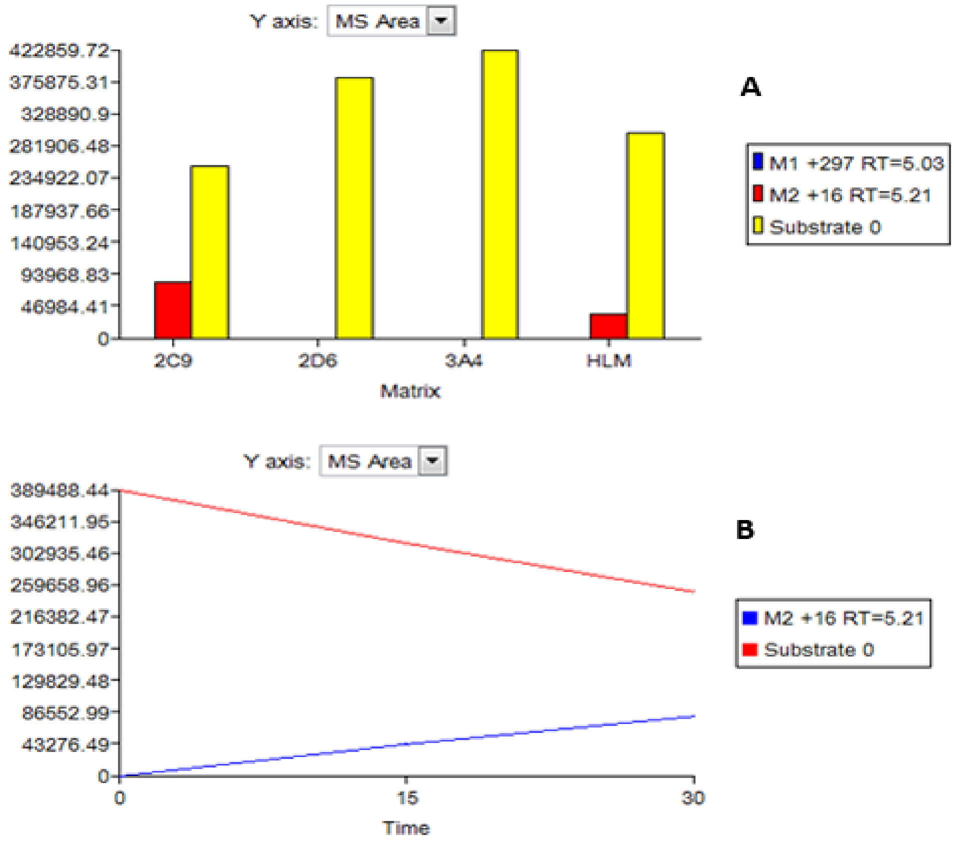


**Figure 5.** (A) Bar plot showing the different species detected among the experiment for all the cytochromes tested and the HLM for Tolfenamic. (B) Parent (red) and metabolite detected concentrations represented against the three times sampled for 2C9 (0, 15, 30 min).

For Tolfenamic (**Figure 5**), one main metabolite was detected (shown on **Table 2**, 4th entry). It was possible to confirm which Cytochrome was responsible to degrade the parent compound (CYP2C9) and in addition, the *in silico* prediction was also confirmed.

### Lumiracoxib

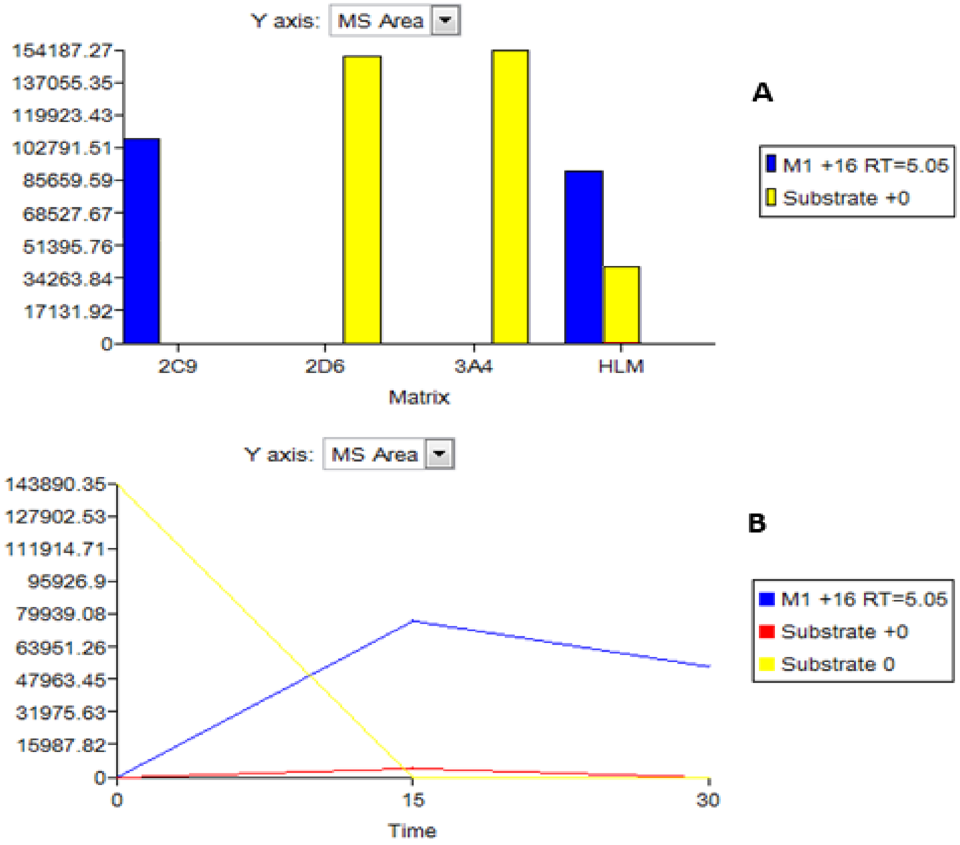


**Figure 6.** (A) Bar plot showing the different species detected among the experiment for all the cytochromes tested and the HLM for Lumiracoxib. (B) Parent compound (yellow) and metabolites detected concentrations represented against the three times sampled for 2C9 (0, 15, 30 min).

Again, the experimental result (**Figure 6**) validates for Lumiracoxib the *in silico* prediction performed with the planned workflow in this chapter; one main metabolite was detected (shown on **Table 2**, 3rd entry). Furthermore, it was possible to identify the CYP responsible for the parent compound metabolism (CYP 2C9).

### Vedaprofen

In this case it was not possible to detect any metabolite due to the impossibility to detect the parent compound (or any metabolite) in the mass spectrometer. For this reason, in this case it was not possible to validate the prediction made via MetaSite.

### Diflunisal & Iododiflunisal

In this pair of molecules (**Figure 7** and **8**), one main metabolite was predicted for Diflunisal (**Table 2**, 1st entry) using the MetaSite package. In the case of Iododiflunisal, no metabolite was predicted due to the protection provided by the crucial position of the iodine atom in the molecular scaffold, conferring metabolic resistance to this promising TTR amyloid inhibitor.

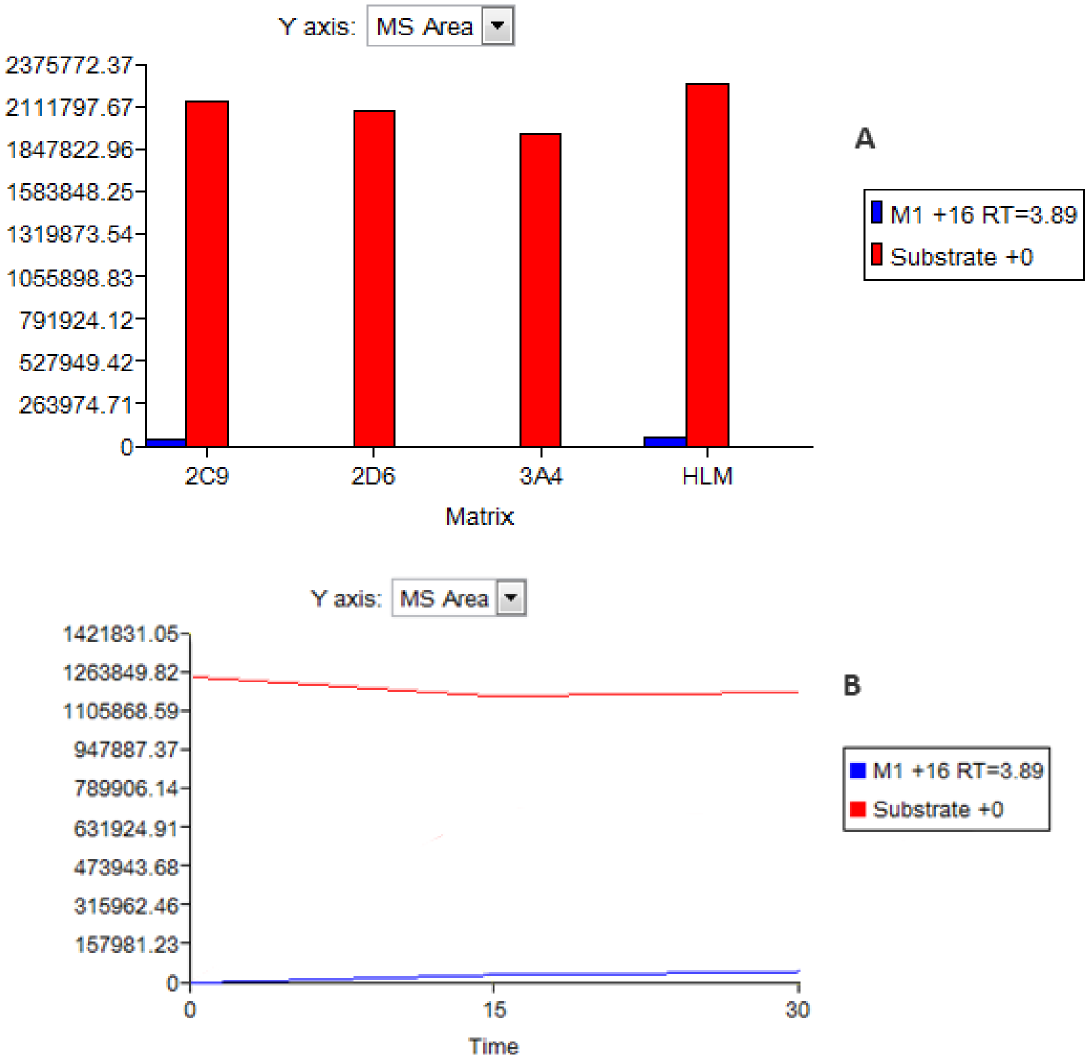


**Figure 7.** (A) Bar plot showing the different species detected among the experiment for all the cytochromes tested and the HLM for Diflunisal. (B) Parent compound (red) and metabolite detected concentrations represented against the three times sampled for 2C9 (0, 15, 30 min).

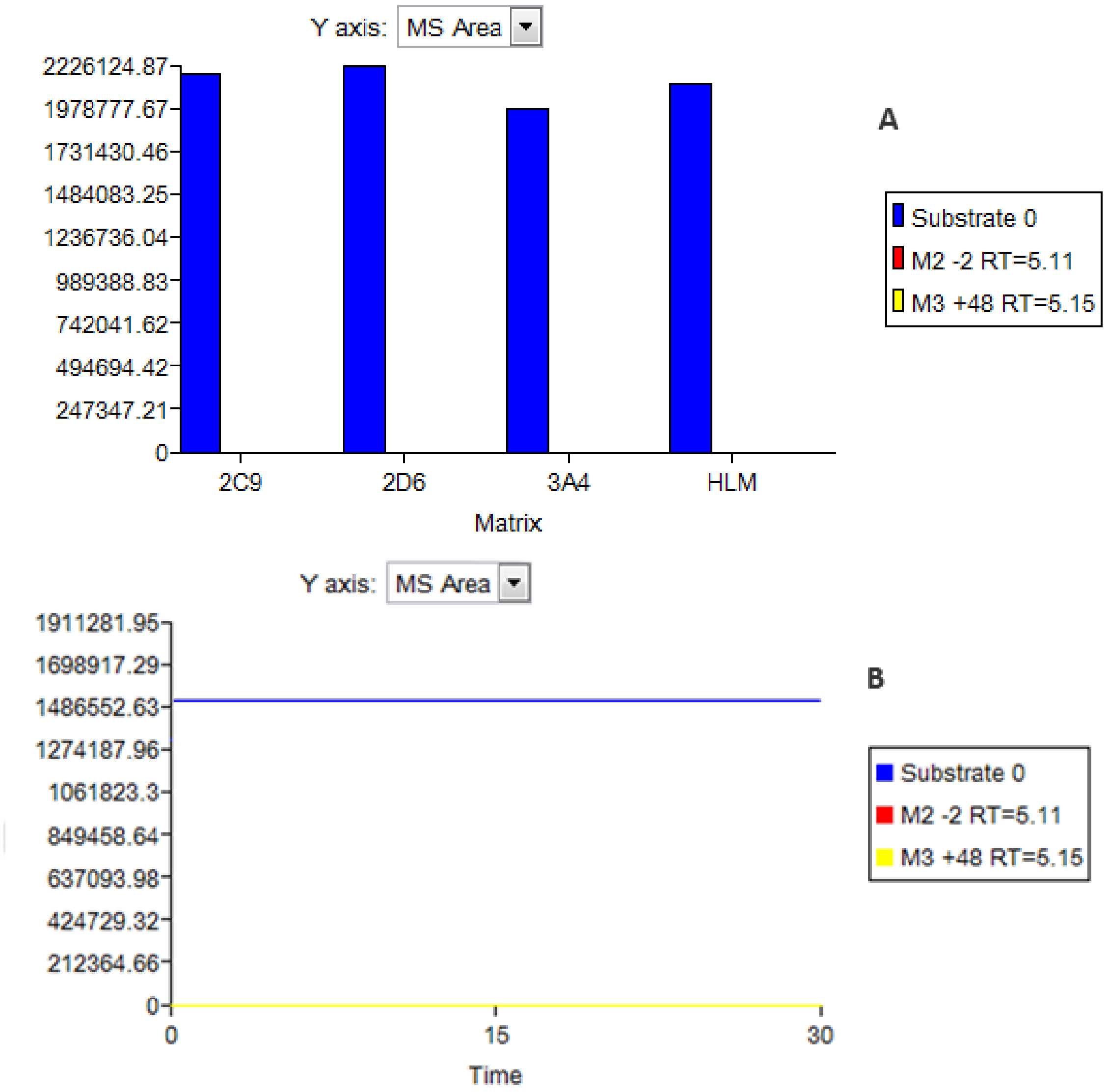


**Figure 8.** (A) Bar plot showing the different species detected among the experiment for all the cytochromes tested and the HLM for Iododiflunisal. (B) Parent compound (blue) and metabolite detected concentrations represented against the three times sampled for 2C9 (0, 15, 30 min).

In summary, for **5** of the compounds tested, it was possible to confirm the predictions obtained from the *in silico* workflow. In addition, Tafamidis and Iododiflunisal were determined as stable compounds from the point of view of Cytochrome P450 metabolism resistance. Furthermore, it can be assumed that Iododiflunisal, a high potent amyloid inhibitor[25], is even more stable than Tafamidis (a registered orphan drug for a TTR-mediated amyloidosis), showing a nice profile for all the cytochromes and HLM tested.

It has been also demonstrated that the MetaSite package is a powerful drug discovery tool in order to predict the most relevant SOM and possible metabolites for a given compound.

Most of the TTR compounds experimentally analyzed in this study have shown isoform selectivity for CYP2C9.

**Further work**

With the aim to take advantage of the predictive/experimental metabolism analysis procedures described in this chapter, a new drug design workflow has been implemented.

As it has been mentioned before, any advance in the pharmacokinetic early optimization analysis will be really important in the drug discovery process. For this reason, considering the metabolic pathways found analyzing the 26 TTR active compounds, a docking experiment on the wtTTR protein structure was performed, using all the metabolites determined using MetaSite for the 26 molecules selected.

The objective was to find any possible metabolite showing a good binding profile against TTR wild type (wtTTR) tridimensional structure model, in order to evaluate new amyloidogenic inhibition pathways derived from each parent compound.

The focus of many drug design processes is usually the optimization of the potency of a given set of lead compounds for a protein target. The optimization of the pharmacokinetic parameters are usually considered in later stages of the process. But a parallel question in the drug design process could be: “can we focus a drug discovery campaign from a different point of view, considering in the first place items like the metabolic stability of the compounds under study?”.

The proposed answer to this question is “yes”, taking into account that with technologies such as the one presented in this chapter, it is possible to predict the metabolic profile for a given compound and, at the same time, to predict which kind of substitutions can be introduced in the molecule to improve its stability vs. the Cytochrome CYP P450 enzymes, and avoid in such manner undesirable degradations of the molecule (resulting in lack of potency or generation of toxic effects associated with the compound metabolites).

Considering this point of view, the set of 26 TTR active compounds were analyzed in order to get analogs with a different chemical structure, through decorations of the compound’s scaffolds. These structural modifications were designed to be stable against the cytochrome-mediated metabolism, and a new workflow to start the lead optimization step for TTR protein has been established, taking as starting point compounds with a high metabolic stability.

Some preliminar docking experiments in the wtTTR protein structure (still unpublished and not yet validated through chemical synthesis/biological activity analysis) have been carried out with these metabolically optimized TTR ligand analogs, and several virtual hits have been obtained, opening new avenues to drug discovery and development of TTR active and metabolically stable compounds for amyloidosis treatment, that could be eventually driven to preclinical and clinical phases.

As a result of this combined computational workflow (prediction of metabolic pathways and design of metabolically stable analogs of TTR ligands + docking of TTR analogs into the TTR protein structure active site) a database of these TTR analogs for most of the active compounds in the TTR Consortium library has been generated. Each entry in this annotated database contains: a) the parent structure of the active TTR ligand and its proposed metabolites; b) the predicted metabolic stability (or liability) of the parent TTR ligand and of its proposed metabolites; and c) the predicted theoretical affinity towards the TTR fibillogenesis assay of the parent TTR ligand and of its proposed metabolites. This annotated database and its data-intensive contents could eventually be prioritized to obtain new compounds that are synthetically accessible, and to carry out experimental activity assays and experimental metabolic stabilization studies of the synthesized compounds, in order to generateoptimized TTR ligands.

## Bibliography

(1) Lu, C.; Li, A. P. *Enzyme Inhibition in Drug Discovery and Development: The Good and the Bad*; John Wiley & Sons, Inc.: Hoboken, NJ, USA, **2009**; p. 854.

(2) Guengerich, F. P. Cytochrome p450 and chemical toxicology. *Chemical Research in Toxicology* **2008**, *21*, 70–83.

(3) Guengerich, F. P. Common and uncommon cytochrome P450 reactions related to metabolism and chemical toxicity. *Chemical Research in Toxicology* **2001**, *14*, 611–650.

(4) Singh, A.; Anand, M. K.; Sharma, R.; Khan, S. P.; Sachan, N. K.; Singh, A. K. Cyp450 Enzyme Mediated Interaction. *Journal of Pharmacy Research* **2012**, *5*, 1896–1901.

(5) Cruciani, G.; Carosati, E.; De Boeck, B.; Ethirajulu, K.; Mackie, C.; Howe, T.; Vianello, R. MetaSite: understanding metabolism in human cytochromes from the perspective of the chemist. *Journal of Medicinal Chemistry* **2005**, *48*, 6970–6979.

(6) Korzekwa, K. R.; Grogan, J.; DeVito, S.; Jones, J. P. Electronic models for cytochrome P450 oxidations. *Advances in Experimental Medicine and Biology* **1996**, *387*, 361 – 9.

(7) de Groot, M. J.; Ackland, M. J.; Horne, V. a; Alex, a a; Jones, B. C. A novel approach to predicting P450 mediated drug metabolism. CYP2D6 catalyzed N-dealkylation reactions and qualitative metabolite predictions using a combined protein and pharmacophore model for CYP2D6. *Journal of Medicinal Chemistry* **1999**, *42*, 4062–70.

(8) Mancy, A.; Broto, P.; Dijols, S.; Dansette, P. M.; Mansuy, D. The Substrate Binding Site of Human Liver Cytochrome. *Biochemistry* **1995**, 10365–10375.

(9) Chen, H.; de Groot, M. J.; Vermeulen, N. P. E.; Hanzlik, R. P. Oxidative N-Dealkylation of p-Cyclopropyl-N,N-dimethylaniline. A Substituent Effect on a Radical-Clock Reaction Rationalized by Ab Initio Calculations on Radical Cation Intermediates. *The Journal of Organic Chemistry* **1997**, *62*, 8227–8230.

(10) Darvas, F. Predicting metabolic pathways by logic programming. *Journal of Molecular Graphics* **1988**, *6*, 80–86.

(11) Pastor, M.; Cruciani, G.; McLay, I.; Pickett, S.; Clementi, S. GRid-INdependent descriptors (GRIND): a novel class of alignment-independent three-dimensional molecular descriptors. *Journal of Medicinal Chemistry* **2000**, *43*, 3233–3243.

(12) Goodford, P. J. A Computational Procedure for Determining Energetically Favorable Binding Sites on Biologically Important Macromolecules. *Journal of Medicinal Chemistry* **1985**, *28*, 849–857.

(13) Rendic, S.; Di Carlo, F. J. Human cytochrome P450 enzymes: a status report summarizing their reactions, substrates, inducers, and inhibitors. *Drug Metabolism Reviews 29*, 413–580.

(14) Wester, M. R.; Yano, J. K.; Schoch, G. a; Yang, C.; Griffin, K. J.; Stout, C. D.; Johnson, E. F. The structure of human cytochrome P450 2C9 complexed with flurbiprofen at 2.0-A resolution. *The Journal of Biological Chemistry* **2004**, *279*, 35630–35637.

(15) Yano, J. K.; Wester, M. R.; Schoch, G. a; Griffin, K. J.; Stout, C. D.; Johnson, E. F. The structure of human microsomal cytochrome P450 3A4 determined by X-ray crystallography to 2.05-A resolution. *The Journal of Biological Chemistry* **2004**, *279*, 38091–38094.

(16) Iid, P.; Iiia, P.; Rienzo, F. D.; Fanelli, F.; Menziani, M. C.; Benedetti, P. G. D.; Chimica, D.; Emilia, R.; Campi, V.; Modena, I.- Theoretical investigation of substrate specificity for cytochromes P450. *Database* **2000**, *14*, 93–116.

(17) Macdonald, J. I.; Dickinson, R. G.; Reid, R. S.; Edom, R. W.; King, a R.; Verbeeck, R. K. Identification of a hydroxy metabolite of diflunisal in rat and human urine. *Xenobiotica* **1991**, *21*, 1521–1533.

(18) Medicines, V.; Unit, E. COMMITTEE FOR VETERINARY MEDICINAL PRODUCTS. *EMEA* **1997**, 1996–1998.

(19) Rordorf, C. M.; Choi, L.; Marshall, P.; Mangold, J. B. Clinical pharmacology of lumiracoxib: a selective cyclo-oxygenase-2 inhibitor. *Clinical Pharmacokinetics* **2005**, *44*, 1247–1266.

(20) Sidelmann, U. G.; Christiansen, E.; Krogh, L.; Cornett, C.; Tjørnelund, J.; Hansen, S. H. Purification and 1H NMR spectroscopic characterization of phase II metabolites of tolfenamic acid. *Drug Metabolism and Disposition* **1997**, *25*, 725–731.

(21) Yamazaki, R.; Kawai, S.; Mizushima, Y.; Matsuzaki, T.; Hashimoto, S.; Yokokura, T.; Ito, A. A major metabolite of aceclofenac , 4 ¢ presses the production of interstitial pro-collagenase / proMMP-1 and pro-stromelysin-1 / proMMP-3 by human rheumatoid synovial cells. *Inflammation Research* **2000**, *49*, 1521–1533.

(22) Sato, J.; Yamane, Y.; Ito, K.; Bando, H. Structures of mefenamic acid metabolites from human urine. *Biological & Pharmaceutical Bulletin* **1993**, *16*, 811–812.

(23) Johansson, I. M.; Anlér, E.-L.; Bondesson, U.; Schubert, B. Isolation of meclofenamic acid and two metabolites from equine urine — a comparison between horse and man. *Journal of Pharmaceutical and Biomedical Analysis* **1986**, *4*, 171–179.

(24) Bowman, R. E.; Brunt, K. D.; Godfrey, K. E.; Kruszynska, L.; Reynolds, A. A.; Thrift, R. I.; Waite, D.; Williamson, W. R. N. Syntheses of Flufenamic Acid Metabolites I and II and Other N-Aryl- anthranilic Acids. *Journal of the Chemical Society, Perkin Transactions 1* **1973**, 1–4.

(25) Mairal, T.; Nieto, J.; Pinto, M.; Almeida, M. R.; Gales, L.; Ballesteros, A.; Barluenga, J.; Pérez, J. J.; Vázquez, J. T.; Centeno, N. B.; Saraiva, M. J.; Damas, A. M.; Planas, A.; Arsequell, G.; Valencia, G. Iodine atoms: a new molecular feature for the design of potent transthyretin fibrillogenesis inhibitors. *PloS One* **2009**, *4*, e4124.

# CONCLUSIONS

The main conclusions of this thesis, associated to the proposed objectives, are:

a) A chemico-biological database containing the chemical structures and biological activities of around 800 transthyretin (TTR) ligands has been implemented.

b) A computational workflow to obtain TTR ligand fingerprints has been developed, and the application of this workflow to the repurposing of marketed antiinflammatory drugs has delivered **3** compounds as new TTR stabilizers.

c) A computational workflow to obtain a TTR-protein structure based pharmacophore has been developed, and the application of this workflow to a database of flavonoid compounds has delivered **one** compound as a new TTR stabilizer.

d) A computational workflow to obtain Ligand Efficiency Indices (LEIs) retrospective analysis of the TTR Chemico-Biological Space (CBS) has been developed, and its application to the TTR Consortium database has allowed the evaluation of the efficiency of TTR ligands in the different TTR chemical structural classes.

e) A computational workflow of a TTR-protein structure based ligand docking protocol has been developed, and its application combined with the LEIs retrospective workflow has allowed the design of new more efficient TTR ligands structurally related to the most active TTR ligand, Iododiflunisal.

f) A combined computational/experimental workflow for the analysis of the metabolic stability of TTR ligands, has been developed, and its application to compounds with TTR affinity (including drugs and a selected subset from our database) has allowed to rationalize their pharmacokinetic and metabolic properties, and the design of new derivatives with optimized metabolic stability.

g) A chemico-biological database containing the chemical structures of TTR active ligands, and of their predicted metabolites, as well as the theoretical affinities of all these compounds through the TTR docking protocol, has been implemented.

# RESUMEN

## 1. Introducción

### 1.1 Amiloidosis

Las amiloidosis son enfermedades asociadas a la alteración o daño producido en las funciones de un órgano determinado debido a la acumulación de depósitos extracelulares, conocidos como amiloides. Estos depósitos son producidos por la desorganización de una proteína (o fragmentos de ella) que ha perdido su plegamiento natural, perdiendo por lo tanto su función.

Actualmente, se han identificado unas 20 proteínas humanas como precursoras de amiloides relacionadas con una amplia variedad de enfermedades. La proteína Transtirretina (TTR) es una de ellas.

### 1.2. Transtirretina

La proteína TTR es una proteína homotetramérica que tiene como función primaria transportar la hormona tiroxina (T4) a través de la barrera hematoencefálica, y actúa de transportadora de emergencia de dicha hormona T4 en plasma. Así mismo también es la principal transportadora de la vitamina A mediante la formación de un complejo con la *Retinol-Binding Protein*.

La proteína TTR presenta un comportamiento amiloidogénico, presentando más de 100 mutaciones distintas asociadas a diversas amiloidosis hereditarias. El mecanismo de agregación y de formación de los depósitos amiloides no se conoce con detalle, pero está aceptado que el proceso global consta de varias etapas: 1) cambios conformacionales en la forma monomérica de la proteína; 2) agregación de estos monómeros modificados para formar distintas especies oligoméricas (no fibrilares y solubles); 3) formación de protofibras y fibrillas y, por último, 4) elongación y formación de las fibras maduras (que posteriormente acaban formando los depósitos amiloides en forma de placas)[1]. Algunas de las enfermedades raras que son causadas por este proceso son, por ejemplo: la polineuropatía familiar amiloide (FAP), la cardiomiopatía familiar amiloide (FAC), la amiloidosis selectiva del sistema nervioso central (CNSA) o la amiloidosis senil sistémica (SSA).

Actualmente, el trasplante de hígado es la única terapia disponible para FAP y FAC, con resultados no siempre positivos y en muchos casos insuficiente o totalmente inefectiva en los casos de mutantes de TTR relacionados con el sistema nervioso central, donde también se sintetiza la proteína.

Basada en la hipótesis de la estabilización cinética del tetrámero usando ligandos con alta afinidad por el centro de unión de la proteína, que sustituyan a la hormona y estabilicen el tetrámero, se ha desarrollado una nueva aproximación terapéutica que se muestra prometedora y ha demostrado ser capaz de prevenir la disociación inicial de la proteína, requerida para la formación de las fibras.

Gracias a la colaboración y participación de un grupo multidisciplinar, ha sido posible obtener y evaluar la actividad biológica de 800 compuestos para la estabilización cinética del homotetrámero de TTR. Tras el estudio de diversas estructuras cristalográficas depositadas en el Protein Data Bank (www.rcsb.org), ha sido posible demostrar que el centro de unión de TTR presenta 3 pares de depresiones llamadas "Halogen Binding Pockets" (HBP) que son capaces de reconocer selectivamente átomos de yodo en frente de otros halógenos[2]. Teniendo en cuenta este concepto y sabiendo que muchos antiinflamatorios no esteroides (NSAID) son capaces de unirse y estabilizar la proteina TTR[3,4], se han podido obtener familias de derivados de NSAIDs yodados, con afinididades muy elevadas por la proteina y con una capacidad estabilizadora muy prometedora.

### 1.3. Objetivos de la presente tesis

Los principales objetivos de este proyecto de investigación están centrados en el diseño de nuevos (y mejores) estabilizadores de TTR. Para ello, se han puesto a punto las siguientes metodologías:

a) Generar una base de datos químico-biológica que compila todos los compuestos generados por el Consorcio de TTR (tanto los históricos creados a lo largo del proyecto iniciado en el año 2000, como los nuevos sintetizados) y que contiene todas las estructuras y actividades biológicas relativas a la proteína TTR.

b) Explorar la posibilidad de utilizar técnicas de *repurposing* aplicadas al descubrimiento de nuevos inhibidores de la disociación de TTR partiendo de fármacos ya existentes, centrando el estudio en compuestos anti-inflamatorios.

c) Diseñar nuevos compuestos flavonoides capaces de unirse a TTR y estabilizarla mediante diseño basado en estructura.

d) Incorporar el análisis mediante el uso de los *Ligand Efficiency Indices* (LEI) como nueva herramienta para diseñar nuevos compuestos con una eficiencia optimizada para la proteína TTR.

e) Desarrollo computacional de una estrategia combinada (predicitva y experimental) basada en el estudio de la estabilidad metabólica de un grupo de ligandos de TTR, como posible herramienta para la optimización de compuestos de nuestra base de datos interna en base a la obtención de mejores propiedades farmacocinéticas.

## 2. *Repurposing* basado en Fingerprints

Desde hace años, las compañías farmacéuticas has aumentado la inversión de forma significativa en el proceso de descubrimiento y optimización de fármacos, aplicando nuevas técnicas para ello, como por ejemplo, el diseño de compuestos basado en la estructura de las proteínas diana, química combinatoria, *high-throughput screening* (HTS), etc. Sin embargo, el impacto de estas mejoras no ha sido lo suficientemente significativo en el aumento de la eficiencia del proceso de descubrimiento de fármacos.

Por ello, se han consolidado otras estrategias para reducir el tiempo y dinero invertido en el proceso. Una de ellas es el reposicionamiento de fármacos (o *repurposing*).

Esta aproximación se basa en buscar una nueva indicación terapéutica para un fármaco existente en el mercado (o retirado del mismo por cualquier razón, excepto por motivos de seguridad para el paciente). El hecho de conocer el perfil farmacocinético y de seguridad de un fármaco en el mercado, hace que sea posible su desarrollo para otra indicación reduciendo la inversión necesaria, en un proceso de desarrollo del fármaco potencialmente más rápido y con riesgos mínimos en cuanto a seguridad y toxicidad.

Con el objetivo de aplicar este tipo de metodología para el descubrimiento de compuestos estabilizadores de la proteína TTR con posibilidades de llegar al mercado, se ha establecido una estrategia basada en ligandos utilizando un modelo de *fingerprints* para ello.

Los *fingerprints* son vectores 2D/3D que contienen información de la presencia o ausencia de determinadas propiedades o elementos en una molécula.

Usando este tipo de vectores puede codificarse información relativa a elementos topológicos de una molécula, conectividad química y conformaciones tridimensionales.

Esta metodología nos ha permitido realizar una búsqueda basada en la similaridad con 3 compuestos estabilizadores de TTR muy potentes sobre un grupo de 120 fármacos en el mercado, extraídos de *Prous Science Integrity DataBase* (*http://integrity.prous.com/*)[5,6].

Para llevar a cabo este análisis se utilizaron 3 tipos distintos de *fingerprints*: MACCS (basados en la estructura de los compuestos a analizar), TAF (basados en la tipología de los átomos presentes) y GpiDAPH3 (basados en los elementos farmacofóricos presentes), obteniendo 9 potenciales candidatos, de los cuales, tras la evaluación *in vitro*, 3 compuestos que han llegado al mercado para otras indicaciones (Lumiracoxib, Tolfenamic y Aceclofenac), son capaces de estabilizar el tetrámero e inhibir la formación de fibras, confirmando de este modo, las predicciones basadas en la similaridad y demostrando la utilidad de la aproximación basada en el *repurposing*.

## 3. Modelización basada en farmacóforo

Un farmacóforo se define como el conjunto de elementos estéricos y electrónicos necesarios para asegurar una interacción intermolecular óptima de un compuesto con una diana biológica concreta, de forma que se bloquee o se active una determinada respuesta. La modelización basada en esta tecnología permite el hallazgo de nuevos compuestos con un perfil de actividad interesante para la diana terapéutica en cuestión.

Teniendo en cuenta que es bien conocido que muchas flavonas (y derivados)[7,8,9] muestran una gran afinidad hacia TTR y son capaces de estabilizar el tetrámero *in vitro*, y que dichos tipos de compuestos se unen preferentemente al canal de unión de la proteína, se ha llevado a cabo una aproximación de modelización basada en farmacóforo para encontrar flavonas que puedan ser de utilidad para la estabilización del homotetrámero de TTR.

Una base de datos de flavonas y flavonoides disponibles comercialmente, extraída del proyecto MMsINC[10], se ha analizado usando un modelo farmacofórico construido a partir de una estructura cristalográfica siguiendo los siguientes pasos: a) identificación y extracción del farmacóforo relativo a la estructura escogida (PDB ID: 1KGJ, http://www.rcsb.org); b) análisis conformacional de los compuestos seleccionados (usando algoritmos de dinámica molecular, lowMD); c) clasificación de los confórmeros en base a su encaje dentro del modelo farmacofórico obtenido; d) visualización y selección de los *hits* virtuales obtenidos[11].

Siguiendo esta metodología, se han podido priorizar 4 nuevos compuestos con estructura flavonoide, candidatos a ser nuevos inhibidores de la amiloidogénesis asociada a TTR. Tras el análisis *in vitro*, se ha podido confirmar que uno de los compuestos adquiridos y ensayados, posee una actividad antiamiloidogénica moderada.

## 4. Ligand Eficiency Indices (LEIs)

El concepto de eficiencia de un ligando (LE), que relaciona la energía de enlace (ΔG) de un compuesto con la proteina de interés, con el tamaño de dicho ligando (el número de átomos exceptuando hidrógenos) fue introducido por Hopkins et al. en 2005[12]. El uso de los LE que asocia la afinidad (expresada en forma de $K_i$, $IC_{50}$ o similares) con el tamaño molecular (peso molecular, número de átomos, NHEA o número de átomos pesados, etc) está bien establecido, y dos nuevas definiciones de índices de eficiencia de ligandos (LEIs o *Ligand Efficiency Indices*) se introdujeron en 2005: BEI (índice de eficiencia de enlace, que establece una relación entre la potencia y el peso molecular de una determinada molécula) y SEI (índice de eficiencia de superficie, que relaciona la potencia de una molécula con su área superficial polar, PSA). La combinación de ambos índices para el análisis retrospectivo del espacio químico-biológico, ha demostrado ser una poderosa herramienta para el estudio de trayectorias de optimización en proyectos de descubrimiento de fármacos[13,14].

Con el objetivo de optimizar los compuestos recogidos en la bases de datos disponibles en el consorcio (tanto sintetizados como adquiridos), evaluados *in vitro* contra la proteína TTR, se ha explorado una nueva vía, basada en este concepto de los índices de eficiencia de ligandos, para la obtención de nuevos inhibidores de amiloidosis con mayor eficiencia, centrada en un área concreta del espacio químico-biólogico de la proteina TTR. Para ello se ha seleccionado una familia muy prometedora de inhibidores de amiloidogénesis, ya publicados[5], que ha sido incorporada a un mapa bidimensional, usando las representaciones de los LEIs (nBEI vs NSEI), combinando las 3 variables cruciales: potencia, peso molecular y PSA. Los mapas bidimensionales obtenidos permiten llevar a cabo un análisis retrospectivo de la serie de compuestos en detalle y, así mismo, permiten sugerir posibles nuevos caminos de optimización para obtener posibles candidatos más eficientes como ligandos TTR.

Además, usando herramientas computacionales (*docking*, modelos de farmacóforo y cálculos teóricos de $K_i$) en combinación con los Indices de Eficiencia de Ligando, se ha podido establecer una metodología prospectiva que permite idear nuevos compuestos, teóricamente, más eficientes y por lo tanto candidatos más firmes para llegar a fases avanzadas del proceso de descubrimiento de fármacos (preclínica o fases clínicas)[15]. De este estudio *in silico*, tras obtener una lista de propuestas sintéticas asumibles, se han obtenido 4 compuestos activos, siendo la mayoría incluso más eficientes que las moléculas que se han utilizado como punto de referencia del estudio.

## 5. Estudios *in silico* de Estabilidad Metabólica

Como se ha comentado con anterioridad, como parte de un proyecto de investigación multidisciplinar, y con el objetivo principal de obtener más y mejores inhibidores amiloidogénicos para la proteina TTR, se han sintetizado alrededor de 800 moléculas, compiladas todas ellas en una base de datos de creación propia.

Con la intención de añadir valor a algunos de nuestros candidatos más prometedores y poder plantear un proyecto de diseño y optimización de compuestos basados en la mejora de las propiedades farmacocinéticas de nuestros compuestos, y en colaboración con el grupo del Profesor Gabriele Cruciani de la Università degli Studi di Perugia, en Italia, se ha llevado a cabo un estudio tanto computacional como experimental del perfil farmacocinético de un subgrupo de 26 inhibidores de TTR de diversas familias químicas. El objetivo de este estudio es el de poder establecer un modelo predictivo de las propiedades fisicoquímicas, farmacocinéticas y de la estabilidad metabólica de este grupo de moléculas escogidas, en términos de absorción, interacción con las diferentes isoformas de los citocromos CYP450 a estudiar, posibles posiciones de metabolismo en la estructura de los ligandos TTR, y otros parámetros relacionados.

### 5.1. Estudio *in silico*

Los 26 compuestos escogidos para llevar a cabo el estudio presentan una buena actividad inhibidora de fibrilogénesis. Muchos de ellos han alcanzado fases clínicas avanzadas para otras indicaciones o son fármacos en el mercado.

Para llevar a cabo las predicciones y estimaciones sobre el posible perfil metabólico de los compuestos, se ha utilizado el paquete de software MetaSite[16]. Dicho programa permite al usuario visualizar el perfil metabólico completo de una determinada molécula e identificar las posiciones metabólicas (en inglés SOM o "*Sites of Metabolism*") más probables. La metodología utilizada se basa en la comparación de dos grupos de descriptores moleculares, uno para caracterizar el citocromo P450 que se esté considerando y otro para el substrato potencial (es decir, la molécula introducida en el sistema por el usuario). Ambos grupos de descriptores crean la "huella digital" que representa al enzima y al substrato, respectivamente.

Tras estos cálculos se puede estimar y definir la accesibilidad del substrato al grupo hemo de la enzima CYP450 que se está considerando. Este proceso permite al sistema hacer una predicción de cuáles serán las posiciones metabólicas (SOM) más probables basada en la integración de una función de probabilidad correlacionada con la energía libre del proceso completo.

También se ha llevado a cabo una predicción de la selectividad de un determinado substrato hacia un grupo de diversas isoformas de citocromos, gracias a otro software (CYP Consortium Models) desarrollado también en el grupo de investigación del Dr. Cruciani. Este programa permite evaluar cuales serán las isoformas responsables de degradar un compuesto a través de un proceso y unos algoritmos muy similares a los utilizados en MetaSite.

Tras este doble filtro, se ha podido establecer una primera priorización de los compuestos en base a la estabilidad metabólica teórica. Teniendo en cuenta el hecho de que, algunos de los compuestos seleccionados son fármacos o moléculas en fase clínicas avanzadas, se ha llevado a cabo una comparación entre la información obtenida tras las predicciones efectuadas, y la información recopilada de la bibliografía sobre los metabolitos aislados y caracterizados, permitiendo establecer un pequeño grupo de validación para las estimaciones de estabilidad metabólica *in silico*.

Tras analizar toda la información recopilada de esta primera fase del estudio, se han realizado los experimentos de estabilidad metabólica con 7 de los compuestos analizados: Tafamidis (el primer fármaco lanzado al mercado para el tratamiento de una de las amiloidosis de TTR), Meclofenámico, Tolfenámico, Lumiracoxib, Vedaprofeno, Diflunisal y Yododiflunisal.

### 5.2. Estudio *in vitro*

Con el objetivo de confirmar y validar el estudio computacional realizado, se han llevado a cabo los ensayos de estabilidad metabólica desarrollados en el grupo del Dr. Cruciani, consistentes en la incubación a diferentes tiempos de los compuestos con las diversas isoformas de los citocromos (en este caso 2C9, 2D6 y 3A4). Tras las mencionadas incubaciones, se han inyectado en un equipo de cromatografía líquida conectado a un detector MS-MS, de las distintas muestras que proporcionaron los datos necesarios para determinar que metabolitos se han formado en cada caso. Además, al realizar los ensayos a distintos tiempos, ha sido posible llevar a cabo un tipo de representación gráfica muy similar a una cinética de reacción donde ha sido posible ver la evolución del compuesto “padre” y sus diversos metabolitos “hijos” a lo largo de la ventana de tiempo seleccionada.

Para llevar a cabo los experimentos de estabilidad metabólica con los inhibidores seleccionados, se ha seguido el protocolo detallado a continuación:

El compuesto a analizar se diluye en 10 ml de una mezcla $H_2O$:$CH_3CN$ (50:50) para obtener una solución de concentración 0.25 mM.

En un tubo Micronic de 1.2 ml, se añaden 234 µL de solución tampón (fosfato 0.01 M pH=7.4). A continuación se añaden 6.3 µl de una solución de HLM UltraPool™ HLM 150 (cod. 452117 de BD Bioscience) **o** 25 µL de CYP3A4 + Reductasa + b5 Supersomes – 0.5 nmol (BD Bioscience) **o** 25 µL de CYP2C9 + Reductasa + b5 Supersomes – 0.5 nmol (BD Bioscience) **o** 25 µL de CYP2D6 + Reductasa + b5 Supersomes – 0.5 nmol (BD Bioscience).

Seguidamente, 5 µL de la solución de compuesto 0.25 mM preparada anteriormente se adicionan a la mezcla enzimática y se activa la reacción con 5 µL de solución de NADPH-GS (para 1 ml de solución, 15 mg NADPH, 13.8 G-6-P, 6 µL of G-6-P deshidrogenasa, 1000 µL of $MgCl_2$ 48 mM). La mezcla resultante se incuba a 37 ºC durante 0, 15 y 30 minutos.

Cuando se finaliza la incubación, se centrifugan las muestras durante 10 min, a 4 ºC (10000 r.p.m) y cada muestra es inyectada en un equipo de cromatografía LC-MS-MS (2 µL de muestra, 95:5 $H_2O$-formiato de amonio/$CH_3CN$, 0.3 mL/min, 9 min).

Los resultados preliminares obtenidos han permitido confirmar que el uso de herramientas predictivas destinadas a la evaluación y determinación de la estabilidad metabólica que permiten obtener moléculas resistentes a la degradación metabólica, podría llegar a ser una estrategia de optimización de compuestos con actividad estabilizadora para TTR a tener en cuenta en futuros proyectos relacionados con esta proteína y las enfermedades relacionadas con ella. Así mismo, los resultados obtenidos en el estudio experimental de la degradación metabólica, han confirmado las predicciones realizadas a lo largo del estudio teórico y han permitido la apertura de una nueva estrategia basada en la optimización *a priori* de la estabilidad metabólica de compuestos que puedan ser estabilizadores del tetrámero de TTR y aplicables para el tratamiento de amiloidosis relacionadas con la proteína, de forma más efectiva y rápida.

## 6. Conclusiones

Tras los estudios realizados a los largo de este proyecto de tesis, las principales conclusiones asociadas a los objetivos propuestos son:

a) Se ha creado una base de datos que contiene alrededor de 800 ligandos de la proteína TTR.
b) Se ha desarrollado una metodología computacional basada en el uso de fingerprints y se ha aplicado al reposicionamiento (*repurposing*) de fármacos anti-inflamatorios en el mercado, obteniendo 3 nuevos compuestos capaces de estabilizar la proteína.
c) Se ha desarrollado una metodología computacional basada en el uso de farmacóforos. Dicha estrategia se ha aplicado sobre una base de datos de compuestos tipo flavonoides, permitiendo la identificación de un nuevo estabilizador de TTR.
d) Se ha llevado a cabo un análisis retrospectivo del espacio químico-biológico de la proteína TTR, usando los *Ligand efficiency índices* (LEIs). Ello ha permitido evaluar la eficiencia de los ligandos conocidos para TTR.
e) Se ha desarrollado una nueva aproximación al diseño de inhibidores de amiloidogénesis a través de un protocolo de *docking* combinado con la tecnología de los LEIs. Ello ha permitido diseñar nuevos y más eficientes estabilizadores de TTR basados en uno de los compuestos más potentes descritos (Yododiflunisal).
f) Se ha desarrollado una metodología combinada (computacional/experimental) para el análisis de la estabilidad metabólica de un grupo de ligandos de TTR

(incluyendo entre ellos fármacos y otros compuesto de interés). Esto ha permitido racionalizar sus propiedades farmacocinéticas (en términos de metabolismo) y diseñar nuevo derivados con una estabilidad optimizada.

g) Se ha creado una base de datos que contiene las estructuras de los ligandos activos seleccionados para este estudio, junto a sus metabolitos determinados de forma computacional y experimental, así como las afinidades teóricas (calculadas en base a un protocolo de *docking*).

## Bibliografía

(1) Foss, T. R.; Wiseman, R. L.; Kelly, J. W. The pathway by which the tetrameric protein transthyretin dissociates. *Biochemistry* **2005**, *44*, 15525–15533.

(2) Blaney, J. M.; Blaney, J. M.; Jorgensen, E. C.; Connolly, M. L.; Ferrin, T. E.; Langridge, R.; Oatley, S. J.; Burridge, J. M.; Blake, C. C. F. Computer graphics in drug design: molecular modeling of thyroid hormone-prealbumin interactions. *Journal of Medicinal Chemistry* **1982**, *25*, 785–790.

(3) Munro, S. L.; Lim, C.-F.; Hall, J. G.; Barow, J. W.; Craik, D. J.; Topliss, D. J.; Stockigt, J. R. Drug Competition for Thyroxine Binding to Transthyretin (Prealbumin): Comparison with Effects on Thyroxine-Binding Globulin. *Journal of Clinical Endocrinology & Metabolism* **1989**, *68*, 1141–1147.

(4) Adamski-Werner, S. L.; Palaninathan, S. K.; Sacchettini, J. C.; Kelly, J. W. Diflunisal analogues stabilize the native state of transthyretin. Potent inhibition of amyloidogenesis. *Journal of Medicinal Chemistry* **2004**, *47*, 355–374.

(5) Mairal, T.; Nieto, J.; Pinto, M.; Almeida, M. R.; Gales, L.; Ballesteros, A.; Barluenga, J.; Pérez, J. J.; Vázquez, J. T.; Centeno, N. B.; Saraiva, M. J.; Damas, A. M.; Planas, A.; Arsequell, G.; Valencia, G. Iodine atoms: a new molecular feature for the design of potent transthyretin fibrillogenesis inhibitors. *PloS One* **2009**, *4*, e4124.

(6) Razavi, H.; Palaninathan, S. K.; Powers, E. T.; Wiseman, R. L.; Purkey, H. E.; Mohamedmohaideen, N. N.; Deechongkit, S.; Chiang, K. P.; Dendle, M. T. a; Sacchettini, J. C.; Kelly, J. W. Benzoxazoles as transthyretin amyloid fibril inhibitors: synthesis, evaluation, and mechanism of action. *Angewandte Chemie (International ed. in English)* **2003**, *42*, 2758–2761.

(7) Baures, P. W.; Peterson, S. a; Kelly, J. W. Discovering transthyretin amyloid fibril inhibitors by limited screening. *Bioorganic & Medicinal Chemistry* **1998**, *6*, 1389–1401.

(8) Green, N. S.; Foss, T. R.; Kelly, J. W. Genistein, a natural product from soy, is a potent inhibitor of transthyretin amyloidosis. *Proceedings of the National Academy of Sciences of the United States of America* **2005**, *102*, 14545–14550.

(9) Maia, F.; Rosa, M.; Kijjoa, A.; Pinto, M. M. Saraiva M.J.; Damas A.M. The binding of xanthone derivatives to transthyretin. *Biochemical Pharmacology* **2005**, *70*, 1861–1869.

(10) Masciocchi, J.; Frau, G.; Fanton, M.; Sturlese, M.; Floris, M.; Pireddu, L.; Palla, P.; Cedrati, F.; Rodriguez-Tomé, P.; Moro, S. MMsINC: a large-scale chemoinformatics database. *Nucleic Acids Research* **2009**, *37*, 284–290.

(11) D. Blasi, M. Pinto, J. Nieto, G.Arsequell, G. Valencia, A. Planas, N.B. Centeno, J. Q. Drug discovery targeted at transthyretin cardiac amyloidosis: rational design, synthesis, and biological activity of new transthyretin amyloid inhibitors. *Amyloid* **2011**, *18*, 55–57.

(12) Hopkins, A. L.; Groom, C. R.; Alex, A. Ligand efficiency: a useful metric for lead selection. *Drug Discovery Today* **2004**, *9*, 430–431.

(13) Abad-Zapatero, C.; Perišić, O.; Wass, J.; Bento, a P.; Overington, J.; Al-Lazikani, B.; Johnson, M. E. Ligand efficiency indices for an effective mapping of chemico-biological space: the concept of an atlas-like representation. *Drug Discovery Today* **2010**, *00*, 1–8.

(14) Abad-Zapatero, C.; Blasi, D. Ligand Efficiency Indices (LEIs): More than a Simple Efficiency Yardstick. *Molecular Informatics* **2011**, *30*, 122–132.

(15) Blasi, D.; Arsequell, G.; Valencia, G.; Nieto, J.; Planas, A.; Pinto, M.; Centeno, N. B.; Abad-Zapatero, C.; Quintana, J. Retrospective Mapping of SAR Data for TTR Protein in Chemico-Biological Space Using Ligand Efficiency Indices as a Guide to Drug Discovery Strategies. *Molecular Informatics* **2011**, *30*, 161–167.

(16) Cruciani, G.; Carosati, E.; De Boeck, B.; Ethirajulu, K.; Mackie, C.; Howe, T.; Vianello, R. MetaSite: understanding metabolism in human cytochromes from the perspective of the chemist. *Journal of Medicinal Chemistry* **2005**, *48*, 6970–6979.

# ANNEXES

## List of publications

(1) D. Blasi, M. Pinto, J. Nieto, G.Arsequell, G. Valencia, A. Planas, N.B. Centeno, J. Q. Drug discovery targeted at transthyretin cardiac amyloidosis: rational design, synthesis, and biological activity of new transthyretin amyloid inhibitors. *Amyloid* **2011**, *18*, 55–57.

(2) Pinto, M.; Blasi, D.; Nieto, J.; Arsequell, G.; Valencia, G.; Planas, A.; Quintana, J.; Centeno, N. B. Ligand-binding properties of human transthyretin. *Amyloid* **2011**, *18*, 51–54.

(3) Abad-Zapatero, C.; Blasi, D. Ligand Efficiency Indices (LEIs): More than a Simple Efficiency Yardstick. *Molecular Informatics* **2011**, *30*, 122–132.

(4) Blasi, D.; Arsequell, G.; Valencia, G.; Nieto, J.; Planas, A.; Pinto, M.; Centeno, N. B.; Abad-Zapatero, C.; Quintana, J. Retrospective Mapping of SAR Data for TTR Protein in Chemico-Biological Space Using Ligand Efficiency Indices as a Guide to Drug Discovery Strategies. *Molecular Informatics* **2011**, *30*, 161–167.

## List of PDB crystallographic structures

*1ICT* Wojtczak, A.; Neumann, P.; Cody, V. Structure of a new polymorphic monoclinic form of human transthyretin at 3 Å resolution reveals a mixed complex between unliganded and T 4 -bound tetramers of TTR. *Acta Crystallographica Section D Biological Crystallography* **2001**, *57*, 957–967.

*3CFN* Lima, L. M. T. R.; Silva, V. D. A.; Palmieri, L. D. C.; Oliveira, M. C. B. R.; Foguel, D.; Polikarpov, I. Identification of a novel ligand binding motif in the transthyretin channel. *Bioorganic & Medicinal Chemistry* **2010**, *18*, 100–110.

*2F8I* Razavi, H.; Palaninathan, S. K.; Powers, E. T.; Wiseman, R. L.; Purkey, H. E.; Mohamedmohaideen, N. N.; Deechongkit, S.; Chiang, K. P.; Dendle, M. T. a; Sacchettini, J. C.; Kelly, J. W. Benzoxazoles as transthyretin amyloid fibril inhibitors: synthesis, evaluation, and mechanism of action. *Angewandte Chemie (International ed. in English)* **2003**, *42*, 2758–2761.

*2QGC* Johnson, S. M.; Connelly, S.; Wilson, I. a; Kelly, J. W. Biochemical and structural evaluation of highly selective 2-arylbenzoxazole-based transthyretin amyloidogenesis inhibitors. *Journal of Medicinal Chemistry* **2008**, *51*, 260–270.

*3ESO* Johnson, S. M.; Connelly, S.; Wilson, I. a; Kelly, J. W. Toward optimization of the second aryl substructure common to transthyretin amyloidogenesis inhibitors using biochemical and structural studies. *Journal of Medicinal Chemistry* **2009**, *52*, 1115–1125.

*3B56* Gales, L.; Almeida, M. R.; Arsequell, G.; Valencia, G.; Saraiva, M. J.; Damas, A. M. Iodination of salicylic acid improves its binding to transthyretin. *Biochimica et Biophysica Acta* **2008**, *1784*, 512–517.

*1KGJ* Muzioł, T.; Cody, V.; Wojtczak, A. Comparison of binding interactions of dibromoflavonoids with transthyretin. *Acta Biochimica Polonica* **2001**, *48*, 885–892.

*1OG2* Williams, P. A.; Cosme, J.; Ward, A.; Angove, H. C.; Jhoti, H. Crystal structure of human cytochrome P450 2C9 with bound warfarin. *Nature* **2003**, *424*, 464–468.

*1R9O* Wester, M. R.; Yano, J. K.; Schoch, G. a; Yang, C.; Griffin, K. J.; Stout, C. D.; Johnson, E. F. The structure of human cytochrome P450 2C9 complexed with flurbiprofen at 2.0 Å resolution. *The Journal of Biological Chemistry* **2004**, *279*, 35630–35637.

*1W0E* Williams, P.; Cosme, J.; Vinkovic, D. M.; Ward, A.; Angove, H. C.; Day, P. J.; Vonrhein, C.; Tickle, I. J.; Jhoti, H. Crystal structures of human cytochrome P450 3A4 bound to metyrapone and progesterone. *Science* **2004**, *305*, 683–686.