DEPARTAMENT OF LIFE SCIENCES

FACULTY OF SCIENCE AND TECHNOLOGY
COIMBRA UNIVERSITY

High-Resolution Saturation Transfer Difference (STD) ^{1}H NMR techniques and Docking simulations applied to Protein-ligand interactions

David Miguel G. F. Dias

2011

DEPARTAMENT OF LIFE SCIENCES

FACULTY OF SCIENCE AND TECHNOLOGY
COIMBRA UNIVERSITY

High-Resolution Saturation Transfer Difference (STD) ^{1}H NMR techniques and Docking simulations applied to Protein-ligand interactions

Dissertation presented to the Coimbra University, as a requirement to obtain the degree of Master of Science in Biochemistry, conducted under the scientific supervision of Professor Doctor Carlos Geraldes (Coimbra University).

David Miguel G. F. Dias

Acknowledgements

I want to thank Professor Carlos Geraldes for the given opportunity to explore and work on the NMR field in these last years, first as an undergraduate student and secondly as a M.Sc. student. Thank you for all the discussions and guidance. Thank you for all personal and professional contribute in my life, you are a true inspiration.

To João M. C. Teixeira, for all the initial support, all the hours of patience and all the knowledge transmitted, and the long weekends on the spectrometer.

I want to thank Professors David Parker, Mauro Botta and Jesús Jiménez-Barbero for all the collaborative projects and the inherent contribution to my thesis.

To Dr. Margarida Castro, for the good atmosphere, and for all the deposited trust in my, present and future, work and in myself.

For reasons that not need to be mentioned here, I thank Rui Traquete, João Couto, Gonçalo Pereira, Marco Castro, João Marques, Diogo Marques, Mário Pereira and Isaura Martins. A very special thanks to Ricardo Davide, Alexandre Fonseca and Bruno Martins.

To Liliana, for all the love and respect.

Finally, I want to thank my family and closest friends for all the understanding and support on my long last weekends away, and for having me as I truly am. You all know: “that’s the fun I have!”

“*No man is so foolish but he may sometimes give another good counsel, and no man so wise that he may not easily err if he takes no other counsel than his own. He that is taught only by himself has a fool for a master.”*

Hunter S. Thompson

Table of Contents

LIST OF ABREVIATIONS

NMR, Nuclear Magnetic Resonance

MRI, Magnetic Resonance Imaging

CA, contrast agent

STD, Saturation Transfer Difference

HSA, Human Serum Albumin

SAR, structure activity relationship

QSAR, quantitative structure activity relationship

HSQC, Heteronuclear Single Quantum Correlation

gCOSY, gradient Correlated Spectroscopy

GEM, Group Epitope Mapping

BOPTA, 1-(tert-butylamino)-3-(3,4-dihydro-2H-thiochromen-8-yloxy)propan-2-ol hydrochloride

DTPA-Cholate, diethylenetriamine pentaacetic-Cholate

NAPHTO-EGTA, ethylene glycol tetraacetic acid-NAPHTO

WaterLOGSY, Water Ligand Observed Spectroscopy

tr-NOE, transferred Nuclear Overhauser Effect

DPFGSE, Double Pulse Field Gradient Spin Echo

K_D, dissociation constant

A_{STD}, STD amplification factor

High-Resolution Saturation Transfer Difference (STD) ^{1}H NMR techniques and Docking simulations applied to Protein-ligand interactions

Abstract The study of ligand–receptor interactions using high-resolution NMR techniques, namely the saturation transfer difference (STD), is presented for the recognition process between different complexes. From diagnostic to therapeutic complexes, these approaches were used for direct determination of positive binding, followed by the estimation of the correspondent dissociation constant (K_D), GEM characterization and site preference by competitive experiments. To get a three-dimensional view of the interaction similar docking simulations were performed for all the ligands.

The enantioselective binding of the (SSS)-Δ isomer of an yttrium (III) tetraazatriphenylene complex to 'drug-site II' of human serum albumin (HSA) was detected by the intensity differences of its STD ^{1}H NMR spectrum relative to the (RRR)-Λ isomer, by the effect of the competitive binder to that site, N-dansyl sarcosine.

Three MRI contrast agents and the K_D for this interaction with HSA was found for all complexes using the initial slopes protocol. La(BOPTA) and La(DTPA-Cholate) interact mainly with HSA´s drug site I while La(NAPHTO) does not posses a specificity for both drug sites. La(DTPA-Cholate) CA analogue was found to have a lower K_D due to the full lodging of its Cholate moiety in HSA drug site I.

Insulin mimetic VO(DMPP) complexes preferred binding to 'drug-site I' of HSA was also detected by STD ^{1}H NMR by the displacement effect of the competitive binder to that site, warfarin.

The protocol of STD and competitive STD followed by docking simulations was found to be a robust tool to characterize protein-ligand interactions.

Keywords: protein-ligand interaction – STD NMR – docking simulations – competitive STD

Transferência de Saturação por Diferença (TSD) usando RMN de ^{1}H de alta resolução e simulações de Docking aplicadas a interações Ligando-Proteína

Resumo O estudo das interações ligando-receptor através de RMN de alta resolução tornou-se muito mais acessível e com um conjunto de aplicações mais vasto com o desenvolvimento da técnica de Transferência de Saturação por Diferença (TSD). O processo de transporte e reconhecimento de possiveis agentes de diagnóstico e terapia foram caracterizados pela sua constante de dissociação (K_D), pelo seu mapeamento de epitopo e a preferência pelo local de ligação à Albumina do Soro Humano (ASH) realizado por TSD competitiva. Um modelo tri-dimensional foi estabelecido para todas as interacções de modo a complementar a informação recolhida por RMN.

A ligação enantioseletiva de dois complexos de Yttrium revelou a preferência do isómero (SSS)-Δ relativamente ao isómero (RRR)-Λ. Usando um competidor especifico para o local II da ASH e recorrendo a TSD competitivo foi estabelecido como lugar de ligação o este mesmo local, reforçando a sua seletividade.

Três agentes de contraste para Ressonância Magnética foram caracterizados usando o protocolo de TSD e o método dos parciais iniciais foi aplicado para a obtenção de constantes de dissociação (K_D). O La(BOPTA) e o La(DTPA-Colato) foram selectivos para o local I, enquanto que o La(NAFTO-EGTA) não parece apresentar especificidade para nenhum dos locais de ligação estudados. O La(DTPA-Colato) devido a conseguir interagir pela porção do Colato com a ASH apresenta uma maior afinidade para o local que o La(BOPTA).

O insulino-mimético HDMPP foi também estudado e assume-se que este possa ser transportado ligado ao local I da ASH, e que todas as suas espécies em solução são capazes de interagir com a ASH.

Este protocolo desenvolvido baseia-se na aquisição de espetros de TSD usando RMN, com e sem competição especifica para o local, finalizado por simulações de docking apresenta-se aqui como uma ferramenta bastante útil para estudar interações de pequenos ligandos e proteínas.

Palavras chave: interação ligando-proteína – TSD usando RMN – simulações de docking– TSD competitivo

1. INTRODUCTION

1. INTRODUCTION

Over the past 25 years, advances in structural biology have profoundly changed the way in which drugs are designed and discovered. Today, rather than treating protein targets as simple black boxes, pharmaceutical scientists are now treating them as complex molecular entities that possess well-defined structures with active sites that can be rationally activated or deactivated with small molecule ligands. Innovations in structure determination methods along with rapid advances in molecular visualization tools have led to the emergence of structure-aided drug design or rational drug design as an integral part of the drug discovery and development process. [1-4]

Protein interactions with ligands, other proteins, or surfaces are controlled by a complex array of intermolecular processes. Such interactions depend both on the specificity in the binding site as well as the non-specific forces outside the binding pocket. This interplay of specific and non-specific forces controls all protein interactions ranging from bimolecular collisions in solutions to adhesion between cells. The complexity of interactions between proteins and flexible target molecules, including other proteins, nucleic acids and small molecules, is often determined by the considerable flexibility of the protein binding sites and by the structural rearrangements that occur upon binding of the associated molecule or ligands. [5,6]

A goal of many biophysical studies is to determine the molecular forces that control biological interactions and to use this information to rationally manipulate protein structure or function by modifying the protein, the interacting ligand, or both.

The forces that control protein behavior and their physical biochemical origins are inferred from thermodynamic parameters and equilibrium binding kinetic measurements. Calculated energies are used to identify the role of the physical and chemical interactions in protein function and behavior. Although detailed calculations are feasible for small molecules, such calculations become limited as the size and complexity of the biological macromolecules increase. Time-dependent forces between soft or mobile species add yet another degree of complexity, while static models of interactions cannot describe the full range of parameters that influence biological behavior. [5-7]

An accurate understanding, at the molecular level of several biological and biochemical processes is indispensable to deal with the specific mechanisms of life itself. Many techniques may enhance this desirable goal, such as x-ray analysis, *in vivo* experiments and NMR-based methods, among others. [2, 8-11]

NMR methods used in this context can be divided into two categories; protein based, and ligand based experiments. Recently, there has been a strong emphasis on so-called ligand-based because these methods are capable of yielding good results versus time ratio, and outputs more advantages than the protein based protocols, that may require isotopically labeled protein samples. [12-14]

Ligand-based methods are attractive because they are broadly applicable, impose few constraints on the composition of the target protein and don't require isotopic labeling of the protein or ligands. Such experiments include diffusion experiments, saturation transfer difference (STD-NMR), NOE pumping, waterLOGSY (water Ligand-Observed via Gradient SpectroscopY), SALMON (Solvent Accessibility, Ligand binding, and Mapping of ligand Orientation by NMR spectroscopy), transferred-NOE (tr-NOE) and INPHARMA. Ligand-based NMR methods have been employed in screening and in lead optimization to a broad spectrum of biological contexts. One key advantage arises from their capability to pick up specific interactions for compounds of relatively low affinity and their ability to provide limited structural information without any need of crystallization or isotopic labeling. [15,16]

The saturation transfer difference (STD) is one of these techniques that screen ligands for binding activity to proteins. The adoption of this model inherits excellent advantages since the total amount of protein needed is very low, in the μM range, it is relatively easy to implement, and also can be associated with other NMR sequences, and can be applied to large molecular weight diagnostic and therapeutic targets. The type of data possible to obtain from this technique includes dissociation constant estimation and group epitope mapping (GEM). Being a very versatile experiment it can be used almost in all problems that need this kind of data unveiled. [17-20]

One of these fields is Magnetic Resonance Imaging (MRI), where these approaches might resolve most of the challenges on improving the contrast agent-protein interactions. [21,22]

MRI is a non-invasive diagnostic technique used to provide anatomical and functional images of the human body which can use paramagnetic complexes able to enhance the image contrast by increasing the nuclear relaxation rates of water protons in the surrounding tissues, differing to other contrast agents (CA's) used in other clinical imaging studies which are directly imaged. Gadolinium (III) polyaminocarboxyl chelates have been under development for their late use, prior to scan, of administrable contrast agents. The use of this ion with seven unpaired electrons, nine coordination sites and favorable electronic relaxation times bound to a ligand non-covalent binding to Human Serum Albumin (HSA) became an issue of great relevance when studying new possible contrast agents for Magnetic Resonance Angiography (MRA) once the lifetime of CA in blood increases when this interaction occurs. [21-23] Given this feature, and being Albumin the most abundant soluble protein in the body of all vertebrates and the most prominent protein in plasma, its ligand and protein binding study has become a powerful tool for the process. [21]

Most of the compounds interacting with HSA are hydrophobic organic anions and the search for new CAs for angiographic applications has been addressed to Gd (III) complexes endowed with lipophilic substituents and residual negative charge. [23]

The full scope of this interaction is completed with virtual ligand screening using *in silico* methods. The growing importance of structure-based drug design is due to a wide range of relatively fast answers, as it can provide prospective leads and also is a practical alternative or complement to high-throughput screening of large compound libraries and, even, affinities of the distinct ligands can be predicted correctly. The docking and scoring problems countered in this endeavor are central to the theory of biomolecular interactions and are ultimately determined by the nature of the underlying binding energy landscape. However, once the desired synergy of adequate conformational sampling combined with accurate evaluation of energetics has been difficult to achieve with any computational models a constant effort is being carried out. [24,25,26]

Docking software, like Autodock Vina, may build a valid 3D model of the interaction used to confirm or discard previous collected data. The premise that a biological function can be described by a similarity in binding sites and ligand interactions with proteins of known function is the basis for the outlined protocol

in this year of work. The resulting co-structure and functional assignment may provide a starting point for a drug discovery effort. [25]

Complete information about the ligand binding to a specific protein region with a quantitative analysis (that may be obtained by STD-NMR) and a 3D model (obtained by docking simulations) fully characterizes the ligand-binding phenomena. Not only a given relationship is unveiled but also similar binding modes may reduce the drug discover effort and timeline.

The studied HSA binding diamagnetic analogues of contrast agents were: La(BOPTA), La(DTPA-Cholate), La(NAFTO-EGTA) (Fig. 1).

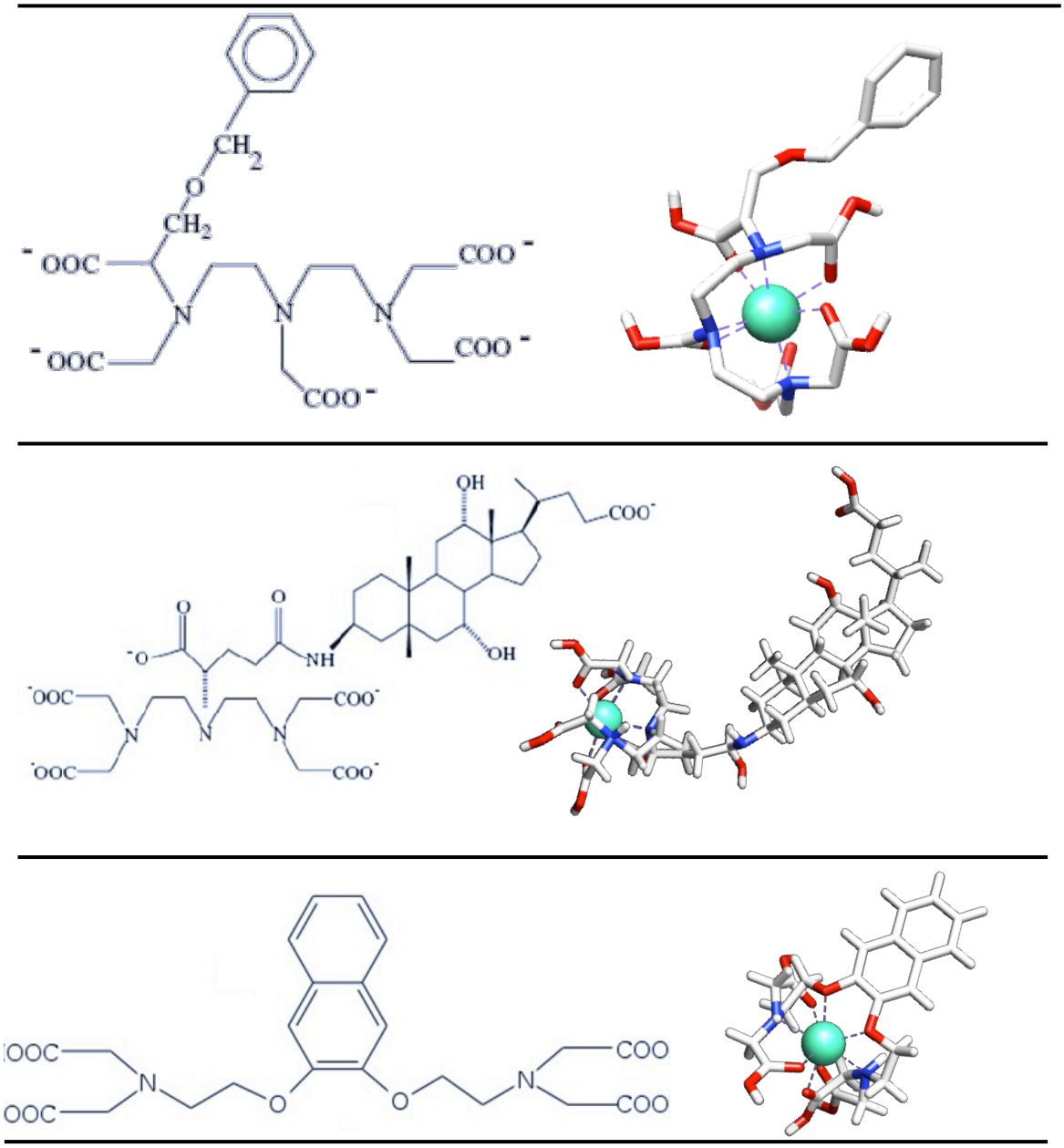


Figure 1. Chemical structures of the studied ligands and 3D view for the La(III) complexes: 1 – BOPTA, 2 – DTPA-Cholate and 3 – NAPHTO-EGTA.

The study of a case of enantiselective binding to HSA was performed with two complexes the RRR-(Λ)-[Y.L^1]$^{3+}$ and SSS-(Δ)-[Y.L^1]$^{3+}$ (Fig. 2) to verify if STD-NMR can distinguish between them and if so where does this type of complex binds using also a competitive assay with N-dansyl sarcosine.

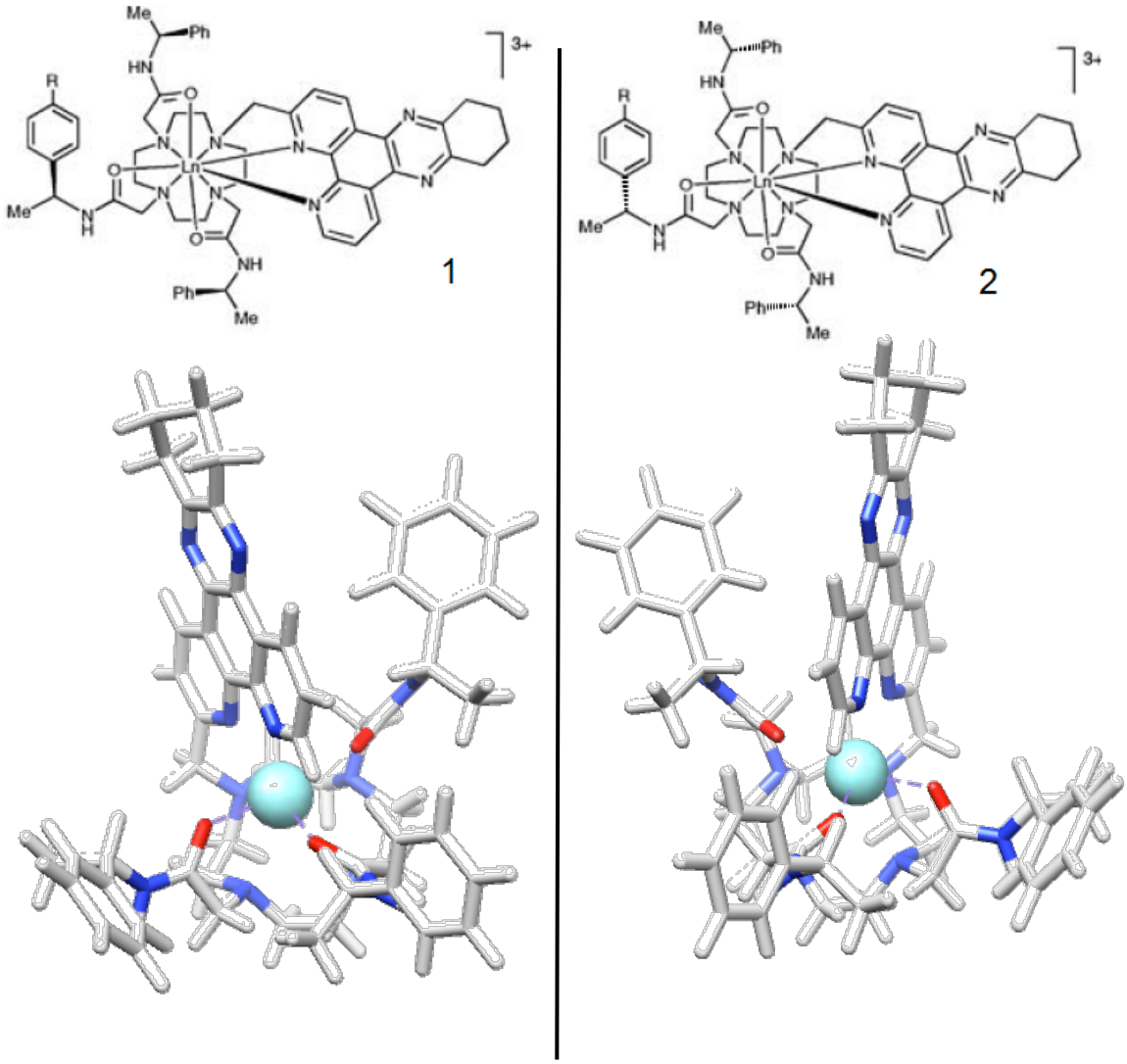


Figure 2. Chemical structures of the studied complexes. 1 – SSS-(Δ)-[Y.L^1]$^{3+}$and 2 – RRR-(Λ)-[Y.L^1]$^{3+}$.

A promising possible insulin mimetic V(IV) complex binding to HSA is also going to be studied to see how it is transported in the blood stream in a HSA-binding model.

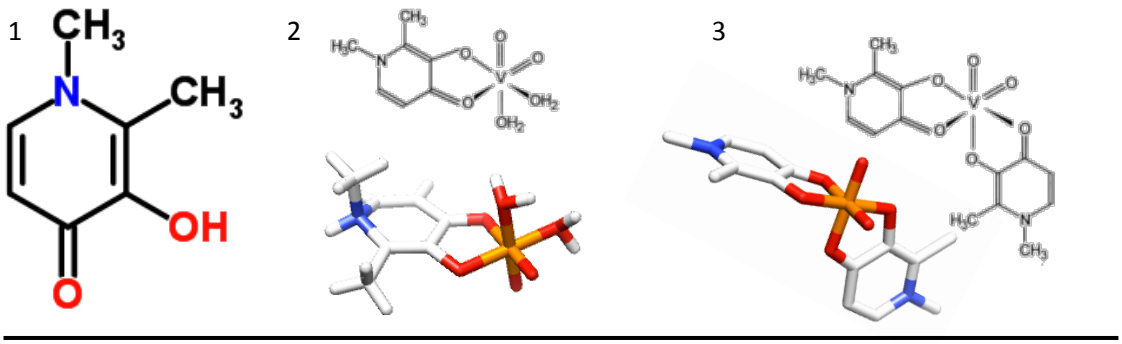
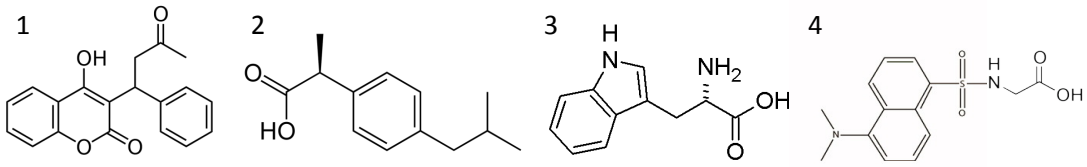


Figure 3. Chemical structures for the insulin mimetic and all solution species. 1 – HDMPP, 2 – 1:1 $VO_2(DMPP)(H_2O)(OH)$ and 3 - 1:2 $VO_2(DMPP)_2$.

The study of all the interaction between the compounds and HSA and where the binding occurs will be performed with the STD-NMR protocol and with the competitive assays using known binders to the HSA binding sites reported by Sudlow [27] (Scheme 1), and later verified using Autodock Vina and Poseview software [28,29] for 3D and 2D models of the interaction, respectively.

In the present work a complete explanation of the STD technique is going to be documented with all the known possible approaches and the interpretation of the given sequence. A few brief technical considerations, according to the spectrometer and software used, are posted in Materials and Methods section.



Scheme 1. Scheme for all binders used for competition. 1- Warfarin, 2- Ibuprofen, 3- L-Tryptophan and 4- N-dansyl sarcosine.

2. OBJECTIVES

2. Objectives

The first objective of this work is to study if STD-NMR can distinguish between the enantiomeric complexes SSS-(Δ)-$[Y.L^1]^{3+}$ and RRR-(Λ)-$[Y.L^2]^{3+}$ in their binding to HSA. Knowing that recent crystallographic studies of HSA complexes have suggested that “drug site II” is the most stereo-differentiating protein binding site, in human serum albumin, we use a known binder for that site, N-dansyl sarcosine, to perform competitive STD experiments.

The second objective of this work was to determine the interaction of selective contrast agents with HSA, providing enough data to study the binding behavior of the ligands, the establishment of the binding epitopes, to disclose the favorable binding site, and to estimate their affinity constants using the binding isotherm of STD initial growth rates method. The CA analogues can be summarized as being: La(BOPTA), La(DTPA-Cholate) and La(NAPHTO-EGTA). To achieve the objective of binding site definition, four known inhibitors were used to, primarily, clarify the binding site of the selected compound. (Warfarin, Ibuprofen and L-Tryptophan)

The third objective was to characterize the interaction of insulin mimetic $VO(DMPP)_2$ complex with is main carrier in the blood stream, HSA.

The fourth, and overall, objective was, after the NMR conclusions, the obtention of a three-dimensional model for the interaction. Since the protein used presents a huge variety of structures already available to any user, suitable models were easy to select.

With the fulfillment of the proposed objectives we intend to implement a valid, and relatively fast, protocol for detecting protein-ligand interactions. Using diagnostic and therapy agents, the objectives represent an ambitious and complete characterization method.

3. THEORETICAL BACKGROUND

3. THEORETICAL BACKGROUND

The characterization of events between a ligand and its receptor is a powerful tool to the development of new drug-based researches and the possibility of knowing binding strength, kinetics, competitive binding, as well as allosteric effects made this long way a lot shorter.

Nowadays the conventional bioassay-based high-throughput screening (HTS) is very popular for the development of alternative and complementary tools. The NMR-based methods have been appearing, in this last years, as very robust to emulate the outputs in new drug discovery, on a broad spectrum of biological processes like antibodies, peptide inhibitors, enzyme substrates, RNA and DNA-protein interactions among many others. [30-32]

When studying the interaction between a ligand and its receptor, or between proteins, many difficulties may arise at any moment, such the large amount of possible ligand candidates differing from each other, and how to gather all the information needed to choose the ones that really matter.

This section has the purpose of briefly reporting the main NMR-based techniques used to study ligand-receptor interactions namely, Water-Ligand Observed via Gradient Spectroscopy (WaterLOGSY), Tranferred NOE's, Structur Activity Relationship (SAR) by NMR and Saturation Transfer Difference (STD) – NMR and how they may overcome some difficulties of detection and characterization of ligand binding processes. [33]

3.1. WATERLOGSY

Popular among NMR-based screening techniques, the WaterLOGSY experiment was first reported by Dalvit et. al [34,35] and its name is derived from Water Ligand Observed via Gradient Spectroscopy.

The WaterLOGSY experiment allows the discrimination between binding and nonbinding ligands in a mixture and consists in the labeling of the bound ligand with a saturation profile. The saturation label is transferred to the binding components of the mixture via bulk water.

Initially, two ways to affect the water magnetization were proposed: a selective saturation and a selective inversion of the solvent magnetization, which

can be achieved by two distinct methods, either by the WEXII proposed by Mori et al. or by the ePHOGSY spin echo scheme proposed by Dalvit. As confirmed by recent literature the ePHOGSY methods become more efficient. [34,38]

In order to apprehend how WaterLOGSY (ePHOGSY) works, the NMR sequence of the method (Fig. 4) has a first section representing the ePHOGSY spin echo, responsible for the inversion of the water magnetization. The narrow and broad rectangular shapes correspond to 90º and 180º hard pulses, respectively. The first 90º pulse sets every magnetization on the transversal plane. The first gradient G1 defocuses the entire magnetization. Then the selective 180º pulse (Φ1) affecting only the water signal, enables water refocusing; the second G1 gradient enhances the defocusing of the off resonance signals and refocus the water resonance. The following 90º pulse leads to complete inversion of the water magnetization. The selective 180º pulse shifts its phase by 90º each scan by phase cycling.

Because of the selective irradiation of the water signal, those resonances that lay close to it may also be affected if the pulse is not fully calibrated and set.

Now, the inverted magnetization needs to be transferred to the bound ligands. To accomplish that, an evolution time, typically 1 to 3 seconds, needed for the NOEs to take effect, is required. For the mixing time (τ_m) a sustained weak gradient G4 gradient is applied to avoid fast relaxation due to radiation dumping. The strong crushing additional G5 is optional and its application accounts for the same purpose as G4. [31,38]

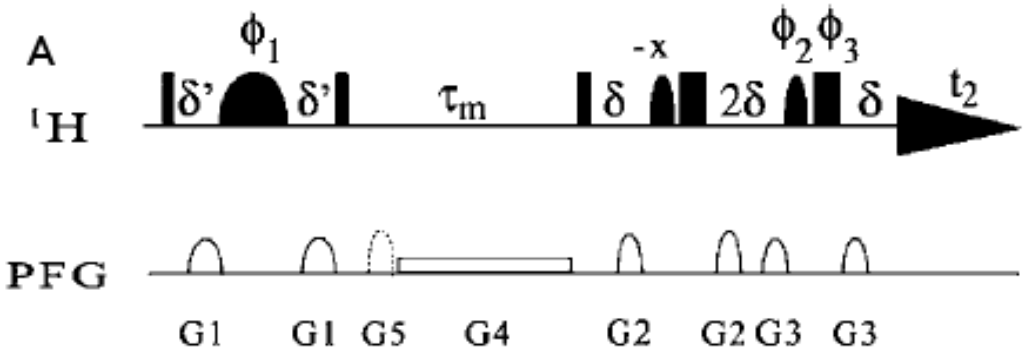


Figure 4. Pulse sequence for the homonuclear one-dimensional ePHOGSY.

There are two possible simultaneous pathways for the inverted magnetization of water to be transferred to the bound ligand. Along with the ligand in the pocket site of the receptor, there are “bound” water molecules whose magnetization is now inverted. The “bound” condition of the water molecules relies on the fact that their residence time is longer than the rotational correlation time of the large molecular complex, otherwise no inversion would occur. The inverted magnetization of these water molecules will pass through direct H-H cross-relaxation to the bound ligand. Because these waters are buried inside the complex, the cross-relaxation between the bound waters and the ligand is negative and tends to bring the ligand magnetization to the inverted state.

The other pathway is through intermolecular exchange of labile carboxyl, amino, a hydroxyl, imidazol, guanidinium and amide proton is from either the receptor binding pocket or somewhere at its surface.

Additionally, free ligands (binders and non-binders) also suffer magnetization transfer from water, which is also via exchangeable protons and posterior spin diffusion. However, in this case, due to the short correlation times of free ligands, the cross-relaxation inside the ligands is positive, yielding a NOE effect characteristic of fast tumbling small molecules.

Also a water selective flip-back can be applied before the observe pulse to allow fast repetition times and an additional hard 180º pulse can be added in the middle of the evolution time for the suppression of artifacts that originate from almost complete relaxation of the protons of the small molecules during the long mixing period. [39]

The resulting spectra (Fig. 5) will show binding and non-binding ligand signals of opposite phase, due to the different signs of the constants that rule the cross-relaxation with short and long correlation time molecules. This allows distinguishing between binders and nonbinding ligands where the pool of spin-inverted binding ligands must overcome the pool of non-inverted-spin ligands on the evolution time of the experiment. For that reason, low ligand excess ratios, of about 10 or 20 should be enough.

While the initial waterLOGSY experiment seemed to be limited to the identification of binders vs non-binders, recent work showed that waterLOGSY can be used to probe for bulk water accessibility to derive the orientation of a

bound ligand. This approach has been termed SALMON (Solvent Accessibility and protein Ligand binding studied by NMR Spectroscopy) and may be found elsewhere. [34,38]

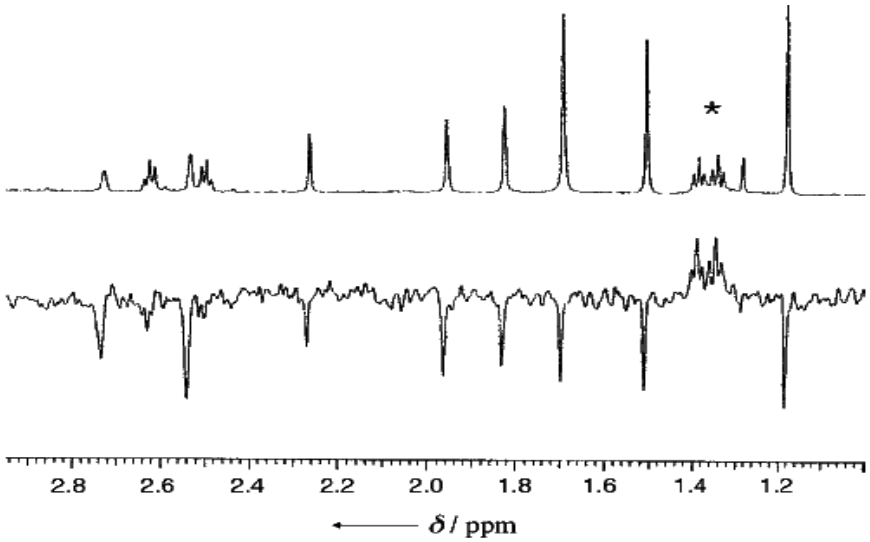


Figure 5. ^{1}H Water-LOGSY spectrum of a library of ten compounds in the presence of a receptor. Signals from the indole derivative with binding affinity for the receptor are easily identified (positive signals). The upper spectrum displays the ^{1}H NMR reference spectrum. The signals belonging to the binding ligand are marked by *.

Being the WaterLOGSY experiment water-dependent, it is limited by the fact that fully deuterated solutions are not welcome, and by its inability to identify strong ligands with slow dissociation rates. However, it represents a interesting NMR approach for primary screening of compounds for binding to the target of interest in the μM range. [39]

3.2. TRANSFERRED NOE'S

The transferred NOE (tr-NOE) effect was originally described by Bothner-By. Peters was the first to propose the tr-NOE use for screening compound mixtures. The tr-NOE (Fig.6) is based on the nuclear Overhauser effect between adjacent spins in the ligand in presence of chemical exchange between the bound and unbound form. [40-42]

NOE effects (NOEs) are often used in determining 3D structures of compounds in solution, where some magnetization from nucleus I is transferred to nucleus S in the mixing period by nuclear relaxation via dipole-dipole coupling.

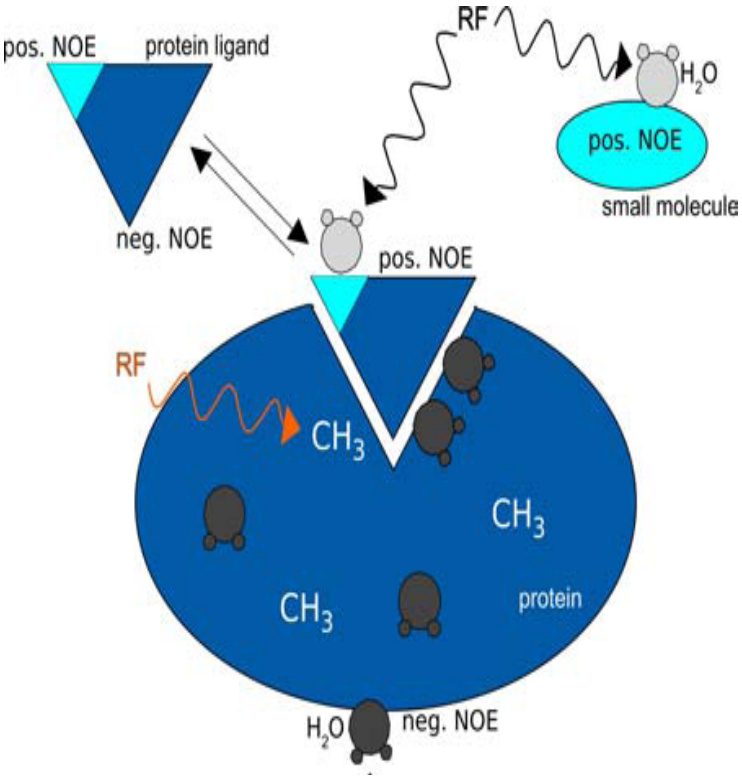


Figure 6. Schematic illustration of the Tr-NOE. Magnetization is transferred from ligand A via the protein back to itself.

The transferred magnetization will precess (at the Larmor frequency) at the chemical shift δI during t1 (Fig. 7), but at the chemical shift δS during mixing time. There will be therefore a peak in the two dimensional spectrum with coordinates (δI,δS), so we may say that these two spins are correlated. Structurally this means that I and S are spatially close (r < 5 Å) and the NOE is a relaxation factor that builds-up during the mixing-time. [8]

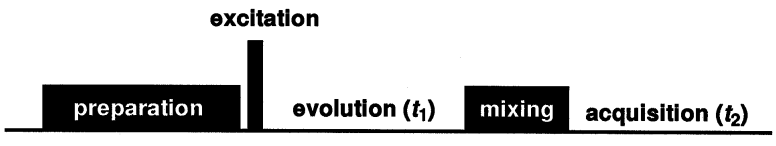


Figure 7 Schematic illustration of the NMR pulse sequence for 2D NOESY experiment.

Low molecular weight (MW) compounds (MW < 2000) have short rotational correlation times (τ_c), therefore presenting positive NOEs. Also, depending on compound shape and field strength, low MW molecules can exhibit

no or small negative NOEs. On the other end, large MW molecules, usually proteins, due to their long τ_c's, show large negative NOEs. Small ligand binding to large molecules leads to a drastic change in ligand NOE's, since the ligand starts tumbling at the large receptor tumbling rate. Within a mixture, those ligands that bind the receptor show NOEs of opposite sign to the cross-binders as they adopt the tumbling behavior of the large receptor – Figure 8.

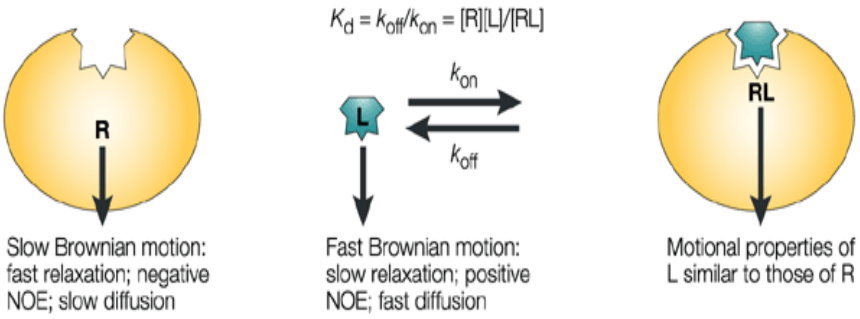


Figure 8. Schematic illustration of transferred NOE's experiment.

Additionally, the absolute magnitude of the NOEs is greater in slow correlation time molecules and its maximum value, which occurs only if dipole-dipole relaxation is featured, is more rapidly achieved. [18]

In tr-NOEs based screens, intramolecular NOEs of compound mixtures are observed via 2D-NOESY NMR spectra in the absence and presence of the receptor. Binding compounds are identified by NOE crosspeaks that have changed sign in the presence of the receptor. It is estimated that tr-NOE is able to probe binding affinities between 100 nM ≤ K_D ≤ 1 mM, once if the residence time isn't long enough no magnetization will happen, and if the binding is too strong no build up of tr-NOE's would be visible in the final spectrum.[43]

Besides the binding confirmation, tr-NOE can provide information about binding pocket environment: once intramolecular tr-NOEs are the key to define bound ligand conformations, intermolecular trNOEs occur between a ligand and a receptor protein, and therefore, in principle, allow the determination of the orientation of bound ligands in protein binding pockets. [19]

For a better understanding of the tr-NOE experiment, we must define two different concepts: the conformation of a bound ligand and its orientation in the binding pocket of the receptor. We can define receptor bound ligand conformation as the three dimensional architecture that the ligand adopts when

bound to its receptor, normally a protein. The orientation concept translates the significant alterations in the orientation of the ligand in the binding pocket that may arise from slight structural modifications.

The receptor-bound conformation of the ligand can be determined either by standard NMR techniques using labeled complexes, where the three dimensional structure of the complex is depicted, or by transferred NOE or transferred cross-relaxation.

3.3. SAR BY NMR

The advent of the SAR by NMR method is given to S. Fesik at Abbott Laboratories. SAR by NMR detects changes in the chemical shifts of the protein target upon addition of ligands (Fig. 9). Weakly interacting compounds that bind at adjacent sites on the protein target are identified and combined with structural information about the orientation of the bound ligands to guide a linked-fragment approach to generate lead drug compounds with greatly increased binding. [44]

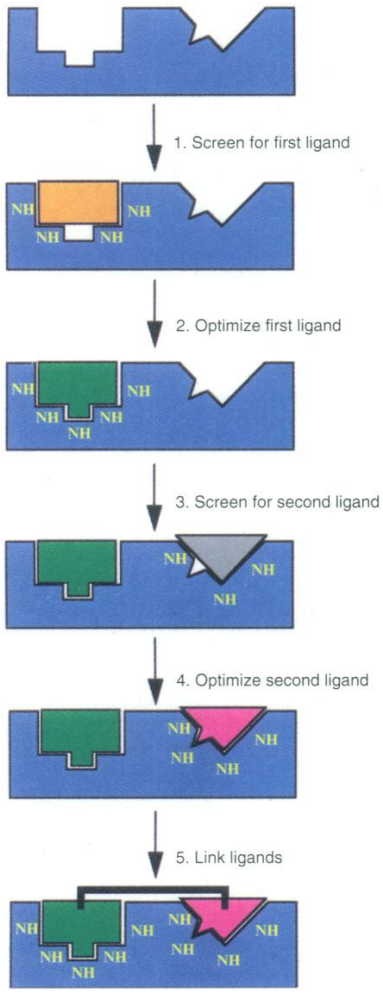


Figure 9. SAR by NMR experimental procedure. The first step screens ligands using the perturbation method with a difference in chemical shift greater than 0.1 ppm for at least two peaks in the spectrum. After hit identification the binding affinity is measured and an optimization is done for this site. The same perturbation method is applied to identify a second low-affinity hit in the presence of the optimized first ligand. Identification and optimization of the second binding site follows.

The first step of the procedure requires screening a library of ligands using the perturbation method with the assumption that a change in chemical shift greater than 0.1 ppm for at least two peaks in the spectrum constitutes a significant change and indicates binding. After hit identification, monitoring chemical shift changes as a function of ligand concentration in a titration experiment is performed to accurately deduce the binding constant. Following identification of a binding partner to the first site, screening of related compounds is conducted to increase and optimize the binding affinity for that site. [45,46]

The same perturbation method is applied to identify a second low-affinity hit in the presence of the optimized first ligand. Following identification and optimization of the second binding site, structural data obtained by additional experiments is used to determine the location and orientation of the protein target in complex with its two low-affinity hits. Maintaining the spatial orientation of the compounds with respect to each other and the protein target, the ligands are synthetically linked together to produce the high affinity ligand as seen on Figure 10 illustrative example. [44]

Binding is detected from the amide chemical shift changes observed in 2D Heteronuclear Single Quantum Correlation (HSQC) spectra. The ^{1}H-^{15}N HSQC experiment exploits the repeating nature of the protein's primary sequence and three-dimensional structure. The amide bond connects the amino acids that compose the protein sequence, creating a repeating series of NH groups that become chemically unique in the context of the protein's tertiary structure. [47]

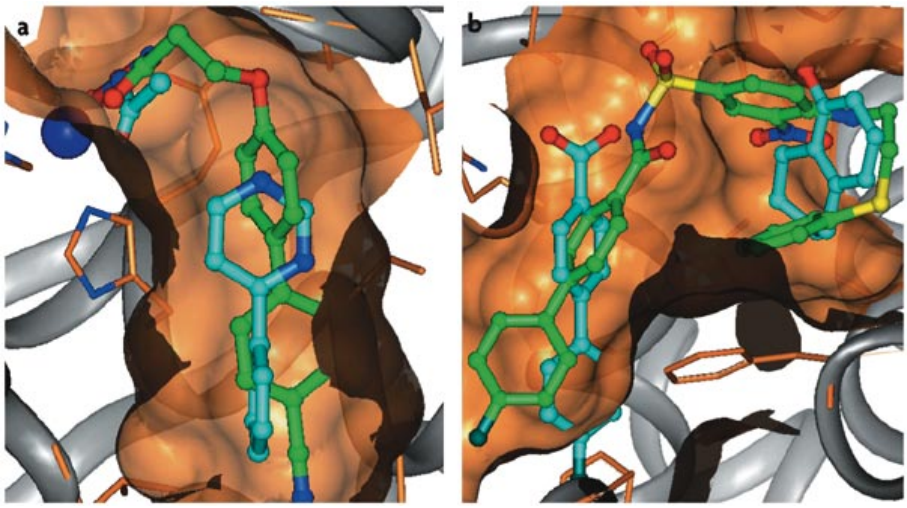


Figure 10. SAR by NMR example. The identified fragment leads are shown with cyan carbons, whereas the linked compounds are denoted with green carbon atoms. *Adapted from [44]*

The experiment only detects those protons directly coupled to nitrogen, so the resulting spectrum lacks any overlap from proton signals associated with carbon (typically aromatics in the relevant ppm range). The spectrum displays one cross peak for amides in the protein at a position characterized by its ^{1}H and ^{15}N chemical shifts that is an already optimized protocol (Fig. 11). In total, the spectrum contains one signal for each residue except proline, which lacks an HN, two paired peaks generated by the asparagine and glutamine sidechain NH_2 groups and additional signals from NH containing side chains. [48,49]

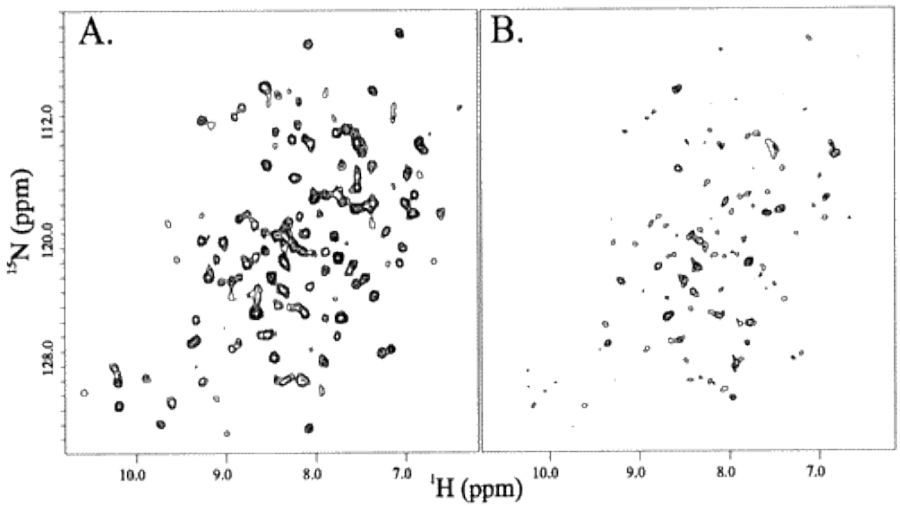


Figure 11 Sensitivity-enhanced of $^{1}H/^{15}N$ HSQC spectra. (A) a dual ($^{15}N/^{1}H$) inverse CryoProbe (Bruker) equipped with a lock and a z-gradient and (B) a conventional triple-resonance ($^{13}C/^{15}N/^{1}H$) inverse probe. *Adapted from [50]*

The HSQC spectrum provides an overall map or fingerprint of the protein target and serves as the basis for assessing ligand binding and the overall fold and behavior of a protein. This experiment is fast and easily interpreted but requires isotopic labeling of the protein with ^{15}N because of its low natural abundance (0.37%). When labeling is not practical, a ^{1}H-^{13}C version of the HSQC can be performed. [50]

As the SAR by NMR method is only applicable to small biomolecules, many small proteins or protein domains may fit in these criteria and may serve as drug targets, so an important role must be given to the methodology. Transferred NOE's are often employed as one of the final steps in the SAR by NMR method, when sequential assignments of the labeled protein target are known, to determine the orientation of two adjacently bound ligands to a protein target.

Overall information is so useful that major drug companies owe a great percentage in their leads to NMR-based screening. [42]

Although the number of compounds screened at a time is limited by the NMR sensitivity, it can be overcome to some extent by using cryoprobe and the method is not amenable to very large proteins but this situation can offset by using selective labeling. [44,45]

3.4. WATER SUPPRESSION – DOUBLE PULSE FIELD GRADIENT SPIN ECHO

Hwang and Shaka (1995) first purposed a scheme for water eradication based on magnetization [51], using a selective excitation as an approach to remove the water signal delivering a pure phase spectra with flat baselines.

Based on the WATERGATE technique, and starting with the simple pulse field gradient spin-echo (SPFGSE) [G - S - G] scheme, where S stands for the unitary transformation caused by the selective pulse and assuming that the gradients are matched in strength and duration, and are sufficiently strong. With further advances they discovered the double pulse field gradient spin-echo (DPFGSE) [G1 - S - G1 - G2 - S - G2] will return any transverse magnetization to it's original position attenuated by the square of the probability that a spin is flipped by S. [52]

The suppression of bulk water is achieved with two hard pulse sequences 90^{o}_{x} - G1 – 90^{o}_{x} – T – 90^{o}_{x} - G1 - G2 – 90^{o}_{x} – T – 90^{o}_{x} - G2 with t=500 μs, and without the phase roll and other optimizations a good water suppression may be expected. The DPFGSE scheme also features the substantial loss of intensity of the peaks with chemical shifts near to the water resonance due to a gradual spin-flip profile of the selective pulse used as an inversion sequence. In order to get a better and more selective excitation profile further pulses and delays may be added.

The way the sequence works (see Figure 12 to follow the magnetization) is quite simple: the first gradient pulse defocuses all magnetization, while the second one just refocuses that magnetization which has been selectively inverted, as the remaining magnetization keeps defocusing through this second gradient pulse. Comparing simple and double echo, the excitation profile reflecting inversion profile of the 180º selective pulse of the spin-echo has no

dispersive component, but using a simple rectangular selective pulse, the DPFGSE leads to a better excitation profile, that may be explained by the probability of a spin flipping which dictates the amplitude of the response.

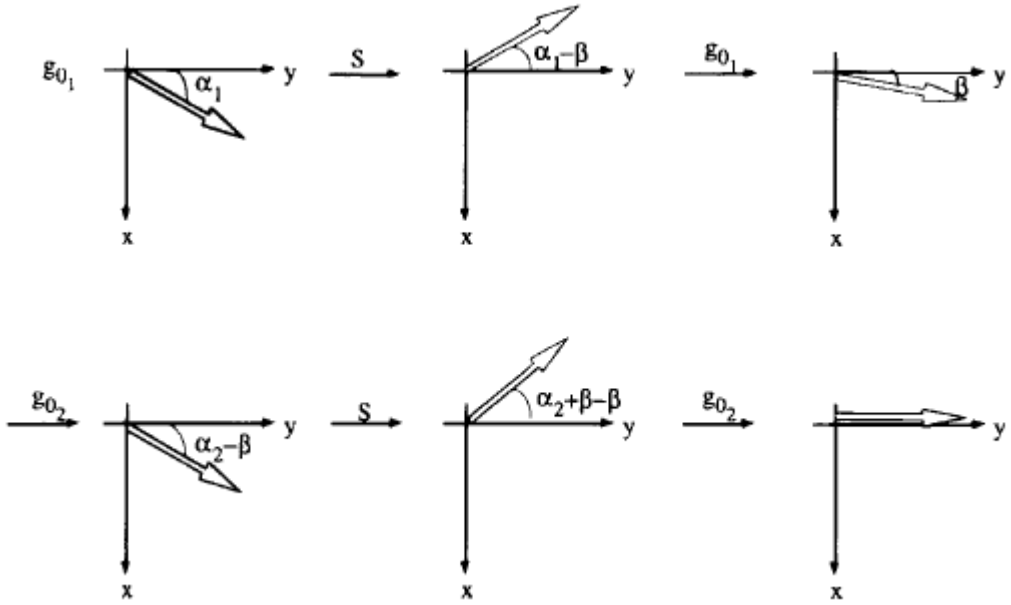


Figure 12. Behavior of transverse magnetization under the application of DPFGSE. The angle β is associated with imperfections of the selective inverting procedure S. The angles α1 and α2 stand for dephasing due to gradients G1 and G2at a given sample location.[12]

After the first echo G – S - G, the inverted magnetization is not correctly refocused, leading to an error β, but after the second echo this phase discrepancy β is added again and therefore compensates the first one and the magnetization is perfectly refocused and returns to its starting point.

The usage of more selective pulses would discard the use of double-echo in the excitation sculpting scheme on the suppression of the water signal, but this scheme can be inputted to an infinite set of NMR experiment sequences.

3.5. SATURATION TRANSFER DIFFERENCE (STD) NMR

The efficiency and simplicity of the STD experiment and its applicability to several 1D and 2D NMR sequences gives a great range of data types to be acquired, giving this technique first reported by Mayer and Meyer a crescent popularity. [17]

The STD protocol gives a spectrum resulting from the subtraction of two spectra: a 1D ^{1}H reference spectrum (off resonance), where the saturation is applied to a region where no resonances appear, usually at 30 ppm; and a 1D ^{1}H (on resonance) where selective saturation is applied to the receptor during a certain saturation delay, where no resonances appear but the ones from the receptor, usually at 0 or slightly negative ppm's, but if no ligand peaks exist in the low field region, the saturation field may be placed there. In this way we guarantee the same amount of input energy in both cases. Only the compounds in the mixture that effectively bind to the receptor remain visible in the so called STD spectrum. This characteristic presents the possibility to discriminate from a mixture of many possible binding compounds to a unique receptor, those which effectively bind the receptor and besides this screening (of many different potential binding compounds) the scrutiny of the ligand binding epitope is possible. [17-19, 53,54]

The easy and fast STD-NMR assay experiment could use as receptor a protein or macromolecular complex of high molecular weight, as well as ribozymes or any other possible receptor compound allowing the use of the STD-NMR experiment in drug-design projects as the main proteic targets are usually large dimensioned. A slow tumbling and large correlation times (T_c) by large dimensioned proteins and fast rates of relaxation by the proteic system of protons tightly coupled by dipole-dipole interactions allows the quantitative interpretation of the spectra. [55]

After a narrowed region of the broad protein resonance is selectively saturated, the saturation is spread to all protein protons via spin diffusion and the bigger the protein and denser its proton matrix is the more efficient the saturation applied is transferred. In a mixture of multiple ligands the complex ligand-receptor assembly with those ligands who are capable of bind which will suffer intermolecular transfer of magnetization by relaxation via the nuclear overhauser effect because when the complex is formed, the small ligand acquires the properties of the large macromolecular complex, having high correlation times and relaxation rates. This mean that the ligand protons progressively become saturated, and so, when the bound ligand molecule dissociates from the complex the saturation is brought to the solution in the free state ligand. Having very short correlation times and small relaxation rates, the ligand that actually binds to the

receptor keeps the saturation for longer periods of time, making possible the creation of a crescent pool of saturated free ligands in the solution which will rise up in a well resolved STD spectrum – Figure 13 and 15.

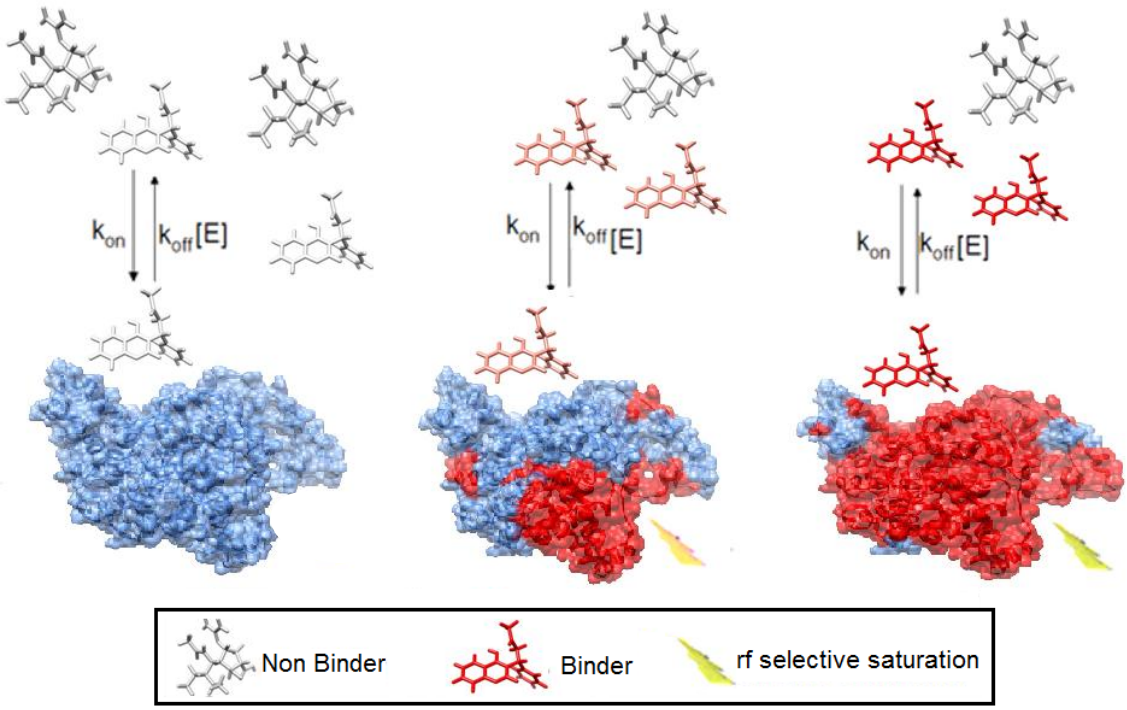


Figure 13. Detection of binding using the STD technique experiment. Frequency selective irradiation (lightning bolt) causes selective ^{1}H saturation (red shade) of the target receptor. Irradiation is applied for a sustained interval during which saturation spreads throughout the entire receptor via ^{1}H-^{1}H cross-relaxation (spin diffusion). Saturations is transferred to binding compounds during their residence time I the receptor binding site. The number of ligands having experienced saturation transfer (hits) increases as more ligand exchanges on and off the receptor during the sustained saturation periods. Non-binding compounds are un-affected.

This assembly of a ligand-receptor complex being the basis of the STD protocol the determination of K_D values using this technique is possible if the values lie between 10^{-3} and 10^{-8} M. For K_D's values higher than 10^{-3} M the ligand interaction is not tight enough magnetization transfer to occur, once the ligand dissociates before it gets saturated, and therefore no STD effect will be visible, on the other hand if the K_D value is below 10^{-8} M the ligand remains on the complex and does not contributes to the pool of saturated labeled ligands, resulting in a slow ligand-receptor complexation turnover, on the established STD timescale. [18]

The pool of saturated labeled ligands grows as new unlabeled ligands bind to the saturated receptor and bound ligands dissociate from the complex.

Therefore a large ligand excess increase the probability of new unlabeled ligand becoming available to bind the receptor, and not the already labeled ones, so an increase of saturated labeled ligands released into solution is expected.

The saturation from the receptor is transferred to the compounds via cross-relaxation, and, obviously, the protons closest to the receptor become more saturated than those that don't bind directly to the receptor and just have a role on complex stabilization. This saturation transference is proportional to $1\backslash R^6$ where R is the distance between two protons one for the receptor and one for the ligand, so the shorter the spatial distance from the proton to the receptor, the higher the intensity in the final STD spectrum is, and the proton with small or no contact with the receptor only receive a saturation by spin diffusion leading to much reduced signals – Figure 14. This profile makes possible the mapping of the binding epitope on the ligand after the ligand-receptor complex becomes dissociated, also known as Group Epitope Mapping (GEM) which, as well as other STD-NMR potentialities, will be discussed further.

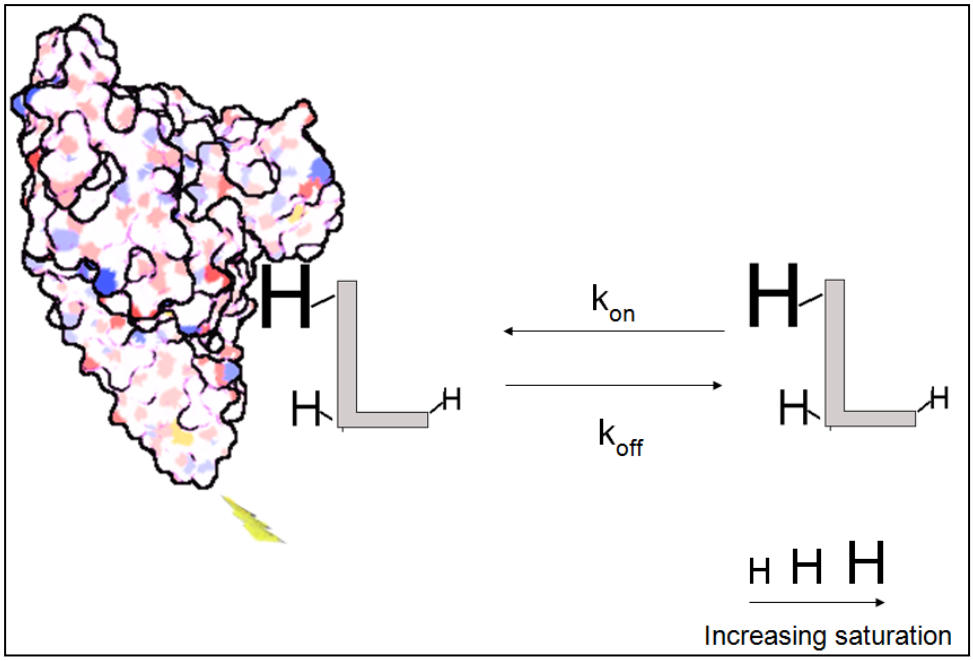


Figure 14 Illustration of GEM for a ligand in a fast exchange between the bound and free state. Irradiation of the protein at a resonance where no ligand signals are present leads to a selective and very efficient saturation, which is transferred to the binding parts of the ligand by intermolecular saturation transfer. Here, groups presented by larger protons symbolize groups in closer contact with the protein and the smallest the ones with few or o contact. Therefore, the degree of saturation of the individual protons of a small ligand molecule reflects the proximity of these to the protein.

As was stated before, to execute a STD experiment we need an on-resonance spectrum, whose peak intensities are referred as I_{SAT}, and after the delivery of saturation, whereas the resonances that suffered saturation transfer will decrease their intensity in the spectrum and resonances from non-binding ligands or remotely binding to the receptor maintain the same intensity. The off-resonance spectrum acquired in the same conditions with the saturation delivered far from the spectral region observed, gives a standard proton spectrum with peak intensities termed as I_0. The subtraction of these two spectra gives a proton spectrum with peak intensities referred as I_{STD}. As only the binding ligands will show positive peaks on the difference spectrum the STD effect is quite simple and easy to observe – Figure 15.

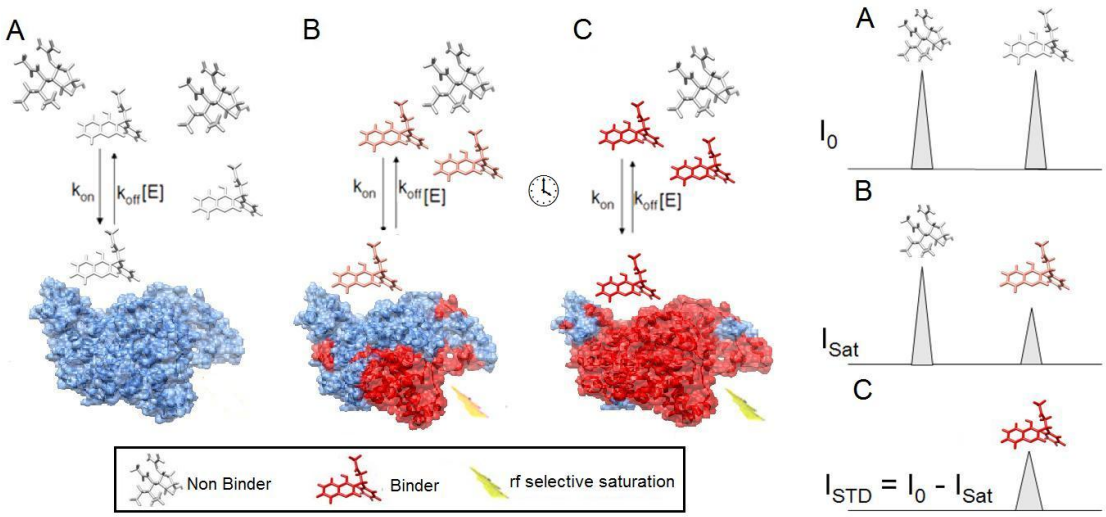


Figure 15. Schematic diagram depicting difference spectroscopy in the STD experiment. A - an off-resonance experiment (reference) applies rf irradiation away from the receptor and ligand protons (blue shade) havind the sectra intensity called I_0. B – an on-resonance experiment, the rf irradiaction selectively saturates receptor and any binding compounds (red shade), this manifests the decreased signal intensity called I_{SAT}. C – the STD response is the spectral difference $I_{STD} = I_0 - I_{SAT}$, which yelds only resonance of the receptor and binding compounds.

The NMR pulse sequence – Figure 16 - responsible for the reported effect presents a first pulse block sculpting to selective saturate the protein in a user defined time, with a train of Gaussian shaped pulses applied n times during the saturation transfer delay (xferdly) usually set to 40. [18] The strength of the pulses referred to deliver a good STD performance must be of 86 Hz [$(\gamma/2\pi)\ B_1$ = 86 Hz].

The sequence also features a delay time d1 before the saturation assigned to an additional relaxation delay set according to the sample

specifications; after the Gauss cascade of pulses the observer pulse is applied, the hard 90^0 pulse applied to whole spectral width region will hit all ligands and receptor protons, and by the difference characteristic of this technique, also the saturated receptor will show peaks on the STD spectrum. To avoid this non desirable situation, and for ligand peaks to emerge more quickly, resolved and clearly from noise, we eliminate protein signals by adding a T_{1p} filter (R2 filter), also known as spin-lock or trim pulse that stop the evolution of the system magnetization and makes a complete transverse relaxation of the receptor, since signal protons in large molecules return to the equilibrium faster than in small ligands. Thus, at the acquisition time given only the ligand resonances remain phased and will appear in the final spectrum. The calibration of this trim pulse depends on the time needed to a complete relaxation of the receptor protons and the length must not exceed it because otherwise we will progressively decrease ligand intensity with further relaxation.

Completing our sequence we have the DPFGSE excitation scheme (already described) to eliminate the water signal from the spectrum and the subtraction of the on and off-resonance spectra is accomplished by phase cycling.

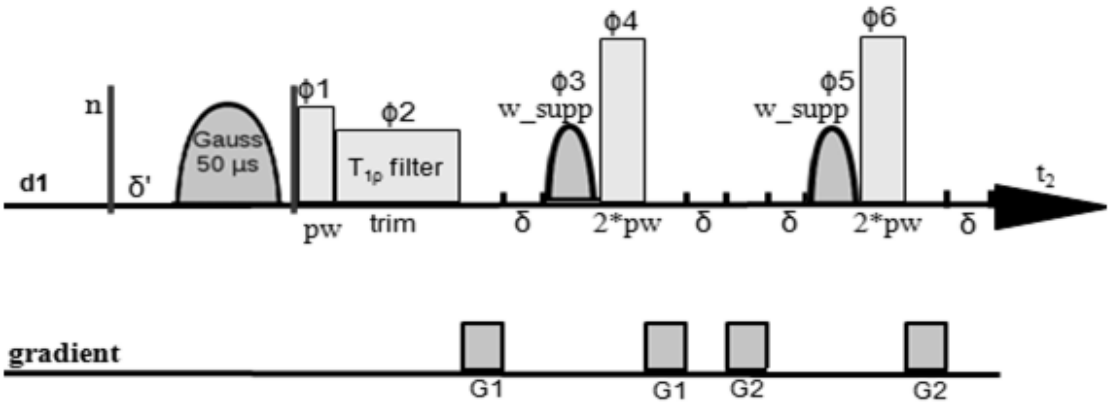


Figure 16. Pulse sequence for the 1D STD NMR spectra with the additional DPFGSE step for H_2O suppression. The subtraction is performed after every scan by phase cycling.

This flowing subsection will present all the experimental potentialities of the Saturation Transfer Difference technique, how to handle spectra for the maximization of the data obtained and what new perspectives are being offered to the scientific community over the last few years.

This brief summary of the approaches used in this current work aims at the explanation of how and how far STD-NMR can be stretched and accomplish success in its protocol.

3.5.1. SCREENING OF LARGE LIGAND LIBRARIES

Since the very first references on STD-NMR, the possibility to rapidly screen up vast libraries of potential ligands for a common receptor was made an hallmark. It was possible, from a given pool of possible ligands, to identify the ones that actually bind to a common receptor. They will be the exclusive ones to appear either on 1D or 2D STD experimental spectra, helping in such way to a more effective the drug design and optimization.

This STD approach was first introduced by Vogtherr and Peters, who functionally screened a library of 20 randomly chosen O-Methylated galactopyranosides and successfully discriminate them, from a common sample containing Sambuncus nigra agglutinin (SNA), and even with no separation or purification. Despite the hard task of full NMR assignment and discriminating of overlapped peaks, ultimately STD revealed itself to be a robust method to detect specific and targeted oriented ligands. More recently it was used to study non-covalent interactions between different glycostructure transforming enzymes and select substrates and products (Fig. 17). Undoubtedly, STD-NMR became a methodology with great interest in many different scientific research areas, due to its effectiveness and short timescale results, especially between large biomolecules and carbohydrates. [56]

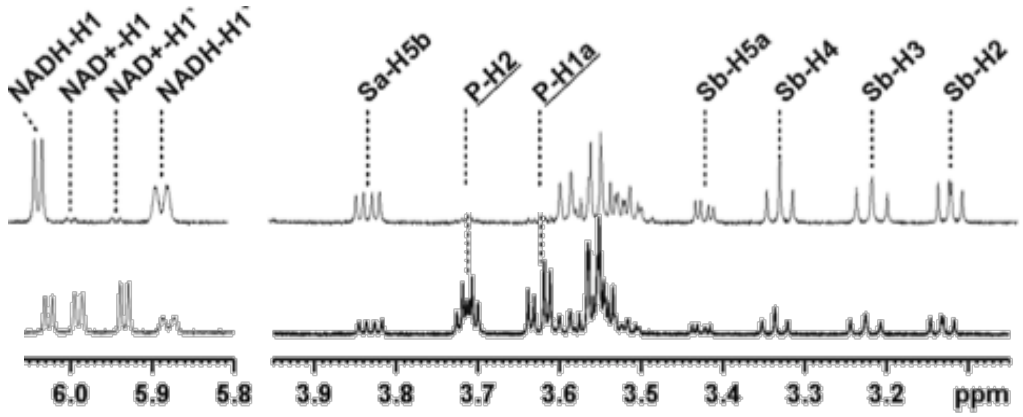


Figure 17. STD NMR spectra recorded during CtXR (10 mM) catalyzed and NADH (10 mM) consuming reduction of xylose (10 mM) to xylitol. After a ^{1}H NMR spectrum recorded at the start. *Adapted from [56]*

3.5.2. TWO-DIMENSIONAL (2D) STD

The uni-dimensional STD sequence can be added to any 2D NMR experiment, and we can enhance the STD characteristics on any 2D experiment, one all need is to compile the sequence.

Since peak intensities reflect ligand-protein proton proximity, it is possible to determine the binding epitopes, once the closest protons will appear firstly with the zooming in on the 2D spectrum. By keeping zooming in, the less intense peaks will start to show up too, characterizing the epitope.

The possibility of such STD associations makes this technique very useful in the scrutiny of ligand-receptor interactions at an atomic level, and more recently the first example of peptide–protein heteronuclear two-dimensional (2D) STD-NMR was reported. [57]

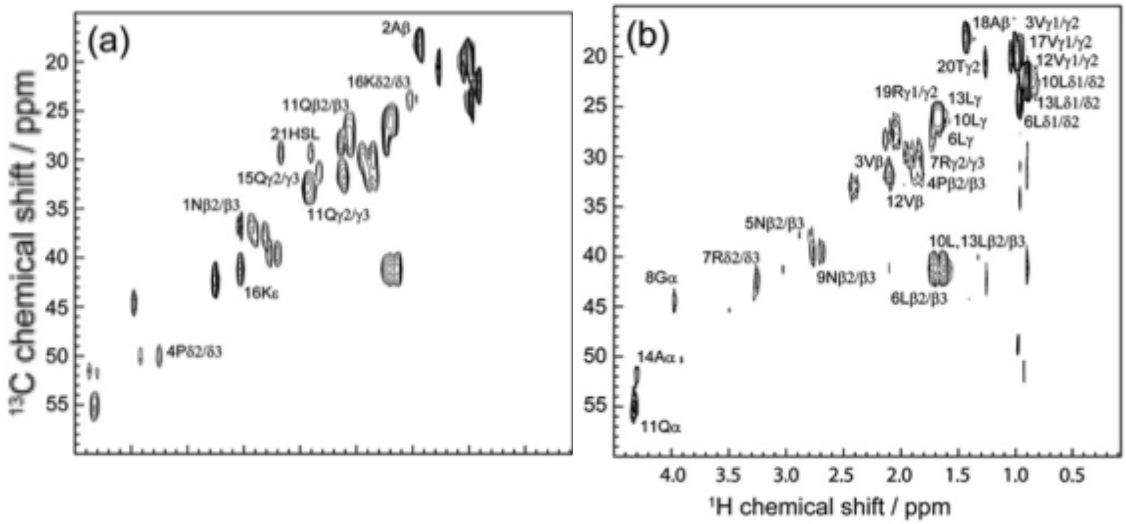


Figure 18. ^{13}C-STD-HSQC data for A20-FMDV2/integrin $\alpha_v\beta_6$ inter- action. (a) 2D STD NMR control (off-resonance) spectrum, and (b) ^{13}C-STD-HSQC difference spectrum. Labels in (a) show minor interacting residues identified in (b), whereas labels in (b) show significant interactions. [57]

The first 2D STD-NMR study demonstrated with glycopyranosides and agglutinin reported no 2D approach to peptides. So the possibility of 2D heteronuclear edited STD NMR study provides information beyond that available from 1D ^{1}H STD NMR by greatly reducing signal overlap. The previous 1D STD NMR experiments suffered indeed from resonance overlap that makes it very difficult to obtain any significant detail beyond identification of each amino acid involved in the interaction. Therefore, 2D STD-HSQC NMR acquisition (Fig. 18), was employed to observe more detail of the interaction. This technique has

proved to be extremely powerful in the study of protein–peptide interactions, especially in the case of interaction of a 21-mer polypeptide with the 186 kDa heterodimeric integrin $\alpha_v\beta_6$, where complete ^{13}C and ^{1}H NMR assignments were obtained for the peptide under identical conditions to the STD experiments to allow STD analysis. 2D ^{13}C-edited STD NMR experiments use a modified HSQC experiment with the ability to add or subtract transients depending on whether ^{1}H saturation is on-resonance at -3.0 ppm or off-resonance at -30.0 ppm following a 2 s selective saturation pulse.

The 2D method has identified 12ValHb and 13LeuHg1/g2 as possessing the largest STD amplification factors with the primary interaction involving 6Leu, 10Leu, 12Val and 13Leu. These observations differ from the original 1D data where 12Val was judged to have a lower percentage STD amplification than 10Leu, 11Gln and 13Leu, and with 11Gln scoring more modestly in the 2D STD NMR experiment. Furthermore, the C-terminal amino acids 15Gln-20Thr, that were thought to be significant contact sites in the original experiments, are also shown in a more modest light with the 2D data. These differences illustrate the challenge and effects of signal overlap in peptide based 1D ^{1}H STD NMR that 2D heteronuclear STD methodology can circumvent, and a more efficient epitope characterization may be done, as presented (Fig. 19).

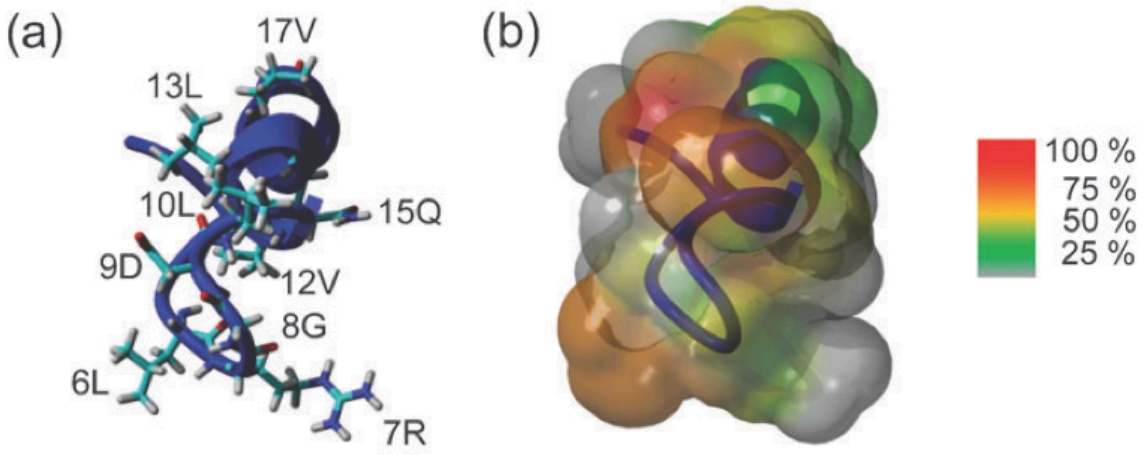


Figure 19. Structure of A20-FMDV210 showing (a) significant amino acids highlighted by 2D STD NMR and (b) the solvent accessible surface colored as a percentage of the maximum STD transfer to highlight significant interactions along the helix. **[57]**

Orientational information regarding peptide–integrin binding can be sought from ^{15}N-edited 2D STD NMR that provides STD information regarding the peptide backbone. 2D ^{15}N-edited STD NMR experiments were performed by collection of separate I_o and I_{sat} datasets in an equivalent manner to protein–protein cross saturation experiments to minimize observed phase artifacts. [58] Phase artifacts are due to low saturation transfer efficiency caused by the relatively large distances between target protein ^{1}H nuclei and bound peptide backbone ^{15}NH groups.

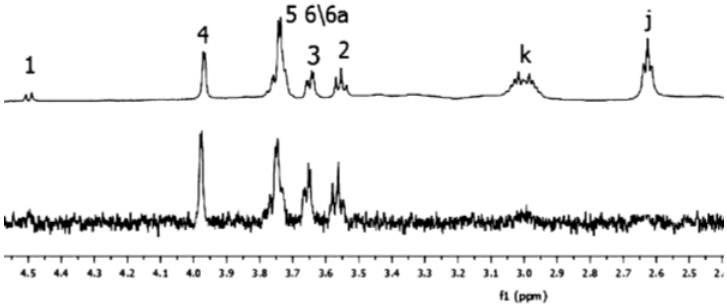
The 2D STD NMR approach has allowed the confirmation the dominant leucine interaction within the Asp-Leu-Xxx-Xxx-Leu (DLXXL) motif in addition to observing minimal transfer to the RGD motif, as well as, for the first time, the significance of the interaction of Val12, suggesting a different peptide orientation for binding to the integrin compared to the TFE (Trifuoretanol) induced structure of the peptide. [57]

As we can see, this successful application of 2D STD NMR has proved that this methodology has potential to understand the molecular details of peptide–protein interactions, particularly involving protein targets without known structures, and it is indeed, a major breakthrough.

3.5.3. CHARACTERIZING THE BINDING EPITOPE

Once a ligand interacts with a binding site on the receptor, different distances separate the ligand protons from it, and those that are physically closest became more saturated due to the STD protocol, which also allows discriminating between the closest protons using the intensity manifested on the final STD spectra. After the peak assignment is done the peaks with the greater percentage of intensity reduction are the ones closest to the protein. Using the fraction I_{STD}/I_0, being I_0 intensity on the reference spectrum and I_{STD} the intensity of the different peaks in the STD spectrum, we normalize the peak with greater reduction to 100 % and other peaks follow this normalization. [60]

Such protocol may be used to map different kinds of interactions, and in this way one knows how the ligand binds to its receptor.



Protons	La(DOTAGal)
H2	94
H3	100
H4	93
H5 H6[a]/6[b]	64
k	11
j	10

Figure 20. One-dimensional ^{1}H NMR (top) and saturation transfer difference (STD) NMR (bottom) spectra of a 0.82 mM La(DOTA- Gal) and Summary of the characterized binding epitopes with saturation transfer difference (STD) values relative to H3 as a percentage. [61]

To obtain enough data for a GEM characterization a single STD spectrum is needed, and if the signal to noise ratio and the saturation transfer is good this normalization protocol delivers very accurate and quick maps of interaction of ligands with the binding site. So it is possible to say what proton and/or what region of the ligand binds primarily to the receptor to be studied. [61]

3.5.4. DISSOCIATION CONSTANTS ESTIMATION

The direct estimation of K_D values is also possible with this technique, although the values have to be between 10^{-3} M and 10^{-8} M for reasons already explained in the previous section.

The STD signal area increases directly with the increase of the number of molecules that received the saturation transfer in the given delay, thus, one of the ways of increasing this effect is to provide a great ligand excess, so that the probability of a saturation-free ligand binding the receptor will be higher, obtaining a greater pool of saturated ligands free in the solution.

The outdated way of obtaining a K_D value of a ligand is via titration of the ligand with a constant concentration of the receptor, and a new parameter was introduced by Mayer and Meyer, the STD amplification factor (A_{STD}). This parameter makes possible the higher accuracy for the real value magnitude. This STD parameter is defined by the product of fractional saturation of that given proton by the ligand excess over the protein, giving a more trusty value of the direct protein-receptor interaction. So, if we get an A_{STD} of 4, it means that the

STD signal intensity in the ligand is 4 times higher than proton in a protein. Being the A_{STD} a normalization it allows a direct comparison between corresponding resonances of a ligand in two different samples, and may also be used as a measure of the amplification of the protein observed in the STD signals of the ligand. The I_0-I_{SAT} is the intensity of the signal in the STD spectrum and fractional STD effect given by (I_0-I_{SAT})/I_0, I_0 being the absolute area of the signal in the reference spectrum and I_{SAT} the absolute area of the same signal in the on resonance spectrum, which may increase or decrease.

Since the STD signal intensity increases with the ligand excess, it might be hard to understand why the absolute value could decrease, so the explanation follows; although more molecules of ligand suffer the saturation transfer, at high ligand concentrations the ratio between saturated and total ligands decreases. The multiplication of (I_0-I_{SAT})/I_0 by the ligand excess (ligand concentration over protein concentration) gives the A_{STD} (see Equation 1). Plotting A_{STD} against total ligand concentration results in a standard one-site drug-responsive type of curve. This protocol makes possible the screening of saturation profiles on binding sites.

$$A_{STD} = \frac{I_0 - I_{SAT}}{I_{SAT}} \times Ligand\ Excess \qquad \textbf{(1)}$$

As the ligand excess increases, the A_{STD} become higher and therefore the discrimination between protons that directly contribute to ligand binding and those who are only involved in the stabilization and have little or no contact with the protein, was made easier. Nevertheless a higher saturation delay also gives more saturated ligand molecules, leading to a greater A_{STD}, but with implications discussed on the next section.

Lepre and Peng studied the determination of K_D values based on titration. The I_{STD} can be written as I_{STD}= Cα_{STD}[EL], where C is a proportionality constant that makes the appropriate unit conversions, α_{STD} is a dimensionless scaling factor that represents the maximum STD amplification. The reference I_0 is just proportional to L_T, and so the fractional STD effect can be written as I_{STD}/I_0 = ε [EL]/L_T = α_{STD} P_B = P_B^E and ε is the ligand excess (L_T/E_T), and we have equation (2) where P_B is the bound ligand fraction [EL]/L_T and P_B^E is the fractional occupation of the binding site by ligand L. [31]

$$A_{\mathrm{STD}} = \varepsilon \frac{I_{\mathrm{STD}}}{I_0} = \frac{\alpha_{\mathrm{STD}}[L]}{[L] + K_{\mathrm{D}}} \quad \textbf{(2)}$$

This equation reminds the well-known equation of Henri-Michaelis-Menten enzymatic reaction rate v_0:

$$V_0 = \varepsilon \frac{V_{\max}[S]}{[S] + K_{\mathrm{m}}} = \frac{\alpha_{\mathrm{STD}}[L]}{[L] + K_{\mathrm{D}}} \quad \textbf{(3)}$$

So to estimate the K_D value for a given ligand-receptor interaction it is only necessary to fit the A_{STD} values across the titration of the ligand resonances to equation (2), the lowest K_D value is taken as the more accurate result.

Another method developed by Cheng and Prusoff using a known inhibitor on a competitive STD experiment also made possible the determination of the dissociation constant. Using equation (3) where IC_{50} is the value where 50% of the protein is inhibited by I with a known K_D value, the ligand K_D is readily obtained in a single experiment. [62]

$$K_{\mathrm{D}} = \frac{K_{\mathrm{i}}[L]}{IC_{50} - K_{\mathrm{i}}} \quad \textbf{(4)}$$

The STD-NMR technique once again gives different possibilities to be acknowledgement of the nature of the binding site, and secondary binding sites or allosteric changes must be a never to forget factor, being an approximation the results obtained are quite relevant.

3.5.5. The Binding Isotherm of STD Initial Growth Rates

Direct approaches have failed to give correct values of equilibrium dissociation constants (K_D) from STD NMR spectroscopic titrations, since it has been demonstrated that the magnitudes of the determined constants depend on the particular STD user defined signal of the ligand chosen to build the

corresponding binding isotherms. Leading to of a relatively large uncertainty that might be inherent to the determination of K_D by STD NMR spectroscopy, precluding the use of this technique for accurate measurements of protein–ligand affinities in solution.

In 2010 Angulo and Nieto presented a protocol to accurately solve STD NMR quantitative studies of protein–ligand interactions, some of which involving multimodal binding ligands on one binding site, in which STD intensities can be perceptively affected by ligand cross-rebinding processes. They have shown that such multimodal systems can be quantitatively treated by an appropriate analysis of STD NMR experiments based on STD initial growth rates. Thus, the experimental system can be deconvoluted as a simple sum of the contributions from each mode.

This approach is justified by considering the effects of fast ligand rebinding during the receptor saturation in the STD experiment. If there is a certain probability that a given ligand molecule reenters into the binding site after a preceding binding event (and this rebinding is fast enough related to the relaxation properties), then the ligand spin populations would be partially perturbed due to the previous transfer step, and hence, its capacity to receive more transfer of saturation from the receptor will be different from that from a fresh ligand molecule. Consequently, the amount of magnetization transferred to the free ligand would be smaller. The rebinding process in STD has not been considered before, probably because typical experimental conditions for standard STD NMR spectroscopy (i.e., for binding detection, screening, and determination of group epitope mapping) disfavor ligand rebinding because it uses high ligand-to-protein ratios. In contrast, protein–ligand titration experiments can involve large fractions of bound ligand, in the low ligand-to-protein ratio regions of the experimental isotherm, increasing then the likelihood of rebinding events. Indeed, we demonstrate herein that rebinding is one of the main causes of errors in the determination of affinities by STD NMR spectroscopic titration experiments.

The apparent K_D obtained be previous titration methods were prone to output a wide range of values. In the dependence of the saturation time of the binding of chitobiose to WGA using A_{STD} from STD NMR spectroscopic titrations, as a function of the saturation time (t_{sat}) it was reported a range from 300 to 730 μM, at 1 and 4 seconds, respectively. For the proton chosen in the case of L-

Tryptophan binding to bovine serum albumin (BSA) the values obtained, for example, at saturation time of 4 seconds, ranged from 3300 to 1060 μM. So a resolution to normalize parameters, such as, saturation time and fraction of bound ligand were planned and proposed by them. [33]

To take in order these considerations a protocol was done by constructing a binding (Langmuir) isotherm as a function of the ligand concentration in the sample using the initial growth rates of the STD amplification factors (A_{STD-0}), instead of the A_{STD} factors at a given saturation time. The initial growth rate corresponds to the A_{STD} value at the limit of zero t_{sat}, when virtually no ligand turnover takes place, therefore, avoiding the potential effects of fast protein–ligand-rebinding processes. The growth of STD-AF values with saturation time can be appropriately described by a monoexponential asymptotic equation, equation 5.

$$A_{STD}(t_{Sat}) = A_{STD.\,Max}[1 - \exp(-K_{Sat}.\,t_{Sat})] \qquad \textbf{(5)}$$

in which $A_{STD-max}$ represents the maximum A_{STD} achievable for a given proton (i.e., for very long saturation times), and k_{sat} is its saturation rate constant. After this first fit, the initial slope (A_{STD-0}) is easily obtained by the product $A_{STD-max}$ by k_{sat}. Afterwords, and to finalize the protocol, the initial slopes are represented as a function of the ligand concentration, and the mathematical fit to a Langmuir isotherm ($y=B_{max}\,X/\,(K_D+X)$) is carried out. [33]

The analysis of the saturation rate, k_{sat}, provides additional evidence for the presence of fast rebinding. This parameter, that determines how fast the plateau of the A_{STD} build-up curve is reached, is not constant throughout the titration, decreasing with ligand concentration (Figure 21). This means that a steady-state is reached at shorter saturation times for low ligand concentrations, as a consequence of the decrease of maximum STD-AF caused by rebinding.

Moreover, different build up rates may lead to a detection of most saturated proton, when it is not, but the one, after a long saturation time, that reached a higher plateau, due to spin diffusion or cross relaxation. [33]

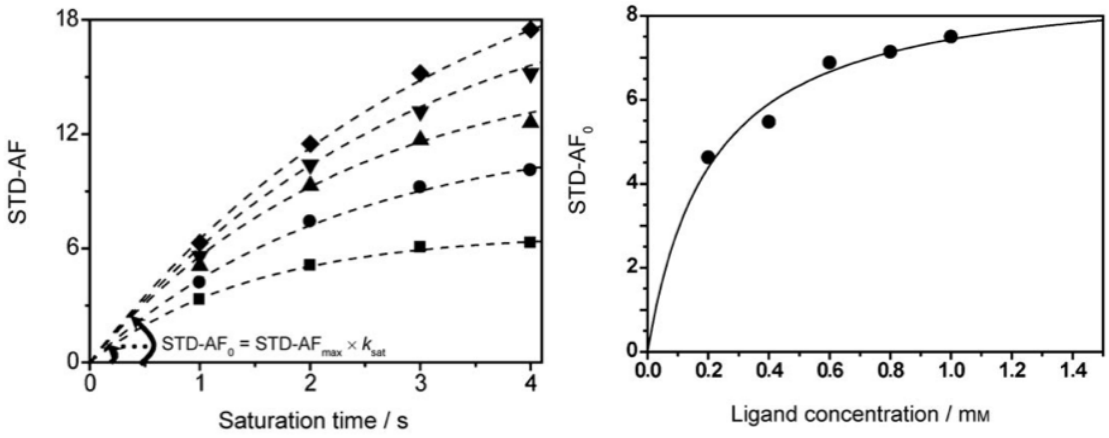


Figure 21. The binding isotherm of STD-AF initial growth rates approach. The new protocol for protein–ligand affinity measurements by STD NMR spectroscopy is illustrated for the L-tryptophan–BSA system. a) For each ligand (l-Trp) concentration (0.2 (■), 0.4 (●), 0.6 (▲), 0.8 (▼), and 1.0 mm (◆)), A_{STD} values) are obtained at different saturation times (1,2, 3, and 4 s) and fit by using the equation 5. b) The initial slopes are represented as a function of the ligand concentration, and the mathematical fit to a Langmuir isotherm, delivering the thermodynamic K_D value.

The similarity of K_D values obtained from the ones reported by different analytical methodologies gives a very good status for this protocol, which, and after taking a second look, takes 4 times the original time in a original titration, but yields accurate proof of the importance of following it for further studies of protein-ligand interactions in a quantitative manner.

3.5.6. COMPETITIVE STD

The first order kinetics adopted by the STD amplification factor on the ligand-receptor interaction revealed a new assay based on the interaction of two ligands with different dissociation constants to the same receptor, the so-called “competitive STD”. First of all the limitation on K_D ranges known for the standard STD may be overcome based on the reduction of the I_{STD} of the ligand fitted on the given range (10^{-3} M to 10^{-8} M). The I_{STD} will be as high as the amount of protein available to dock the ligand and transfer the saturation applied. So, in a mixture of competing ligands, the available binding pocket will decrease and the ligand with the lowest K_D, which will leave less binding pockets available, will decrease the relative intensity of the ligand with lowest affinity.

The protocols adopted for the estimation of small dissociation constants include an initial STD spectrum acquisition with the ligand in study, then addition

of determined concentration of the competitor and finally to obtain a new STD spectrum in the same experimental conditions. Accurate K_D values have been determined by competition studies, in which the unknown K_D of a ligand is determined by monitoring its displacement from the protein binding site while a reference ligand of known affinity is titrated over the same sample, using the Cheng–Prusoff equation. [62]

The unknown constant of the competitor will be calculated based on the reduction of the I_{STD} of the ligand with the known K_D using equation 6.

$$K_I = \frac{C_I(K_D - iK_I)}{i(C_L + K_D)} \qquad (6)$$

The fact that the discribed protocol will lead to a small amount of time needed to depict the stated principles helped STD-NMR technique to become a very powerful tool in this type of study of ligand-receptor interaction, especially when compared with non-spectroscopic assays. After the, sometimes troubled, task of peak assignment is done, the potentialities clearly rise up in this challenging study.

This competition STD NMR approach can also be used for compound library screening of ligands over a wider affinity range including high-affinity ligands that would be missed by the STD NMR method.

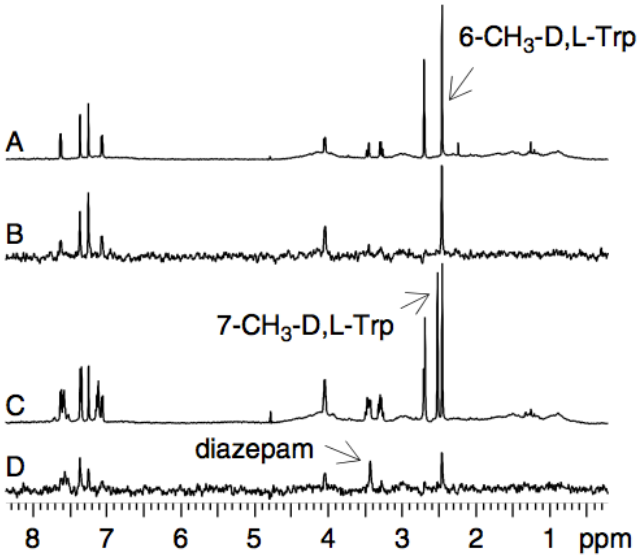


Figure 22. Application of competition STD NMR to screen a compound mixture. (A) 1D proton spectrum of a weak ligand, 6-CH_3-D,L-Trp ◃95 μM) used as an STD indicator, in the presence of 5 μM HSA; (B) the corresponding STD NMR spectrum; (C) 1D proton spectrum of a compound mixture including 6-CH_3 -D,L-Trp (95 μM), 7-CH_3 -D,L-Trp (90 μM) and diazepam (15 μM) with 5 μM HSA; (D) corresponding STD NMR spectrum. [63]

In the example given by Wyss et al (Fig. 22), a substantial reduction in the STD signal intensities of 6- D,L-CH_3-Trp was observed, which indicates the presence of an active ligand in the mixture. A new STD signal appeared at 3.45 ppm and matches the methyl signal of diazepam, identifying diazepam as a ligand. The absence of STD signals from 7-CH_3-D, L-Trp is consistent with the fact that it does not interact with the protein. The observed fractional reduction of the STD signal intensities of the 6- CH_3-D,L-Trp methyl group in the presence and absence of diazepam was 0.41. The dissociation constant of diazepam was estimated to be 2.4 μM, which agrees well with the value of 2.6 μM reported in the literature. We can see that the STD signal intensities of 6-CH_3-D,L-Trp were reduced, the STD signal for the methyl signal of diazepam at 3.45 ppm was observed and no STD signals of 7-CH_3 -D,L-Trp were detected. [63]

3.5.7. STD ADVANTAGES AND LIMITATIONS

The STD technique is based on the NOE effect and comparing it directly with the tr-NOE technique that used 2D NOESY experiments which take typically 4 hours against the 0.5 hour of STD acquisition we can see clearly that, as a screening technique, 1D STD overcome tr-NOE. Moreover the crowded 2D NOESY makes the discrimination of inverted peaks a difficult task.

The WaterLOGSY experiment when used in a mixture of several ligands with overlapped chemical shifts makes also the spectroscopist burden heavy because it becomes very hard to distinguish on each ligand, especially if compared with STD, where only binding ligands manifest their presence in the final spectrum.

The total amount of protein required for an acquisition may vary, depending in protein cost and the capacity of delivering good STD spectra with protein concentrations in the nano molar range an extra positive point to this technique can also added.

The STD methodology is based on the dipole-dipole interactions of the saturation transfer from receptor. This will work just fine on a receptor with a dense proton matrix, such as proteins, but if the work uses for example RNA strands as a receptor, which has a low intrinsic proton density, STD may not

result so effectively. The WaterLOGSY method may be used to overcome this aspect, since the water molecules that surround RNA strand compensate the lack of proton density and therefore is will still be possible to maintain the spin-diffusion and cross-relaxation needed for a good final spectrum.

We can never forget the STD applicability to an almost unlimited number of NMR techniques and having the cascade of selective pulsed added on any NMR experiment we will have the corresponding sequence effect with the inclusion of the STD features.

The dissociation constants are just an estimation of the real value and other approaches may help to this determination below the range values stated throughout this work.

Finally we can say that STD methodology allow us the identification of ligand binding epitope, and will work for any large protein with a dense proton matrix, it doesn't require an isotope labeling, has a very reasonable timescale for the data obtained and even the limit on K_D ranges may be enlarged.

3.6. DOCKING SIMULATIONS

The computational techniques used in drug discovery utilize rapid *in silico* assessment of large libraries of chemical structures in order to identify those structures that are most likely to bind to a drug target, typically a protein, receptor or enzyme.[64]

Significant progress has been made in structure based drug design by pharmaceutical companies at different stages of drug discovery, such as identifying new hits, enhancing molecule binding affinity in hit-to-lead, and reducing toxicities in lead optimization. *In silico* methods based on molecular and quantum mechanics, such as docking, molecular dynamics and *ab initio* chemical reactivity calculations, bring us closer to understand drug metabolism and predict drug–drug interactions, predict drug regioselectivity, stereoselectivity, reactive metabolites, induction, inhibition and mechanism-based inactivation, as well as their implementation in hit-to-lead drug discovery.

The goal of drug discovery is to find the best medicines to prevent, treat and cure diseases quickly and efficiently. To fulfill this goal, computational tools

have helped medicinal chemists modify and optimize molecules to become potent drug candidates, have led biologists and pharmacologists to explore new disease genes and novel drug targets, and have been also guiding drug metabolism scientists to achieve better pharmacokinetic profiles and avoid drug toxicities. [64]

These *in silico* approaches have been widely applied to predict drug absorption, distribution, metabolism, excretion and toxicity (ADMET). The optimization of chemical space in early discovery stage using *in silico* tools will shorten the total drug discovery cycle time and at the same time enhance the late-stage drug survival rate (Fig. 23).

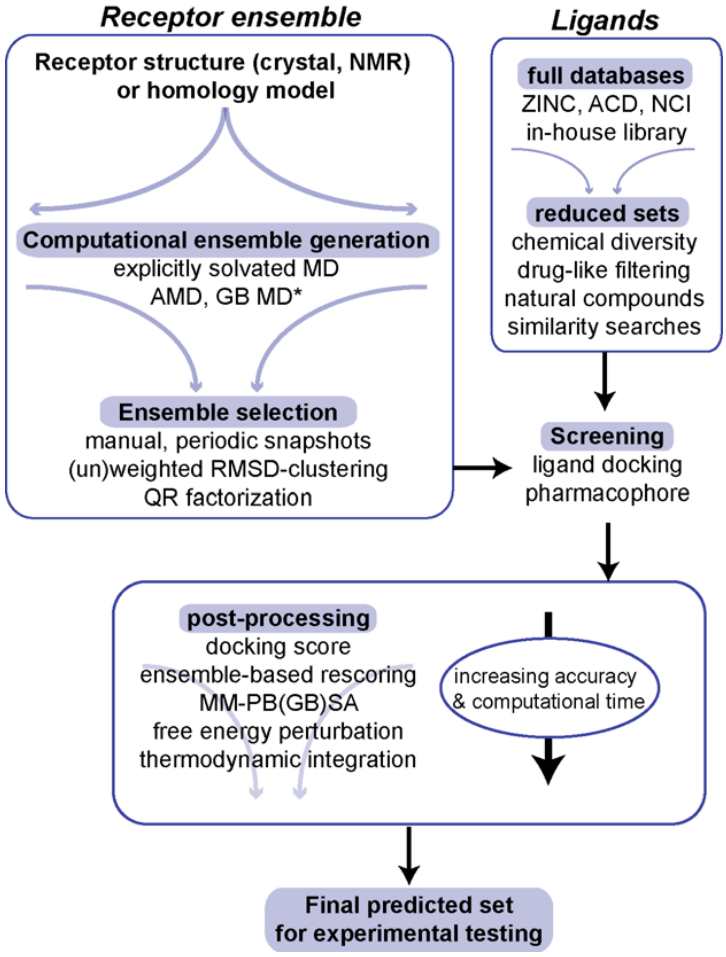


Figure 23. General workflow for ensemble-based virtual screen experiment. Blue arrows indicate size of data sets (i.e. increasing or decreasing) at each step; * denotes emerging methods that have not yet been tested. (AMD: accelerated molecular dynamics, GB MD: generalized Born molecular dynamics, RMSD: root-mean-square-deviation, ZINC – ZINC Is Not Commercial, ACD: Available Chemical Database, NCI: National Cancer Institute, MM-PB(GB)SA: Molecular Mechanics – Poisson-Boltzmann (Generalized Born) Surface Area).

In silico drug metabolism prediction methods are ligand-based such as building pharmacophore and QSAR quantitative structure–activity relationship (QSAR) modeling, before structure-based drug design emerges.[65] QSAR modeling still plays a big role in pharmaceutical industry because of the significant growth of high-throughput screening data. In QSAR, neural networks, classification methods such as recursive partitioning, decision trees, and support vector machines are constructed with topological, structural and electronic descriptors. ADMET properties are thus predicted on the basis of these descriptors, such as lipophilicity and solubility. Usually QSAR predictions work particularly well with structurally similar compounds, whereas molecules from other regions of the chemical space can cause outliers. Furthermore, QSAR is limited by the quality of the available experimental data. QSAR can sometimes provide hints about the active site if according descriptors appear in the regression equation, such as steric constraints and lipophilicity. Pharmacophore models are constructed to overlay structures of all ligands also to simulate the spatial and chemical properties of the binding site. [66]

Pharmacophore models, structure-based docking and molecular dynamics are all mechanism-based approaches. Recently, the prediction power has been improved in building 3D-QSAR [67] – Figure 24 - pharmacophore models with spatial atomic descriptors in consideration. These descriptors include molecular interaction fields, electronic properties and shapes of active sites. Indeed, structure-based modeling by combining molecular and quantum mechanics, enables us to predict substrate affinity, lability and metabolic pathways. In practice, mass spectrometry, nuclear magnetic resonance spectroscopy, proteomics, metabolomics, and other advanced bioanalytical techniques have been improved significantly, which can help us validate these *in silico* predictions quickly.

Water molecules often play an essential role in drug–enzyme interactions: the displacement of water molecules from binding site is one main source of binding free energy. [68] Water molecules were found in the binding pockets of drug-metabolizing enzymes, which can make favorable interactions with either ligands or protein active site residues. Therefore, active-site water molecules are often included in computational methods as one of the strategies to predict drug metabolism.

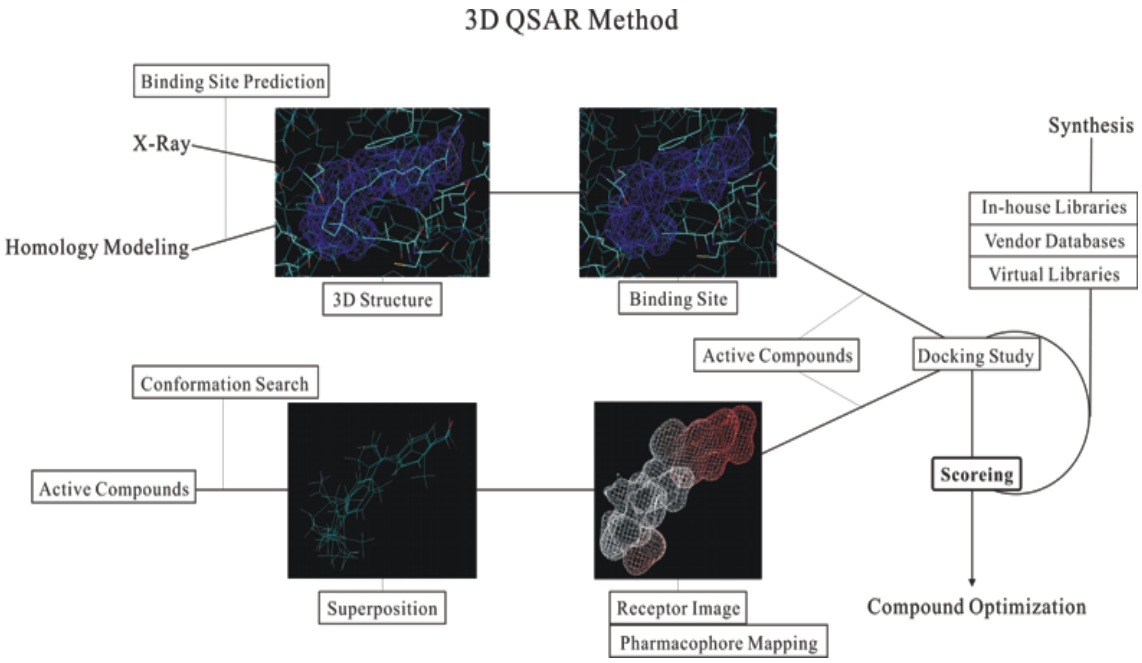


Figure 24. General workflow for 3D QSAR method.

A continuous time-course structural determination would be an ideal method to visualize the dynamic change of active site residues, but still, not possible in the current experimental settings. Applying this strategy, the accuracy of docking algorithms was significantly enhanced. [68]

More rigorous molecular dynamics based approaches to include full receptor flexibility are now being applied in larger-scale virtual screening applications, thanks in part to steady increases in computational power and the development of new and improved methods improving computational efficiency of the underlying strategies. [69]

Some approaches also approximate binding free energy by adding up individual contributions of different interactions. However, the individual energy terms are derived from physico-chemical theory and are not determined by fitting to experimental affinities. In most cases, gas phase molecular mechanical contributions are combined with solvation free energies. Evaluation of solvation energy is a challenge both in terms of computational demands and accuracy. The methods used to calculate solvation include implicit solvent methods like Possion-Boltzmann and surface generalized Born methods. The gas phase energy calculations depend on the type of the force field, for example, AMBER and CHARMM.

Protein flexibility and the dynamics of intermolecular interfaces can regulate binding affinity and specificity in molecular recognition. It has been suggested that structural stability and flexibility during molecular recognition are associated with the ruggedness of the underlying binding energy landscape and can be related to various functions, such as specificity or permissiveness in recognition. Hierarchical approaches incorporating both ligand and protein flexibilities have contributed to recent progress in ligand-protein docking. Such procedures include multistage docking approaches and a hierarchy of energy functions that aim to capture the subtleties of protein flexibility upon ligand binding. Two of the best programs, and more widely used, for these exhaustive and accurate simulations are Autodock and HADDOCK.

3.6.1. AUTODOCK VINA

The method used in this work for virtual ligand screening, was Autodock Vina. This recently developed program improves approximately two orders of magnitude speed-up compared with the molecular docking software previously developed by the Scripps Research Institute, AutoDock 4.2, while also significantly improving the accuracy of the binding mode predictions (Fig. 25). These advantages clearly place this program as an excellent docking software. [25]

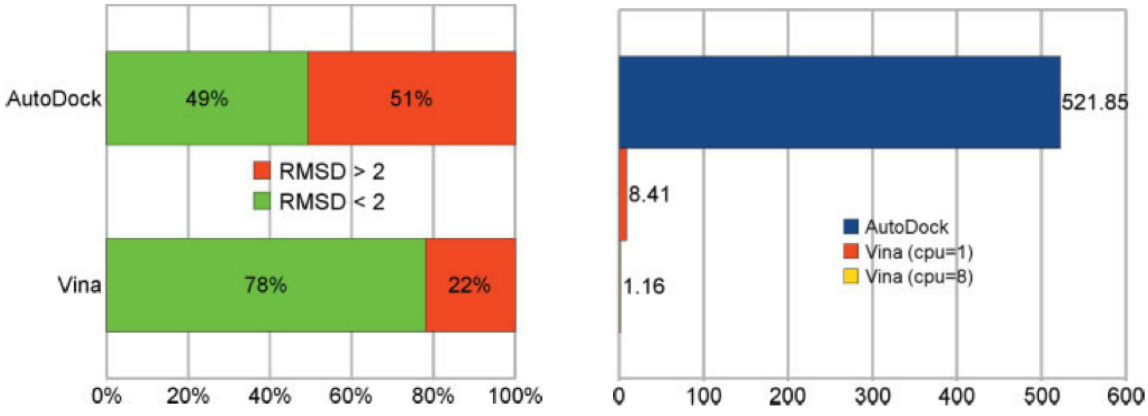


Figure 15. The fraction of the 190 test complexes for which RMSD < 2 Å was achieved by AutoDock and Vina; B) Average time, in minutes per complex, taken by AutoDock, single-threaded Vina and Vina with eight-way multithreading. The time for AutoDock does not include the necessary initial AutoGrid run. [25]

Among the assumptions made by Autodock Vina is the commitment to a particular protonation state of and charge distribution in the molecules that do not

change between, for example, their bound and unbound states. Additionally, docking generally assumes much or all of the receptor rigid, the covalent lengths, and angles constant, while considering a chosen set of covalent bonds freely rotatable (Fig. 26).

Importantly, although molecular dynamics directly deals with energies (referred to as force fields in chemistry), docking is ultimately interested in reproducing chemical potentials, which determine the bound conformation preference and the free energy of binding. It is a qualitatively different concept governed not only by the minima in the energy profile but also by the shape of the profile and the temperature. [25]

Docking programs generally use a scoring function, which can be seen as an attempt to approximate the standard chemical potentials of the system. When the superficially physics-based terms like the 6–12 van der Waals interactions and Coulomb energies are used in the scoring function (based on AMBER force field), they need to be significantly empirically weighted, in part, to account for this difference between energies and free energies.

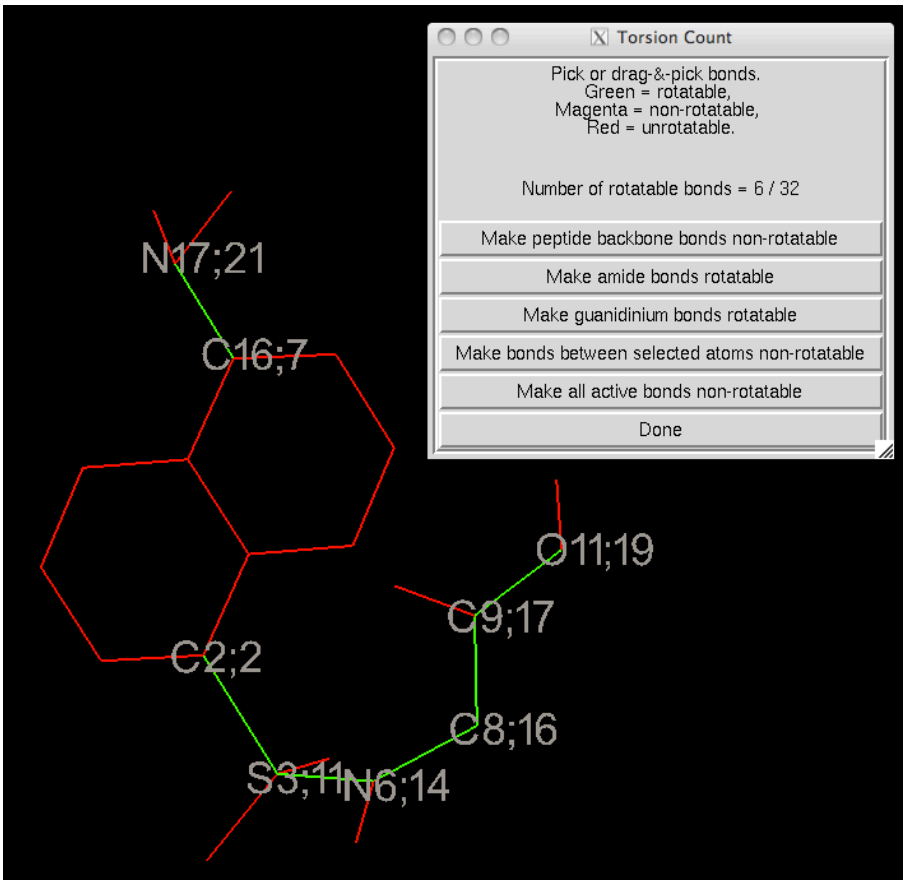


Figure 26 View of autodock Tools for dansyl sarcosine when picking the torsion tree for molecular structure.

Additionally, manually choosing the atom types for grid maps, calculating grid map files with AutoGrid, and additional time waiting for this results is no longer necessary, as Vina calculates its own grid maps quickly and automatically. It also clusters and ranks the results without exposing the user to intermediate details (Fig. 27).

A frequently encountered source of issues for AutoDock 4 users has to do with the fixed data structure sizes in the program: the maximum number of atoms, the maximum number of rotatable bonds, the maximum grid map size, etc. These limits are fixed at compile time, and setting them higher might waste memory and time. To change these limits to meet their needs, the users are required to alter them in the source code and recompile AutoDock 4—a task too daunting and error-prone for many users. In contrast, Vina does not have any such limitations, for practical purposes. The program adapts itself to the input.

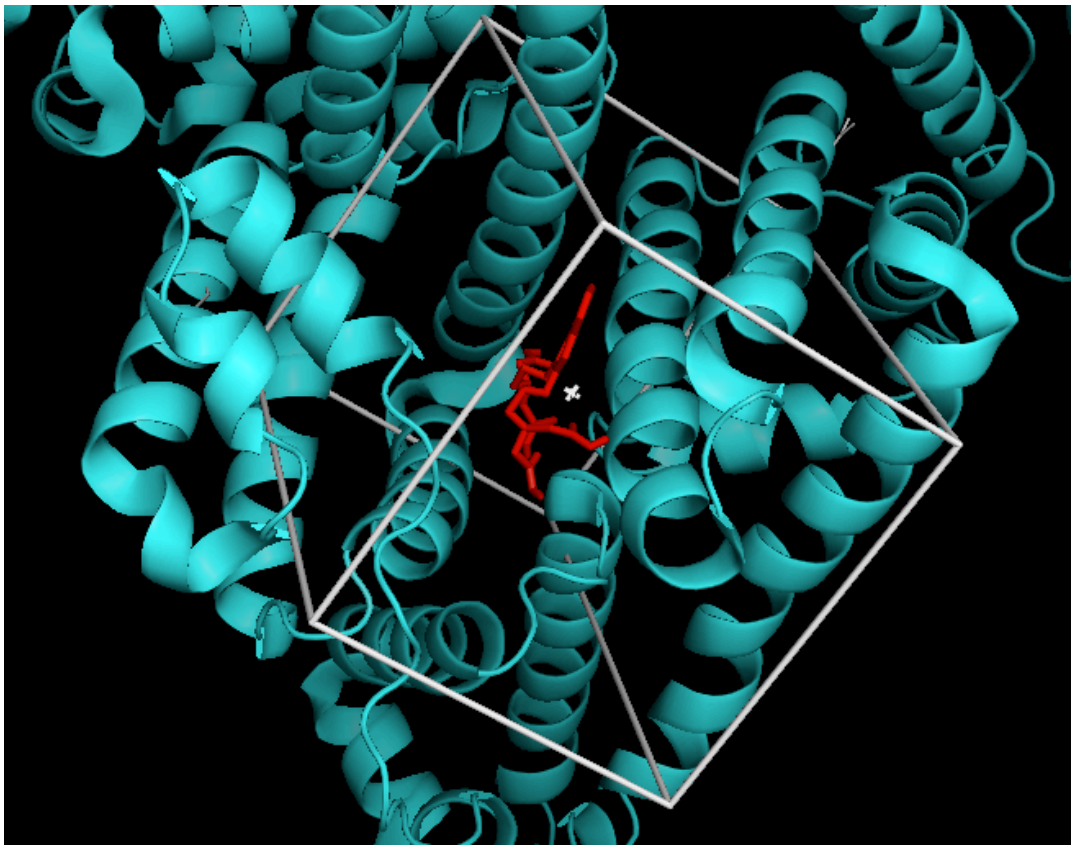


Figure 27. View from Autodock Tools when selecting the grid box to be used in docking simulations.

All Autodock Vina docking simulations steps can be summarized as:

- Protein or receptor input in pdb extension and polar hydrogen atoms are added to it and the nonpolar hydrogen atoms are merged. This can now be saved as receptor.pdbqt.
- Ligands are open as pdb files and automatically are added gasteiger charges, merges non-polar hydrogen, and detects the aromatic carbons. We can choose the torsion tree to respect our previous knowledges. This can now be saved as the ligand.pdbqt file.
- After detecting the protons prone to establish the interaction the Grid Box is built, and large enough to accommodate the ligand to move freely.
- Finally all pdbqt files are established to configure the run. We have the receptor, the ligand and the grid dimensions as well as its coordinates. Run the software and expect good results

4. MATERIALS AND METHODS

4. Materials and Methods

4.1 SOLUTIONS

The ligands were complexed with La^{3+}, which was chosen in order to replace and simulate the Gd^{3+}, due to its diamagnetic characteristics and similar radius, 195 and 180 pm respectively. Human Serum Albumin was purchased from Sigma-Aldrich (ref. A3782), essentially fatty acid free, ~99% (agarose gel electrophoresis), lyophilized powder. Ibuprofen and Warfarin were also purchased from Sigma-Aldrich. All solutions were prepared in the absence of buffer, using deuterium oxide 99,9% ^2D, also from Sigma-Aldrich.

La(BOPTA), DTPACholate and NAFTO-EGTA were received, respectivelly from Professors Silvio Aime and Mauro Botta from Dipartimento di Chimica, Università di Torino, Itally and Dipartamiento di Scienze dell' Ambiente e della Vita, Università del Piemonte Orientale “Amedeo Avogadro”, Alessandria, Italy. (DTPA-Cholate) and (NAPHTO-EGTA) were mixed with $LaCl_3$ in a 5% ligand excess solution. The pH was adjusted to 10 using NaOD and DCl stock solutions, while La(BOPTA), which was already prepared and crystallized, and the pH adjusted to 7 using NaOD and DCl stock solutions. The complex solution was prepared in 99.9% deuterated water solvent and non-buffered. It was achieved a clean homogeneous solution was obtained in all cases.

The SSS-(Δ)-$[Y.L^1]^{3+}$ and RRR-(Λ)-$[Y.L^2]^{3+}$ complexes, and N-dansyl sarcsosine were received from Professor David Parker from Department of Chemistry, Durham University, UK, already in complexed form with Yttrium (III), the pH was adjusted to 7 using NaOD and DCl stock solutions.

HDMPP, Vanadate solutions and vanadium-DMPP complexes were received already prepared, at pH around 7, by Prof. Dr. Margarida Castro from Department of Life Sciences, Coimbra University.

4.2. NMR STUDIES

All ^1H NMR spectra were acquired using a 3 mm pulse field gradient (PFG) inverse probe (Idpfg_3mm) on a Varian NMRS-600 NMR Spectrometer working at 599.72 MHz. For each sample 1D ^1H spectra were acquired using all the same parameters used for the STD NMR spectra, with exception to the number of scans. For ^1H spectra 128 to 256 scans were performed while for the STD NMR

spectra 1024 to 2048 scans were collected. The acquisition time was set to 2 seconds, the repetition time was 5 seconds, in the STD experiments a selective saturation pulse of 5 seconds was applied over the protein. For the binding isotherms an array of selective saturation pulses was done from 0.5 seconds to 5 seconds. A line broadening of 0.75 Hz, corresponding to the digital resolution was used in all spectra for analysis. All spectra were analyzed using MestreNova Software v5.1.1-3092.

For plotting procedure all data treatment was done in GraphPad Prism.

4.2.1. The gradient-selected COrrelation SpectroscopY (COSY)

COSY is a 2D NMR technique that gives correlations between J-coupled signals by incrementing the delay between two 90^{o} proton pulses. The resulting 2D spectrum is generally displayed as a contour plot, which is similar to a topographical map. When looking at a contour map, we are actually looking down at a cross-section (slice) of a 3D-image of an NMR spectrum. The usual 1D spectrum is traced on the diagonal of the plot and any peaks that are not on the diagonal represent cross-peaks or correlation peaks that are a result of J-coupling. Thus, by simply tracing a rectangle using the diagonal and cross-peaks as vertices you will know which protons are coupled to each other. Standard COSY experiments require phase cycling to remove unwanted signals and thus can be quite time consuming. This can be largely circumvented using gradient-selected COSY (gCOSY), which utilizes pulsed field gradients to destroy unwanted z-magnetization and hence their associated signals (axial peaks). Depending on concentration quality gCOSY spectra can be acquired in as little as 5 minutes.[71]

4.2.2. Acquiring 2D gCOSY

First we need to acquire a 1D spectrum and set the spectral width and the offset frequency and then when transform it in a gCOSY in Experiments > 2D > Homonuclear Correlation Experiments > gCOSY. The number of scans for each FID, in the directly detected dimension (F2) must be inputted regarding different solutions, we may increase the number if we have a dilute sample. The number

of increments is responsible of defining the resolution in the indirectly detected dimension (F1). The more increments and , the higher the resolution is.
Change 90^{o} pulse to a 45^{o} pulse if you want to decrease the size of the diagonal and emphasize active coupling partners. We can also change “tau” = propagation time that is usually 0, but we can set it to 0.2 to emphasize long-range couplings.
If presaturation is desired, set “satmode='y'” and verify that select the frequency value previously found in the 1D setup of presaturation.
After the collection and process of the 2D spectra we can Increase number of levels on contour plot using the command “dpcon(15,1.2)”. The “dpcon” flag is for displaying the contours. The first number (15, in this case) is the number of contour lines (default is 4). The second number (1.2, in this case) is the relative spacing intensity (default is 2).

4.3. STD-NMR STEP by STEP

4.3.1. Compiling a new sequence in the system.

The BIOpack with the DPFGSE and the full NMR sequence came in a .tar.gz extension, a normal compacted file from Linux OS. The content is easily extracted and we will have the corresponding folders: maclib, manual, parlib, psglib and templates. We need to copy their content to the corresponding folders in our sytem (normally the path is User/vnmrsys/, this is the main folder for any given user). Then join User/vnmrsys/psglib/ folder, and open the terminal with the command line of the OS. To generate and successfully complete the sequence one must type “seqgen name of the sequence to be installed.c”. This will automatically place all of the needed files in our corresponding folders and VNMR will be able to recognize all the sequence parameters. A message will indicate the success of the compilation. Open VNMR and the sequence is working, if the compilation was successfully achieved.

4.3.2 Setting up the experiment

4.3.2.1 Calibration of DPFGSE for water suppression

The most common way to calibrate the DPFGSE sequence is to load the ^{1}H NMR experiment with the active flag for water suppression. For this, go to

Experiments > User Defined > 1D DPFGSE water suppression. To visually display the sequence (Fig. 28) one must type “dps” (display sequence) on the command line or go to Acquisition tab and select ”sequence” and automatically the NMR sequence of the experiment is displayed.

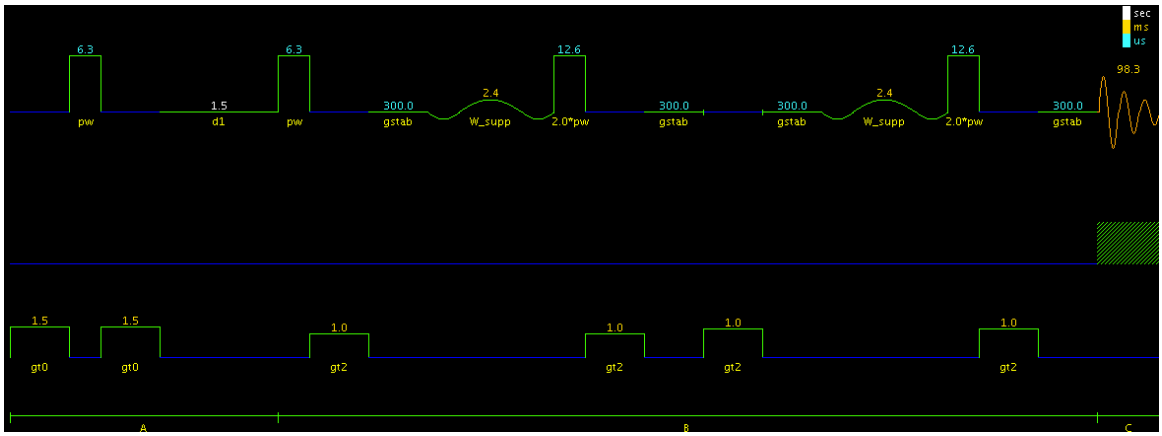


Figure 28. 1D DPFGSE NMR sequence as displayed in VNMRJ v2.3A software. Values indicated in the Figure are those defined by default.

Explaining what is visually outputted, the sequence represents the G1 – soft180^{o} – hard180^{o} – G1 - G2 –soft180^{o}–hard180^{o} – G2 scheme reported by Hwang and Shaka. [51]

The selective pulse necessary for selective removal of water resonances is termed, in our system, “W_supp”. As depicted in the above Figure 28, the word W_supp appears below the selective pulse, meaning that this is the selected pulse shape to be applied. However, if it's the first time this sequence block (DPFGSE) is used (say we were running carbon acquisition), it is necessary to generate the selective wave since it does not exist by default in the system. The software will generate the pulse, calibrating its power to the length defined by the user and accordingly to the parameters defined in the probe file for 90^{o} pulse width (pw) and our transmitter power (tpwr). The selective pulse length is defined in the width box of the same menu. A pulse width around 2 ms is referenced as a good value. [51] A too short pulse width will be translated in insufficient suppression and a too long pulse width will generate J-coupling artefacts in the final spectra. So if one has any doubts we might as well perform an array for this parameter.

To accomplish this go to the Acquire tab > Pulse Sequence menu and click the Recreate Water Refocusing Shape button, after selecting the pulse width for the pulse. When the pulse is generated, a big yellow message is displayed informing the user that in order to achieve good water suppression only the fine power for water reference shape ("wrefpwrf") parameter could be calibrated. The menu and the message are depicted in Figure 29. An array calibration protocol, similar to the one used for 90^{o} pulse calibration, fulfills the needs of this approach, if desirable. So we search for the value that yields the less intense water signal. It is recommended from NMR technicians and developers that steady state scans be performed prior to acquisition, for this sequence 16-32 should be enough. After correct calibration, the same parameters can be used in every sequence that allows DPFGSE, obviously if the same probe is used.

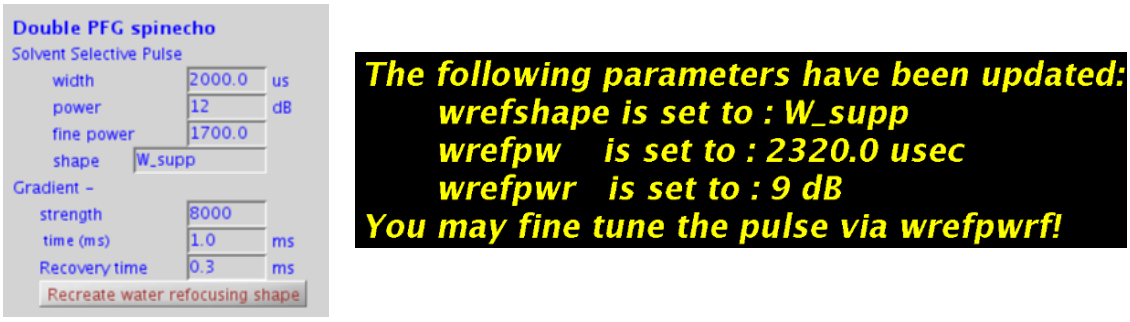


Figure 29. Menu to DPFGSE and the message saw by the user as displayed in VNMRJ v2.3A software.

4.3.2.2 The STD experiment

To load the Saturation Transfer Difference experiment go to Experiments > User Defined > Saturation Transfer Difference. We can also display the sequence by typing "dps", the display will show the corresponding NMR sequence.

First, define the water suppression parameters with the values previously calculated and set all the other basic parameters of acquisition to their previously optimized values.

4.3.2.2.1. Selective Saturation

4.3.2.2.1.1 Pulse Shape

A Gaussian shaped pulse is standardly defined as selective irradiation. Nevertheless, to change the pulse shape type “satshape='shape'” in the command line, where shape is the name of the file that defines the shape without the .RF extension. Be sure that the shape you select is already created and placed in the correct directory. If a non-existent shape is selected the software will output no error, and the user might be misled.

4.3.2.2.1.2 Pulse Power

An important task to accomplish here is the calibration of the Gaussian pulse to (γ/2¶) B1 = 86 Hz, according to Mayer and Meyer. In our system we add a small doubt that was “How to convert Hz in dB?”

So, and with the help from Varian technician Péter Sándor, we discover the “attval” parameter. Once in the command line type “attval(pw90,tpwr)”, where “pw90” is and “tpwr” are our own parameters arrayed, if we have a pw of 7 and a tpwr of 54 the typing will be attval(7,54) and it will output a Table similar to Figure 30.

Attenuator Value	PW90(usec)	B1(Hz)
63	2.5	100675.3
62	2.8	89725.1
4	2215.9	112.8
3	2486.3	100.6
2	2789.8	89.6
1	3130.2	79.9
0	3512.2	71.2
-1	3940.9	63.4
-2	4421.8	56.5
-3	4961.5	50.4
-4	5567.0	44.9

Figure 30. Table to calculate the pulse width in microseconds for a B1 of 86 Hz as displayed in VNMRJ v2.3A software.

Search in the Table for the attenuator value that gives a B1 field of 86 Hz. Then, type satpwr = B1hz, where B1 in Hz is the value determined by the Table.

If desired, the pulse width of selective irradiation can be modified with the satpw parameter.

4.3.2.2.1.3 Selective saturation Frequency

In order to correctly select the desirable frequency to saturate, first acquire a normal ^{1}H NMR spectrum, which will show the complete profile of the sample, ligands and receptor. Using an 86 Hz Gaussian pulse during 30 ms, a region of ±15 Hz at 50% amplitude is affected. It is important that resonances are not affected at ±150 Hz away from the selected frequency otherwise we could not observe the saturation transference effectively.

Placing the cursor in a region of the spectrum where **only protein resonances** are hit, and no ligand overlay, is the only way we to guarantee a successful protocol. Then type “movetof” to that region in the command line and ask the “tof” value typing “tof?”. This value is outputed in Hz and will be the one that defines the STD saturation frequency. The off-resonance reference saturation frequency can be as it is by default, about 16000 Hz, around 30 ppm. One should not forget to replace the transmitter offset to its original value, after choosing the selective region of saturation.

4.3.2.2.1.4 Spin-Lock Calibration

Load the “attval” Table again (Fig. 30) and it assess the spectral width, in Hz, occupied by the protein. This value is the desired B1 field in Hz. Remember considering that half of B1 affects the left side of the transmitter offset while the other half corresponds to the right side. Correlate the obtained B1 value with the attenuator value and this will be the power applied by the spin-lock pulse.

Using the 1D DPFGSE water suppression experiment we can now complete the calibration of the spin-lock filter, since this experiment also has the spin-lock pulse flag. This flag is termed “trim_flg” and can be activated/deactivated. The Trim parameter defines the spin-lock pulse width in seconds. This means that it must be input in seconds, if 30 ms is the desired value then type in the command line trim=0.03. Now, it is only necessary to acquire multiple spectra with different spin-lock lengths to find out what is the minimum amount of time required to full protein relaxation, for different protein samples. Remember to acquire some scans prior to final acquisition, > 16 are desirable, in order to avoid several artifacts.

After the calibration is complete just define the same values in the STD experiment.

Having completed this protocol, do not forget to also define steady state scans in the STD experiment, > 64 are desirable.

4.4. COMPETITION by STD-NMR

The possible competition for the same binding site of HSA was studied by direct displacement of the ligands with the widely known HSA ligands, Ibuprofen, Warfarin, L-Tryptophan and N-dansyl sarcosine. These direct competition experiments were used primarily to determine the binding site of these CAs in the HSA protein according to the level of the displacement relative to HSA and ligand alone.

4.5. DOCKING Simulations

The molecular docking simulations were performed using Autodock Vina 1.1.1. 1 from the Scripps Research Institute. Autodock Tools 1.5.4 (revision 29) was used to prepare all the needed “pdbqt” files for Autodock Vina.

The HSA structure used for the chiroptical probes was the one with the entry 2XVQ.pdb, because this X-ray structure contains the N-dansyl sarcosine, for which is known to interact with drug site II, and therefore, the adequate model to study the possible binding site for our ligands.

To have the most accurate results over HSAs drug site I and II, two X-ray structures were used, both complexed, with the known binders Warfarin and Ibuprofen respectively. The entry code in pdb database is 2BXD for drug site I and 2BXG for drug site II. With this protocol docking simulations were accurately directed to both binding sites and the results are much more accountable than with random structures.

From all structures, containing the different known binders, the chain A was selected and polar hydrogen atoms were added to the HSA and its nonpolar hydrogen atoms were merged.

The structure of the chiroptical probes were obtained from DFT energy minimization calculations performed previously with Gaussian 03, while all of the remaining coumpounds were built in Gaussian 03 software after a clean structure spatial arrangement.

After opening the ligands in Autodock Tools, the program automatically adds gasteiger charges, merges non-polar hydrogens, and detects the aromatic

carbons. Rotatable bonds were set according to the stereochemistry of the molecule, maintaining the chelate cage of the contrast agent, and all the other possible rotatable bonds were left this way.

All calculations for fixed protein flexible ligand docking were done using the Lamarckian Genetic Algorithm (LGA) method. The GRID box parameters used were:

- For 2XVQ, center x = 0, center Y = 5, center Z = -24, dimension 30 x 30 x30 and spacing of 1 Angstrom, with a total of 29791 total grid points per map.

- For 2XBD, center x = 5, center Y = -9, center Z = -7, dimension 20 x 26 x 30 and spacing of 1 Angstrom, with a total of 20181 total grid points per map.

- For 2XBG, center x = -3, center Y = -15, center Z = -9, dimension 26 x 26 x 30 and spacing of 1 Angstrom, with a total of 22599 total grid points per map.

The grid box was set around the ligands binding structure to fully cover the entire binding site and accommodate ligands to move freely. The best pose was chosen with the lowest docked energy, after the docking search was completed.

USFC Chimera (version 1.5.2 build 32411) was used to read the Autodock results and to obtain the images, views presented in this work, as well as the videos presented in the thesis defense.

For the 2D interactions PoseviewWeb 1.97.02, [28,29] was used to represent all possible interactions and fully characterize them. Chain A from all the structures was used and the best poses for all ligands in “mol2” files (obtained from autodock, the pdbqt files that are those carry the binding modes for the ligands, we just save them with this extension and run the 2D simulation with the program). These kinds of extension retain all the Cartesian coordinates like the .pdb files, so the program known where in space is the ligand interacting with the protein.

5. ENANTIOSELECTIVE BINDING OF A LANTHANIDE (III) COMPLEX TO HSA STUDIED BY ^{1}H STD NMR TECHNIQUES

5. Enantioselective binding of a lanthanide(III) complex to human serum albumin studied by ^{1}H STD NMR techniques

5.1 Abstract

The enantioselective binding of the (SSS)-Δ isomer of an yttrium (III) tetraazatriphenylene complex to 'drug-site II' of human serum albumin (HSA) was detected by the intensity differences of its STD ^{1}H NMR spectrum relative to the (RRR)-Λ isomer, by the effect of the competitive binder to that site, N-dansyl sarcosine, upon the STD spectrum of each isomer, in the presence of HSA and by 3D docking simulations.

5.2 Introduction

Highly emissive complexes of the lanthanide (III) ions show great promise as chemosensors for potential application in studying biological systems.[72-80] Much current research focuses on the relationship of structure to various aspects of chemical and biological behaviour, in order to improve promising complexes and to design new complexes for specific purposes.[81-86]

In order for a lanthanide (III) complex to have potential utility as a biological probe, it must retain its selectivity and response in the cellular environment, notably in the presence of the high intracellular concentrations of proteins. Recent studies of potential luminescent lanthanide probes have investigated the effect of protein binding on their in vitro behaviour. [86-88] One such investigation yielded a novel example of dynamic helicity inversion following enantioselective protein binding: the (SSS)-Δ enantiomer of $[Ln.L]^{3+}$ binds to human serum albumin (HSA) with a resulting helicity change as measured by circularly polarised luminescence (CPL). [88] In contrast, the (RRR)-Λ enantiomer associates more weakly with the protein and does not exhibit this inversion. This behaviour was unique to the complex structure, and was not replicated by other complexes bearing either the same chromophore or pendant arms. Here, we report efforts to define the regions of the complex, which bind to the protein by employing saturation transfer difference (STD) NMR techniques.

STD NMR requires the use of a diamagnetic metal ion, which precludes the direct study of Eu(III) or Tb(III) complexes. The diamagnetic yttrium (III) is a

good model for Eu(III) and Tb(III), with an ionic radius that is only 0.02 Å smaller than Tb(III) in nine-coordinate systems. Y(III) complexes were synthesized from the (RRR) and (SSS) isomers of L from yttrium acetate, using analogous procedures to those described previously for Eu(III), Tb(III) and Gd(III) complexes of L.[75]

The aim of this study was to see if STD-NMR could distinguish this enantiomeric pair of isomers, and map the epitope of the binding. Also a 3D model obtention for the interaction is aimed after competitive STD.

5.3 Results and Discussion

The full assignment for the complexes was performed (Fig. 31) to allow the epitope mapping. Complete analysis of gCOSY spectra (Appendix 1) was achieved to fulfill this demand.

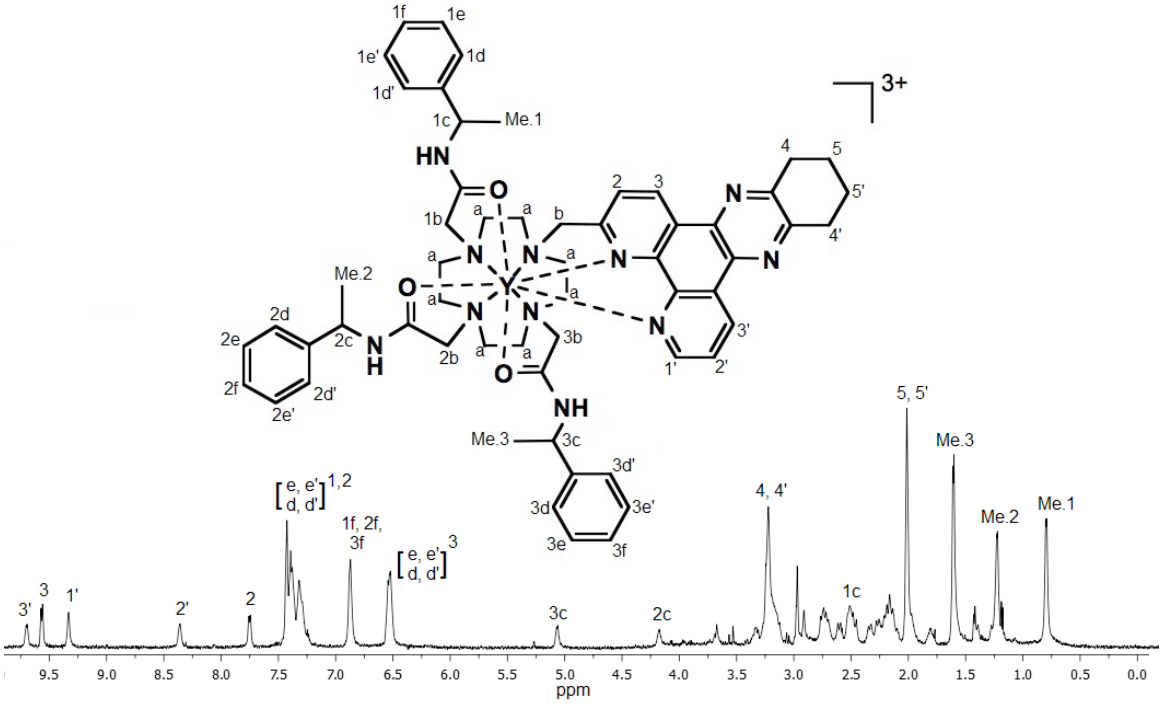


Figure 31. ^{1}H NMR assignments for the *(SSS)*- Δ [Y.**L**]$^{3+}$ complex

We can see that the third pendant arm is the one with the highest value for chemical shifts both for the phenyl and methyl resonances. This shift is consistent with a deshielding ring current effect associated with a through-space interaction involving another aromatic system.

Figure 32 represent the ^{1}H NMR spectra of S-(Δ)-$[Y.L^{1}]^{3+}$ and RRR-(Λ)-$[Y.L^{2}]^{3+}$ respectively. Despite some resonances that we considered as impurities, both spectra are very similar.

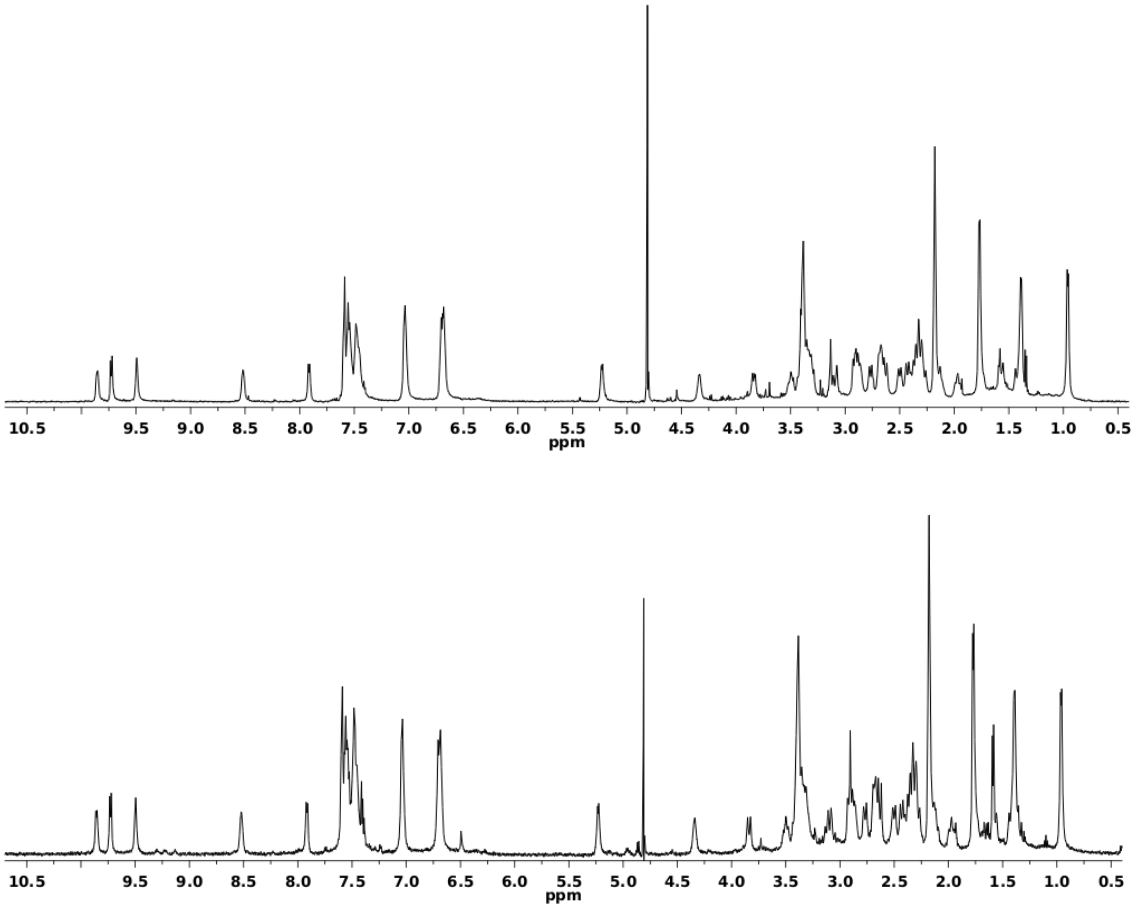


Figure 32. The ^{1}H DPFGSE spectrum of a mixture of 500 μM of *(SSS)*- Δ [Y.**L**]$^{3+}$ complex (upper) and 500 μM RRR-(Λ)-$[Y.L^{1}]^{3+}$ complex (lower). Trim pulse was used in both experiments to suppress the protein resonances.

Figure 33 represents the STD NMR spectra of the two mixtures, SSS- and RRR-, respectively. At naked eye it is quite difficult to distinguish differences between the spectra, however, when both epitopes are carefully analyzed they show very different results. Table 1 gathers the information about peak picking and the determined binding epitope for the two interactions. Only two peaks have a difference in GEM % values below 10, those are the peaks at 8.53 ppm and at 1.93 ppm. The percentage of saturation was determined referent to the peak at proton 1´(9.86 ppm), following the protocol that the highest STD amplification factor is set as 100%.

Table 1. STD amplification factor (A_{STD}) and relative STD (to the most shifted proton) for 900 μM (SSS)-Δ and (RRR)- Λ [Y.L]3+ in the presence of 30 μM HSA.

Proton - ppm	(*RRR*)-Λ $[Y.L]^{3+}$		(*SSS*)-Δ $[Y.L]^{3+}$		% (SSS) – %(RRR)
	A_{STD}	Relative STD (%)	A_{STD}	Relative STD (%)	
1' – 9.6	3.51	100	4.08	100	-
3 – 9.5	2.50	71.2	3.86	94.6	23.4
3' – 9.3	1.46	41.5	2.41	59.0	17.5
2' – 8.35	2.63	75.0	3.24	79.5	4.5
f – 7.04	1.88	53.6	3.19	78.1	24.5
2(e.d) – 6.4	2.48	70.8	3.75	91.8	21.1
4.4' – 3.2	1.65	47.1	2.67	65.4	18.2
5.5' – 2.0	1.74	49.6	2.86	70.0	18.5
Me.2 – 1.6	0.62	17.6	1.03	25.2	7.6
Me.1 – 1.1	1.34	38.2	2.07	50.8	12.5
Me.3 – 0.8	1.21	34.6	2.11	51.7	17.1

Given the HSA binding isotherm for the (SSS)-Δ isomer (log K = 5.1 for a 1:1 binding model) and the observed multi-site binding of the (RRR)-Λ isomer consistent with stepwise formation of various adducts of lower affinity, [88] higher A_{STD} values of the (SSS)-Δ isomer protons are attributed to its stronger protein interaction.

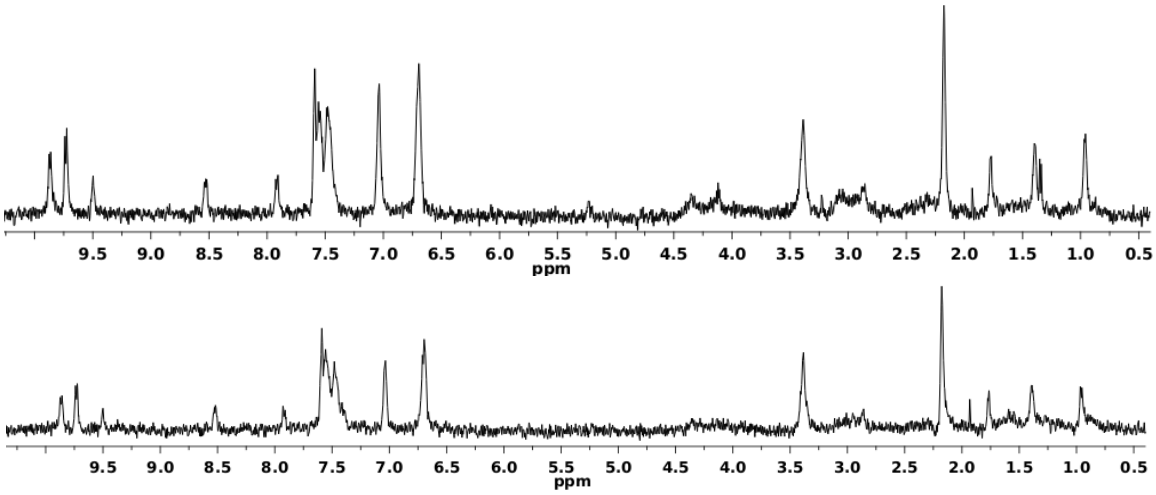


Figure 33. ^{1}H STD NMR spectrum of 1000 μM SSS-(Δ)-$[Y.L]^{3+}$ (upper) and 1000 μM RRR-(Λ)-$[Y.L]^{3+}$ (lower) with 30 μM HSA. Trim pulse was used in both experiments to suppress the protein resonances. Saturation was performed at -0.5 ppm.

The protons that appear in the standard NMR spectrum, but not in the STD spectrum, correspond to the cyclen ring and the amide methylene groups, which do not interact with the protein, consistent with their location in the centre

of the complex, shielded by the pendant arms and chromophore. In addition, protons from the pendant arm that interacts with the chromophore and the cyclohexyl ring of the chromophore do not appear in the STD spectrum, or have very similar ASTD values to the (RRR)-Λ enantiomer, indicating that this moiety is not in the main core of the interaction with HSA.

A competitive test was carried out with N-dansyl sarcosine with both isomers and with 30 μM HSA. Figure 34 represents the STD NMR spectra of SSS-(Δ)-$[Y.L^1]^{3+}$ and RRR-(Λ)-$[Y.L^2]^{3+}$ with N-Dansyl sarcosine. A few differences for the isomers are visible, but after the integration and analys, a more close competition for the binding site by SSS-(Δ) isomer was depicted.

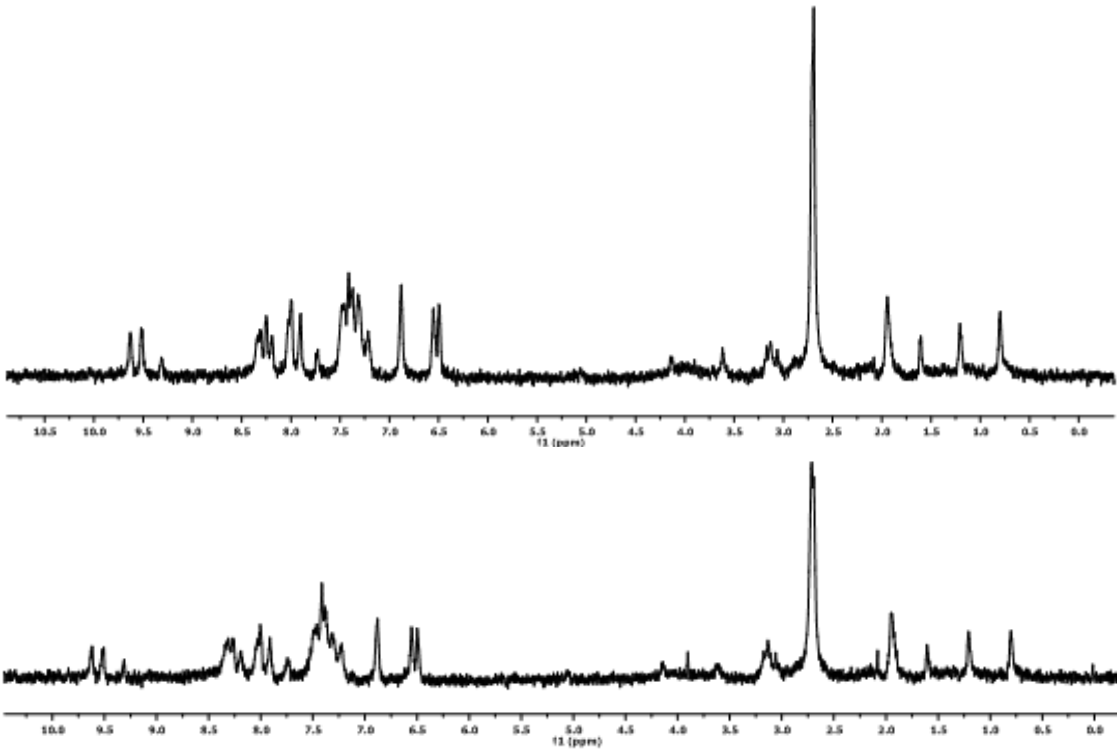


Figure 34. ^{1}H STD NMR spectrum of 900 μM SSS-(Δ)-$[Y.L]^{3+}$ (upper) and 900 μM RRR-(Λ)-$[Y.L]^{3+}$ (lower), with 1000 μM N-Dansyl Sarcosine and 30 μM HSA. Trim pulse was used to suppress the protein resonances. Saturation was performed at -0.5 ppm.

STD amplification factors for the (SSS)-Δ isomer in the presence of N-dansyl sarcosine are vastly different from those measured in the absence of the competitor, while there is less difference for the (RRR)-Λ enantiomer (Table 2). These results corroborate the findings of the relaxivity studies. [88]

Table 2. Change in relative STD (to the most shifted proton) for 900 mM *(SSS)*-Δ and *(RRR)*-Λ [Y.**L**]$^{3+}$ in the presence of 30 mM HSA upon addition of 2 mM *N*-dansyl sarcosine.

% of change after N-dansyl-sarcosine adition	
***(RRR)*-Λ**	***(SSS)*-Δ**
-	-
8.8	-30.5
12.0	-36.8
23.9	24.7
0.5	-62.8
15.4	-37.7
-16.5	-41.1
-13.1	-33.8
-4.9	-14.6
-10.8	-30.6
-0.5	-15.9

In order to provide a 3D model for this interaction, molecular docking simulations with HSA binding site II with both enantiomers was performed using Autodock Vina. [25] This software, however, does not support Yttrium, and thus it was removed, remaining its structurally unchanged cage. Since, due to this limitation, three units of positive charge and its subsequent energy interaction are lost, these simulations are considered a qualitative three-dimensional search and not a quantitative model.

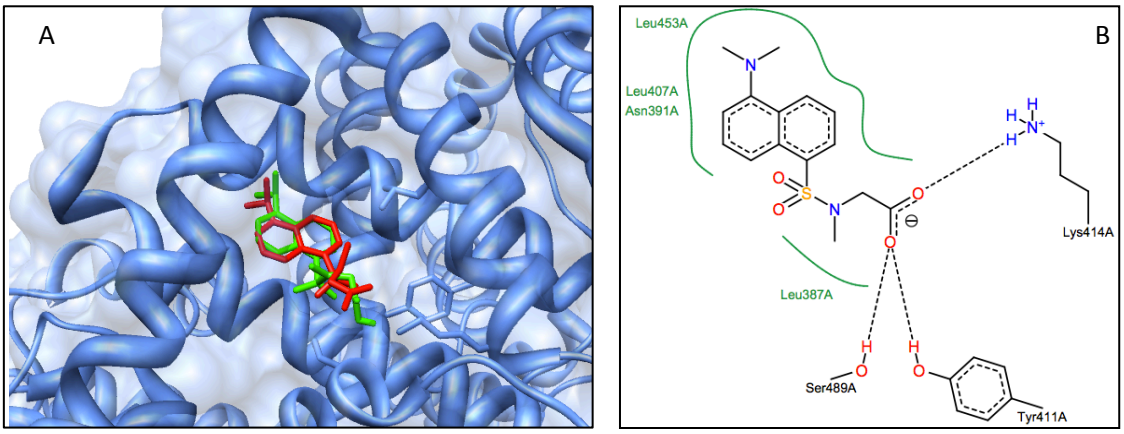


Figure 35. A) N-dansyl sarcosine, a known a binder of HSA drug site II. Overlay of structures from X-ray structure 2XVQ.pdb, (red) and docked with Autodock Vina (green); B) 2D scheme for interactions between N-dansyl sarcosine and site II of HSA, (black dashed lines – hydrogen bonds; green solid lines – hydrophobic interactions)

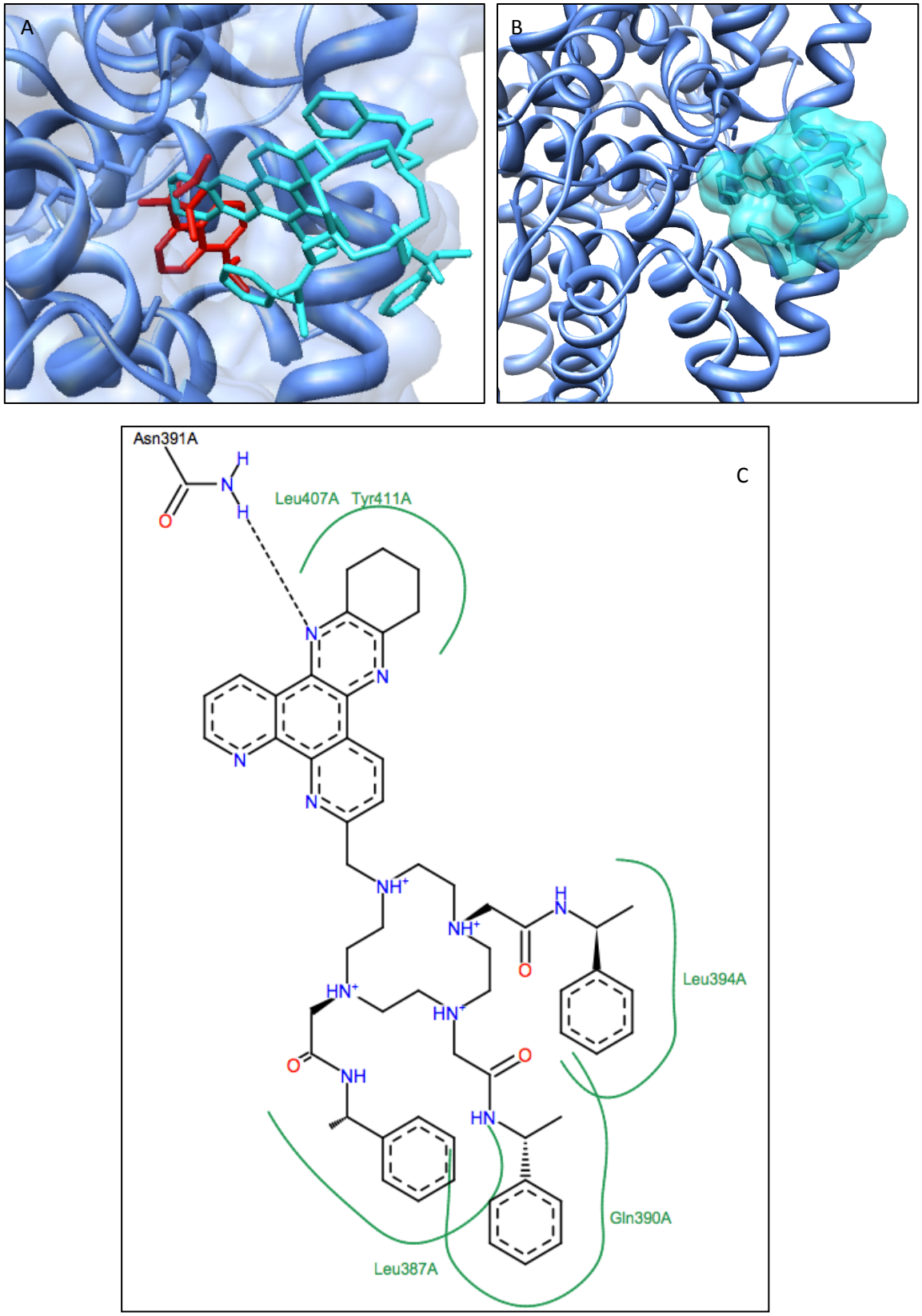


Figure 36. A and B) Docked structure for (SSS)-Δ isomer (cyan) overlaid with and without N-dansyl sarcosine (red), respectively. Indicating a specific competition for this site and a lodging inside the binding site. C) 2D scheme for interactions between (SSS)-Δ isomer and site II of HSA (black dashed lines –hydrogen bonds, green solid lines – hydrophobic interactions; green dashed lines – Pi-cation interactions).

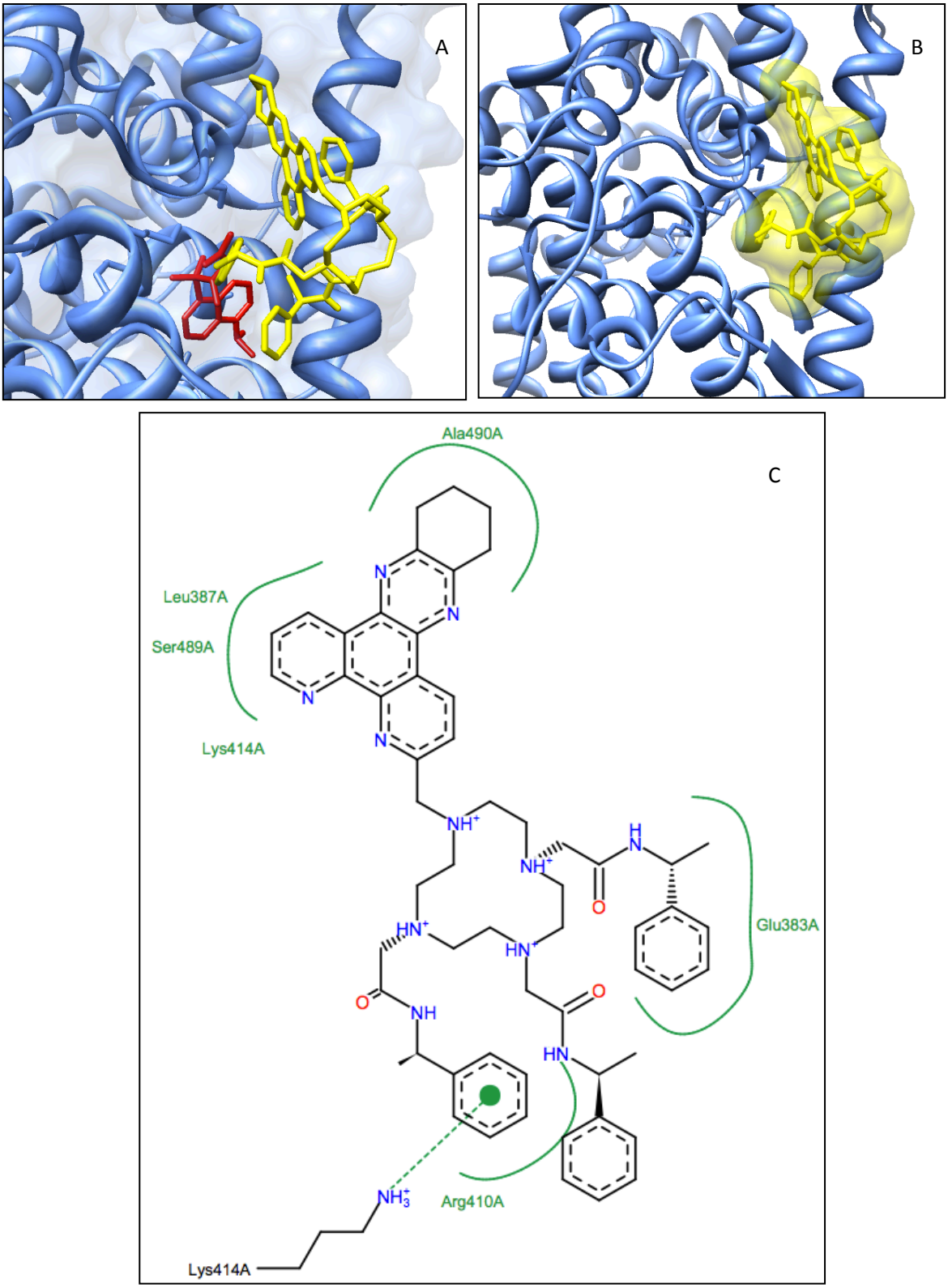


Figure 37. A and B) Docked structure for (RRR)-Λ isomer (yelow) overlaid with and without N-dansyl sarcosine (red), respectively. Indicating a specific competition for this site and a lodging inside the binding site. C) 2D scheme for interactions between (SSS)-Δ isomer and site II of HSA (black dashed lines –hydrogen bonds, green solid lines – hydrophobic interactions; green dashed lines – Pi-cation interactions).

The simulated 3D orientation of N- dansyl sarcosine at the HSA binding site II and the 2D interaction model coincides with the X-ray structure (Figure 35 A,B) and validate this simulation model.

The 3D scheme (Figure 36 A and B) shows that the (SSS)-Δ isomer interacts directly with the drug site II, mainly through the phenylmethyl pendant arms and the aromatic region of the chromophore, as suggested by the STD NMR data. This provides an explanation for the observation that this binding behaviour requires conservation of both the chromophore structure and the pendant arms. A 2D model (Figure 36 C) was also established based on the docking results and using Poseview which highlights the direct competition of the complex with N-dansyl sarcosine for interaction with the same residues, mainly Tyr 411, Asn 391 and Leu 387. This 2D view depicts the hydrophobic contact of the phenylmethyl pendant arms and the chromophore with HSA, which is further stabilized by an hydrogen bond between Asn 391 and a chromophore nitrogen.

However for the (RRR)-Λ isomer 3D simulation, a different interaction is depicted, coherently with the NMR results. The region of the cromophore is not lodged in the same region of N-dansyl sarcosine (Fig. 37), it seems that this interaction is made by the first phenylmethyl arm. Despite the fact that the (RRR)-Λ isomer shares some similar residues of drug site II with the other isomer (also detected by STD-NMR), these 3D simulations support the preference of the protein binding site II for the (SSS)-Δ $[Ln.L]^{3+}$ isomer.

In conclusion, this work has demonstrated the utility of STD NMR experiments in providing detailed structural information about the protein binding interaction. The study confirms that (SSS)-Δ $[Ln.L]^{3+}$ associates selectively with HSA compared to the (RRR)-Λ enantiomer, and identifies the regions of the complex which directly interact with drug site II of the protein.

6. SELECTIVE MRI CONTRAST AGENTS – HSA BINDING MODEL CHARACTERIZATION BY ^{1}H STD NMR

6. Selective Contrast agents – HSA binding model characterization by ^{1}H STD-NMR

6.1 Abstract

STD NMR was used to study the HSA interaction of three MRI contrast agents and the K_D for this interaction with HSA was found for all complexes using the initial slopes protocol. La(BOPTA) and La(DTPA-Cholate) interact mainly with HSA´s drug site I while La(NAPHTO) does not posses a specificity for both drug sites. La(DTPA-Cholate) was found to have a lower K_D than all the other CA´s and the 3D docking simulation model built with X-ray structures of HSA supports all STD NMR findings. The protocol of STD and competitive STD followed by docking simulations was found to be a robust tool to characterize protein-ligand interactions.

6.2 Introduction

It is now well established that the information content of a magnetic resonance imaging (MRI) experiment can be significantly enhanced by the use of a suitable contrast agent (CA). Nowadays about 35% of MRI scans make use of CAs and it is expected that this percentage will further increase with the availability of more sensitive and specific agents. [89–92] A MRI CA is not directly visualized in the image, only its effects on water proton relaxation times are observed. The increased relaxation rates allow the attainment of an intense signal in a short time and a better signal-to- noise ratio by the acquisition of a higher number of measurements.

As unpaired electrons display a remarkable ability to reduce T_1 and T_2, the search for efficient CAs has focused on paramagnetic metal complexes. The metal of choice has been the Gd^{III} ion due to its high paramagnetism (seven unpaired electrons) and for its favourable properties in terms of electronic relaxation.[93] This ion forms very stable complexes with polyamminocarboxylate ligands and Gd-DTPA (Magnevist) was the first MRI CA approved for clinical use. Since then, other Gd^{III}-based CAs similar to Magnevist has been introduced into clinical practice. They have very similar pharmacokinetic properties because they are distributed in the extracellular fluid and are eliminated via glomerular filtration.

The relative ability of a paramagnetic metal complex to act as a MRI CA is expressed by its relaxivity, that is, the property that is a measure of the relaxation enhancement of water protons in solution containing the paramagnetic agent at a 1 mm concentration. [94] Much work has been carried out in the last two decades to improve the relaxivity of Gd^{III}-based agents. It was recognised early on that high relaxivities could be attained for slow-moving systems endowed with fast exchange rates of the coordinated water molecule(s) and suitably long electronic relaxation rates of the unpaired electrons on the metal ion. [91,92,95] On this basis several macromolecular systems have been designed in which the Gd^{III} chelates are either covalently or non-covalently bound to high molecular weight substrates. In particular much attention has been devoted to supramolecular adducts formed by suitably functionalised Gd^{III} chelates and human serum albumin (HSA). Due to their increased retention in the vascular system, these CAs can also be used for 3D imaging of the vasculation in MRA. [95]

In this study three selective contrast agents were studied on their interaction with the main blood carrier HSA, Gd(BOPTA), Gd(DTPA-Cholate) and Gd(NAPHTO-EGTA), using their diamagnetic La(III) analogues. The contrast agents used in this are categorized as contrast agents to magnetic resonance angiography (MRA), and therefore they should probably be carried by HSA. There is the exception on La(BOPA) (MultiHance® from Bracco Diagnostics, Inc.) that is indicated by FDA for intravenous use in magnetic resonance imaging (MRI) of the CNS in adults to visualize lesions with abnormal blood brain barrier or abnormal vascularity of the brain, spine, and associated tissues that can be used for MRA but also for increase contrast in brain tissues and liver.

Direct competitive studies were performed with known inhibitors of albumin binding site I such as Warfarin, and binding site II such as Ibuprofen and L-Tryptophan. This direct competition was used primarily to disclose the binding site of the contrast agent in the HSA protein and with this ligand-protein complex establish the ligand epitope .In order to calculate a good affinity constant approximation and the mode of binding a recent protocol was followed. [33]

A comprehensive and valid model for the interaction was established with docking, for all the studied interactions after having all the quantitative results from NMR.

6.3 RESULTS AND DISCUSSION

6.3.1 La(BOPTA)

Comparing the ^{1}H Wsupp and the STD NMR spectra of a mixture of La(BOPTA) and HSA, are seen in the STD spectrum, almost only those resonances from the aromatic ring, the anchor moiety of the contrast agent. Those resonances near 4.6-4.8 ppm are not visible in the reference spectrum, once they were saturated in the DPFGSE suppression scheme; therefore it is not possible to characterize their degree of saturation in the STD NMR spectrum. However, despite their almost negligible intensity, there are also other resonances from the complex cage present in the STD NMR spectrum.

The aromatic resonances of La(BOPTA) were here treated as a whole due to the fact that they are overlapped. The STD amplification factor (A_{STD}) found for a mixture of 1 mM of La(BOPTA) and 30 μM of HSA (Fig. 38) was 5.04. After the analysis of the STD NMR spectrum of this reference the assays of direct competition were aligned. No GEM was performed due to the fact that only a resonance became visible in the STD NMR spectrum.

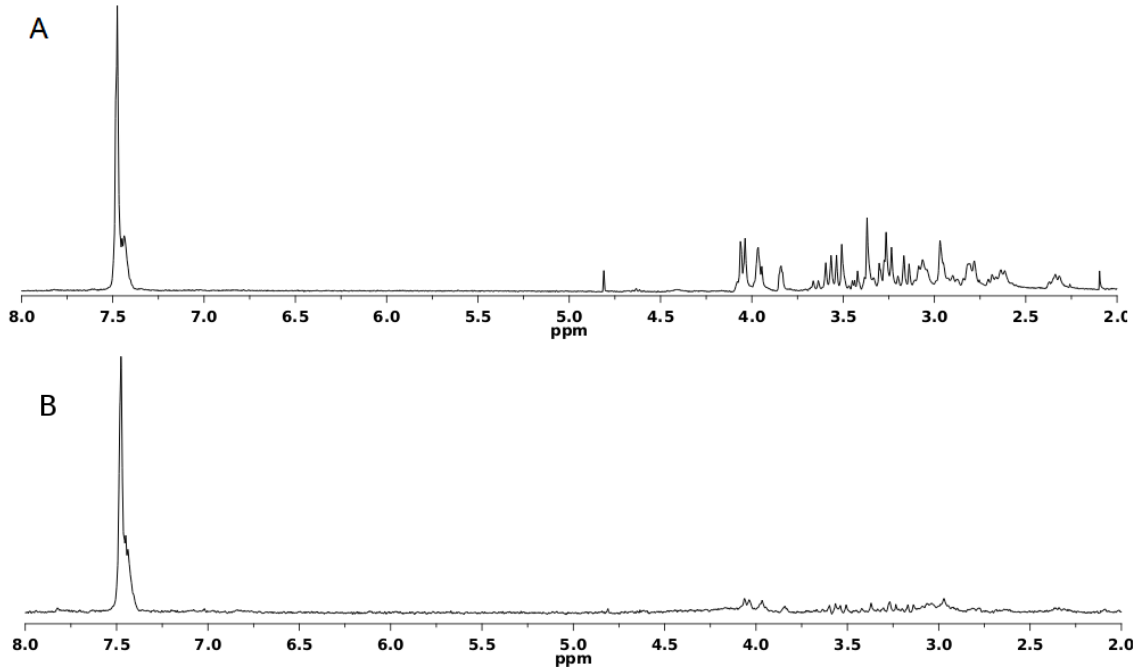


Figure 38. (A) The ^{1}H DPFGSE spectrum of a mixture of 1 mM La(BOPTA), and 30 μM HSA (B) The STD NMR spectrum of the same mixture. Trim pulse was used in both experiments to suppress the protein resonances.

Analysing the ^{1}H Wsupp reference spectrum and the corresponding ^{1}H STD NMR spectrum of a mixture of 1 mM of La(BOPTA), 640 μM of Ibuprofen and 30 μM of HSA (Fig. 39) the A_{STD} of the aromatic signal calculated to be 0.87, and thus it suffered a 83% reduction (since the conditions of the CA and the HSA were maintained, the reduction of the A_{STD} can only be justified by a direct competition for Ibuprofen's binding site).

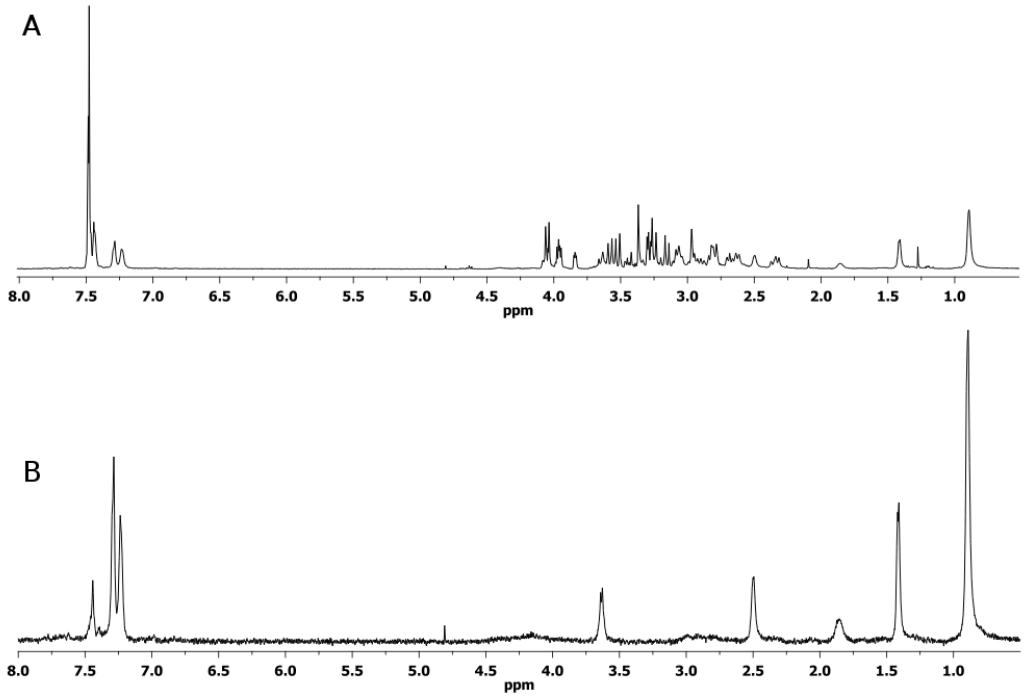


Figure 39. (A) The ^{1}H DPFGSE spectrum of a mixture of La(BOPTA), 1 mM, Ibuprofen 640 μM and HSA, 30 μM (B) The STD NMR spectrum of the same mixture. Trim pulse was used in both experiments to suppress the protein resonances.

In an analogous way the ^{1}H Wsupp reference and STD NMR spectra of a mixture of 1 mM of La(BOPTA), 0.2 mM of Warfarin and 30 μM of HSA (Fig. 40) were used to calculate the A_{STD} of 0.11 resulting of the competition of Warfarin for interaction between the CA and the protein, which means that the presence of the La(BOPTA) resonances in the STD NMR spectrum practically disappeared. In this competition experiment, the obtained A_{STD} for the aromatic moiety of La(BOPTA) suffered a 99% decrease relative to the A_{STD} reference value (Table 3).

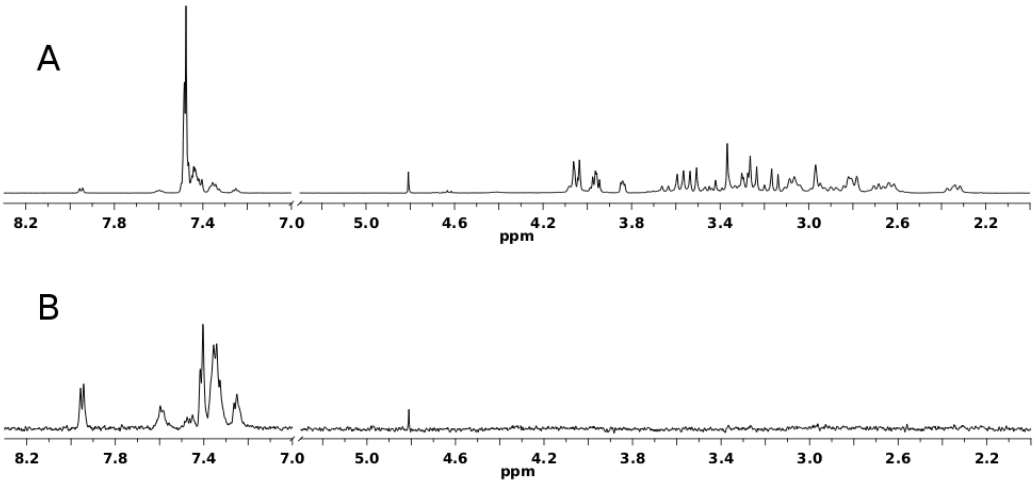


Figure 40. (A) The ^{1}H DPFGSE spectrum of a mixture of La(BOPTA), 1 mM, Warfarin 200 μM and HSA, 30 μM (B) The STD NMR spectrum of the same mixture. Trim pulse was used in both experiments to suppress the protein resonances.

In this competition experiment the A_{STD} obtained along with the observable STD response of the La(BOPTA) in the presence of Warfarin give us an accurate result, and it is correct to assume that the binding of La(BOPTA) happens mainly through the site I of Human Serum Albumin. A competitive experiment between 1 mM of La(BOPTA) and 50 μM of Warfarin in the presence of 30 μM of HSA output a ASTD of 1.31. Although some overlap with the Warfarin resonances is considered, the resonances of the aromatic region of La(BOPTA), despite their lower intensity, are visible.

A control test using a mixture of Warfarin and L-Tryptophan was performed to attest that Warfarin did not displaces the L-Tryptophan from the STD NMR spectra, which is consistent with the information from literature [21]. It is well known that L-Tryptophan binds to the site II of HAS, while Warfarin binds to site I. Therefore, no direct competition was expected. Moreover, this data corroborates the data obtained in the competition of La(BOPTA) and Warfarin.

Thus we performed a competitive test between La(BOPTA) and L-Tryptophan (Fig. 41). The A_{STD} found in these conditions was 3.62, showing a decrease of about 29% relative to the A_{STD} value of La(BOPTA) with HAS (see Table 5).

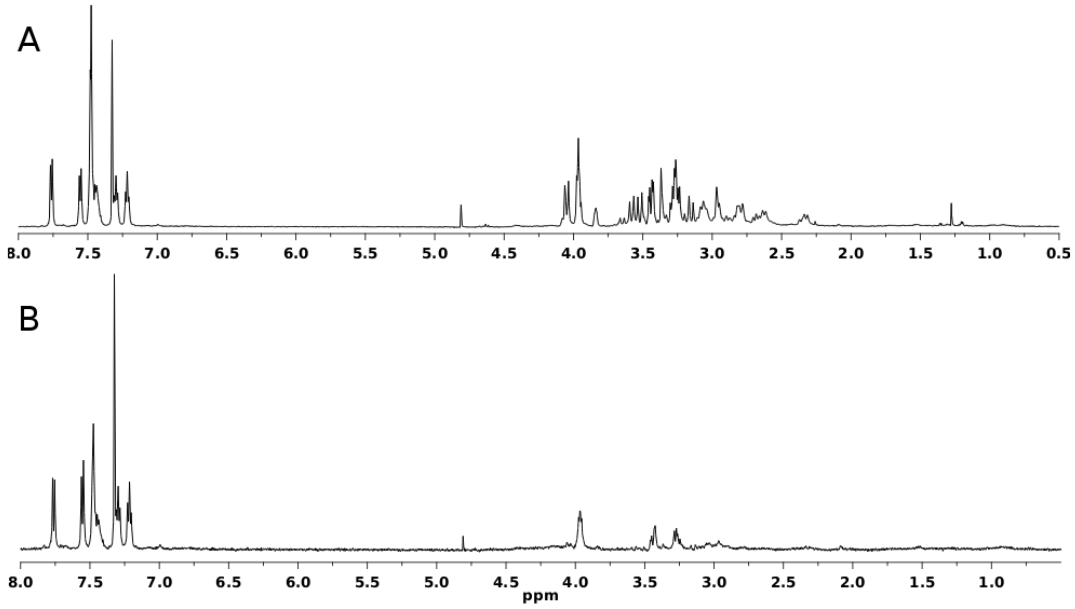


Figure 41. (A) The ^{1}H DPFGSE spectrum of a mixture of La(BOPTA), 1 mM, L-Tryptophan, 1 mM, and HSA, 30 μM (B) The STD NMR spectrum of the same mixture. Trim pulse was used in both experiments to suppress the protein resonances.

Table 3. Data treatment for the competitive STD assays performed for La(BOPTA) and HSA.

La(BOPTA)	ASTD Aromatic ppm	ASTD % reduction
Free 1 mM	5.04	0
Ibuprofen 640 μM	0.87	83
Warfarin 200 μM	0.11	99
L-Tryptophan 1000 μM	3.62	29

The data analysis of the observable STD response of the La(BOPTA) in the presence of Warfarin gives proof that the binding of La(BOPTA) happens mainly through the site I of HSA.

It is known from literature that Ibuprofen binds to site 2 of HSA with a dissociation constant of 0.37 μM. [21] On the other hand, a looser interaction is expected for La(BOPTA), with values of K_D between 5 mM and 0.17 mM (unpublished results). However, it has also been reported that Ibuprofen is able to displace ligands that bind to site I (Warfarin) of HSA, although with a lower affinity constant. If the displacement registered in the Ibuprofen competitive

assays was due to a direct competition for the binding site II of HSA, we would expect a higher intensity of the La(BOPTA) signal in the STD NMR spectra correspondent to the competition with the Warfarin. Therefore, taking to consideration the amount to Ibuprofen present in the competitive experiment and the binding constants from literature, we conclude that the displacement of La(BOPTA) by Ibuprofen is mainly due to a interaction of Ibuprofen with the site I of HSA, either by direct competition of by an allosteric effect.

After this conclusion, was established a one-site binding model for this interaction and calculated the corresponding dissociation constant using STD titration with six different La(BOPTA) concentrations and an array of five saturation times. The data treatment was performed according to the protocol of the Binding Isotherm of STD Initial Growth Rates, described by Angulo, et al. [33] (Fig. 42). The saturation rate and maximum STD values are shown in Table 4.

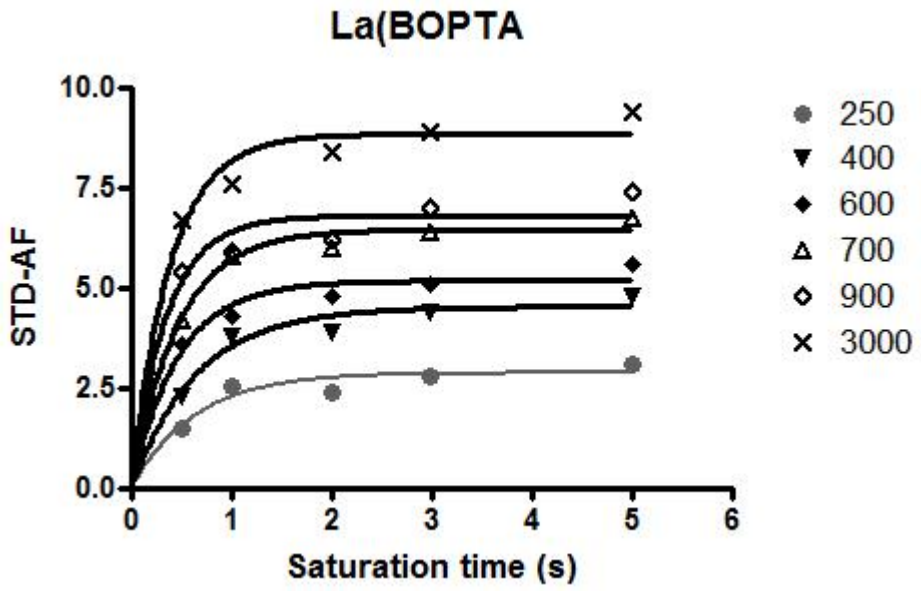


Figure 42. Plot of different concentrations of LA(BOPTA) with, 0,5, 1, 2, 3 and 5s of saturation time of 30uM of HSA with the equation $STD\text{-}AF=STD\text{-}AF_{max}[1-\exp(-k_{sat}*t_{sat})$.

Table 4. Saturation rate, K_{sat}, and the maximum ASTD value (A_{STD}-max) for La(BOPTA).

LA(BOPTA) uM	250	400	500	700	900	3000
A_{STD} -max	2.5	3.68	4.99	4.87	6.98	8.65
K_{sat}	0.77	0.75	0.87	0.71	0.81	0.93
Initial Slope	3.24	4.92	5.74	6.87	7.64	9.3

After plotting the initial slopes into a one-site binding model (Fig. 43) the K_D value of 483 μM was obtained with a R^2 of 0.96 and error margin of 14%. This good result is in agreement with the expected value from other techniques (unpublished results by Aime et al), and it places this new protocol as advantageous to obtain a correct K_D estimation, when compared with a simple ligand titration plot with a given saturation time. The protocol requires five times longer but yields a value about 30% more accurate (for this 483 μM the old protocol gives a 700 μM K_D value).

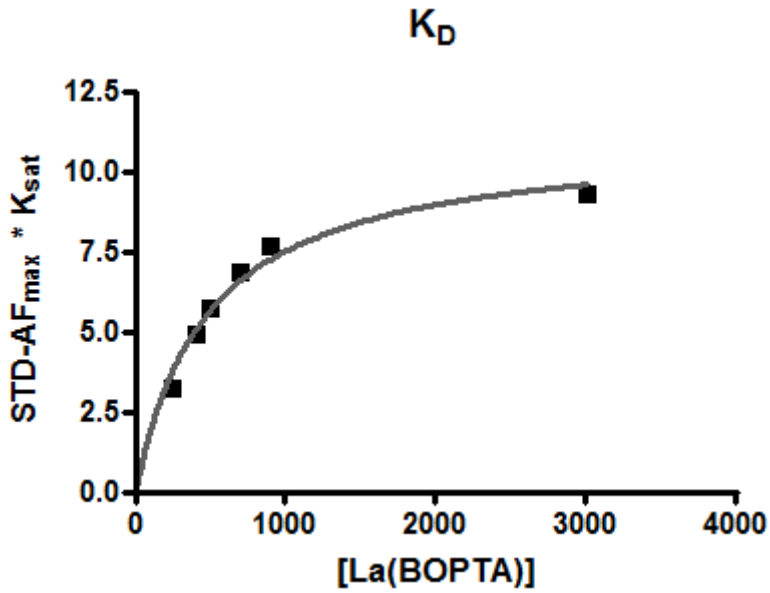


Figure 43. Plot of different concentration of LA(BOPTA) with the one site binding model equation – $Y=B_{max}*X/(K_D+X)$. B_{max} of 10.77 +- 0.8 and a K_D of 483+- 72.83 μM with a R^2 of 0.96.

The quantitative approach obtained by STD NMR was qualitatively represented in a simulated 3D model using Autodock Vina. [25] The structure used for establishing this model was the X-ray structure of HSA complexed with Warfarin, a HSA´s widely known binder for drug site I. With this structure the dynamics of that binding pocket is maintained and the approximate calculation for the CA becomes more closely related to what happen in a real interaction.

To validate the model a simulation was made to see if the program calculated the correct 3D orientation for Warfarin, Figure 44A. The results support the previous conclusions and allow us to establish a coherent model for the binding site I. When the 3D data for La(BOPTA) was completed it was clear

that it was in adjust with the STD NMR data and that the compound was in a direct competition for HSA´s drug site I, mainly through the aromatic moiety of the CA (Fig. 44 C and D).

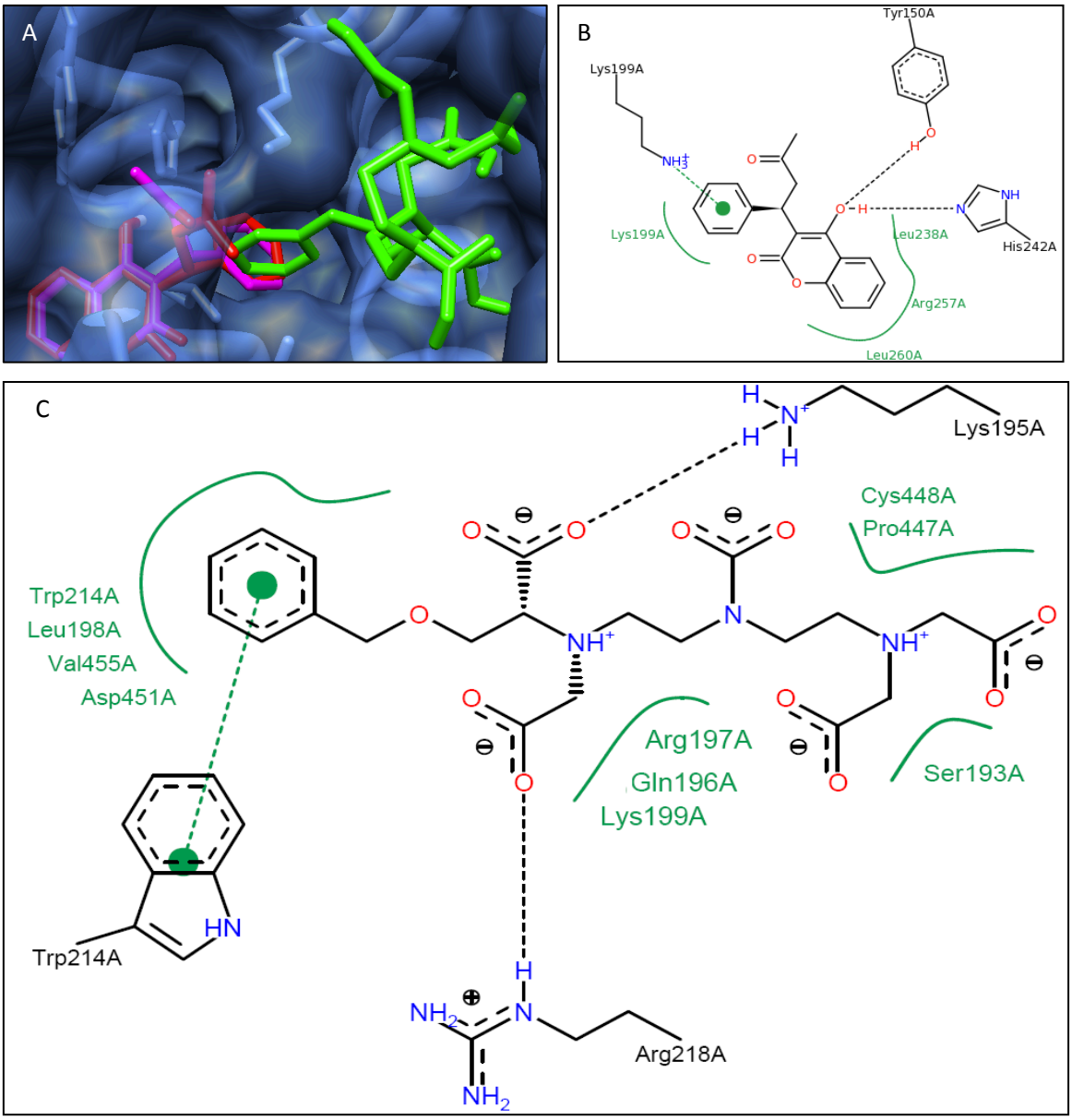


Figure 44. A) Overlay of the representations of Warfarin and BOPTA. In red and purple are, respectively, the warfarin in X-ray structure 2BXD and the simulated one,using Autodock Vina. Depicting almost co-orientation of warfarin and a direct competition by BOPTA. B and C) 2D Interactions between warfarin and BOPTA, with site I of HSA, respectively. 2D calculation were obtained with Poseview Software; black dashed lines – hydrogen bonds; green solid lines – hydrophobic interactions; green dashed lines – Pi-Pi or Pi-cathion interactions;

The main residues involved in this interaction are quite similar for Warfarin and La(BOPTA), indicating direct competition. A huge set of hydrophobic interactions and a pi-pi stacking between the BOPTA aromatic moiety and Trp 214, support the high A_{STD} values for this moiety when compared with the remaining portion of the CA.

In summary, it is clear that the La(BOPTA) interaction with HSA occurs at drug site I and that the full characterization of this interaction can be made by STD NMR. An accurate K_D for the interaction was calculated and docking simulations depicted the HSA´s residues involved, with all the results in perfect agreement.

6.3.2 La (DTPA-Cholate)

Following the same protocol, the La(DTPA-Cholate) compound was studied by STD NMR. Figure 45 represents the ^1H Wsupp and the STD NMR spectra of a mixture of 500 μM La(DTPA-Cholate), and 30 μM HSA. The STD NMR spectrum shows that almost all the resonances from the Cholate moiety are present, of which those that rise the more intense in the spectrum are the sharp resonances at 0.995 and 0.985 (due to their overlapping these resonances were treated as a single one) and 0.751 ppm. This result is concordant with the expectations since most of the anchor group should fit in the binding pocket. The resonances from the reporter group are not visible in the STD NMR spectrum, as in this case, the size of the binding moiety places the chelate group far from the protein surface. Since the sharp resonances from the methyl groups below 1 ppm are the more visible in the whole set of experiments we choose them to evaluate the competition assays that follows. Therefore, the A_{STD} for this interaction was determined to be 0.85 for the peak at 0.99 ppm and 0.91 for the peak at 0.751 ppm, having this reference for the competitive studies.

Although the “OH” protons resonance (see annex 1), around 4 ppm, appear to also interact they have relatively smaller values for A_{STD} to the methyl values and the group epitope mapping was not performed due to the massive overlap in Cholate moiety NMR signals.

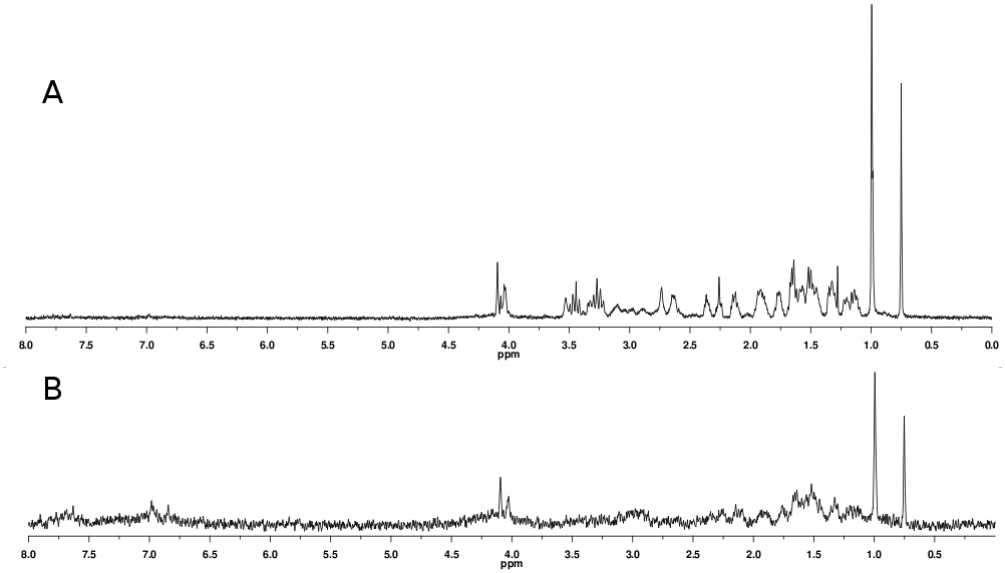


Figure 45. (A) The ^{1}H DPFGSE spectrum of a mixture of 500 μM La(DTPA-Cholate), 0.5 mM, and HSA, 30 μM (B) The STD NMR spectrum of the same mixture. Trim pulse was used in both experiments to suppress the protein resonances.

Three competitive STD experiments between La(DTPA-Cholate) and Ibuprofen were performed. The results, depicted in Figures 46 and 47 correspond to concentrations of Ibuprofen of 500 μM, 700 μM. For these competitions the obtained A_{STD} values were, 0.32, and 0.33, a 65 and 66 % decrease of $_{ASTD}$ value respectively. From the analysis of the STD NMR spectrum (Table 5 and Fig. 47), it is correct to say that with an increase of 200 μM in the Ibuprofen concentration no significant effect was manifested. Some overlapping with the base of the Ibuprofen signal closer to 1ppm might increase the A_{STD} obtained for this interaction. Since the La(DTPA--Cholate) has resonances at the very low ppm region of the spectrum the selective saturation pulse was applied at the aromatic region of the protein, thus interacting with the Ibuprofen resonances. This might be the reason why the Ibuprofen signals appear so intensely in the STD NMR spectrum. At very high concentration of Ibuprofen the signal from the CA is almost null, only residual resonances are visible.

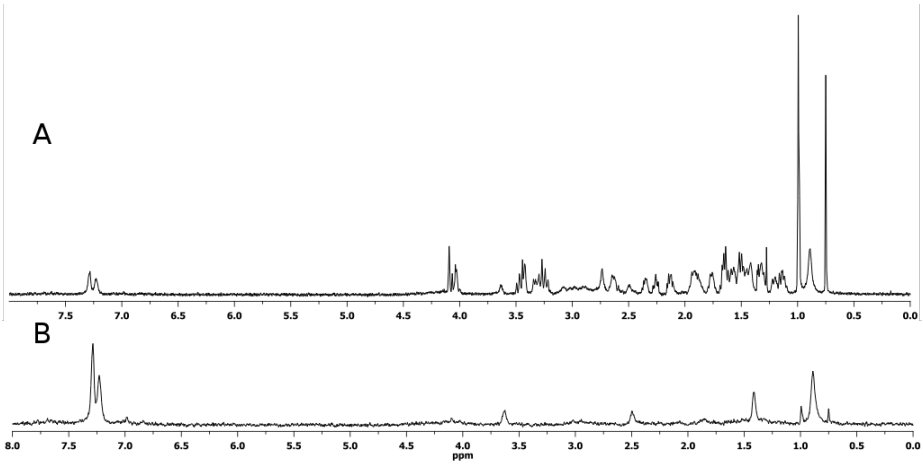


Figure 46. (A) The ^{1}H DPFGSE spectrum of a mixture of La(DTPA-Cholate), 500 μM, Ibuprofen 500 μM and HSA, 30 μM (B) The STD NMR spectrum of the same mixture. Trim pulse was used in both experiments to suppress the protein resonances.

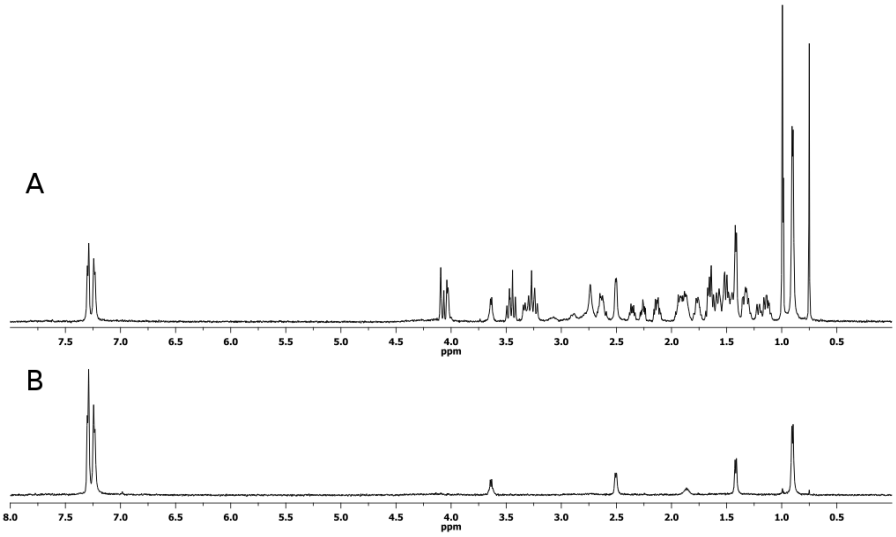


Figure 47. (A) The ^{1}H DPFGSE spectrum of a mixture of La(DTPA-Cholate), 500 μM, Ibuprofen 700 μM and HSA, 30 μM (B) The STD NMR spectrum of the same mixture. Trim pulse was used in both experiments to suppress the protein resonances.

Warfarin experiments showed a greater displacement of the La(DPTA-Cholate) binding, Figures 48 and 49. The A_{STD} for two competitive experiments for 500 μM and 100 μM Warfarin were determined to be 0.18 and 0.19, respectively. The averaged 81% decrease of A_{STD} results show that Warfarin displaces effectively the CA at lower concentrations.

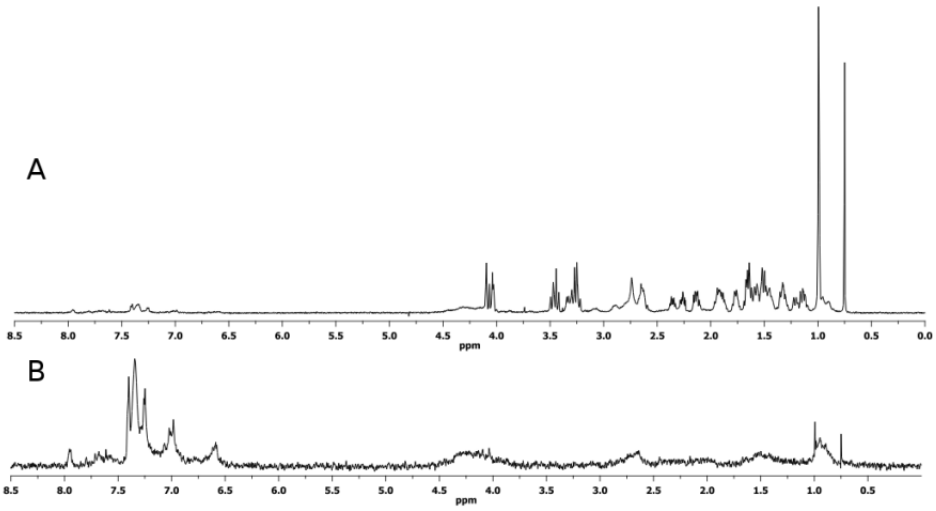


Figure 48. (A) The ^{1}H DPFGSE spectrum of a mixture of La(DTPA-Cholate), 500 µM, Warfarin 100 µM and HSA, 30 µM (B) The STD NMR spectrum of the same mixture. Trim pulse was used in both experiments to suppress the protein resonances.

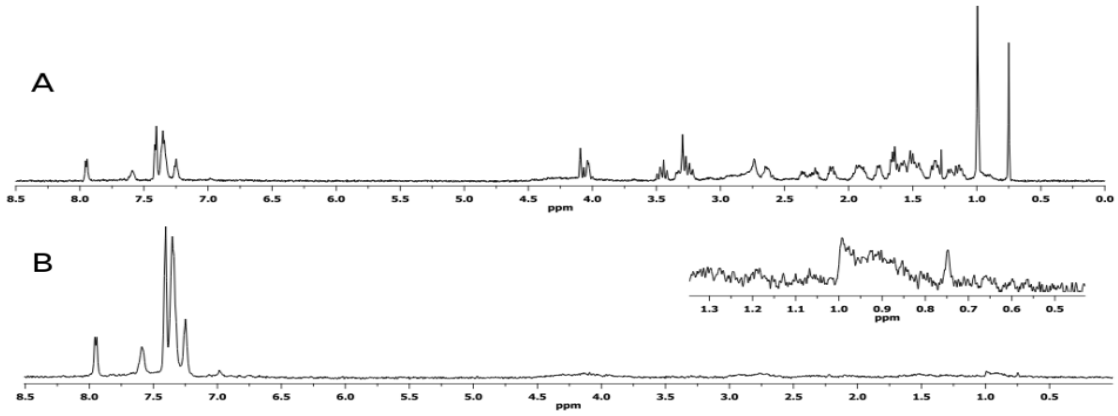


Figure 49. (A) The ^{1}H DPFGSE spectrum of a mixture of La(DTPA-Cholate), 500 µM, Warfarin 500 µM and HSA, 30 µM (B) The STD NMR spectrum of the same mixture. Trim pulse was used in both experiments to suppress the protein resonances.

From the competitive STD assay between the CA in study and L-Tryptophan, Figure 50, was obtained an A_{STD} of 0.92, and as expected only a 8% attenuation of signal intensity, which means that there virtually no displacement in the presence of L-Tryptophan, nor the L-Tryptophan was significantly displaced.

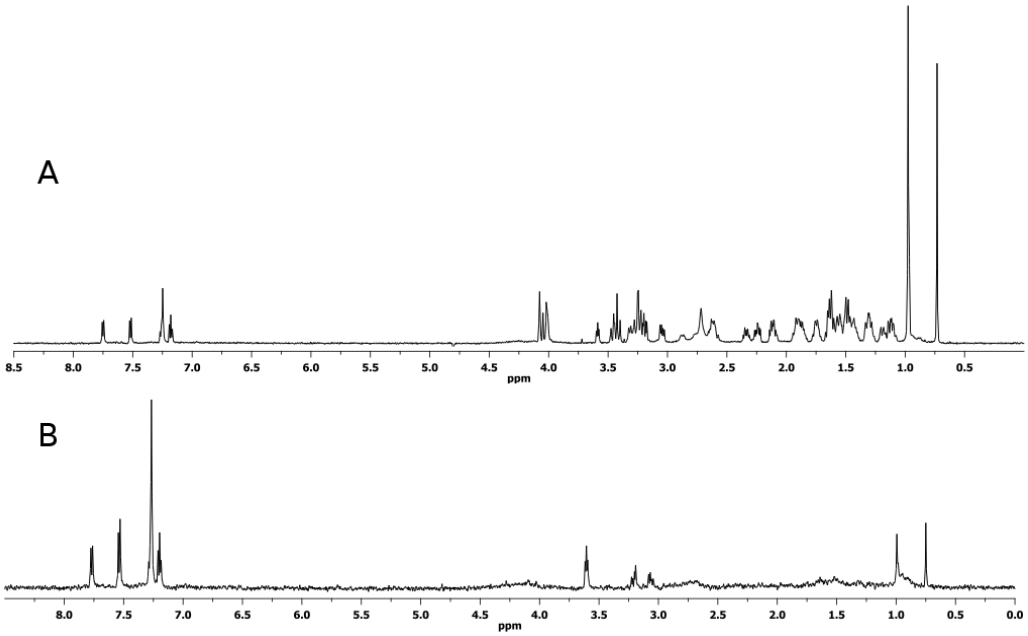


Figure 50. (A) The ^{1}H DPFGSE spectrum of a mixture of La(DTPA-Cholate), 500 μM, L-Tryptophan 500 μM and HSA, 30 μM (B) The STD NMR spectrum of the same mixture. Trim pulse was used in both experiments to suppress the protein resonances.

Table 5. Data treatment for the competitive STD assays performed for La(DTPA-Cholate) for the 0.75 ppm signal.

La(DTPA-Cholate) 500 μM	A_{STD}	A_{STD} reduction (%)
Free 500 μM	0.91	0
Ibuprofen 500 μM	0.32	65
Ibuprofen 700 μM	0.33	66
Warfarin 100 μM	0.19	80
Warfarin 500 μM	0.18	81
L-Typtophan 500 μM	0.93	-2

Since the La(DTPA-Cholate) has resonances at very low ppm region of the spectrum the selective saturation pulse was applied at the aromatic region of the protein, thus interacting with the Ibuprofen resonances. This is the reason why the Ibuprofen resonances appear so intensely in the STD NMR spectrum. At very high concentration of Ibuprofen the signal from the CA is almost null, only residual resonances are visible.

The Warfarin experiments showed a larger displacement of the La(DPTA-Cholate) binding and so we conclude that the binding of this complex occurs at the site I (Warfarin) of HSA. We argue that if the binding of La(DTPA-Cholate) would occur at the site II (Ibuprofen) a significant displacement of the L-Tryptophan should be observed.

To calculate the dissociation constant was used once again the Binding Isotherm of STD Initial Growth Rates protocol. [33] The peak chosen was the one with high A_{STD} value, at 0.75 ppm. Concentrations that best fitted the model are plotted against saturation time in Figure 51 and the saturation rate and maximum STD values are shown in Table 6. These values were used for K_D estimation.

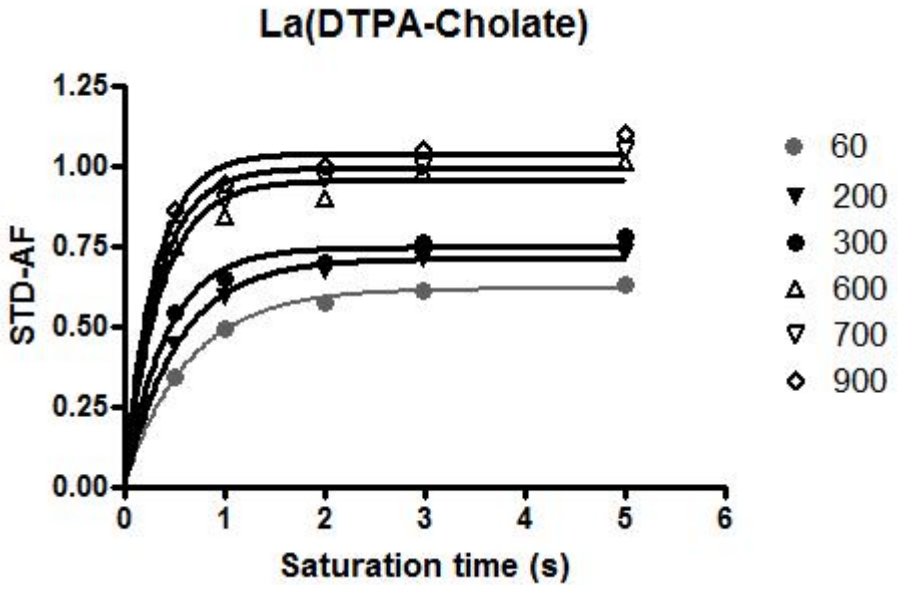


Figure 51. Plot of different concentrations of LA(DTPA-Cholate) with, 0,5, 1, 2, 3 and 5s of saturation time of 30 μM of HSA with the equation $STD\text{-}AF=STD\text{-}AF_{max}[1-exp(-k_{sat}*t_{sat})$.

Table 6. Saturation rate, K_{sat}, and the maximum ASTD value (A_{STD}-max) for La(DTPA-Cholate) for the 0.75 ppm signal.

La(DTPA-Cholate) μM	60	200	300	600	700	900
A_{STD} -max	0.65	0.75	0.79	0.86	0.88	0.92
K_{sat}	0.94	0.78	0.78	0.84	0.84	0.83
Initial slope	0.69	0.96	1.01	1.02	1.05	1.11

The initial slopes were plotted into a one-site binding model (Fig. 52) and a K_D of 36 μM was obtained with a R^2 of 0.98 and error margin of 11%. This result is in agreement with the expected value obtained from other techniques, which was around 20 μM (unpublished results by Aime et al). As we can see, acquiring

some more points on the ascendant portion of the curve, between 60 and 200 μM should improve the result and decrease a little bit more the K_D values, nevertheless the protocol for this parameter is quite successful.

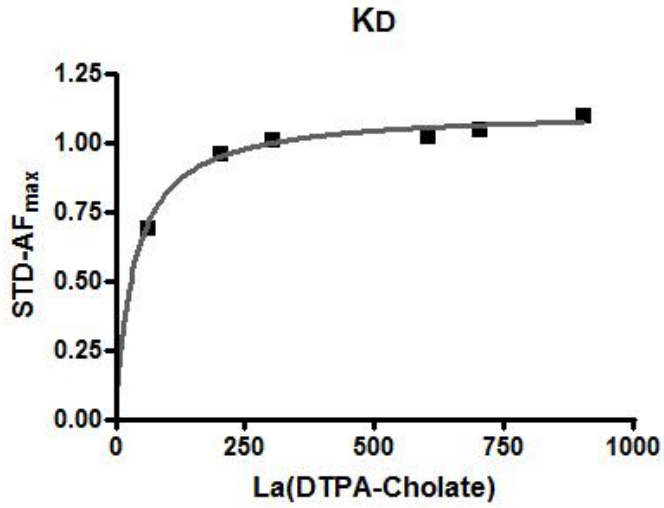


Figure 52. Plot of different concentration of LA(DTPA-Cholate) with the one site binding model equation – $Y=B_{max}*X/(K_D+X)$. B_{max} of 1.117 +- 0.001 and a K_D of 36 +- 3.99 μM with a R^2 of 0.98.

The quantitative approach obtained by STD NMR was three-dimensionally simulated for HSA´s binding site I 3D model. The supported model was able to establish a very clear interaction between the CA and the binding site. STD NMR revealed a variety of relative small and overlaid resonances appearing on the spectra, and now its safe to assume that this effect is due to the embded position for DTPA-cholate in the biding site. The CA into binding site I through its Cholate moiety, explaining also such a low dissociation constant. (Fig. 53)

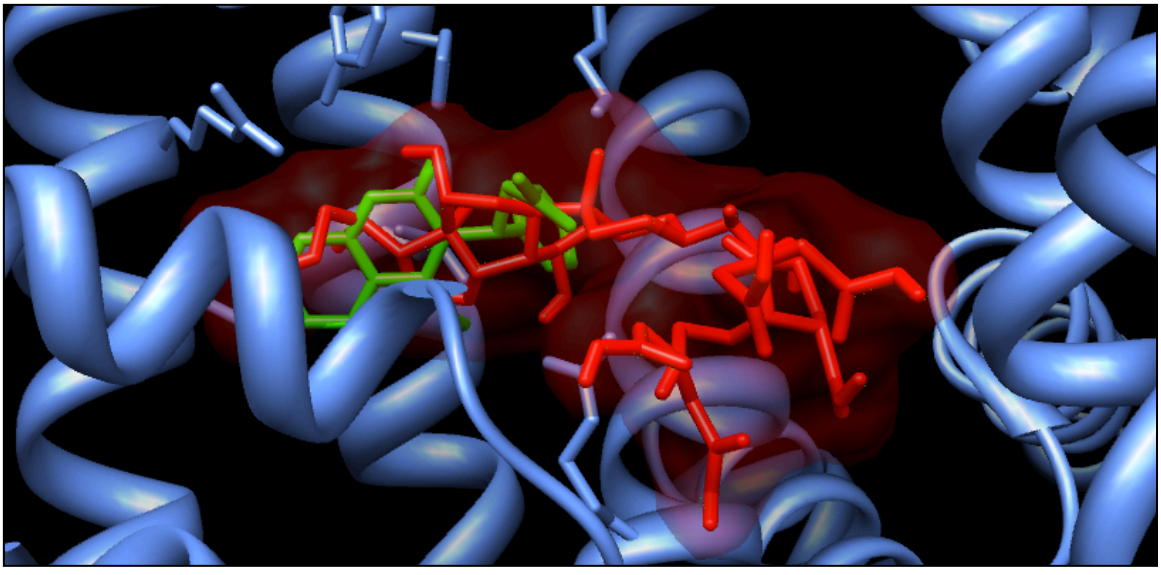


Figure 53. Overlay of the representations of Warfarin (green) and DTPA-Cholate (red) and HSA (cornflower blue). Depicting a direct competition for the binding site.

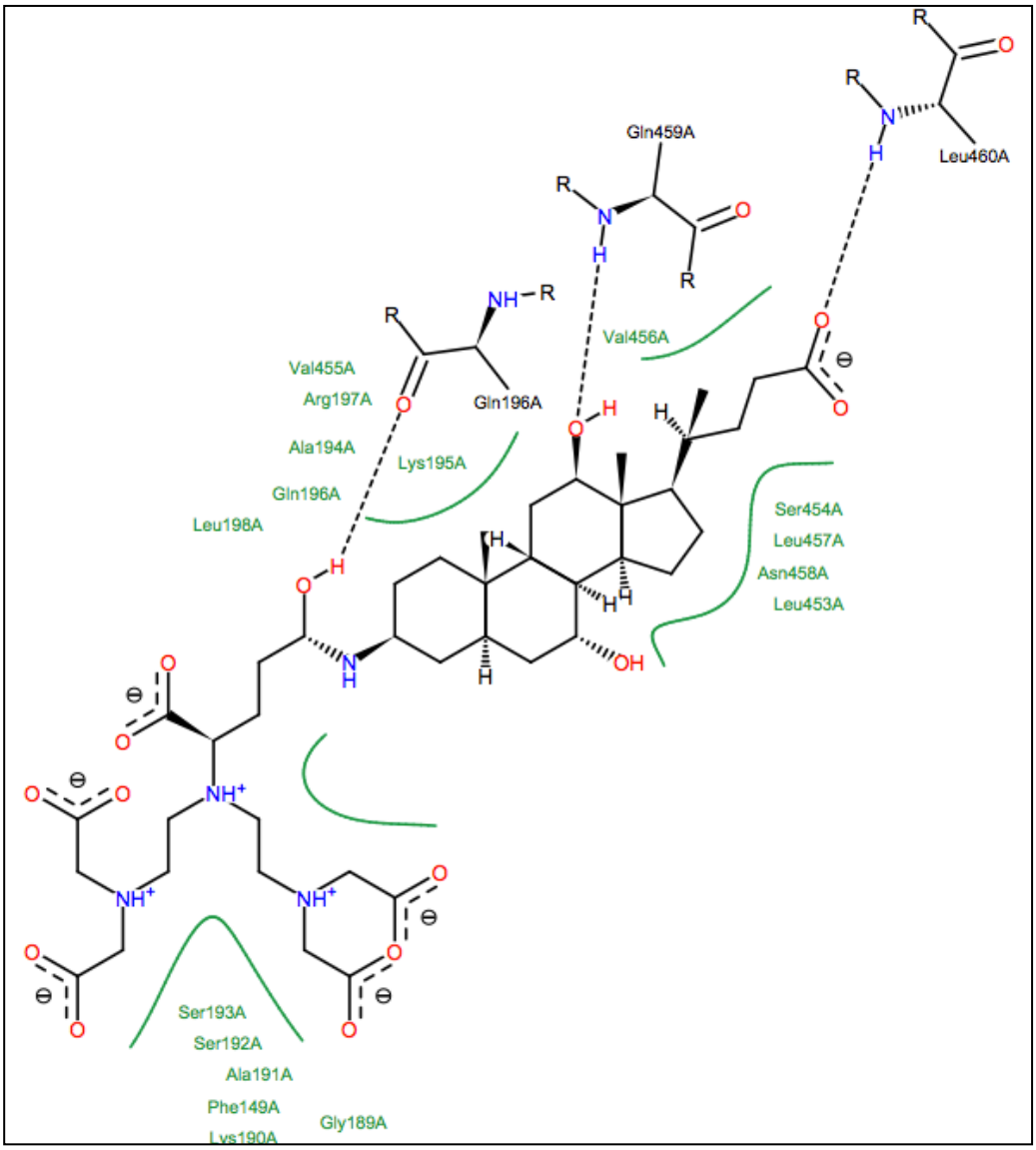


Figure 54. 2D scheme with interactions between DTPA-cholate with site I of HSA. 2D calculation were obtained with Poseview Software (black dashed lines – hydrogen bonds; green solid lines – hydrophobic interactions).

The main residues involved in this large interaction is those already known for this binding site. Analysing Figure 54, a large set of hydrophobic interactions and some hydrogen bounds stabilize the CA-protein interaction, this interaction being also responsible for the good affinity between the studied system.

In summary, it can be concluded that the La(DTPA-Cholate) interaction with HSA occurs via drug site I and the full characterization of this interaction was

made by STD NMR. An accurate K_D value for the interaction was calculated and docking simulations depicted the HSA´s residues involved.

6.3.2 La (NAPHTO-EGTA)

The HSA-CA binding studies follow with La(NAPHTO-EGTA) (Fig. 55) shows the ^{1}H Wsupp spectrum and the STD NMR spectrum of this compound at 500 μM in the presence of 30 μM of HSA. Only the resonances from the naphtalene group are present in the STD NMR spectrum. Moreover, it is clear that the relative intensities between these peaks differ from the ^{1}H to the STD spectrum. The intensity of the singlet at 7.54 ppm is much lower in the STD than in the ^{1}H spectrum, as this singlet corresponds to the two protons of the naftalene that are closest to the chelating group of the molecule (see Annex I for full assignment).

The A_{STD} for this interaction was 1.45 for the peak at the highest ppm value (Table 7). Due to the overlapping of the other proton peaks, and to the fact that it presents 70 % of saturation transfer relative to the high A_{STD} value for the most shifted resonance, this signal was easily picked to follow competition studies and K_D estimation.

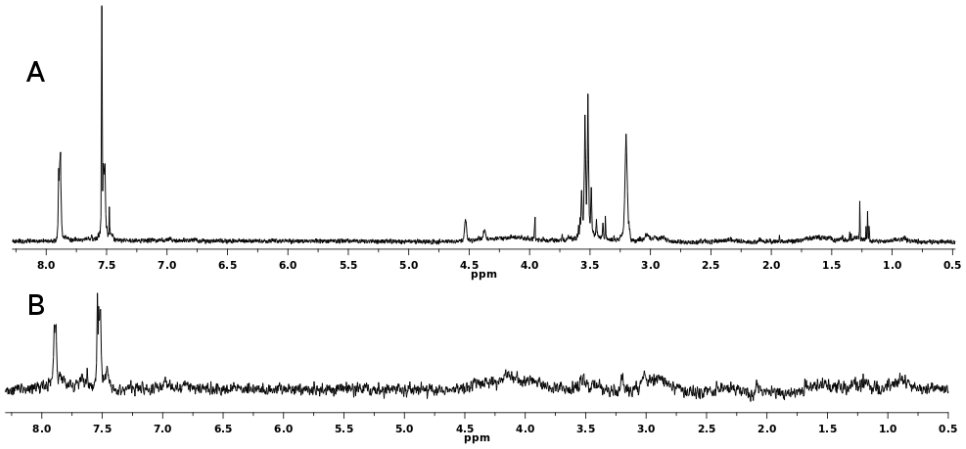


Figure 55. (A) The ^{1}H DPFGSE spectrum of a mixture of 500 μM La(NAPTHO-EGTA), and HSA, 30 μM (B) The STD NMR spectrum of the same mixture. Trim pulse was used in both experiments to suppress the protein resonances.

The competitive assay between an equimolar mixture of La(NAPHTO - EGTA) and 500 μM Ibuprofen with 30 μM HSA shows no reduction of the peak intensities in the STD NMR spectrum for the La(NAPHTO-EGTA) – Figure 56.

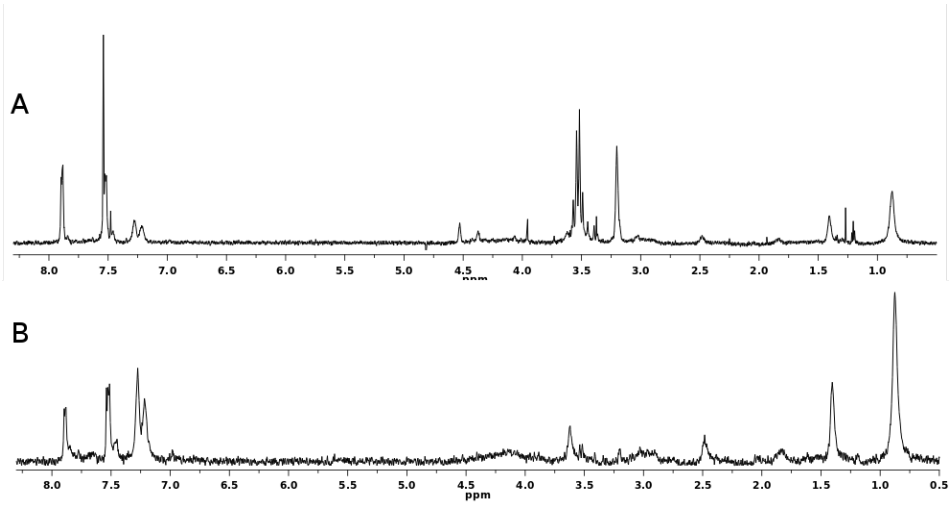


Figure 56. (A) The ^{1}H DPFGSE spectrum of a mixture of 500 μM La(NAPHTO -EGTA), 500 μM Ibuprofen and 30 μM HSA (B) The STD NMR spectrum of the same mixture. Trim pulse was used in both experiments to suppress the protein resonances.

A similar assay of 30 μM HSA with an equimolar mixture of the CA and Warfarin at 500 μM was performed. Again, no reduction in the signals from La(NAPHTO -EGTA) is observed (Fig. 57).

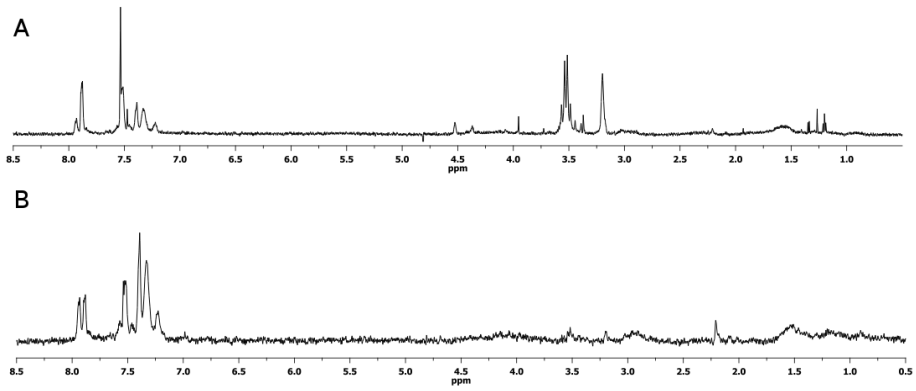


Figure 57.(A) The ^{1}H DPFGSE spectrum of a mixture of 500 μM La(NAPHTO -EGTA), 500 μM Warfarin and 30 μM HSA. (B) The STD NMR spectrum of the same mixture. Trim pulse was used in both experiments to suppress the protein resonances.

Then a competitive test with L-Tryptophan at 500 μM and 30 μM HSA was performed. Here, both L-Tryptophan and La(NAPHTO-EGTA) were present in the STD NMR spectrum with intensities similar to the control test. Moreover, a resonance from the chelate group at 3.8 ppm appears in the STD NMR spectrum (Fig. 58).

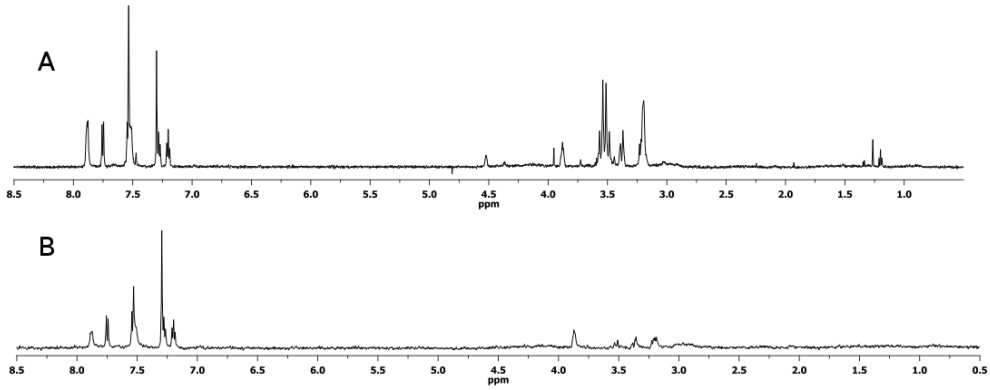


Figure 58. (A) The ^{1}H DPFGSE spectrum of a mixture of 500 μM La(NAPHTO -EGTA), 500 μM L-Tryptophan and 30 μM HSA. (B) The STD NMR spectrum of the same mixture. Trim pulse was used in both experiments to suppress the protein resonances.

As control test a competitive experiment between 500 μM L-Tryptophan and 460 μM Ibuprofen in the presence of HSA 30 μM was performed (Fig. 59). As expected, in the presence of Ibuprofen, the resonances from L-Tryptophan are not present in the STD NMR spectrum.

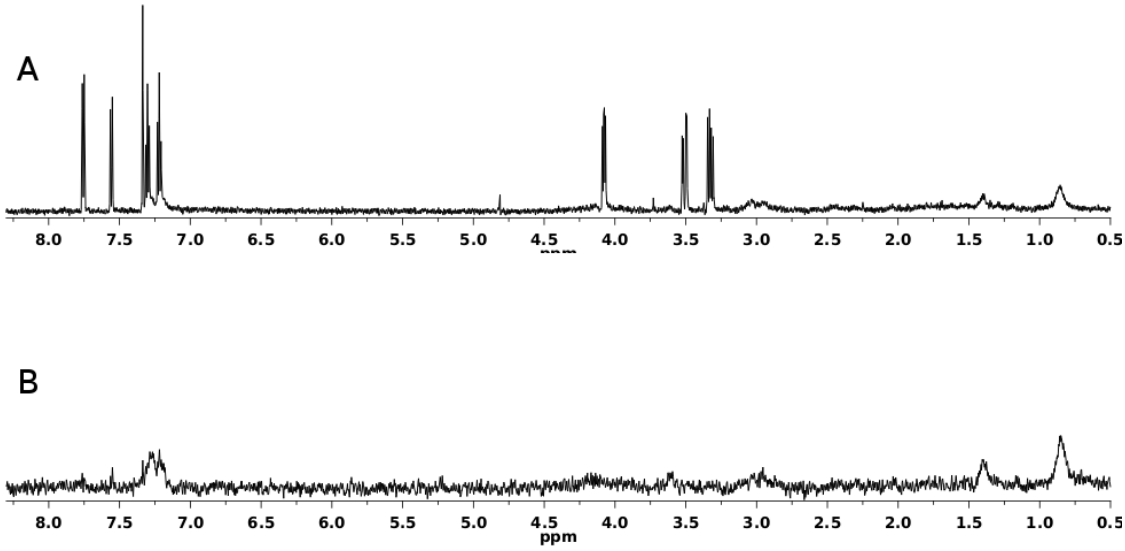


Figure 59. (A) The ^{1}H DPFGSE spectrum of a mixture of 500 μM L-Tryptophan and 460 μM Ibuprofen and 30 μM HSA (B) The STD NMR spectrum of the same mixture. Trim pulse was used in both experiments to suppress the protein resonances.

In a final experiment, containing a mixture of 500 μM of La(NAPHTO EGTA) and L-Tryptophan in the presence of 460 μM of Ibuprofen and 30 μM of HSA (Fig. 60), curiously, the resonances from L-Tryptophan appear in the STD NMR spectrum, although with reduced intensity.

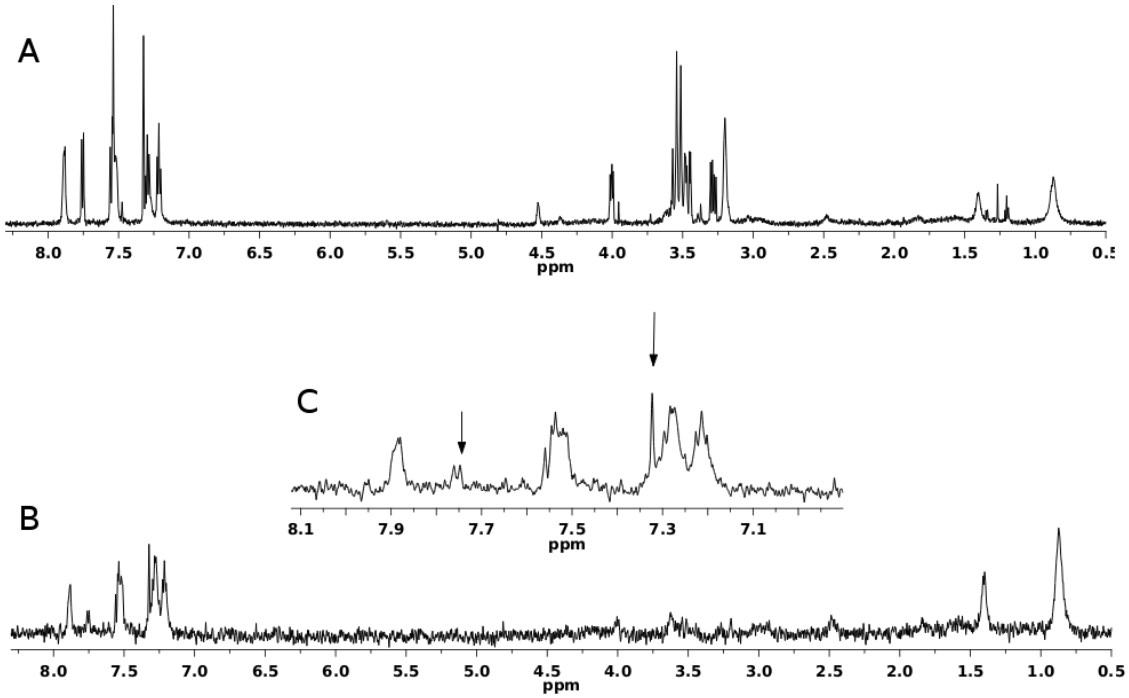


Figure 60. (A) The ^{1}H DPFGSE spectrum of a mixture of 500 μM La(NAPHTO.EGTA), 500 μM L-Tryptophan and 460 μM of Ibuprofen and HSA, 30 μM (B) The STD NMR spectrum of the same mixture. (C) Magnification of STD NMR spectrum from7 to 8.1 ppm, depicting L-Tryptophan resonances. Trim pulse was used in both experiments to suppress the protein resonances

Table 7. Data treatment for the competitive STD assays performed for La(NAPHTO -EGTA)

La(NAFTO EGTA) μM	A_{STD}	A_{STD} increase (%)
Free 500 μM	1.45	
Ibuprofen 500 μM	1.59	110
Warfarin 500 μM	1.9	131
L-Typtophan 500 μM	4.99	344

Considering the results from the competitive assays, where A_{STD} increase after competition (Table 7), we might be in the presence of non-common binding characteristics. Since La(Nafto-EGTA) was not displaced neither by Ibuprofen nor Warfarin and since it favored the binding of L-Tryptophan we might say that La(Nafto-EGTA) may bind to a different site of Human Serum Albumin, a site that is neither site I nor II.

To calculate the dissociation constant with the Isotherm of STD Initial Growth Rates protocol the peak chosen was the one with high A_{STD} value. Although competitive STD show that we are in the presence of a non-competitive binding CA, values fit the one-site binding model. According to Angulo et al

protocol, [33] different concentrations were plotted against saturation time in Figure 61 and the saturation rate and maximum STD values for the protocol are shown in Table 8.

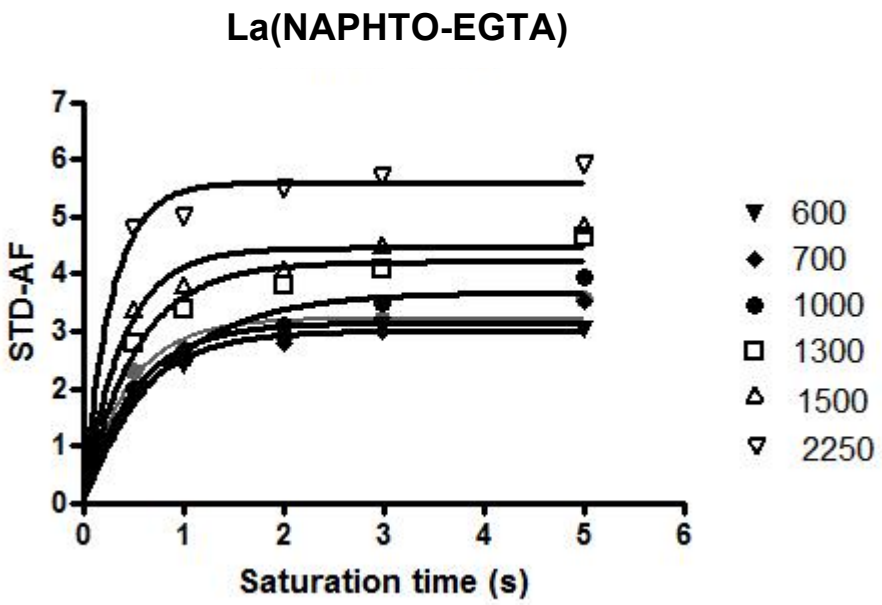


Figure 61. Plot of different concentrations of LA(NAPHTO-EGTA) with, 0,5, 1, 2, 3 and 5s of saturation time of 30 μM of HSA with the equation STD-AF=STD-AF_{max}[1-exp(-k_{sat}*t_{sat}) .

Table 8. Saturation rate, K_{sat}, and the maximum ASTD value (A_{STD}-max) for La(NAPHTO-EGTA) for the 7.9 ppm signal.

La(NAPHTO-EGTA) μM	500	600	700	1000	1300	1500	2250
A_{STD}-max	3.55	3.05	3.51	3.95	4.64	4.87	5.89
K_{sat}	1.18	0.93	0.87	0.81	0.85	0.88	0.9
Initial slope	3.02	3.31	4.04	4.86	5.47	5.56	6.55

The initial slopes were plotted into a one-site binding model (Fig. 62) and a K_D of 1047 μM was obtained with a R^2 of 0.98 and error margin of 12%. This high value is in agreement with a non-specific binding, that represents lower affinity interaction with HSA. As we have plenty of points on the ascendant portion, we are confident of this result. Although a non-specific binding could occur to another unkown site, the results seem to point this way.

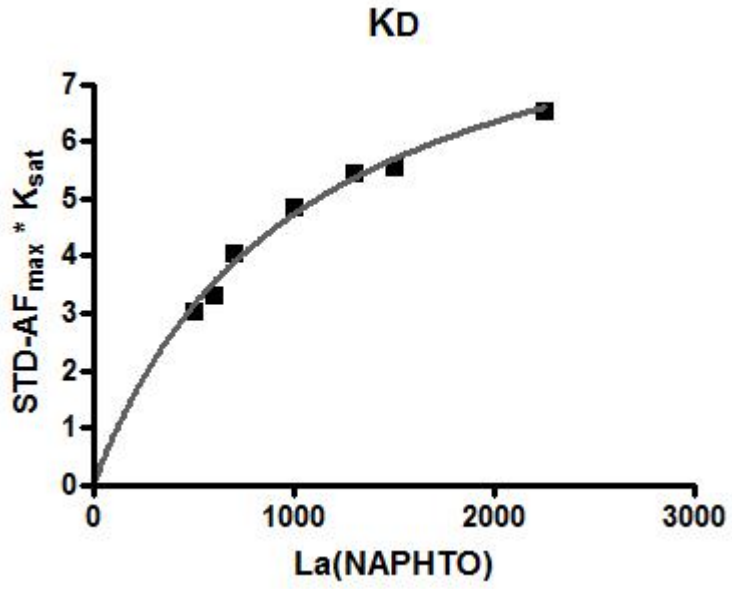
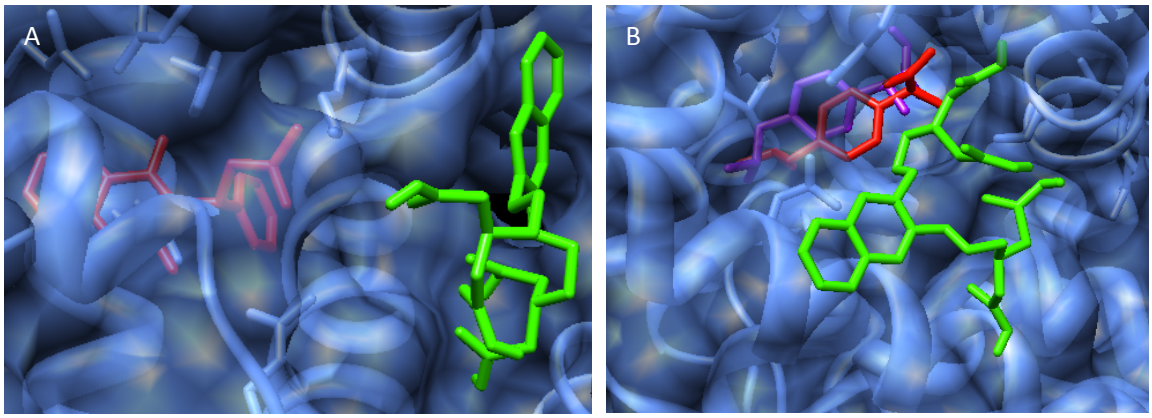


Figure 62. Plot of different concentration of LA(DTPA-Cholate) with the one site binding model equation – $Y=B_{max}*X/(K_D+X)$. B_{max} of 9.65 +- 0.51 and a K_D of 1047 +- 118.1 µM with a R^2 of 0.98.

To corroborate data obtained by STD NMR, both HSA binding sites were three-dimensionally simulated. The model for HSA binding site II was made with HSA´s X-ray structure complexed with Ibuprofen. The simulation, presented in Figure 63 C and D, was not as accurate as the one calculated for drug site I, because the overlay of deposited and calculated structures are not equivalent. Nevertheless, its considered a possible model for this site, once Ibuprofen is relatively close to the X-ray position.



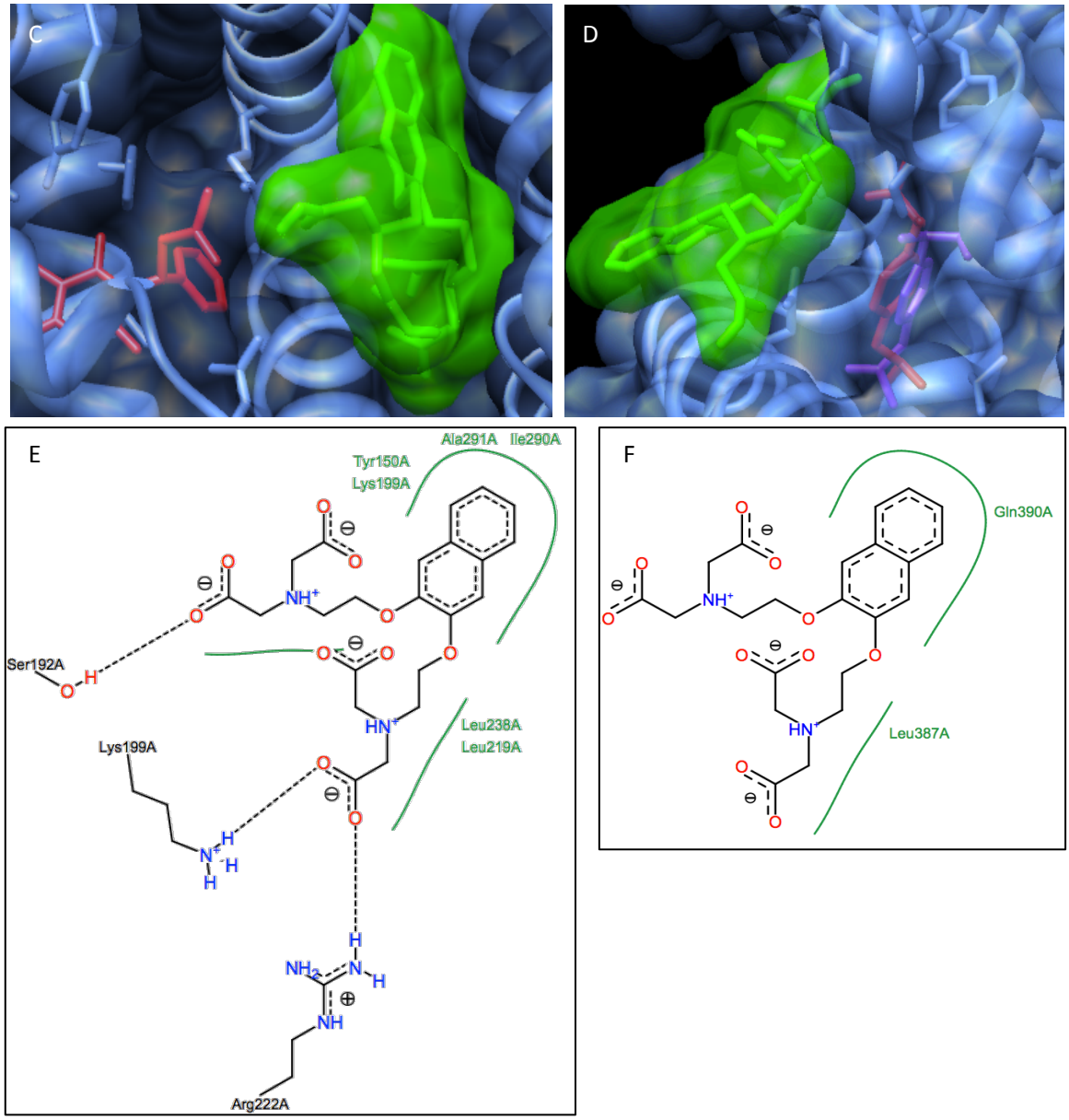


Figure 63. A and B) Overlay of the representations of Warfarin (red) and NAPHTO-EGTA (green) and the molecular surface of the CA (shaded green) depicting no direct competition for drugsite I of HSA. C and D) Overlay of the representations of Ibuprofen from the X-ray structure (red), Ibuprofen simulated (purple) and NAPHTO-EGTA (green) and the molecular surface of the CA (shaded green) depicting no direct competition for drugsite II of HSA. E and F) 2D scheme with interactions between DTPA-cholate with site I and drug site II of HSA, respectivelly. 2D calculation were obtained with Poseview Software (black dashed lines – hydrogen bonds; green solid lines – hydrophobic interactions).

The main forces involved in this interaction are mainly hydrophobic interactions and some hydrogen bonds, but its seen that the NAPHTO moiety of the CA can be docked to the surface of the protein, but with no effective overlap with the drug portion was observed. This finding supports STD NMR results and

one can conclude that no direct competition for the main drug sites of HSA is observed for this CA.

Considering the affinity constant, and according to the results from the competitive assays, we might be in presence of non-common binding characteristics. Since La(NAPHTO -EGTA) was not displaced neither by Ibuprofen nor Warfarin and since it favoured the binding of L-Tryptophan, we might say that La(NAPHTO -EGTA) binds to a different site of Human Serum Albumin, that is neither site I nor II. The La(NAPHTO-EGTA) interaction with HSA occurs, as the NAPHTO moiety is prone to establish hydrophobic interactions at HSA´s molecular surface. The combined results point to a new, or at least different, but specific binding site.

7. STUDY OF THE INTERACTION OF INSULIN MIMETIC VO(DMPP) WITH HSA BY ^{1}H STD NMR TECHNIQUES

7. STUDY OF THE INTERACTION OF INSULIN MIMETIC VO(DMPP) WITH HSA BY ^{1}H STD-NMR TECHNIQUES

7.1 Abstract

Insulin mimetic VO(DMPP) complexes preferred binding to 'drug-site I' of HSA was detected by STD ^{1}H NMR by the displacement effect of the competitive binder to that site, warfarin. The H5 proton of the free ligand was identified as the most interacting proton, and with higher saturation transfer than any other species found in solution. A 3D model simulation for this interaction was established using Autodock Vina.

7.2 Introduction

In the last years Vanadium Compounds have attracted much interest due to their demonstrated pharmacological properties [96]. In particular, their potential use as oral insulin mimetics has been demonstrated by *in vivo* [97,98] and *ex vivo* studies [99], as well as in clinical trials [100].

Thus, intensive research has been carried out to develop Vanadium Compounds to be used as orally administered drugs in the treatment of Diabetes Mellitus (DM), at an effective non toxic dose. Among many synthesized compounds, at the moment, only a few seem to be most promising, and have demonstrated antidiabetic properties [101-102] through indicators of insulin mimetism assessed through *in vitro* and *ex vivo* studies such as glucose uptake rates [103,104], inhibition of free fatty acid (FFA) release [105] and specific protein phosphorylation [100,102]. *In vivo* experiments have confirmed their therapeutic activity and low toxicity [106,107], bioavailability and pharmacokinetics and established the minimum effective dose [108].

In our group the, vanadium complex containing a pyridinone ligand, the bis(1,2-dimethyl-3-hydroxy-4-pyridinonate)oxovanadium (IV), $V^{IV}O(DMPP)_2$ (FIg. 3), has been extensively studied. Its structure in the solid state was already determined [109], the different species which are formed in aqueous solution under aerobic conditions were identified using different techniques, and the respective formation constants were determined [110]. *In vitro* studies with two cell lines, the mice fibroblast SV 3T3 and the human skin fibroblast F26 [111]

were conducted. A human fibroblast cell line, 3T3-L1 [107], as well as *ex vivo* studies with human erythrocytes [112] were also carried out to test its cytotoxicity, and glucose uptake enhancement capacity. Recently its anti-diabetic action was recognized through an *ex vivo* study with isolated primary rat adipocytes, which demonstrated that this compound improves glucose internalization, inhibition of free fatty acids (FFA) release, and has the capacity to phosphorylate specific proteins of the insulin signalling cascade [113]. *In vivo* studies with obese Zucker rats, using Magnetic Resonance Techniques (MRI/MRS), have confirmed the positive effects of $V^{IV}O(DMPP)_2$ on glucose and lipid metabolism, reinforcing its promising anti-diabetic capacity (data to be published in a near future).

However, in spite of all these studies, the mechanism of action of Vanadium Compounds is not yet clarified. A main concern in this field is how the Vanadium Compounds are transported in the blood stream and taken up by the cells. It's important to know the interactions between the administered Vanadium Compounds and endogenous small and/or macromolecules present in serum and to find out if decomposition of the external Vanadium Compounds occurs or a ternary complex will be formed, as well as to investigate what is the role of serum proteins in the transport of Vanadium Compounds [114-118]. This information is crucial to search for a better drug design to improve the therapeutic efficacy of these compounds.

This work reports a study of the interaction in aqueous solution of vandate/HDMPP complexes with human serum albumin (HSA) using the ^{1}H NMR spectroscopic technique Saturation Transfer Difference (STD) [17-19], which provides powerful data about kinetics and binding sites of ligands to a protein, validation of binding epitopes, and estimation of affinity constants.

The vanadate/HDMPP (M/L=1/2) solution contains the 1:1 and 1:2 V(V) species, the same which result from the dissolution of the solid $V^{(IV)}O(DMPP)_2$ in water, under aerobic conditions, as already demonstrated [110]. The use of the vanadate/HDMPP system instead of a aerobic solution of $V^{(IV)}O(DMPP)_2$ avoids the presence of a small amount of paramagnetic species [21], which would not allow the correct use of the ^{1}H NMR spectroscopy and STD technique.

Competitive assays with two known inhibitors of site I and site II of HSA, warfarin and ibuprofen, respectively, were performed primarily to identify the binding sites of the compound.

7.3 Results and Discussion

The spectrum of the free ligand shows the signals corresponding to the two methyl groups and to the aromatic protons H5 and H6. In the spectrum obtained from the vanadate /DMPP solution (1:2) multiple signals for each type of protons are observed, indicating the presence of 3 species in solution, in slow exchange in the NMR time scale: the free ligand, the 1:1 VO_2(DMPP)(H_2O)(OH) and the 1:2 $VO_2(DMPP)_2$ V(V) species, assigned in the Figure 64.

When HSA was added to the vanadate /DMPP (1:2) system, the resonances of the ligand in the three species broaden, as expected due to protein binding and the increase in solution viscosity. From STD spectra, group epitope mapping for the free ligand (Table 9) revealed to be the main interacting proton the H5 to all species, the N-CH_3 and CH_3 resonances had lower values. For the vanadate /DMPP (1:2) system only H6 and H5 were mapped due to overlap of the other two resonances (Table 10).

Table 9. Group epitope mapping (GEM), as percentage, for the proton signals of HDMPP normalized relative to H5.

GEM %	H6	H5	N-CH_3	CH_3
Free HDMPP	88	100	59	73

Table 10. Values of group epitope mapping (GEM), as percentage, for the proton signals of HDMPP free, VO2(DMPP) (H_2O)(OH) and $VO2(DMPP)_2$ species, normalized relative to free HDMPP proton H5.

HDMPP - species	A_{STD}	GEM (%)
H6 - 1:2 and 1:1	23.98	75
H6 - Free	20.85	65
H5 - 1:2	26.58	83
H5 - 1:1	26.26	82
H5 - Free	31.97	100

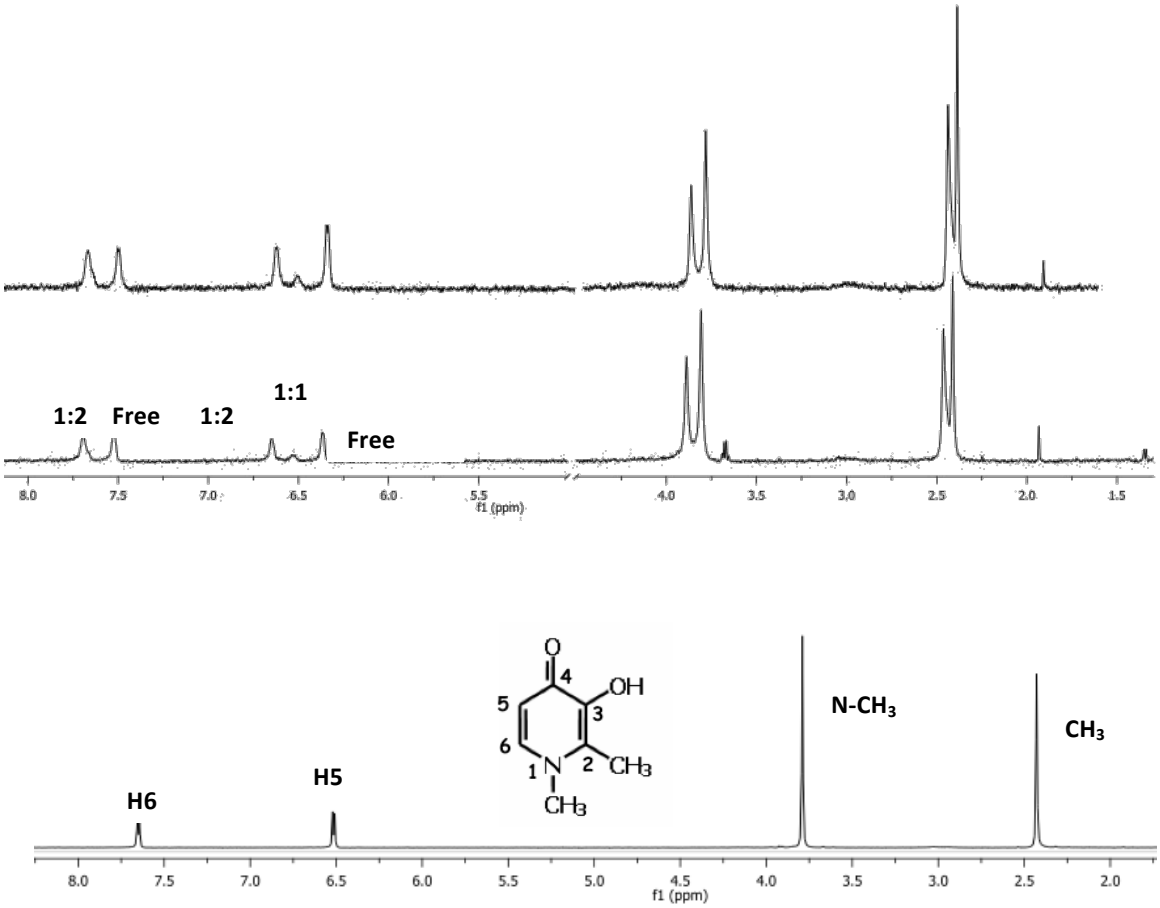


Figure 64. A and B) Overlay of the representations of Warfarin (red) and NAPHTO-EGTA (green) and the molecular surface of the CA (shaded green) depicting no direct competition for drugsite I of HSA.

Analysis of the STD spectrum and determination of GEM values indicate that there is a preference for the binding of the free HDMPP specie, although all the species experiment interaction with HSA.

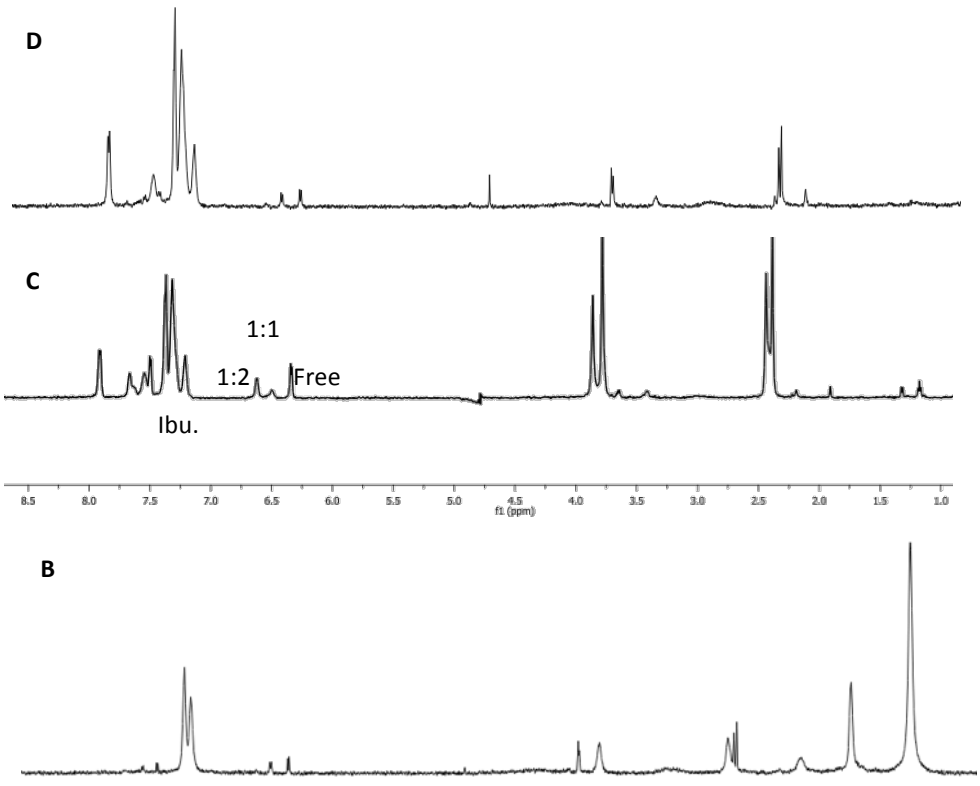
It is expected that the interaction of the ligand HDMPP with HSA occur via non-specific binding at the surface of the protein. However, to check if site I and II of HSA are possible binding sites for this insulin mimetic complex, competitive ^{1}H NMR STD experiments were performed in the presence of warfarin and ibuprofen, known binders to site I and II respectively. The displacement of the HDMPP ligand by these known inhibitors of HSA after being added to the solution of vanadate/DMPP and HSA, reflects the direct interaction of the ligand with these sites.

Analysing ^{1}H STD NMR spectra (Fig. 65) and taking into account the A_{STD} values of Table 11 it is possible to conclude that the most favourable interaction is

with the binding site I of HSA. The H6 and H5 aromatic protons were used in these calculations because they presented the highest values of saturation transfer received from the protein. Although a displacement of the ligand from both site I and site II of HSA was observed, according to the A_{STD} values of the solution with HSA and Hdmpp in the absence and presence of warfarin and ibuprofen, the binding to site I seems to be favoured relative to site II.

Table 11. Values of A_{STD} for H6 and H5 protons for free 0.5 mM VO:HDMPP, 0.4 mM Ibuprofen and 0.1 mM Warfarin. The % displacement reflects the decrease of the A_{STD} value upon the addition of known binders.

	A_{STD}		**% of displacement by competition**	
HDMPP	H6	H5	H6	H5
Free	6.22	16.42		
Ibuprofen	2.41	2.18	61	86
Warfarin	0.016	0.038	99	98



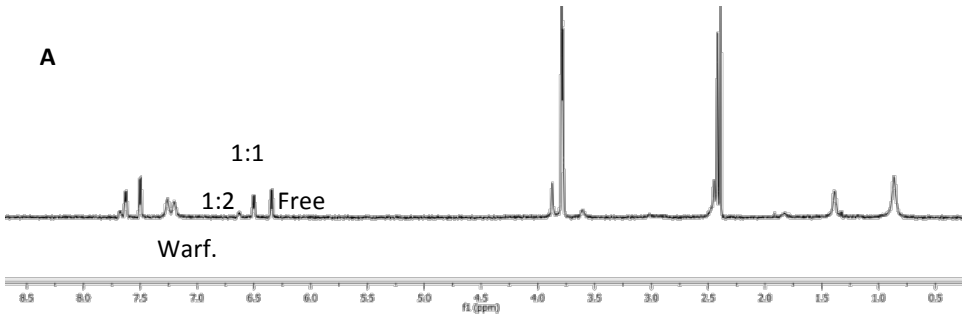
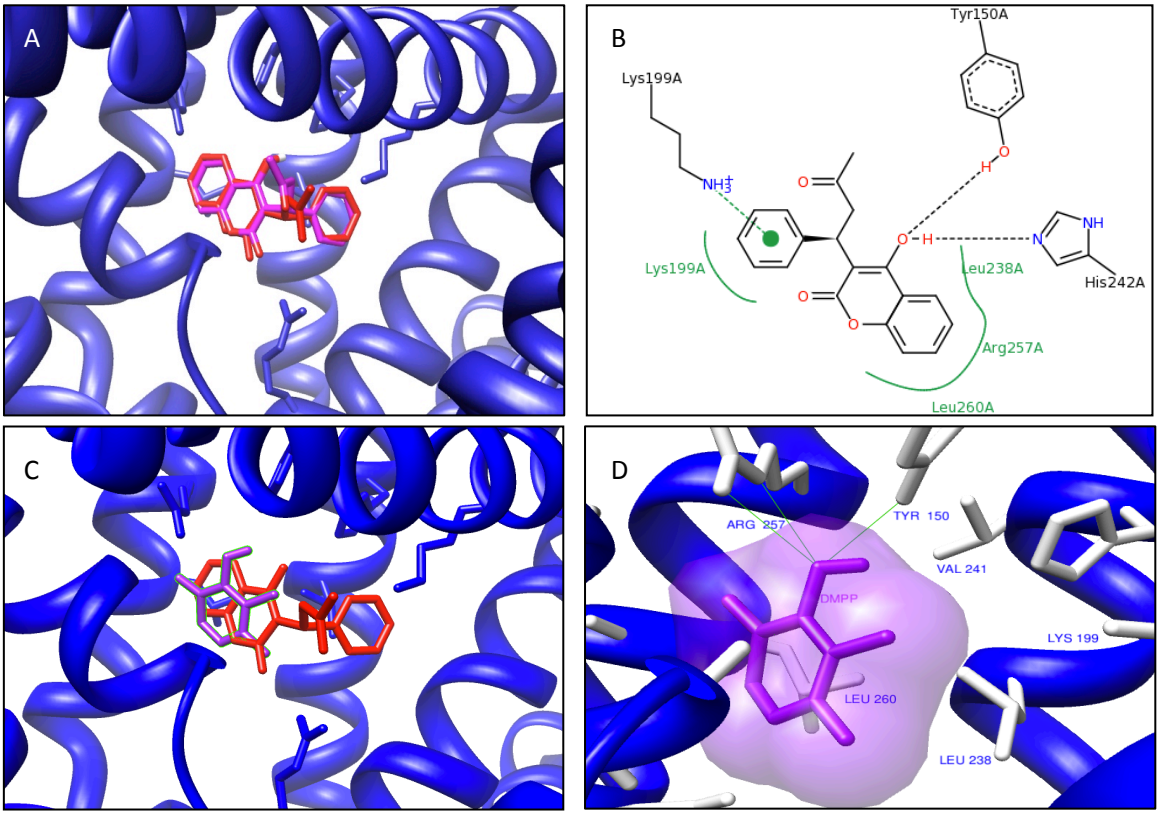


Figure 65. A)^{1}H NMR spectrum of a solution containing vanadate / DMPP (1:2.5), 0.5 mM in vanadate, 0.1 mM Warfarin and 0.030 mM HSA; B) STD spectrum of the same system; C) ^{1}H NMR spectrum of a solution containing vanadate / DMPP (1:2.5), 0.5 mM in vanadate, 0.4 mM Ibuprofen and 0.030 mM HSA; D) STD spectrum of the same system. A selective saturation pulse (5 s) was applied at the 0 ppm region of the protein. The 30 ms spin-lock pulse was calibrated to avoid unwanted protein resonances



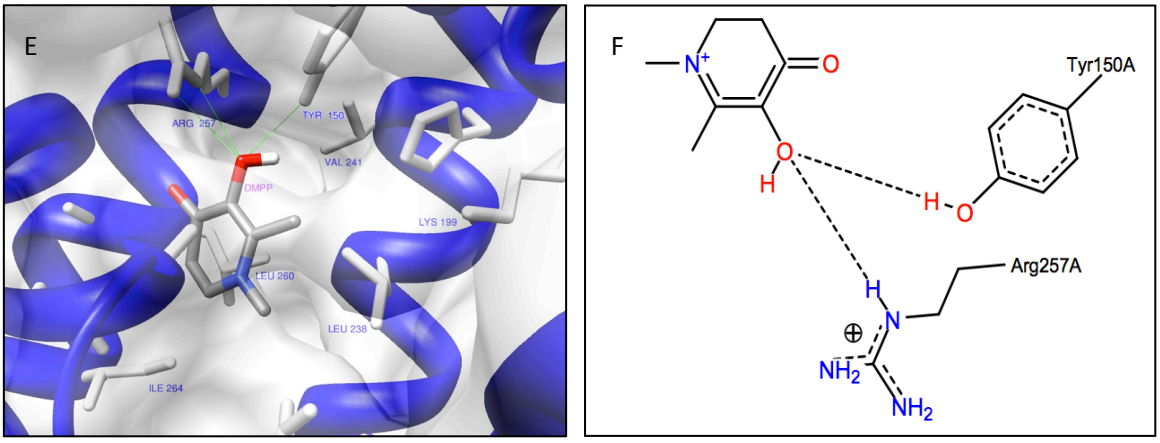


Figure 66. A) In red and purple are, respectively, the X-ray structure of warfarin and the simulated one; B) Interactions between warfarin and site I of HSA; black dashed lines – hydrogen bonds; green solid lines – hydrophobic interactions; green dashed lines – Pi-cation interactions; C) overlay of the structures of warfarin (red and Hdmpp (purple) in HSA drug site I, indicating a specific competiton for this site; D) Representations of the main residues of the protein interacting with the ligand Hdmpp (the residues less than 5 Angstroms from HDMPP (white) and HDMPP molecular surface (shaded purple) and hydrogen bonds (light green) ; E) Molecular surface (shaded white) and HDMPP (coloured by element – nitrogen as blue, oxygen as red, hydrogen as white and carbons as shaded grey)) on drug binding site I; F) Molecular interactions between HDMPP and site I of HSA, calculated using Poseview (black dashed lines –hydrogen bonds).

To get an atomic resolution of this interaction virtual screening was performed using Autodock Vina 1.1.1. The software was used to assess the three dimensional disposition of the binding scenario (Fig. 66).

Molecular docking software allowed to assess the binding pose and probable interactions of the HDMPP ligand within the binding pocket of HSA drug site I, the Warfarin binding site. As we can see the Autodock Vina software is fully capable of producing good results once the calculated pose for warfarin is almost 100% overlapped with the pose observed in the X-Ray model of the deposited structure. The interactions are depicted in figure 66B. The software does not accept vanadium so the docking was done only with the HDMPP alone. This might be a good model since the vanadate compound will for sure establish more interactions in the binding site, but we can see that it is directly interacting with the Warfarin Site (Fig. 66) with the same residues and perfectly lodged in HSA binding site I.

In summary, the interaction is made favourably with the drug binding site I and stabilized in a similar manner as its known binding competitor Warfarin. This serum protein helps to an effective transport and delivery of the vanadium compound. Even though the STD calculation, the free species being the major

interacting protein, the other species also bind, and the presented model is valid to 1:1 and 1:2 species as well. (data not shown - simulations done with Phosphorous instead of vanadium, which favours and stabilizes HSA interaction with the both solution species, 1:1 and 1:2)

8. CONCLUSION AND FUTURE WORK

8. CONCLUSION AND FUTURE WORK

After some considerations made in the Results and Discussion sub-sections, this final conclusion aims to summary all the data obtained, and also to check if all objectives of this work were achieved.

Proposals for future work are also stated here as a means to improve the outputs of this protocol.

8.1 CONCLUSION

The protocol developed during this work revealed to be very fruitful since the enantioselective study is already accepted for publication by *Organic and Biomolecular Cemistry*.

The protocol can, indeed, accurately calculate dissociation constants, describe how the complexes (CA, chiroptical probe, insulin mimetic) interact with the protein HSA, and with competitive STD we can tell what is the region of the protein involved in this interaction and finally the docking simulations help explaning the data obtained by NMR, so the protocol is in full working order.

STD-NMR can indeed distinguish between the enantiomers SSS-(Δ)-$[Y.L^1]^{3+}$ and RRR-(Λ)-$[Y.L^2]^{3+}$ in their binding to HSA and the site is now determined to be HSA´s site II, confirming it as stereo-selective. The 3D model helped to achieve this first objective.

The study of binding of CA´s to HSA revealed a stronger interaction of the La(DTPA-Cholate) CA analogue and was confirmed by docking simulations. This competitive STD technique was able to distinguish were the binding was prone to occur in terms of protein region. While La(NAPHTO) appears to have non-specific binding, the other two CA's analogues mainly interact with HSA´s binding site I.

The Insulin mimetic VO(DMPP) complex preferably binds to drug-site I of HSA and it was obtained that the main interacting specie was the free HDMPP but also the other two species are closely interacting with the protein.

In conclusion, all objectives were achieved on this work, revealing the effective usage of STD NMR followed by docking simulations to study the interaction of therapeutic and diagnostic agents with their respective receptors.

8.2 FUTURE WORK

The K_D values of the contrast agents used can be calculated with competition protocols, either by the determination of the IC_{50} of the binder used with know K_D (Warfarin, Ibuprofen and N-dansyl sarcosine) on the interaction with the Human Serum Albumin used, so with a few effort the constants can be obtained within a week of work. Either by direct displacement or by saturating a binding site, the protocol can establish if the results agree with the reference value or not. This new protocol can also be used in our group in the future.

After determining the binding constants and also knowing what ligand proton interacts with the protein by the GEM results it is, an even better quantitative information about the ligand docking on the protein, could be obtained. So some calculations using HADDOCK, which may help improve the scoring function and very nice results and projections can be made.

Docking simulations using other software could also make possible the quantitative and qualitative simulation for vanadium, Yttrium, lanthanum, and gadolinium compounds.

Since the full characterization for these interactions was performed in this work my future proposal is to try to implement it in other system rather than HSA, perhaps DNA, beta amyloid peptides or even enzymes.

9. REFERENCES

9. References

[1] Ooms, F. *Curr. Med. Chem*. **2000**, 7, 141-158.

[2] Blundell, T.L., Jhoti, H. and Abell, C. *Nat. Rev. Drug Discov*. **2002**,1, 45-54.

[3] Wishart, D. *Curr. Pharm. Biotechnol*., **2005**,6, 105-120.

[4] Wuthrich, K. *Acta Crystallogr. D 51*, **1995**,249-270.

[5] Brooks B., Olafson, B., States, D, Swaminathan, S., Karplus M. *Journal Of Computational Chemistry* . **1983**, 4, 187-217.

[6] Drews, J. *Science,* **2000**,1960-1964.

[7] T., Kaare, Olsen, J., **2009**, Cell. Mol. Life Sci. **2009**, 66, 2231–2247.

[8] Ludwig, C., Guenther, U., *Front. Biosci*. **2009**, 14, 4565-4574

[9] Peng, J., Moore, J., Abdul-Monanan, *N. Prog. Nucl. Magn. Reson*. **2004**, 44, 225-256.

[10] Stockmann, B., Dalvit, C. *Nucl. Magn. Spectrosc.,* **2002**, 41, 187-231.

[11] Baldus, M. *Curr Opin Struct Biol*. **2006**, 618-623.

[12] Jahnke, W., Widmer, H. *Cell. Mol. Life Sci*. **2004**, 61, 580–599

[13] Jhoti, H.,Cleasby, A., Verdonk, M., Williams, G. *Curr Opin Chem Biol*. **2007**,11,485-93.

[14] Skinner, A., Laurence, J. *J Pharm Sci*., **2008**, 4670-4695.

[15] Betz, M., Saxena, K., Schwalbe, H. *Curr Opin Chem Biol*. **2006**,10,219-225.

[16] Sprangers, R., Velyvis, A., Kay, L. *Nat Methods*. **2007**, 4, 697-703.

[17] Mayer, M., Meyer, B., *Angew. Chem. Int. Ed.,* **1999**, 38-12, 1784-1788.

[18] Mayer, M., Meyer, B., *J. Am. Chem.. Soc*., **2001**, 123, 6108-6117.

[19] Meyer, B., Peters, T., *Angew, Chem. Int. Ed*., **2003**, 42 No8, 864-890.

[20] John H. Streiff,J., Juranic, N. Macura, S., Warner, D., Jones, K., Perkins, W *Mol Pharmacol* **2004**, 66,929–935.

[21] Peters, T., All About Albumin – Biochemical, genetics and Medical application. **1996**.

[22] Henrotte, V., Elst, L. V., Laurent, S., Muller, R. *J. Bio. Inorg.Chem.,* **2007**, 12, 929–937.

[23] Aime, S., Chiaussa, M., Digilio, G., Gianolio, E., Terreno, E. *J. Bio. Inorg. Chem.,* **1999**, 4, 766-774.

[24] Seeliger, D., de Groot, B. *J. Comput. Aided. Mol. Des*. **2010**, 24,417–422.

[25] O. Trott, A. J. Olson, *Journal of Computational Chemistry*, **2010**, 31, 455-461.

[26] Vries, S., Dijk, M., Bonvin, A. *Nat Protoc.* **2010**, 5, 883-897.

[27] Sudlow, G., Birkett, D. G., *Mol. Pharmacol*., **1975**, 11, 824-832.

[28] K. Stierand, P.C. Maass, M. Rarey, *Bioinformatics*, **2006**, 22, 1710-1716.

[29] K. Stierand and M. Rarey, *ChemMedChem*, **2007**, 2, 853 – 860.

[30] Wang, Y-S., Liu, D., Wyss, D.F., *Magnetic Resonance Chem*., **2004**, 42, 485-489.

[31] Lepre, C.A., Moore, J.M., Peng, J.W., *Chem. Rev*., **2004**, 104, 3641-3675

[32] Fielding, L., *Prog. Nuc. Magn. Res. Spect*., **2007**, 51, 219-242.

[33] Angulo, J., Enríquez-Navas, P. M., Nieto, P. M. *Chem-Eur. J.* **2010**, 16, 7803-7812. [34] Dalvit, C., Flocco, M., Knapp, S.,Mostardini, M., Perego, R., Stockman, B. J., Veronesi M., Varasi, M., *J. Am. Chem. Soc*., **2002**, 124, 7702-7709.

[35] Dalvit, C., Fasolini, M., Flocco, M., Knapp, S., Pevarello, P. Veronesi, M., *J. Med. Chem.*, **2002**, 45, 2610-2614.

[36] Mori, S. et al, *J. Magn. Reson*. B, **1996**, 11, 182-196.

[37] Mori, S. et al, *J. Biomol. NMR*, **1996**, 7, 77

[38] Dalvit, C. et al, *J. Biomol. NMR*, **2001**, 21, 349-359.

[39] Stott, K. et al, J*. Magn. Reson.,***1997**, 125, 302–324.

[40] Balaram P. et al, *J. Am. Chem. Soc.* **1972**, 94, 4015-4017

[41] Balaram P. Breslow, E. *Biochemistry*. **1973**, 12, 4695-4704

[42] Meyer B. et al, *Eur. J. Biochem*. **1997**,246, 705-709

[43] Kalk, A. and Berendsen, H. J. C. *J., Magn. Reson*., **1976**, 24, 343-366.

[44] Fesik, S. et al, *Science* **1996**,274:1531–1534.

[45] Hajduk P. et al, *J Med Chem*, **2000**, 43,3862–3866

[46] Hajduk, p. et al., *J Med Chem*. **1999**,42:2315-2317

[47] Salvatella X. et al, *Chem Soc Ver* **2003**, 32:365–372

[48] Sanders J. et al, Oxford University Press.**1987** p 308

[49] Evans J. et al, Oxford University Press.**1995** p 444

[50] Fielding L. et al, *Curr Top Med Chem* ,**2003** ,3:39–53.

[51] Hwang, T., Shaka, A. G., *Jour. Magn. Reson*., **1995**, 112, 275-279

[52] Piotto, M., Saudek, M., Sklend, J., *J. Biomol. NMR*, **1992**, 2, 611

[53] Kemper, S., Patel, M. Errey, J., Davis, B., Jones, J. Claridge, T. *J. Magn. Reson.,* **2010**, 1–10

[54] Feher, K., Patrick Groves, P., Batta, G., Jimenez-Barbero, J., Muhle-Goll, C., Kover, K. *J. Am. Chem. Soc*. **2008**, 130, 17148–17153

[55] Kalk, A., Berendsen, H. J. C. J., *Magn. Reson*., **1976**, 24, 343-366.

[56] Brecker, L., Tyla, C., Nidetzky, B. *Carbohydrate Research*, **2008**, 2153–2161

[57] Wagstaff, J., Vallath, S. Marshall, J. Williamson, R., Howard, M. *Chem. Commun*., **2010**, 46, 7533–7535

[58] Shimada, I., Ueda, T., Matsumoto, M., Sakakura, M., Osawa,. M., Takeuchi, K., Nishida N., Takahashi, H. *Prog. Nucl. Magn. Reson. Spectrosc*., **2009**, 54, 123–140.

[59] Meinecke, R.,Meyer, B. *J. Med. Chem*., **2001**, 44, 3059–3065.

[60] Neffe, A. T., Vilang, M., Gruneberg, I., Meyer, B., *J. Med. Chem*., **2007**, 50, 3482-3488.

[61] Teixeira, J., Dias, D., Cañada, F., Martins, J., André, J., Jiménez-Barbero, J. Geraldes, C. *J Biol Inorg Chem*. **2011**, in press

[62] Cheng, Y., Prussof, W., *Biochem Pharmacol*.,**1973**, 22, 3099-3108

[63] Wang,Y., Liu, D., Wyss, D. *Magn. Reson. Chem.* **2004**, 42: 485–489

[64] Sun, H. and Scott,D., *Chem Biol Drug Des* **2010**, 75: 3–17

[65] Waterbeemd H. and Gifford E.,*Nat Rev Drug Discov* **2003**,2:192–204.

[66] de Groot M.J., *Drug Discov Today* **2006**,11:601–606.

[67] Clark, R., *Current Topics in Medicinal Chemistry,* **2009***, 9,* 791-810

[68] Kroemer R., *Curr Protein Pept Sci* **2007**,8:312–328.

[69] de Graaf C., *J Med Chem* **2006**,49:2417–2430.

[70] Amaro, R., *Current Topics in Medicinal Chemistry,* **2010***, Vol. 10,*

[71] Kriwacki, R., Pitner, T. *Pharm. Res*. **1989**, 532-554

[72] Bretonniere, Y., Cann, M. J., Parker, D., Slater, R. *Chem. Commun*. **2002**, 1930-1931.

[73] Bünzli, J.-C. G., Sigel, A., Sigel, H., Eds., Marcel Dekker: New York, **2003**.

[74] Hanaoka, K., Kikuchi, K., Kojima, H., Urano, Y., Nagano, T. *J. Am. Chem. Soc*. **2004**, 126, 12470-12476.

[75] Poole, R. A., Bobba, G., Cann, M. J., Frias, J.-C., Parker, D., Peacock, R. D. Org. *Biomol. Chem*. **2005**, 3, 1013-1024.

[76] Pandya, S., Yu, J., Parker, D. *Dalton Trans*. **2006**, 2757-2766.

[77] Yu, J., Parker, D., Pal, R., Poole, R. A., Cann, M *J. Am. Chem. Soc*. **2006**, 128, 2294-2299.

[78] Vandevyver, C. D., Chauvin, A. S., Comby, S., Bünzli, J.-C. G. Chem. Commun. **2007**, 1716-1718.

[79] Montgomery, C. P., Murray, B. S., New, E. J., Pal, R., Parker, D. Acc. Chem. Res. **2009**, 42, 925-937.

[80] New, E. J., Parker, D., Smith, D. G., Walton, J. W. Curr. Opin. Chem. Biol. **2010**, 14, 238-246.

[81] Poole, R. A., Montgomery, C. P., New, E. J., Congreve, A., Parker, D., Botta, M. Org. Biomol. Chem. **2007**, 5, 2055-2062.

[82] Kielar, F., Law, G.-L., New, E. J., Parker, D. *Org. Biomol. Chem.* **2008**, 6, 2256-2258.

[83] Murray, B. S., New, E. J., Pal, R., Parker, D. *Org. Biomol. Chem.* **2008**, 6, 2085-2094.

[84] New, E. J., Parker, D. *Org. Biomol. Chem*. **2009**, 7, 851-855.

[85] New, E. J., Parker, D., Peacock, R. D. *Dalton Trans*. **2009**, 672-679.

[86] Pal, R., Parker, D., Costello, L. C. *Org. Biomol. Chem.* **2009**, 7, 1525-1528.

[87] Montgomery, C. P., New, E. J., Palsson, L. O., Parker, D., Batsanov, A. S., Lamarque, L. *Helv. Chim. Acta*. **2009**, 92, 2186-2213.

[88] Montgomery, C. P., New, E. J., Parker, D., Peacock, R. D. *Chem. Commun.* **2008**, 4261-4263.

[89] R. B. Lauffer, *Chem. Rev*. **1987**, 87, 901 – 927. [2]

[90] A. E. Merbach, E. Toth, , Wiley, Chichester, **2001**. [3]

[91] S. Aime, M. Botta, M. Fasano, E. Terreno, *Chem. Soc. Rev*. **1998**, 27, 19–29.

[92] P. Caravan, J. J. Ellison, T. J. McMurry, R. B. Lauffer, *Chem. Rev.***1999**, 99, 2293–2352.

[93] H. J. Weinmann, A. Mühler, B. Radüchel in Biomedical Magnetic Resonance Imaging and Spectroscopy (Ed.: I. R. Young), Wiley, Chichester, **2000**, p. 705.

[94] L. Banci, I. Bertini, C. Luchinat, Nuclear and Electronic Relaxation,VCH, Weinheim, **1991**.

[95] S. Aime, M. Botta, M. Fasano, E. Terreno, Wiley, Chichester, **2001**, p. 193

[96] Sakurai et al., *Chemical Society Reviews*, **2008**, 37, 2383-2392.

[97] H. Sakurai, K. Fujii, H. Watanabe, H. Tamura, *Biochem Biophys Res Commun*, 1995, 214,1095–1101.

[98] H. Sakurai, H. Sano, T. Takino, H. Yasui, *J Inorg Biochem*, **2000**, 80, 99–105.

[99] K. Kawabe, Y. Yoshikawa, Y. Adachi, H. Sakurai, *Life Sci*. **2006**,78, 2860-2866.

[100] Y. Schechter, *Lett Pept Sci* 5, **1998**, 319-322.

[101] V.G. Yuen, C. Orvig, J.H. McNeill, Can *J Physiol Pharmacol*. **1993**,71, 263-269.

[102] M. Hiromura, A. Nakayama, Y Adachi, M. Doi, H. Sakurai, *J Biol Inorg Chem*, **2007**, 12, 1275–1287.

[103] Z.W. Yu, P.A. Jansson, B.I. Posner, U.P. Smith, J.W. Eriksson, *Diabetologia* **2010**, 40, 1197-1203.

[104] W.M. Mueller, K.L Stanhope, F. Gregoire, J.L. Evans, P.J. Havel, *Obes Res*. **2000**, 8, 530-539.

[105] Y. Adachi, H. Sakurai, *Chem. Pharm. Bull*. **2004**, 52, 428-433

[106] J. Gätjens, B. Meier, T. Kiss, E.M. Nagy, P. Buglyó, H. Sakurai, K. Kawabe, D. Rehder, *Chemistry*. **2003**, 9, 4924-4935.

[107] H. Faneca, V.A. Figueiredo, I. Tomaz, G. Gonçalves, F. Avecilla, M.C. Pedroso de Lima, C.F.G.C.Geraldes, J.C. Pessoa JC, M.M.C.A Castro, *J Inorg Biochem* **2009**, 103, 601–608.

[108] D.C. Crans, *J. Inorg. Biochem*. **2000**, 80,123–131.

[109] J. Burgess, B. Castro, C. Oliveira, M. Rangel, W. Schlindwein, *Polyhedron* **1997**, 16 789.

[110] M.M.C.A. Castro, F. Avecilla, C.F.G.C. Geraldes, B. de Castro, M. Rangel, *Inorg.Chim.Acta* **2003**, 356, 142-154.

[111] D. Rehder, J.C. Pessoa, C.F.G.C. Geraldes, M.M.C.A. Castro, T. Kabanos, T. Kiss, B. Meier, G. Micera, L. Pettersson, M. Rangel, A. Salifoglou, I. Turel, D.Wang, *J.Biol.Inorg.Chem.* **2002**, 7, 384-396

[112] T.C. Delgado, I. Correia, J.C. Pessoa, J.G. Jones, C.F.G.C. Geraldes, M.M.C.A. Castro, *J Inorg Biochem* 2005, 99, 2328-2339.

[113] Passadouro, M., M.M.C.A.Castro et al. *J Inorg Biochem*, **2010**,104, 987-92.

[114] B.D. Liboiron, K.H. Thompson, G.R. Hanson, E. Lam, N. Aebischer, C. Orvig, *J. Am. Chem. Soc.* 2005, 127, 5104–5115.

[115] Sanna, Daniele, Micera, G., Garribba. E. *Inorg. Chem*., **2009**, 48, 5747–5757.

[116] Sanna, D.; Micera, G.; Garribba, E. *Inorg. Chem*., **2010**, 49, 174–187.

[117] Sanna, D., Buglyo, P., Micera, G., Garribba, E. *J Biol Inorg Chem* **2010**, 15, 825–839.

[118] E.G. Ferrer, A. Bosch, O. Yantorno, E.J. Baran . *Bioorg. Med. Chem*. **2008**, 3878–3886.

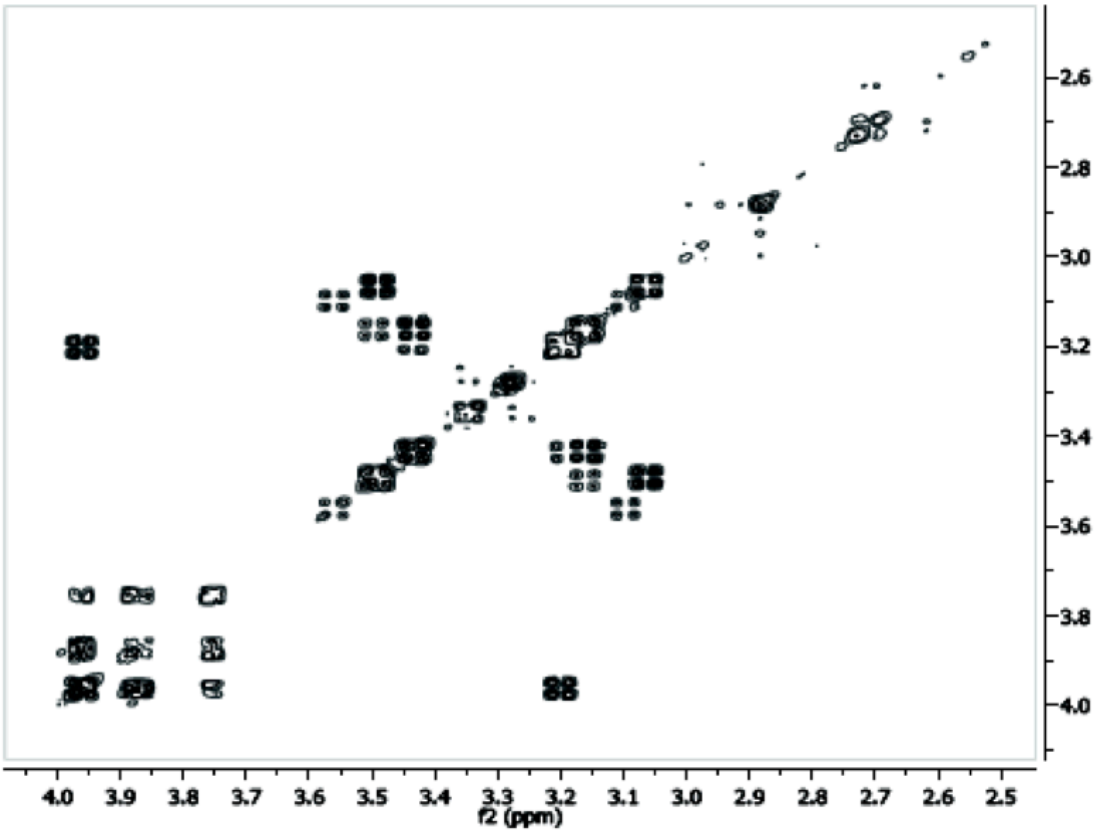
10. APPENDIX

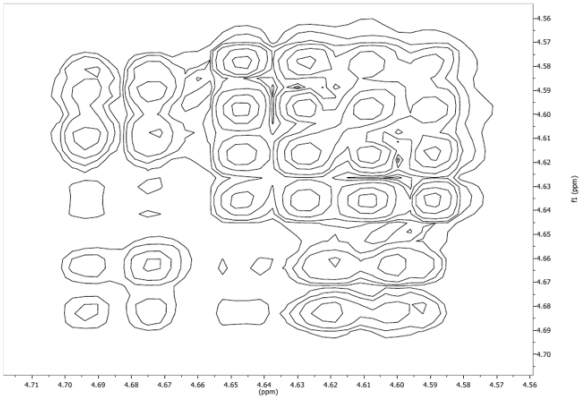
10. Appendix

It is very useful and user-friendly to consult the main characteristic NMR bases on setting up experiments, and achieving not only nice STD spectra but also the knowledge to deal with selective saturation, pulse shape and power, how to select the desirable frequency and many more other keys to run a NMR spectrum.

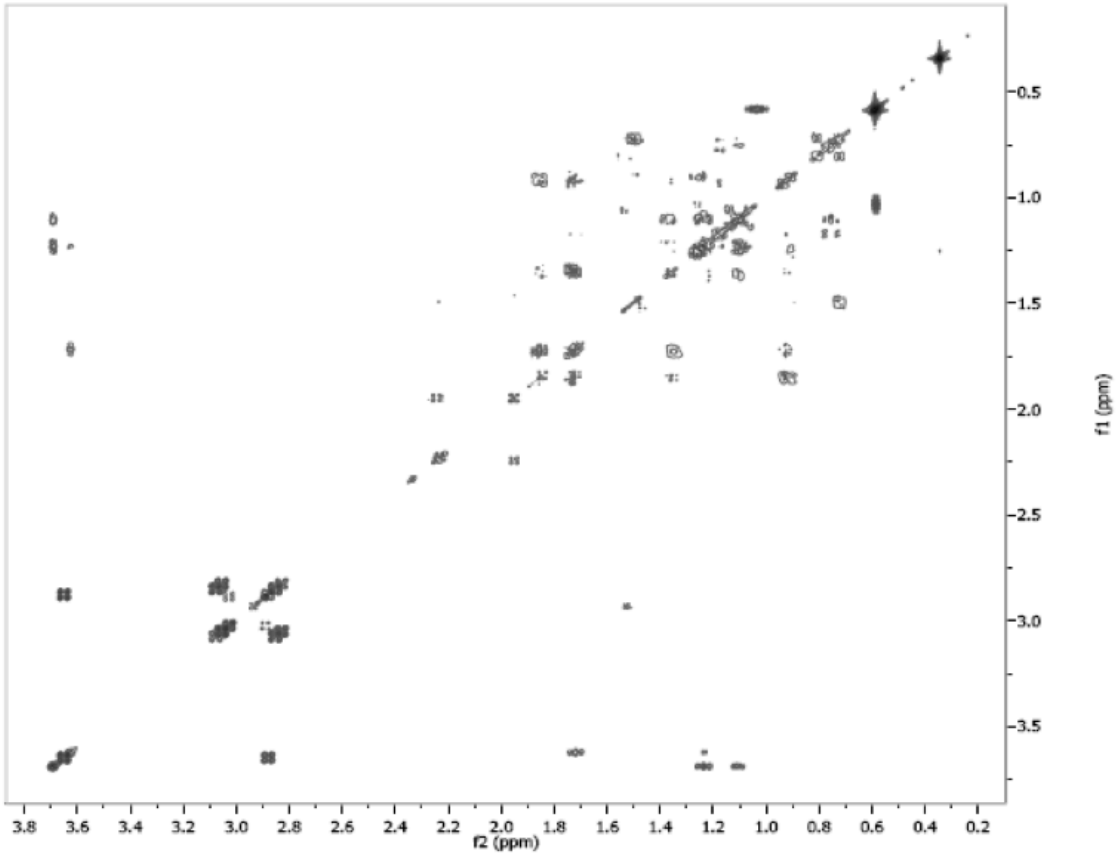
All gCOSY spectra obtained are deposited here as well as the full assignments for thos complexed that were not in current literature.

10.1 gCOSY spectra

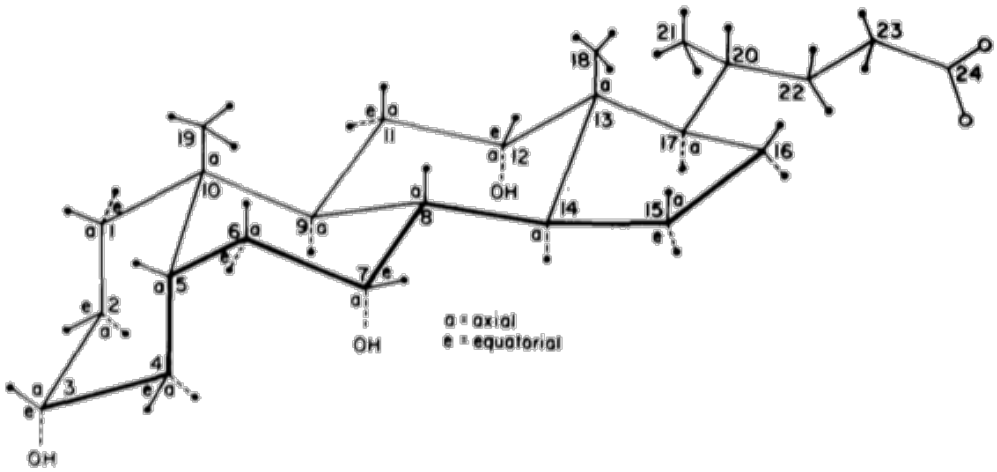


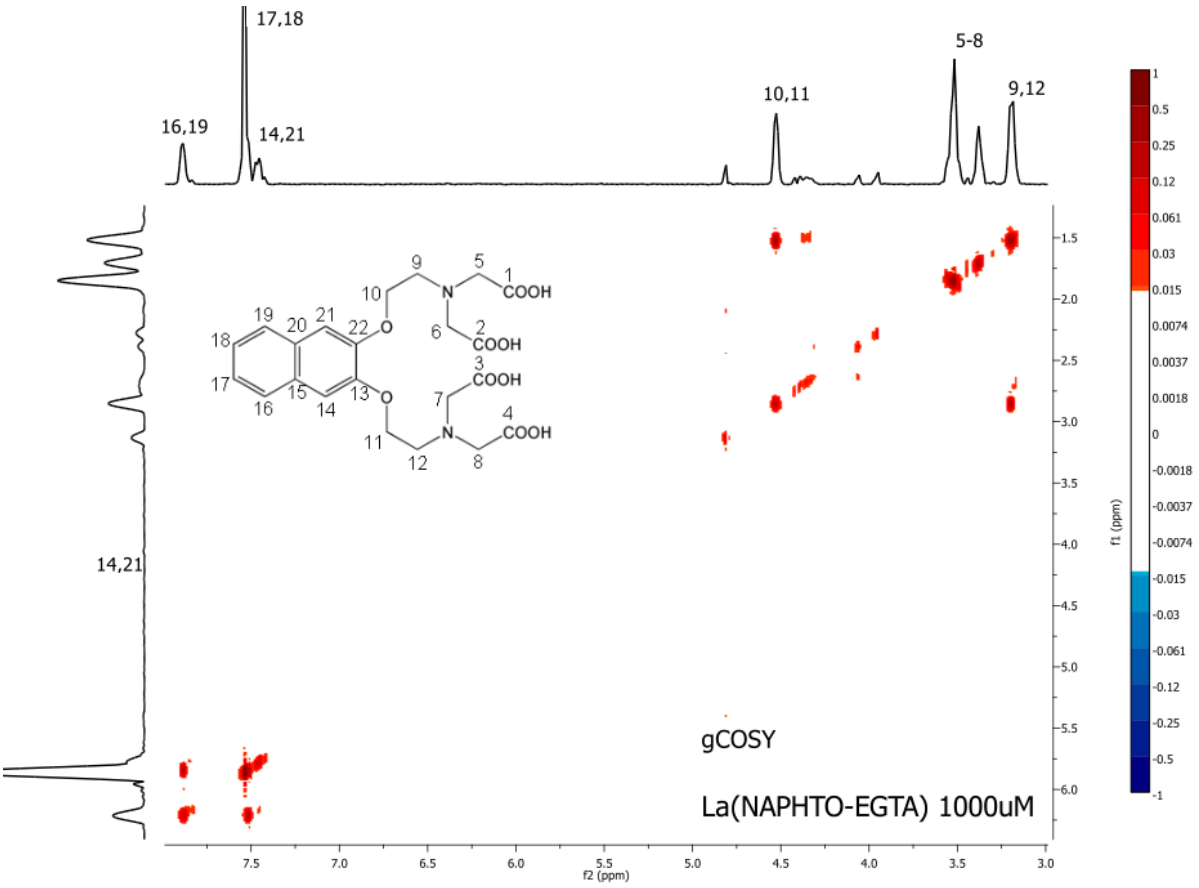


gCOSY obtained for La(BOPTA) aliphatic and aromatic expansion.



gCOSY obtained for La(DTPA-Cholate) being the first two resonance the OH groups and the first two signals the 18 and the overlap of 19 and 21.





gCOSY obtained for La(NAPHTO-EGTA) and full NMR assignment.

10.2 NMR REPORT - NMR study of the Thioflavin T and DO3APIB and interaction in with beta amyloid peptides

Results gathered and perspectives

The experiments were conducted in Coimbra, VNMR 600, in Caparica, Bruker Avance 600 Cryo, and in ITQB (CERMAX), Bruker Avance 800.

The first objective was to apply STD NMR at the beta amyloid peptides and also in another objective reproduce the results of J. Pept. Sci. 2009; 15: 435–441 (Detection of interactions of the β-amyloid peptide with small molecules employing transferred NOEs) and then resume the study of DO3APIB.

Here is presented first all spectra acquired of STD-NMR, and then transferred NOE and some doubts and considerations.

STD - NMR

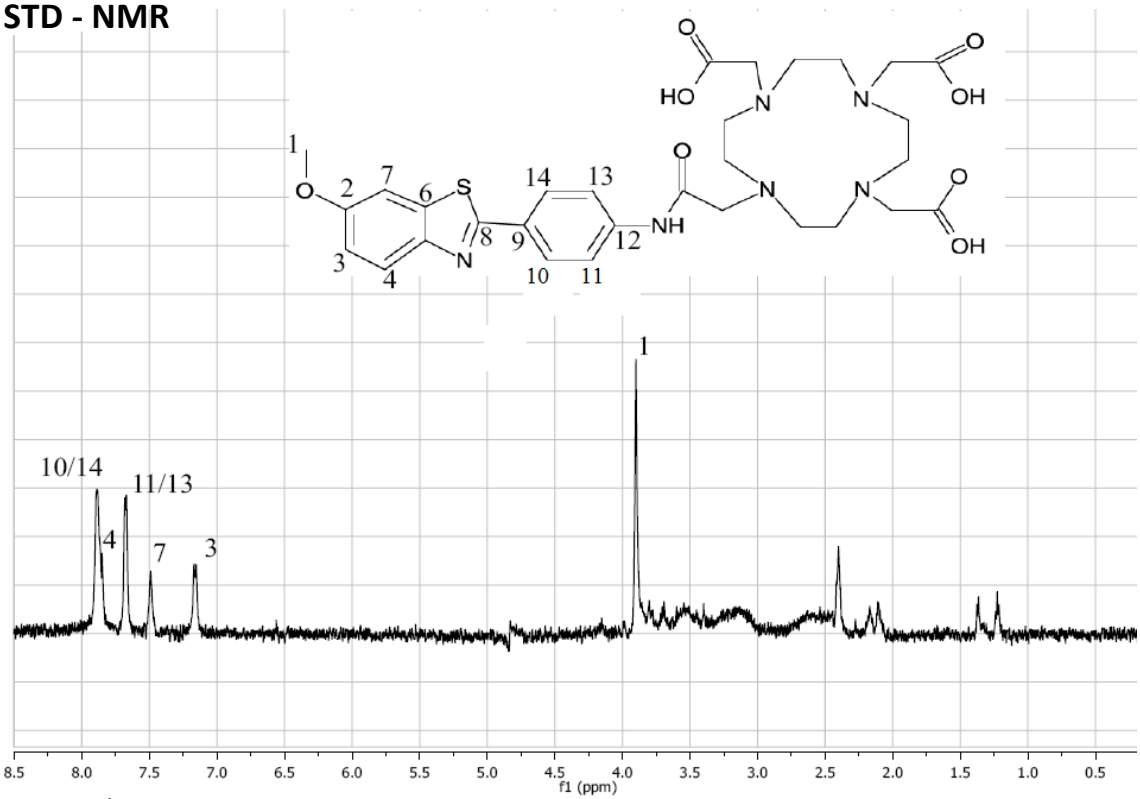


Figure 1: ^{1}H DPFGSE spectrum of DO3APIB, 1 mM, and the NMR assignment for DO3APIB resonances.

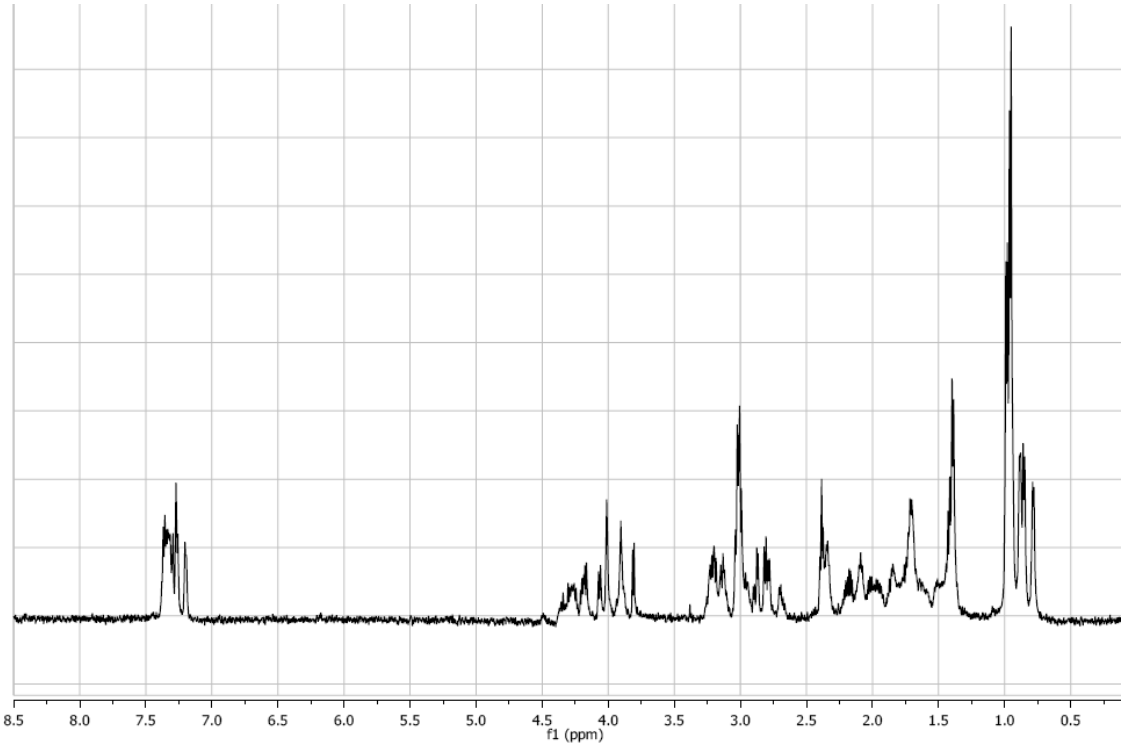


Figure 2: ^{1}H DPFGSE spectrum of beta amyloid 12-28 100 uM.

As seen on normal proton spectra no problem achieving good signal from the beta-amyloid peptide.

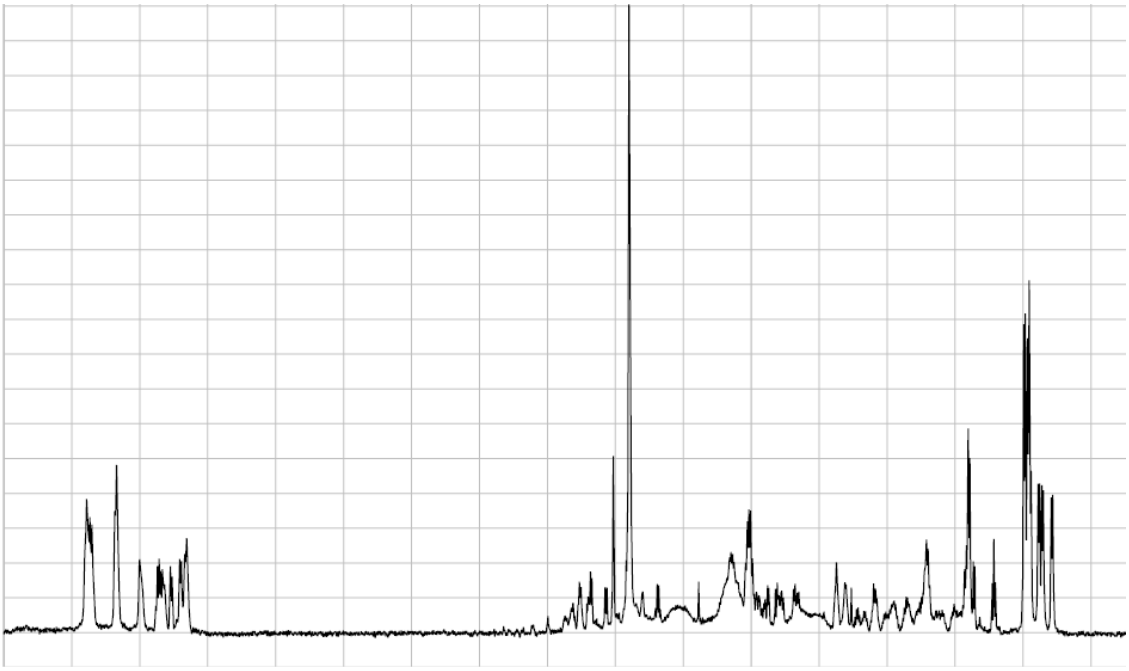


Figure 3: ^{1}H DPFGSE spectrum of beta amyloid 12-28 100 uM and DO3APIB 1000 uM.

STD Spectra for the mixture was acquired with different saturation resonances trying to get only protein signal, and TRIM pulse was exhaustively calibrated and no real STD effect was observed, for example figure 4.

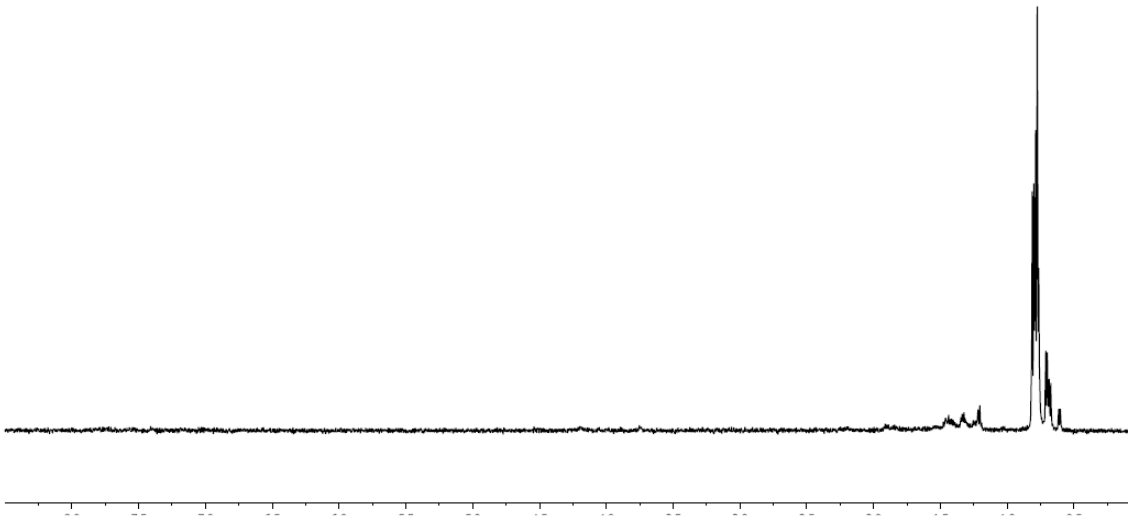


Figure 4: ^{1}H STD NMR spectrum of a mixture of DO3APIB, 1 mM, beta amyloid 12-28, 100 uM. Saturation at 0.85 ppm. Trim pulse was used in experiments to suppress the protein resonances.

Is clearly visible that the Trim pulse cannot successfully suppress the protein resonances because when we saturate at 0.85 ppm (only protein resonances) the STD spectra should be a baseline and no signals should appear, what doesn't happen. Mostly due to the small weight and length of the peptide, the relaxation rate is similar to the small compounds and no STD effect is visible for this shorter beta amyloid peptide.

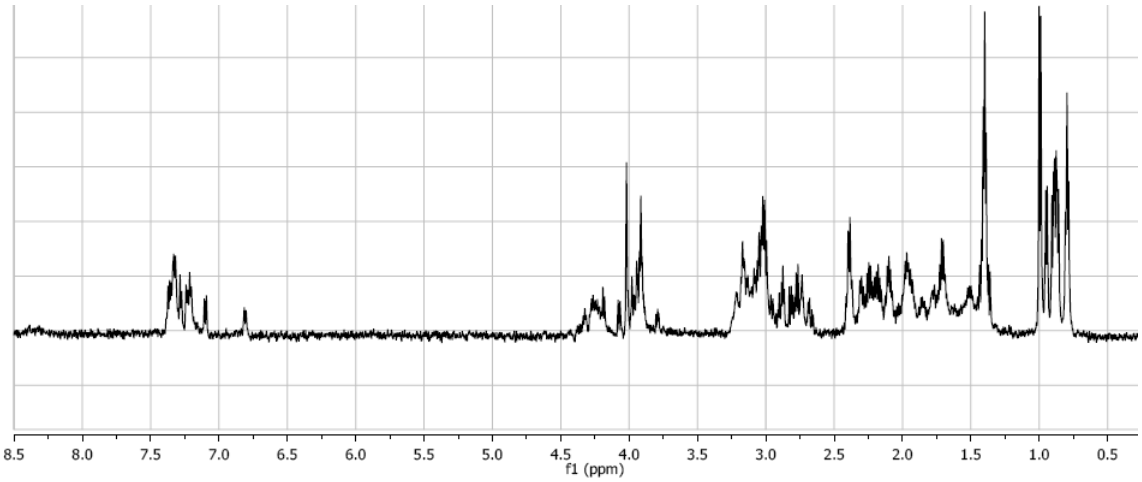


Figure 5: ^{1}H DPFGSE spectrum of beta amyloid 1-28 100 uM

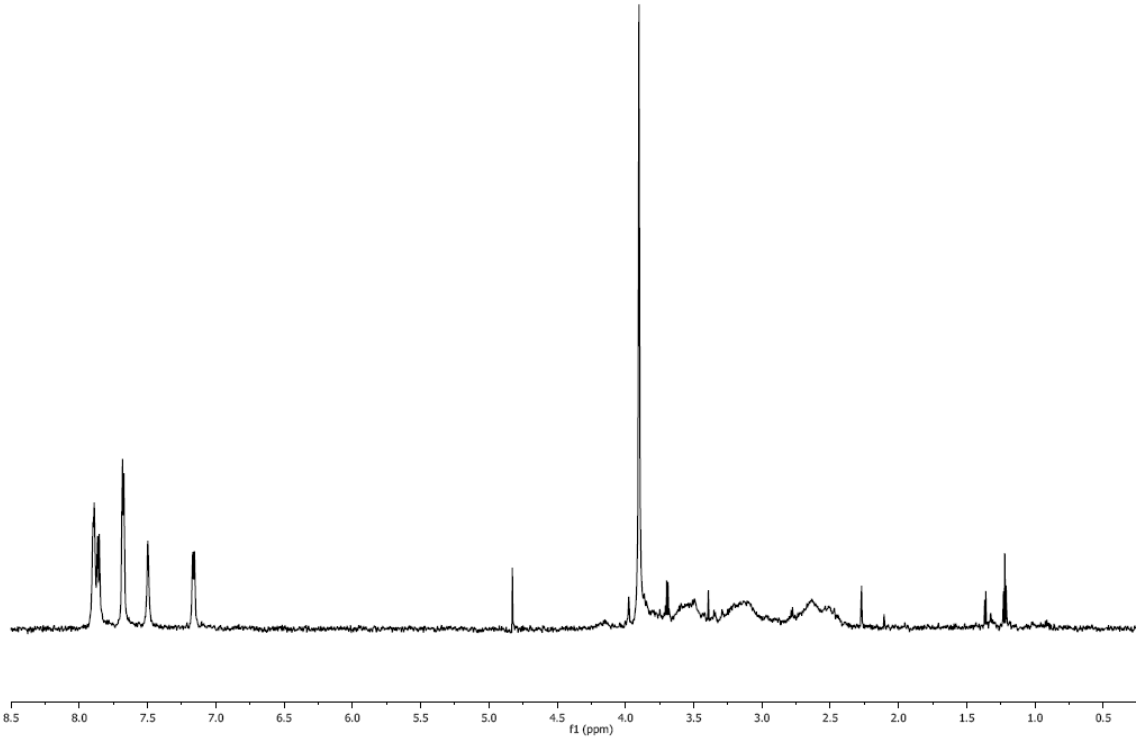


Figure 6: ^{1}H DPFGSE spectrum of beta amyloid 1-28 100 uM and DO3APIB 1000 uM.

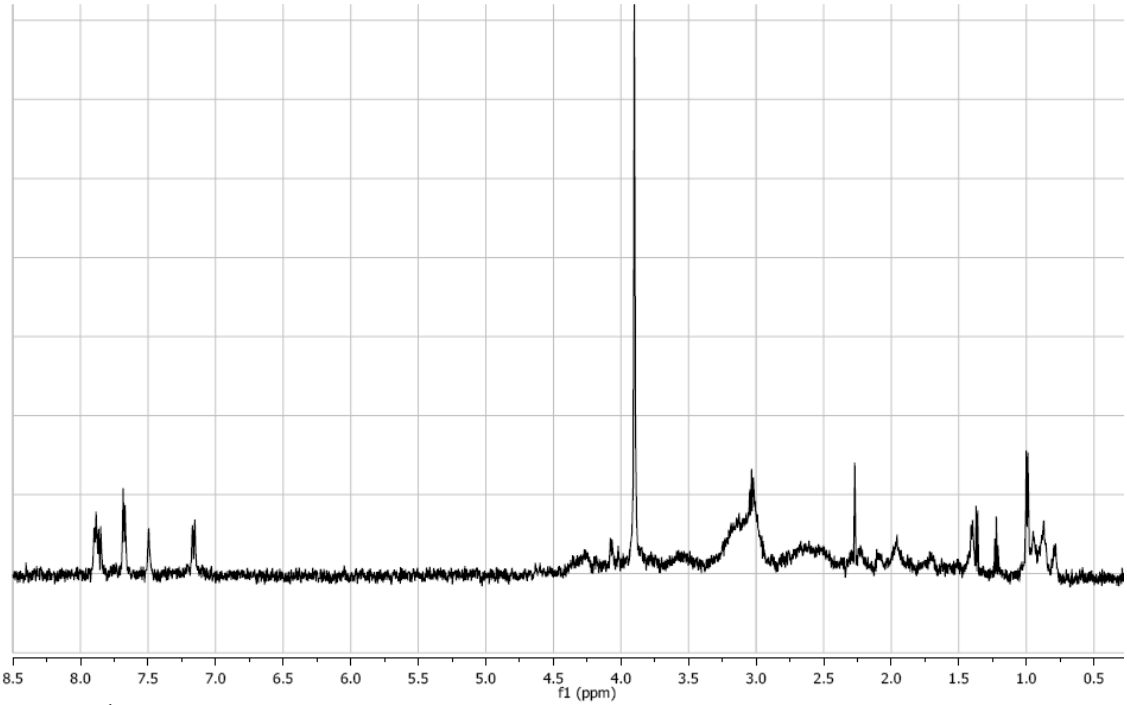


Figure 7: ^{1}H STD NMR spectrum of a mixture of DO3APIB, 1 mM, beta amyloid 1-28, 100 uM. Saturation at 0.85 ppm. Trim pulse was used in experiments to suppress the protein resonances.

With beta amyloid 1-28 also no real STD spectra is achieved. Once the resonances where the saturation was made remain in a baseline, and the signals of the ligand appearing here may be suffering other effect than STD saturations transference

from the peptides, such as the beta amyloid peptide are also appearing, but indeed, the result is far better than with the 12-28 peptide.

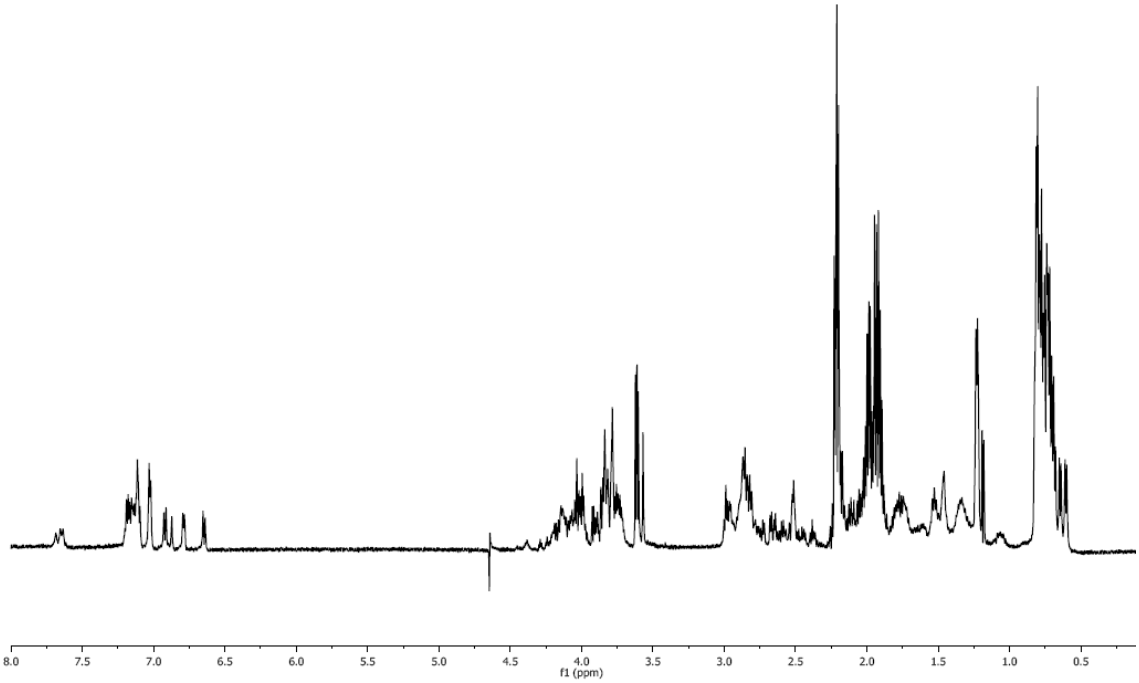


Figure 8: ^{1}H DPFGSE spectrum of beta amyloid 1-40 100 uM.

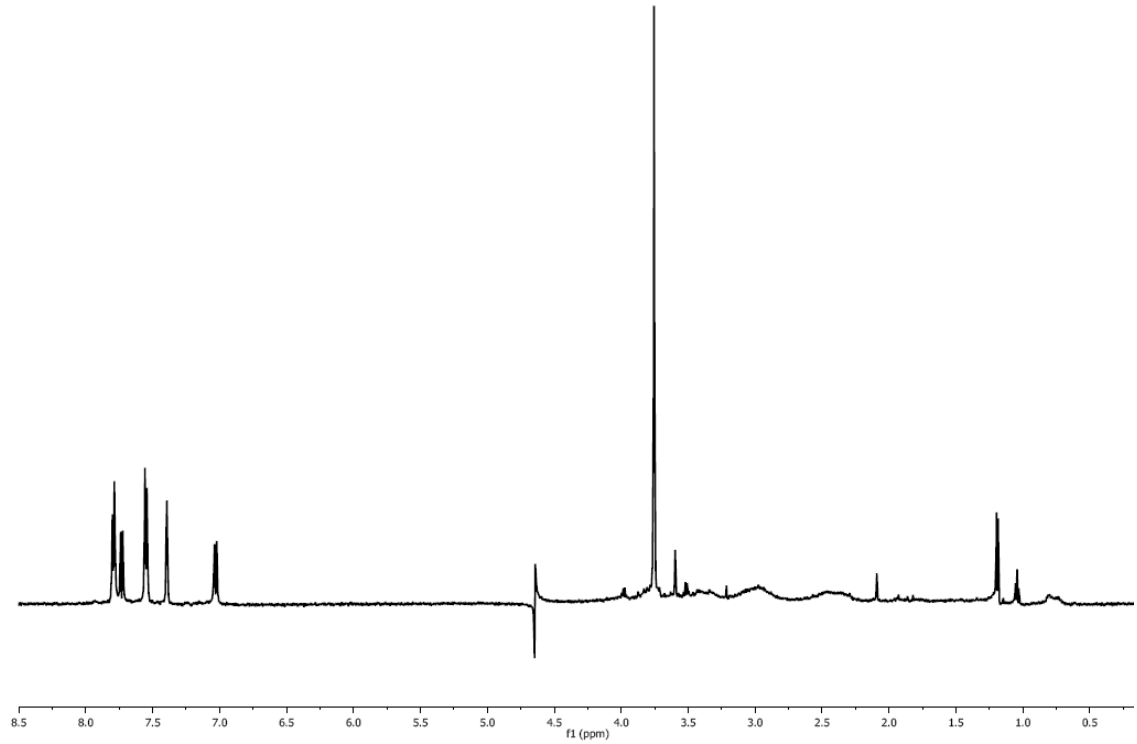


Figure 9: ^{1}H DPFGSE spectrum of beta amyloid 1-40 100 uM and DO3APIB 1000 uM.

When we have the high length peptide the Trim pulse really works on the suppression of peptide signals, we almost see the DO3APIB resonances and the impurities of the compound, the broadened signals of the peptide represent the success of a good suppression with DPFGSE scheme that we are using.

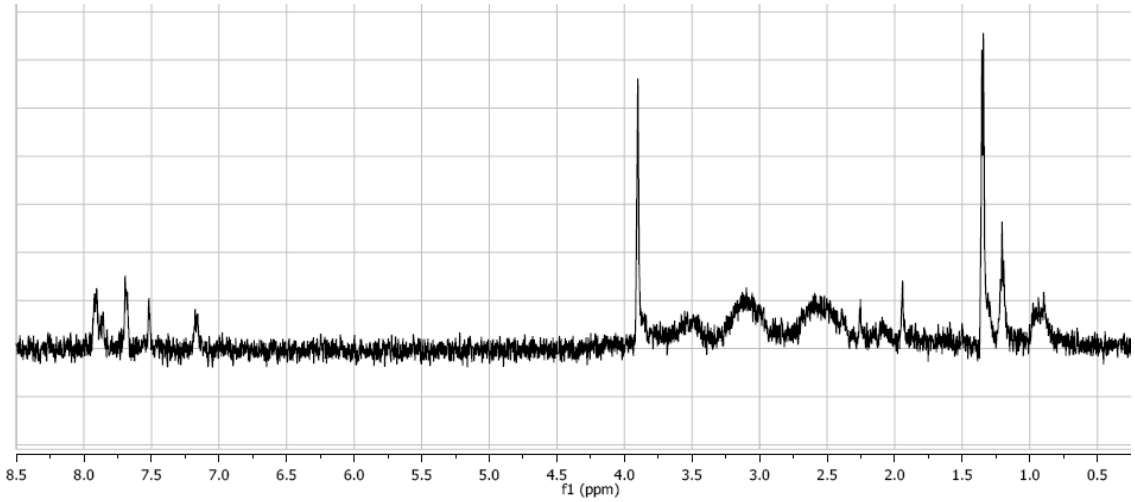


Figure 10: ^{1}H STD NMR spectrum of a mixture of DO3APIB, 1 mM, beta amyloid 1-40, 100 uM. Saturation at 0.85 ppm. Trim pulse was used in experiments to suppress the protein resonances.

The result from STD is coherent with a good suppression once we have broadened and short peaks for the peptide signals, and in here we can see also the DO3APIB resonances appearing quite well on STD spectra. The peaks present in 1.2, 1.4 ppm are ligand impurities carried from the synthesis, also showing somehow a STD effect, or just because they are near the 0.85 ppm were we saturate, once the selective saturation scheme is specific for 0.2 ppm is not likely but no sure for now.

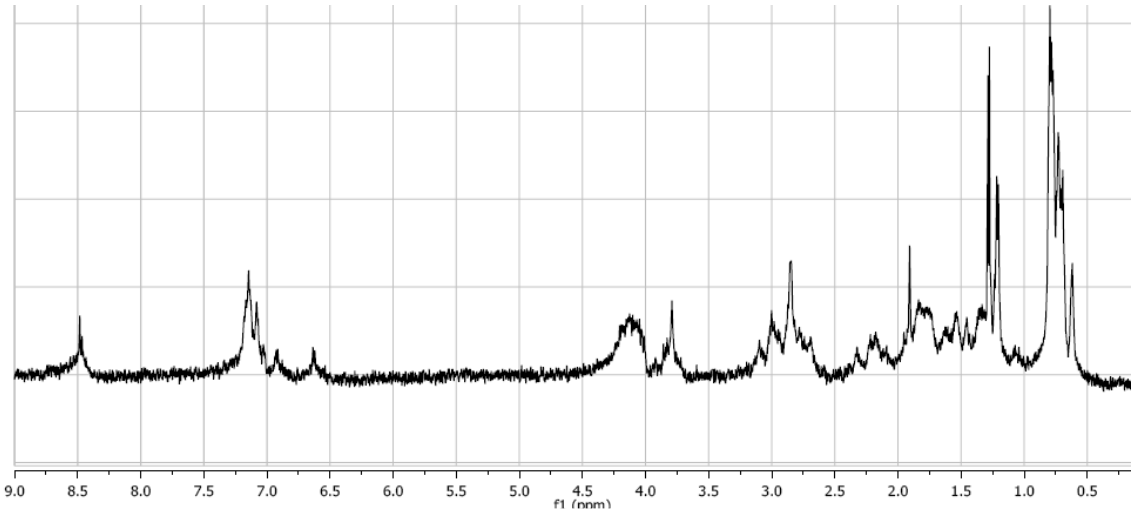


Figure 11: ^{1}H DPFGSE spectrum of beta amyloid 1-42 100 uM.

The full length peptide shows a much less resolved spectra and with broadened signals, indicating that it resembles more like a large target, with relaxation rates, size and weight suitable for STD NMR protocol.

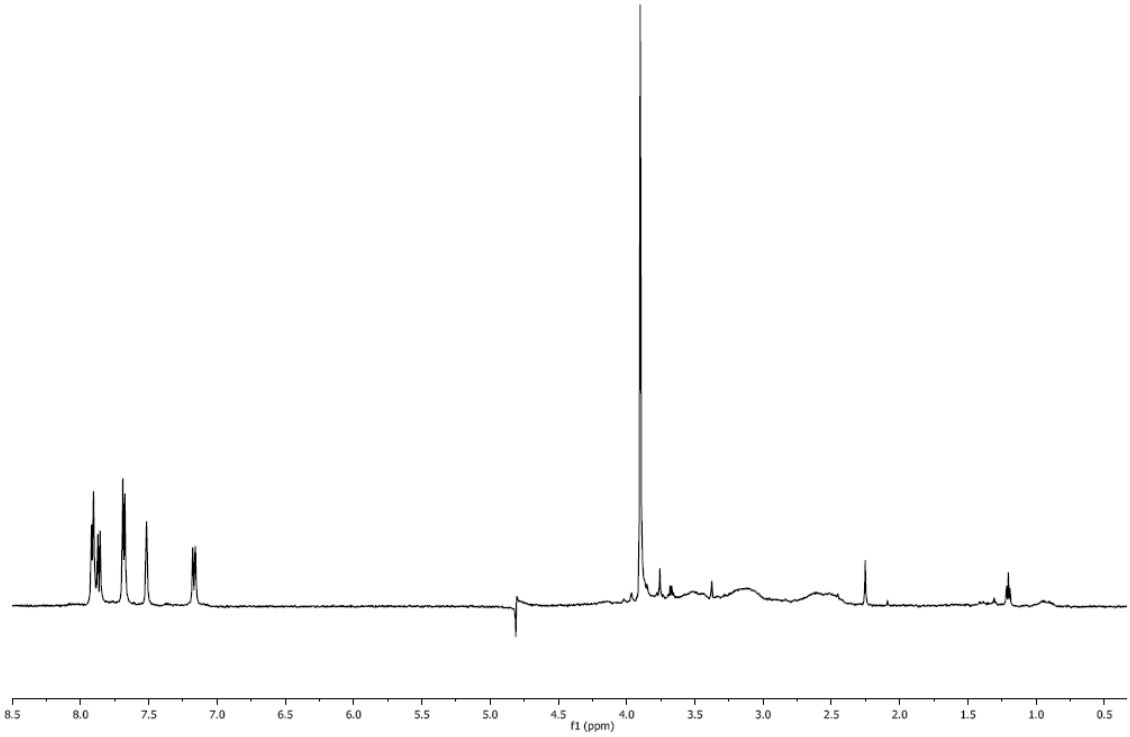


Figure 12: ^{1}H DPFGSE spectrum of beta amyloid 1-42 100 uM and DO3APIB 1000 uM.

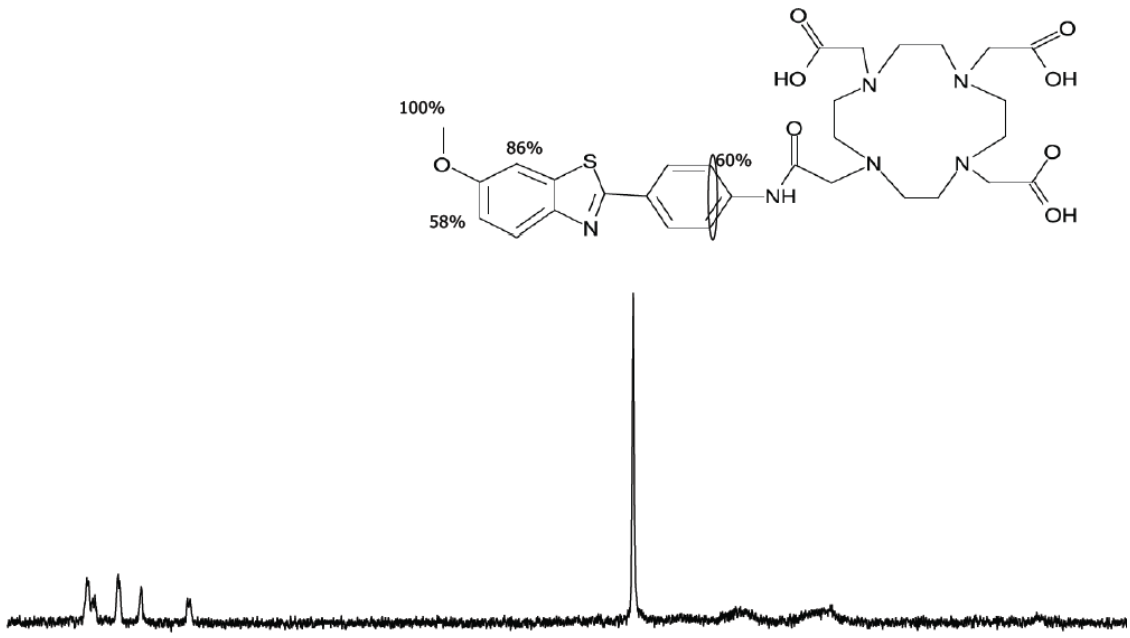


Figure 13: ^{1}H STD NMR spectrum of a mixture of DO3APIB, 1 mM, beta amyloid 1-42, 100 uM. Saturation at 0.85 ppm. Chemical structure representing the characterization of the binding epitope for DO3APIB. Trim pulse was used in experiments to suppress the protein resonances.

The STD spectra acquired with 1-42 beta amyloid peptide stands for a perfect STD NMR spectra were no peptide resonances appear and the ligand signals are clearly appearing and after epitope mapping we see that protons assigned as number 1, 3, 7 and 11/13 (see figure 1) are the ones receiving more saturation transference indicating that the

binding is occurring thru the methyl and also the aromatic rings of DO3APIB. The STD result evidences first of all the binding of the compound, second we have the characterization of the epitope for the interaction, and now it is only missing the protein site involved in the interaction to fully understand this molecular process.

To further proposals I would say that more DO3APIB should be synthesized in order to remove the impurities, if possible, and kinetic constant dissociation may be gathered as well in the STD protocol, established here in Coimbra. I believe the 1-40 and 1-42 peptides are the most likely candidates to continue this work.

Transferred NOE's – Thioflavin T

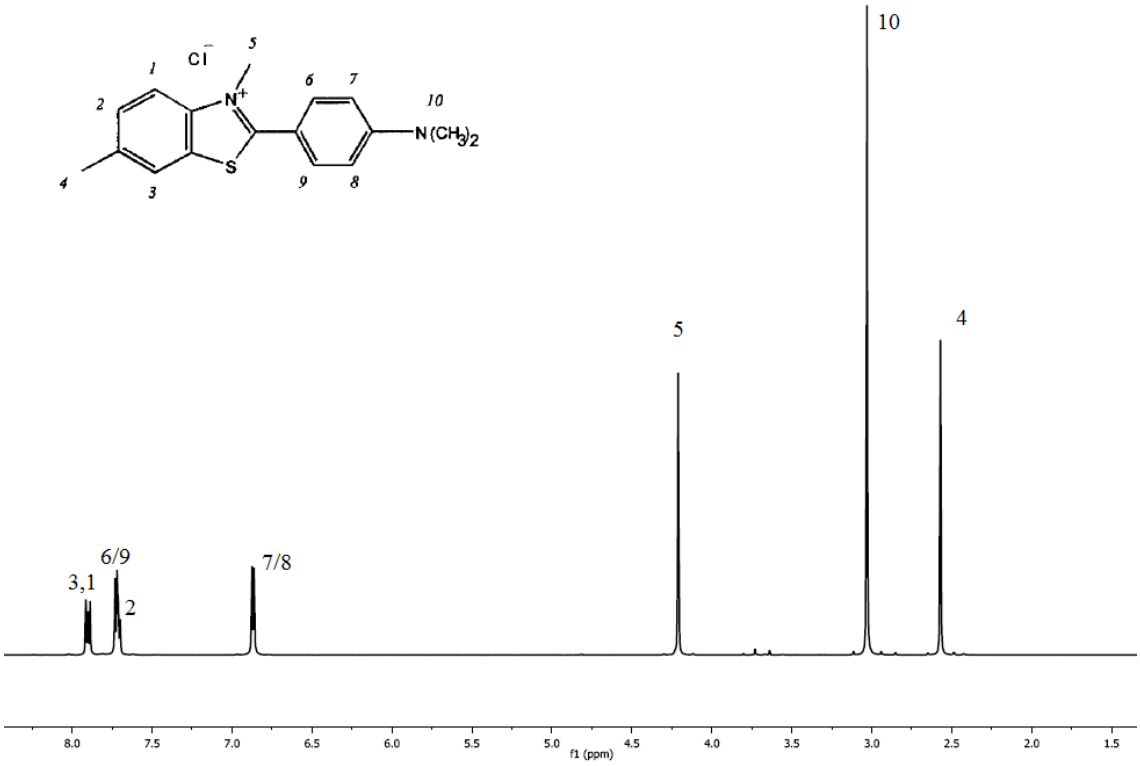


Figure 14: ^{1}H DPFGSE spectrum of a mixture of Thioflavin T, 6 mM, and the NMR assignment for Thioflavin T resonances.

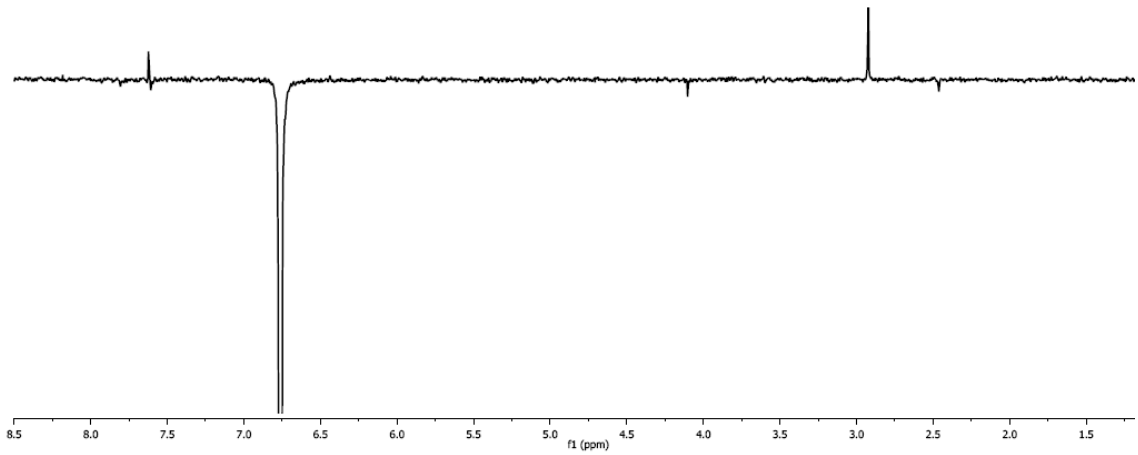


Figure 15: 1D NOESY spectrum of Thioflavin T, 600 uM, in D2O with 800 ms of mixing time. Selective inversion of protons 7/8.

As shown on the paper when we have the Thioflavin T alone, we see a positive NOE on resonances 10 and 6/9 due to the proximity in this dye. When we add a protein, in this case, a peptide, we expet that the NOE became negative if the compound interacts with the peptide because the compound will get the large molecule properties.

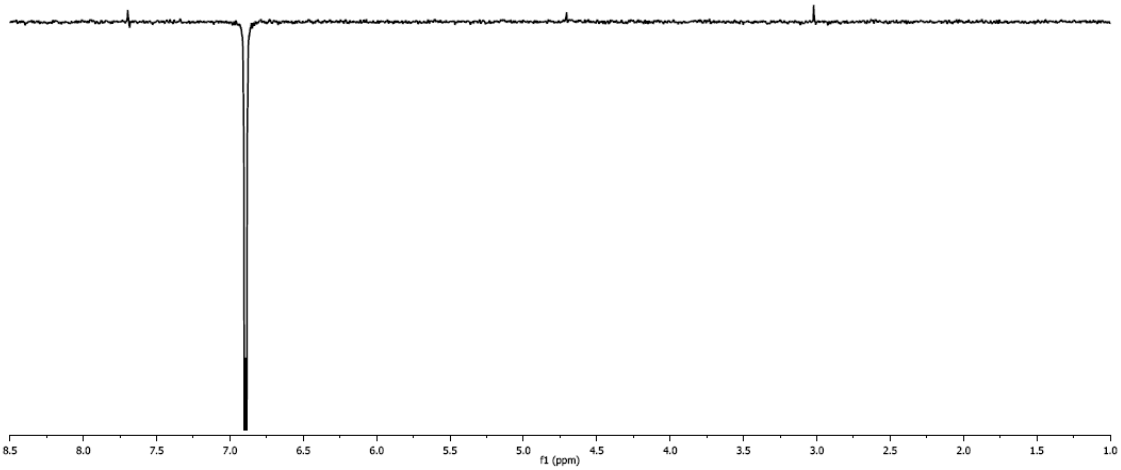


Figure 17: 1D NOESY spectrum of Thioflavin T, 600uM and beta amyloid 1-28, 100 uM, in D2O with 150 ms of mixing time. Selective inversion of protons 7/8.

As we can see the NOE signal maintains the positive signal, and even with lower mixing times (50, 60, 90 ms) they still positive. This fact represents first, a non-binding scenario, secondly, that the beta amyloid used is not large enough for the NOE cross relaxation to take place, or we are in the point that NOE is negative for the magnetic field of the spectrometers used (600 CRYO and 800).
When we see the proton spectrum we see that everything is going as it should for this mixture, see figure 18.

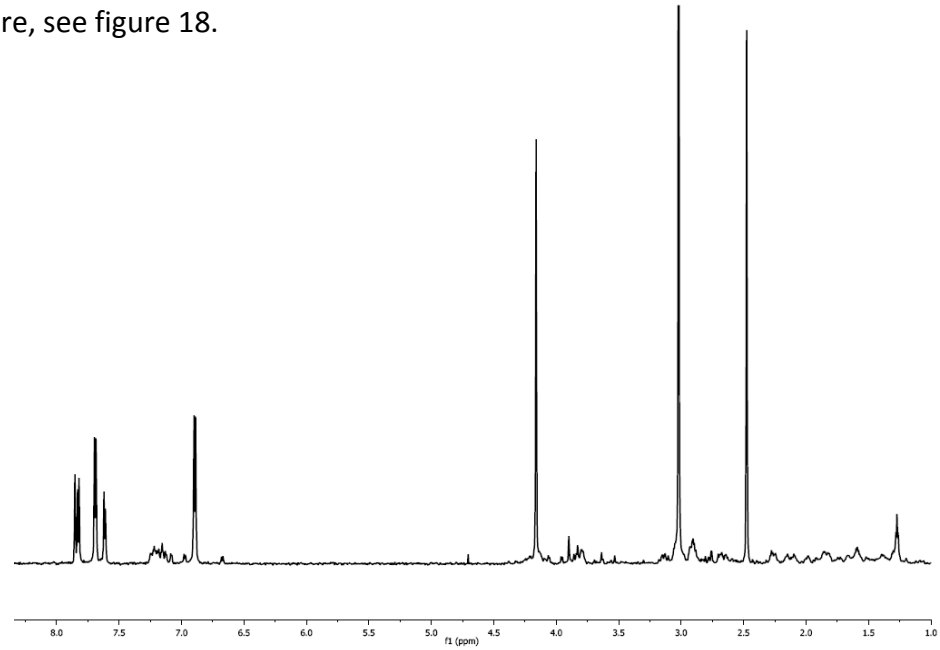


Figure 17: ^{1}H spectrum of Thioflavin T, 600uM and beta amyloid 1-28, 100 uM, in D2O.

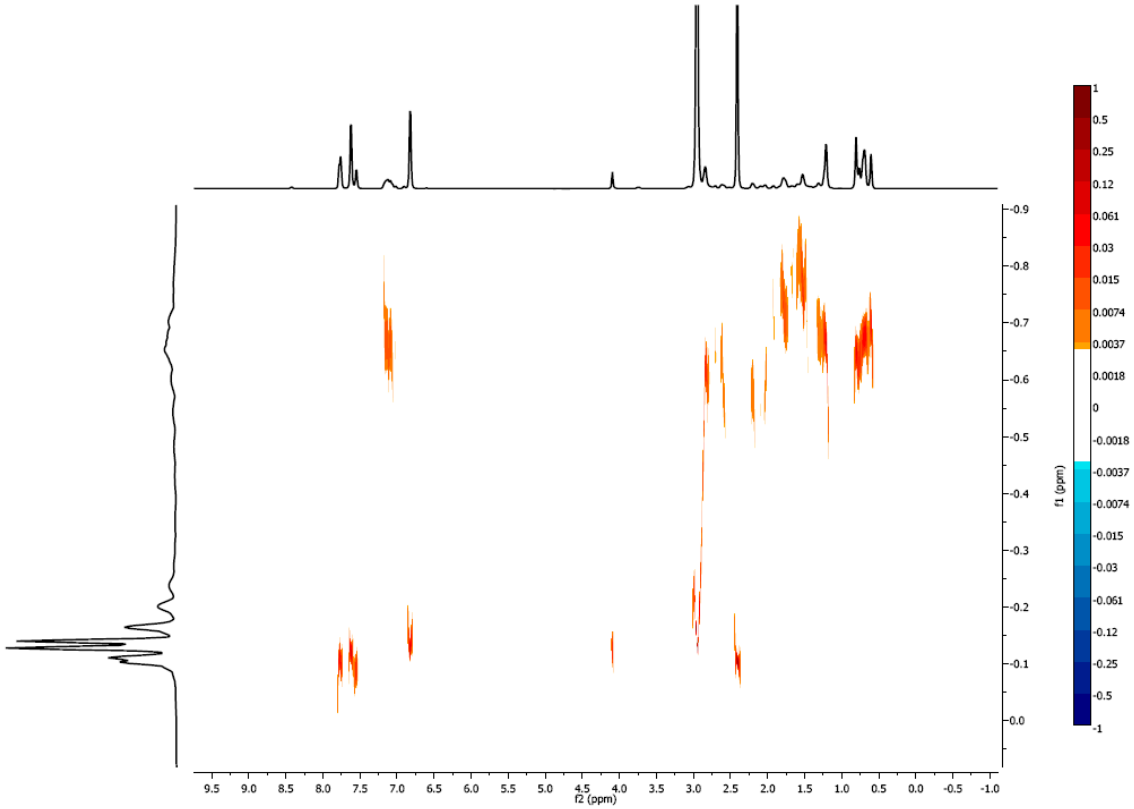


Figure 17: 2D DOSY spectrum of Thioflavin T, 600uM and beta amyloid 1-28, 100 uM, in D2O.

DOSY spectra gives us the difference in terms of coefficient of diffusion of the two components of this mixture, beta amyloid with higher values (0.7) and Thioflavin T lower (0.15). The dragging around 3 ppm indicates the overlap of the signals in a 1D spectrum. At this moment is still missing a 2D DOSY spectrum of the beta amyloid alone, as a reference, and severall DOSYs with different know molecular wheights to get a standard curve for the given molecular weights, and see what are the weights correspondent for this diffusion coefficients.

Transferred NOE's – Thioflavin T

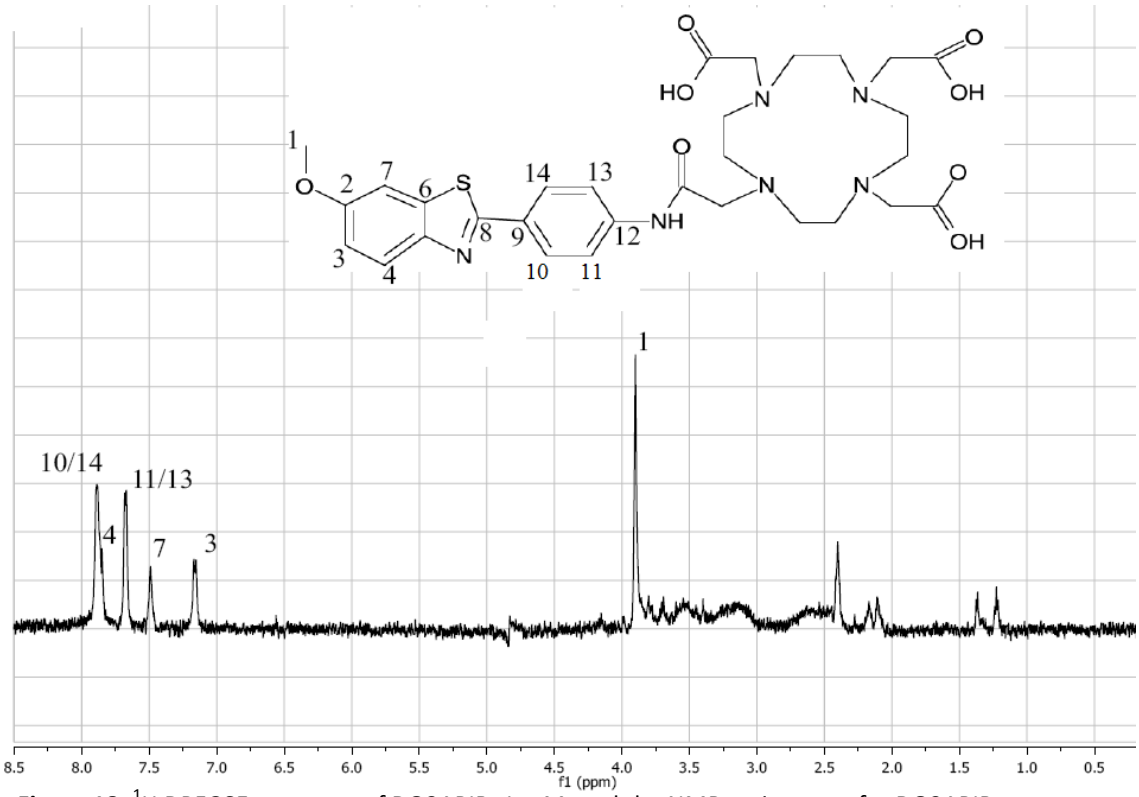


Figure 18: ^{1}H DPFGSE spectrum of DO3APIB, 1 mM, and the NMR assignment for DO3APIB resonances

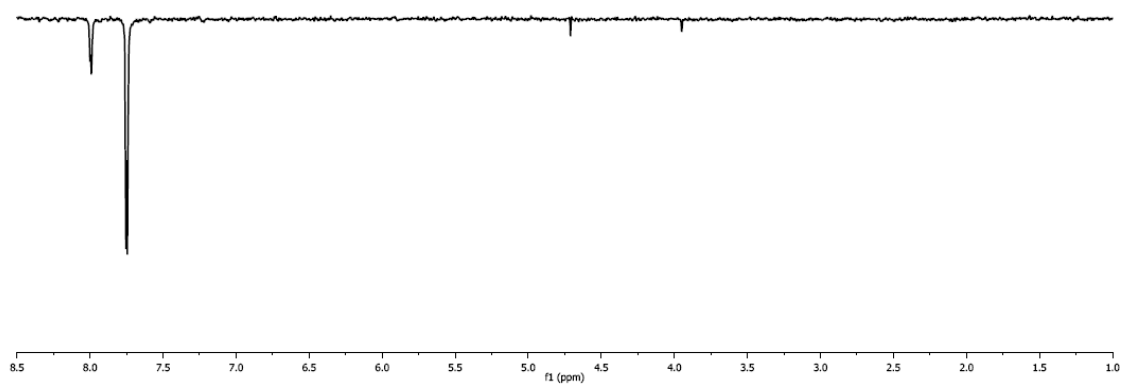


Figure 18: 1D NOESY spectrum of DO3APIB, 600 uM, in D2O with 750 ms of mixing time. Selective inversion of protons 11/13.

As expected protons 11/13 will present NOE effect on protons 10/14, but instead of assuming a positive after inversion of 11/13, no, the NOE is negative. For all the other inversions we also see this negative NOE as well. This might be because of the viscousity of the DO3APIB ligand.

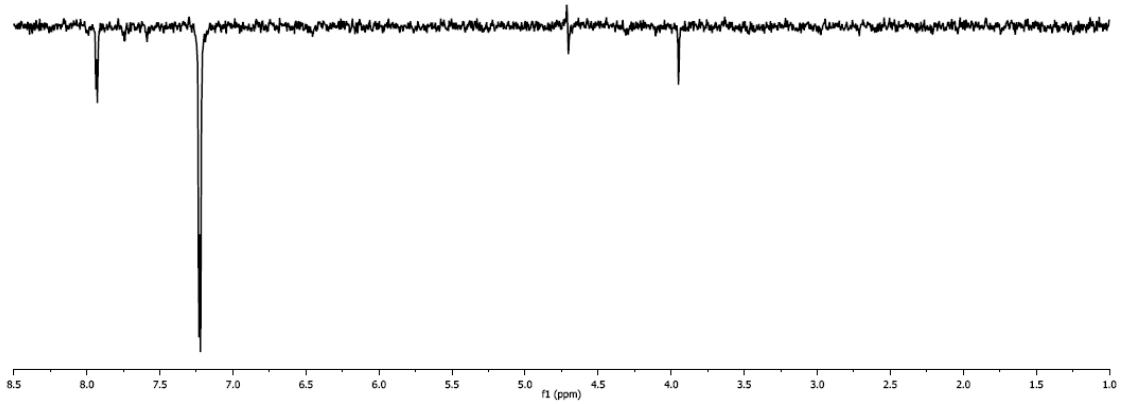


Figure 19: 1D NOESY spectrum of DO3APIB, 600 uM, in D2O with 750 ms of mixing time. Selective inversion of proton 3.

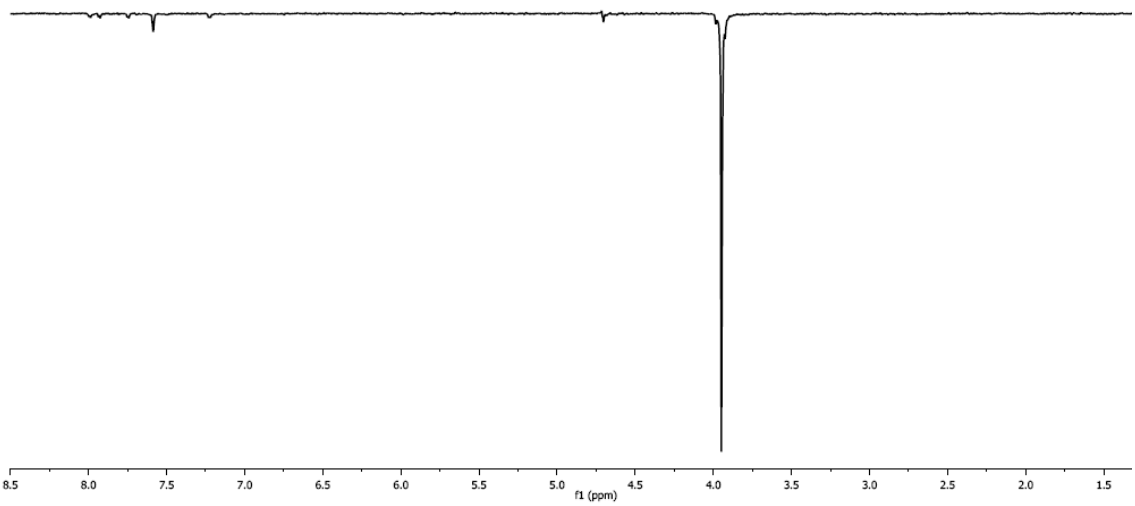


Figure 20: 1D NOESY spectrum of DO3APIB, 600 uM, in D2O with 750 ms of mixing time. Selective inversion of proton 1.

Both inversions present the correct NOE effect but with a negative signal instead of positive. So when we add the beta amyloid peptide the negative NOE remain and no conclusions may be taken from this results.

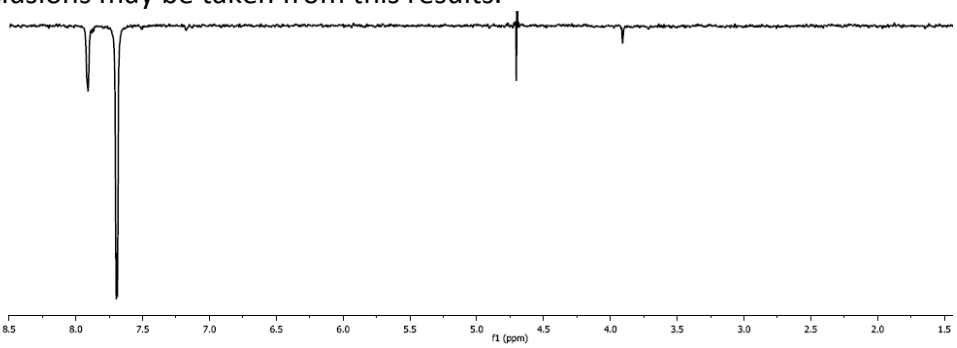


Figure 21: 1D NOESY spectrum of DO3APIB, 600 uM, and beta amyloid 1-28, 100 uM, in D2O with 250 ms of mixing time. Selective inversion of proton 11/13.

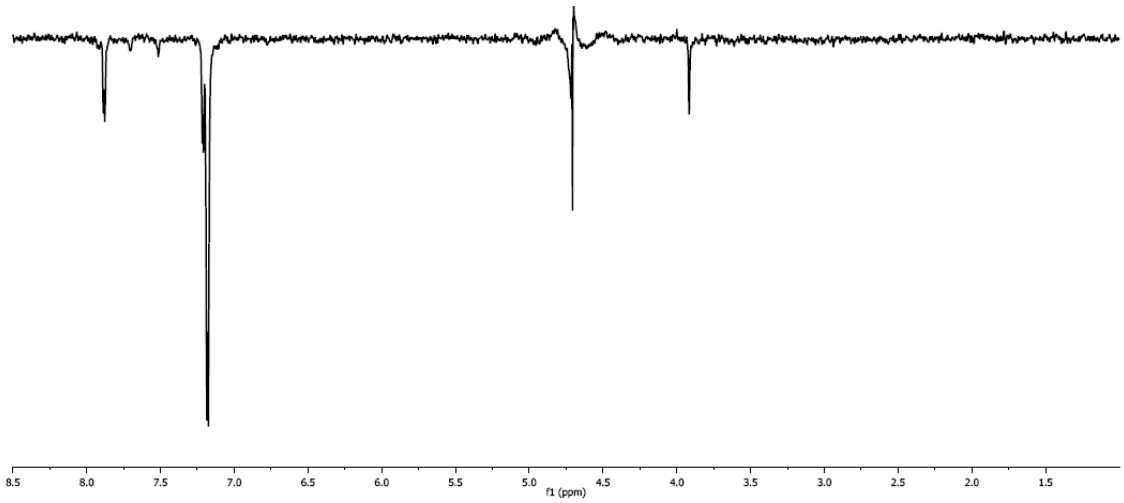


Figure 22: 1D NOESY spectrum of DO3APIB, 600 uM, and beta amyloid 1-28, 100 uM, in D2O with 250 ms of mixing time. Selective inversion of proton 3.

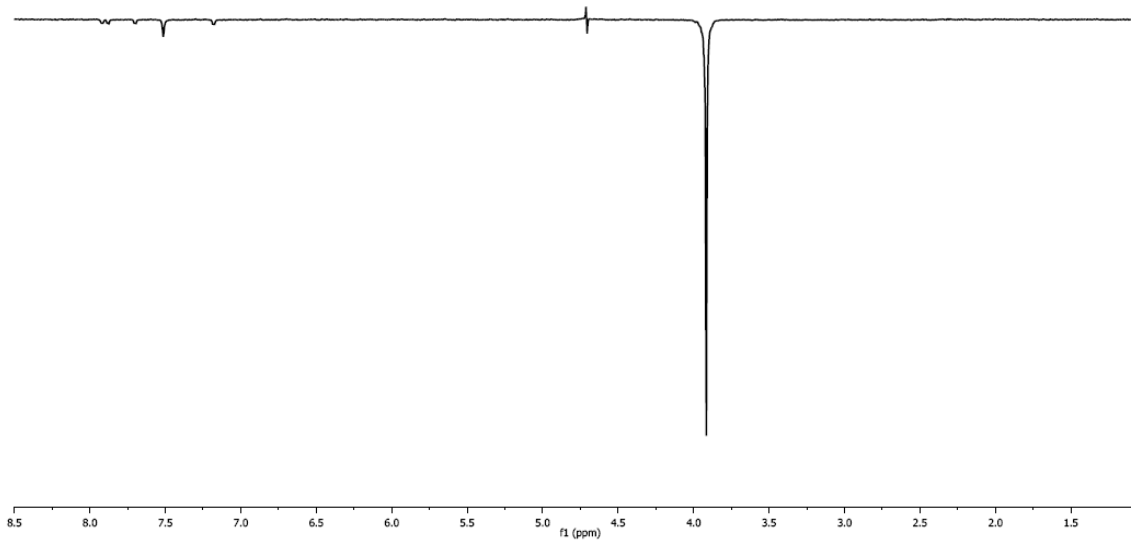


Figure 23: 1D NOESY spectrum of DO3APIB, 600 uM, and beta amyloid 1-28, 100 uM, in D2O with 250 ms of mixing time. Selective inversion of proton 1.

When we acquire the 2D NOESY spectrum for DO3APIB we can see positive NOEs. This is not in agreement with the 1D results obtained. And also a 2D NOESY spectrum for the mixture of DO3APIB and beta amyloid peptides need to be obtained.

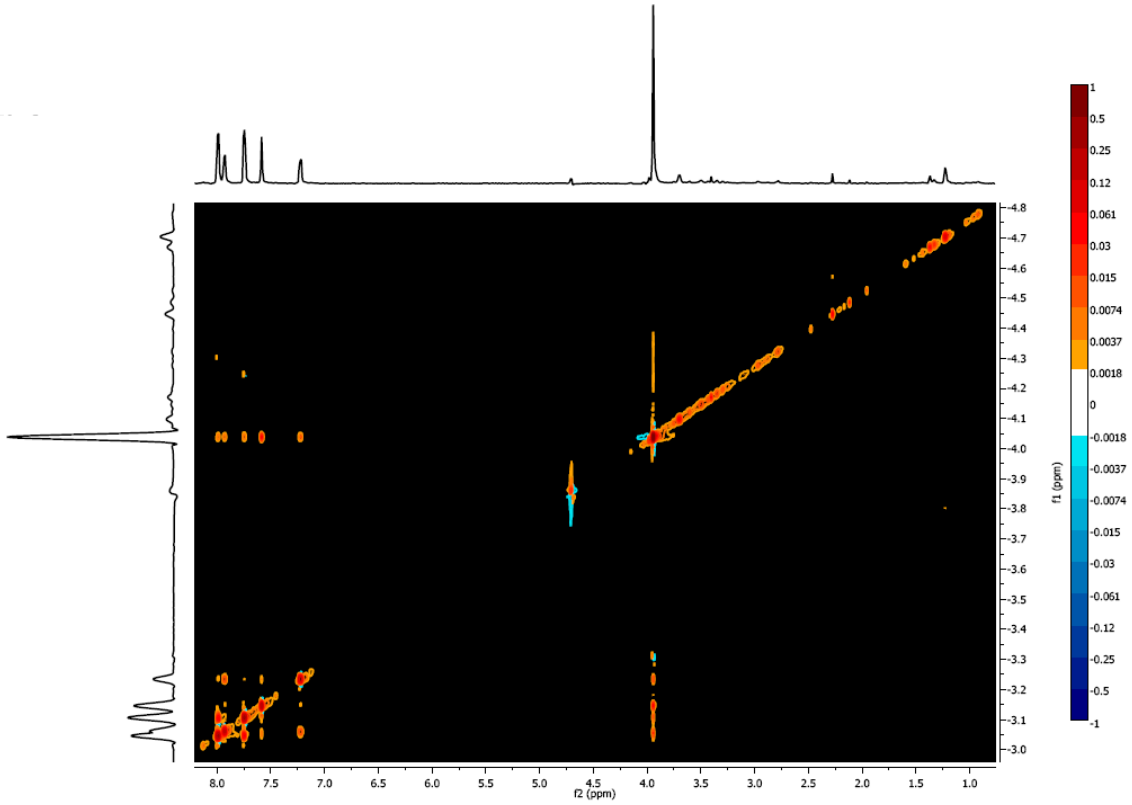


Figure 24: 2D NOESY spectrum of DO3APIB, 1000 uM, in D2O with 600 ms of mixing

Following the same thought for DOSY experiments we also performed 2D DOSY for DO3APIB. Same here for the standard curve to obtained molecular weights and also the need to determine viscousity.

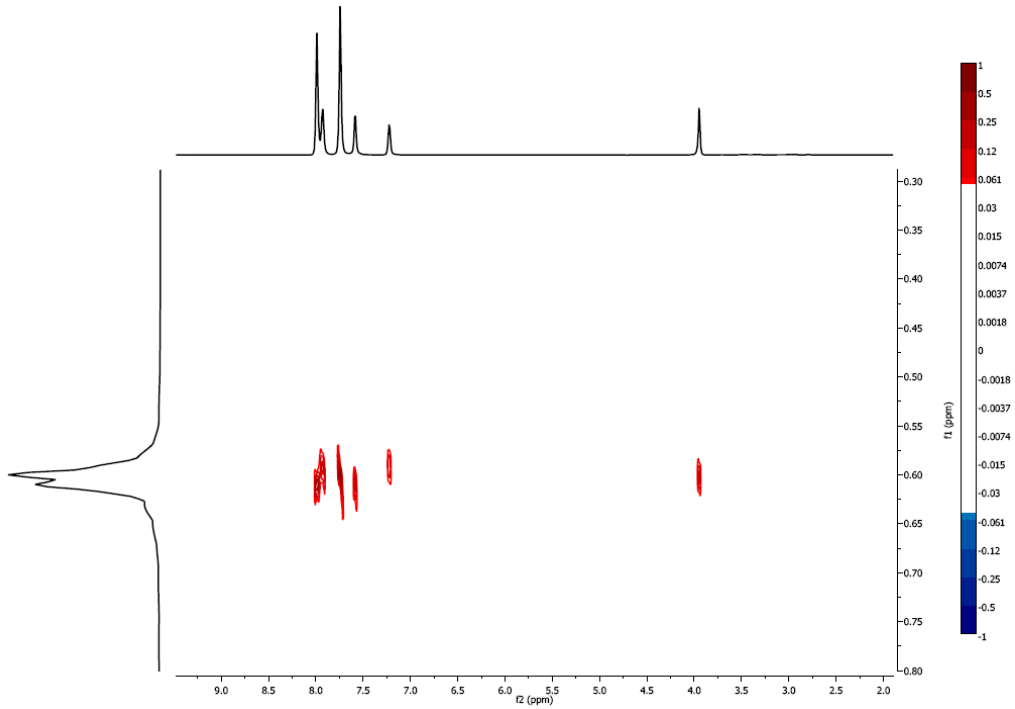


Figure 25: 2D DOSY spectrum of DO3APIB, 1000 uM.

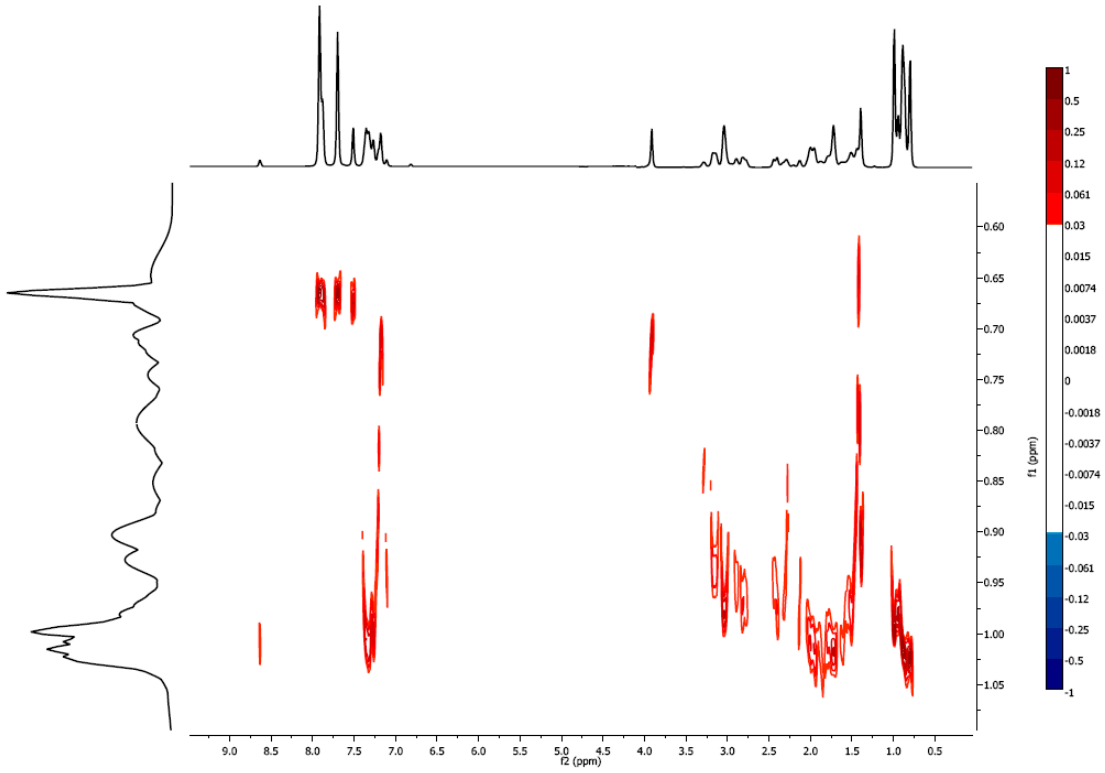


Figure 26: 2D DOSY spectrum of DO3APIB, 1000 uM, and beta amyloid 1-28, 100 uM, in D2

10.3 PUBLICATIONS

J Biol Inorg Chem (2011) 16:725–734
DOI 10.1007/s00775-011-0773-z

ORIGINAL PAPER

The interaction of La^{3+} complexes of DOTA/DTPA glycoconjugates with the RCA_{120} lectin: a saturation transfer difference NMR spectroscopic study

João M. C. Teixeira · David M. Dias · F. Javier Cañada · José A. Martins · João P. André · Jesús Jiménez-Barbero · Carlos F. G. C. Geraldes

Received: 28 December 2010 / Accepted: 19 March 2011 / Published online: 3 April 2011

Abstract The study of ligand–receptor interactions using high-resolution NMR techniques, namely the saturation transfer difference (STD), is presented for the recognition process between La(III) complexes of 1,4,7,10-tetrakis (carboxymethyl)-1,4,7,10-tetraazacyclododecane monoamide and diethylenetriaminepentaacetic acid bisamide glycoconjugates and the galactose-specific lectin *Ricinus communis* agglutinin (RCA_{120}). This new class of Gd(III)-based potential targeted MRI contrast agents (CAs), bearing one or two terminal sugar (galactosyl or lactosyl) moieties, has been designed for in vivo binding to the asialoglycoprotein receptor, which is specifically expressed at the surface of liver hepatocytes, with the aim of leading to a new possible diagnosis of liver diseases. The in vitro affinity constants for the affinity of the divalent La(III)–glycoconjugate complexes for RCA_{120}, used as a simple, water-soluble receptor model, were higher than those of the monovalent analogues. The combination of the experimental data obtained from the STD NMR experiments with molecular modelling protocols (Autodock 4.1) allowed us to predict the mode of binding of monovalent and divalent forms of these CAs to the galactose 1α binding sites of RCA_{120}. The atomic details of the molecular interactions allowed us to corroborate and supported the interaction of both sugar moieties and the linkers with the surface of the protein and, thus, their contribution to the observed interaction stabilities.

Keywords Ligand–receptor binding · Glycoconjugates · Saturation transfer difference NMR spectroscopy · MRI contrast agents · Protein–ligand interaction

J. M. C. Teixeira · D. M. Dias · C. F. G. C. Geraldes (✉)
Department of Life Sciences,
Faculty of Science and Technology,
Center of Neurosciences and Cell Biology,
University of Coimbra,
P.O. Box 3046,
3001-401 Coimbra, Portugal
e-mail: geraldes@bioq.uc.pt

F. J. Cañada · J. Jiménez-Barbero
Department of Chemical and Physical Biology,
CIB-CSIC, Ramiro de Maeztu 9,
28040 Madrid, Spain

J. A. Martins · J. P. André
Centro de Química,
Campus de Gualtar,
Universidade do Minho,
Braga, Portugal

Introduction

Molecular recognition events are of paramount importance in chemistry, biology and biomedicine. A large variety of techniques allow the elucidation of binding events between a ligand and its receptor. As key examples, enzyme-linked immunosorbent assay [1], immunoblotting, radioimmunoassay [2], affinity chromatography [3] and surface plasmon resonance experiments (Biacore) [4] are nowadays commonly employed for this task. In recent years, NMR-based techniques [5] have become increasingly popular when filling in the existing gap for characterization of molecular binding processes at high resolution. Transferred nuclear Overhauser effect spectroscopy [6], nuclear Overhauser effect pumping [7] and water–ligand observed via gradient spectroscopy (WaterLOGSY) [8, 9] are particular and powerful examples of such approaches. Among them, the saturation transfer difference (STD) NMR technique is probably one of the most popular and robust methods [5, 10–14]. This technique allows characterization of

ligand binding through intermolecular saturation transfer and, moreover, allows screening of ligand libraries [11], as well as calculation of affinity constants and mapping the binding epitope [5, 12–14]. In combination with ligand–protein docking studies, it may also help to derive a consistent 3D model of the intermolecular complex [15–21].

It is obvious that many diseases share a thin line with molecular recognition events and that targeting specific receptors is one of the approaches that may be employed to prevent, understand and control diseases. The development of magnetic resonance imaging (MRI) contrast agents (CAs) specifically targeted to different tissues has become a priority, and is a most profitable approach in this context. In particular, and within possible targets, the asialoglycoprotein receptor (ASGPR) is a lectin-type protein only found at the surface of hepatocytes and macrophages [22–25], having a determinant role in the targeting of exogenous compounds to the liver tissues, either for diagnosis or for therapy. On the basis of this knowledge, a new class of CAs has been developed recently with the intent for them to be selectively taken up by the hepatic ASGPR [26–28]. 1,4,7,10-Tetrakis(carboxymethyl)-1,4,7,10-tetraazacyclododecane (DOTA)-like chelators were attached to sugar moieties, galactosyl, glucosyl or lactosyl residues, by pendant arms containing aliphatic chains and amide bonds, resulting in monovalent or multivalent glycoconjugate derivative agents. After the development of the DOTA-based glycoconjugates [26], diethylenetriaminepentaacetic acid (DTPA) bisamide based glycoconjugates were also devised and studied [27]. In both types of CAs, DOTA- and DPTA-based chelates, the structural characteristics are similar: a central reporter group complexing a paramagnetic metal centre (Gd^{3+} for MRI, $^{153}Sm^{3+}$ for γ scintigraphy) with high kinetic and thermodynamic stability and long linear or branched arms with terminal sugar moieties as targeting groups (Fig. 1).

Carbohydrate–protein interactions are relatively weak binding processes. Nevertheless, affinity enhancement is achieved through multiple and simultaneous interactions of glycosides (multivalency) with their lectin receptors, a process known as the cluster glycoside effect [29–32]. In this way, higher valencies of the glycosides produce a synergistic effect in affinity constants when they bind to proteins (i.e. tetraglycosides > triglycosides > diglycosides > monoglycosides) [33, 34]. However, the way in which the sugar-based ligands interact with their lectin receptors in order to increase the binding affinity is still controversial and, most probably, strongly related with the particular ligand structure [31]. There are two main mechanisms by which the cluster glycoside effect may take place: *intramolecular* or *intermolecular* interactions. The intramolecular binding mode is characterized by the binding of multiple sugar moieties, within the same glycoside molecule, to multiple binding sites at the same lectin receptor. Therefore, this binding mode is also termed "chelate-type binding", as the glycoside simulates a

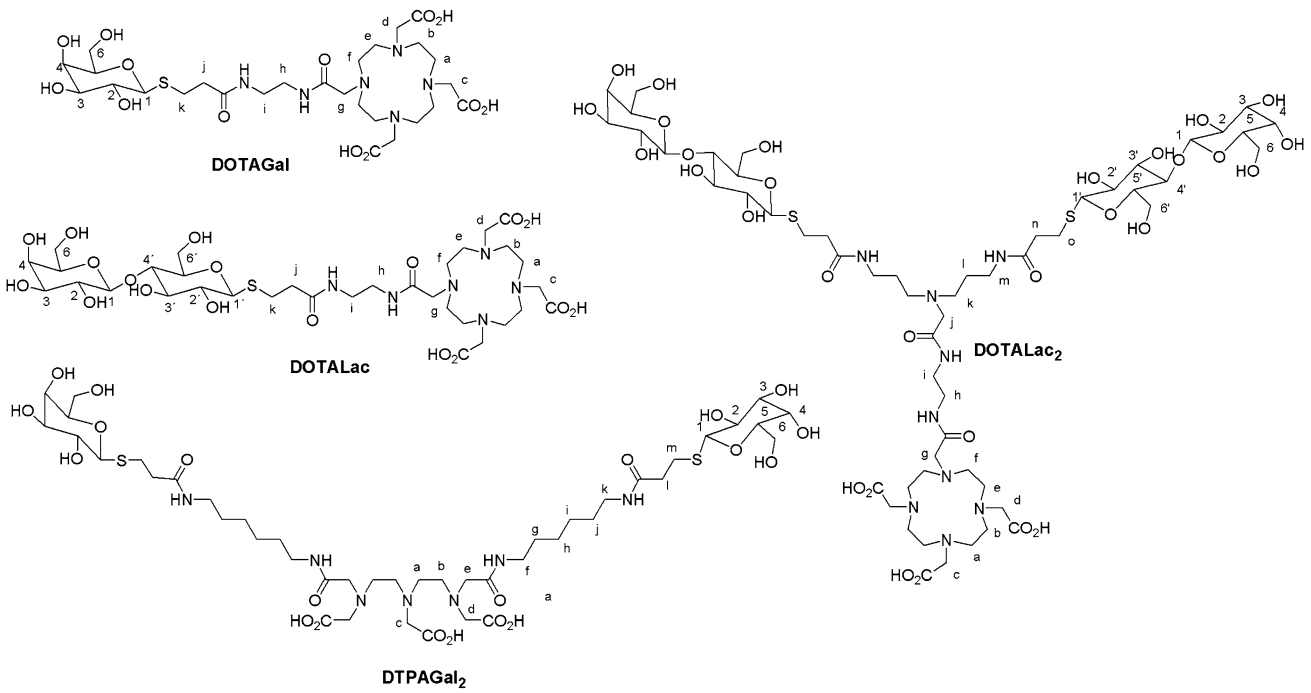


Fig. 1 Chemical structures and proton numbering scheme of the 1,4,7,10-tetrakis(carboxymethyl)-1,4,7,10-tetraazacyclododecane (DOTA) and diethylenetriaminepentaacetic acid (DTPA) glycoconjugates

chelate motif imprisoning the *metal* atom, in this case the lectin. Moreover, to make this type of interaction possible, the binding sites on the protein surface should be close enough to each other to allow simultaneous spanning of the interacting sites by the ligand binding moieties. At the same time, the ligand arms must be significantly long to reach the different binding sites. A conjugation between the protein and the ligand morphology must support this type of interaction. There are also other properties that favour the intramolecular binding mode, such as the presence of a hydrophobic linker, which may promote this binding mode by enhancing interactions between the linker and the protein surface [31, 35, 36]. On the other hand, under the same conditions, a hydrophilic linker could favour an intermolecular binding mode, which takes place when a single multivalent glycoside molecule binds to more than one protein molecule, a process that may finally lead to a precipitate.

In this context, we present the lectin binding features of this new class of DOTA- and DTPA-based glycoconjugate CAs and explore the affinity of the multiderivative glycoconjugate agents for a model lectin receptor. To accomplish this task, four different diamagnetic La(III) chelate analogues (diamagnetic chelates were needed in order not to quench the STD NMR effect) of the Gd(III) compounds of this class of CAs were studied, namely the monovalent La(DOTAGal) and La(DOTALac) and the divalent La(DOTALac$_2$) and La(DTPAGal$_2$) (Fig. 1). STD NMR was chosen to study the binding of the DOTA and DTPA glycoconjugates to a well-known lectin, *Ricinus communis* agglutinin (RCA$_{120}$), that was used as a model of the hepatic ASGPR. Although the carbohydrate recognition domain of the H1 subunit of ASGPR [37] would be a better model system, RCA$_{120}$ was used as a simple model of the membrane ASGPR in this proof-of-principle study of the method employed because it has galactose binding affinities in the same range as ASGPR [38], therefore being largely used for binding assays of galactose derivatives [39, 40]. RCA$_{120}$ water solubility also favours the in vitro NMR studies. RCA$_{120}$ is a dimeric lectin, consisting of two non-covalently bound ricin-like monomers. In turn, each ricin-like moiety is composed of two covalently linked heterochains, chain A and chain B. Chain A is responsible for the catalytic effect that gives this protein its toxic character, whereas chain B is the lectin domain, responsible for sugar affinity. Every B chain has, in principle, two sugar binding sites, dubbed 1α and 2γ [41–44]. However, the exact number of accessible binding sites in each B chain of RCA$_{120}$ was ultimately confirmed by calorimetric assays to be only one (1α), and its identity was revealed by site mutations [45–48].

We present a study, at the molecular level, of the mode of binding of these glycoconjugate derivatives compounds to RCA$_{120}$ by using a combination of STD NMR data and molecular modelling protocols, namely docking calculations.

Materials and methods

Samples

The DOTA and DTPA glycoconjugate derivatives and their La(III) complexes were synthesized and characterized as described previously [26, 27]. The La(III) complexes were dissolved as 99.9% D_2O/10% phosphate-buffered saline solutions. RCA$_{120}$ was isolated as previously described [49]. The protein was dissolved in 99.96% D_2O (purchased from Sigma-Aldrich) in the absence of buffer. The protein concentration ranged from 10 to 25 μM depending on the compound studied and the expected affinity for the protein, in order to achieve a large range of ligand excess. The concentrations of the ligands were selected to obtain most of the A_{STD} points at the beginning of the saturation curve, with ligand excess raging from 10 to 50, and a point of large ligand excess, over 200, was also obtained to define the "plateau" region of the curve. The concentrations of the protein and ligand were not constant for every compound, and were defined according to the desired ligand excess ratio and the quantities available.

NMR studies

All ^{1}H NMR spectra were acquired using a 5-mm pulse field gradient (PFG) triple resonance inverse probe using a Varian VNMRS 600 MHz NMR spectrometer working at 599.72 MHz. For each sample a 1D ^{1}H spectrum was obtained, and the spectral assignments from the literature [12, 26, 27] were used after they had been confirmed by 2D gradient correlation spectroscopy (gCOSY) spectra (data not shown). STD NMR spectra were then acquired, where the double PFG spin echo (DPFGSE) sequence [50] was used for water suppression. Since in our NMR system the STD NMR spectra are acquired directly from phase cycling, the 1D ^{1}H NMR spectra were used as off-resonance references in order to calculate the STD amplification factor [12]. All spectra were acquired using the same parameters: equal spectrometer gain value, spectral window of 8 kHz, number of scans varied between 128 and 256 for 1D ^{1}H spectra and between 1,024 and 2,048 for the STD NMR spectra, a previously calibrated spin-lock filter ($T_{1\rho}$) of 30 ms was used to remove protein resonances, the acquisition time was 1 s and the repetition time was 3.5 s. STD experiments were performed using a saturation delay of 2.5 s. To compare the reference spectra with the STD NMR spectra, the different number of acquisitions was normalized according to Eq. 1:

$$\text{Relative STD}(\%) = \frac{I_{\text{STD}} \times 2 \times \text{scans}_{\text{reference}}}{I_0 \times \text{scans}_{\text{STD}}}, \quad (1)$$

where I_{STD} is the peak intensity of the STD NMR spectra and I_0 is the intensity of the peaks in the ^{1}H reference spectra. Then, the peak intensities were normalized to the STD amplification factor (A_{STD}) (Eq. 2):

$$A_{STD} = \text{relative STD} \times \text{ligand excess}. \quad (2)$$

Binding studies

Affinity constant (K_{D}) estimation was performed by studying the build-up behaviour of the STD NMR spectra in conditions of constant protein concentration and increasing ligand concentration. The K_{D} values were estimated by fitting the plotted data points to a one-site binding model [13, 16] (Eq. 3):

$$A_{\text{STD}} = \frac{\alpha_{\text{STD}} \times [\text{L}]}{[\text{L}] + K_{\text{D}}}, \quad (3)$$

where α_{STD} is the maximum A_{STD} and [L] is the total ligand concentration. Plots and fits were obtained using GnuPlot version 4.2-3.

Docking calculations

Automated docking was performed using Autodock 4.1 [51] and the Lamarckian genetic algorithm [52] as a searching procedure. The Protein Data Bank (PDB) file corresponding to the protein used, RCA_{120}, was 1RZO. The protein model was kept rigid, and torsions were allowed only at the ligand level. Owing to the large number of torsions and the size of the ligands, only the sugar moiety and the adjacent arms were docked. The reporter groups (DOTA and DTPA) were removed and, in the case of divalent ligands, just one of the arms was considered. La(III) ligand chelates were designed in three dimensions using Maestro (Schrödinger) [53]. For glycoconjugate derivatives, grid maps were constructed using 50 × 50 × 50 points, with a grid box point spacing of 0.303 Å and centred some points below the binding site. The size of the initial random population was set differently for each compound, 150 for La(DOTAGal) and 50 for La(DTPAGal$_2$). The other parameters were set common for all runs, the maximum number of generations was 27,000, the elitism was 1, the probability that a gene would undergo a random change was 0.02 and the crossover probability was 0.80. Fifty docking runs were performed. The maximum number of generations was reached for these calculations and the total number of evaluations was kept around 1.1×10^7. For galactose calculations, grid maps were constructed using 40 × 40 × 40 points, with a grid box point spacing of 0.336 Å and centred at the ligand that was placed in the binding site in the PDB file by default. The size of the initial random population was 50 individuals, the maximum number of energy evaluations was 1.5×10^7, the maximum number of generations was 40,000, the elitism was 1, the probability that a gene would undergo a random change was 0.02 and the crossover probability was 0.80. Fifty docking runs were performed. The results were clustered using a root-mean-square deviation (RMSD) cutoff of 0.5 Å.

Results and discussion

Figure 2 represents both the 1D ^{1}H and the STD NMR spectra of the four La(III)-complexed glycoconjugates in the presence of RCA_{120}. Resonances from the ^{1}H spectra were assigned on the basis of data in previous publications [12, 26, 27] and on gCOSY analysis. The resonances are identified in Fig. 2, following the proton numbering schemes shown in Fig. 1. The sugar resonances are the main visible resonances in the STD NMR spectra, thus proving that these DOTA/DTPA branched glycoconjugate derivatives specifically interact with RCA_{120} through the sugar moieties. Owing to the nature of the STD experiment, it is possible to characterize the binding epitope of the ligand for a particular interaction. Table 1 summarizes the saturation profiles (relative STD) determined for the sugar and linker protons of the four compounds studied H protons refer to the protons from the galactosyl residue, and H′ protons refer to those of the glucosyl residues, in the case of lactosyl derivatives. The evaluation of the saturation profile of the six (12) groups of protons from the galactosyl (lactosyl) residues proves that the sugar protons that remain closer to the protein are always H3 and H2 of galactosyl residues, with H4 also experiencing great transfer of saturation [49]. On the other hand, protons H5 and H6/6′ seem to remain somehow further from the protein surface, since weaker STD effects were observed. The percentage of transferred saturation for protons H5/5′, H66^a/6′6$^{a\prime}$ and H3/3′ in the case of lactosyl derivatives was not considered because these peaks are superimposed on each other. The binding epitope revealed for these glycoconjugates is in agreement with the expected epitope on the basis of previous studies for this type of interaction [12, 29]. The anomeric protons were not considered in the group epitope mapping evaluations because the DPFGSE water suppression scheme dramatically affected its intensity. It is noteworthy that additional STD effects were observed for the linker protons (see Fig. 2), which will be discussed later.

After normalization of the acquired data points to the STD amplification factors (A_{STD}) [12], the plots of A_{STD} versus ligand concentration (μM) were drawn and fit to a one-site binding model. Figure 3 represents the two plots

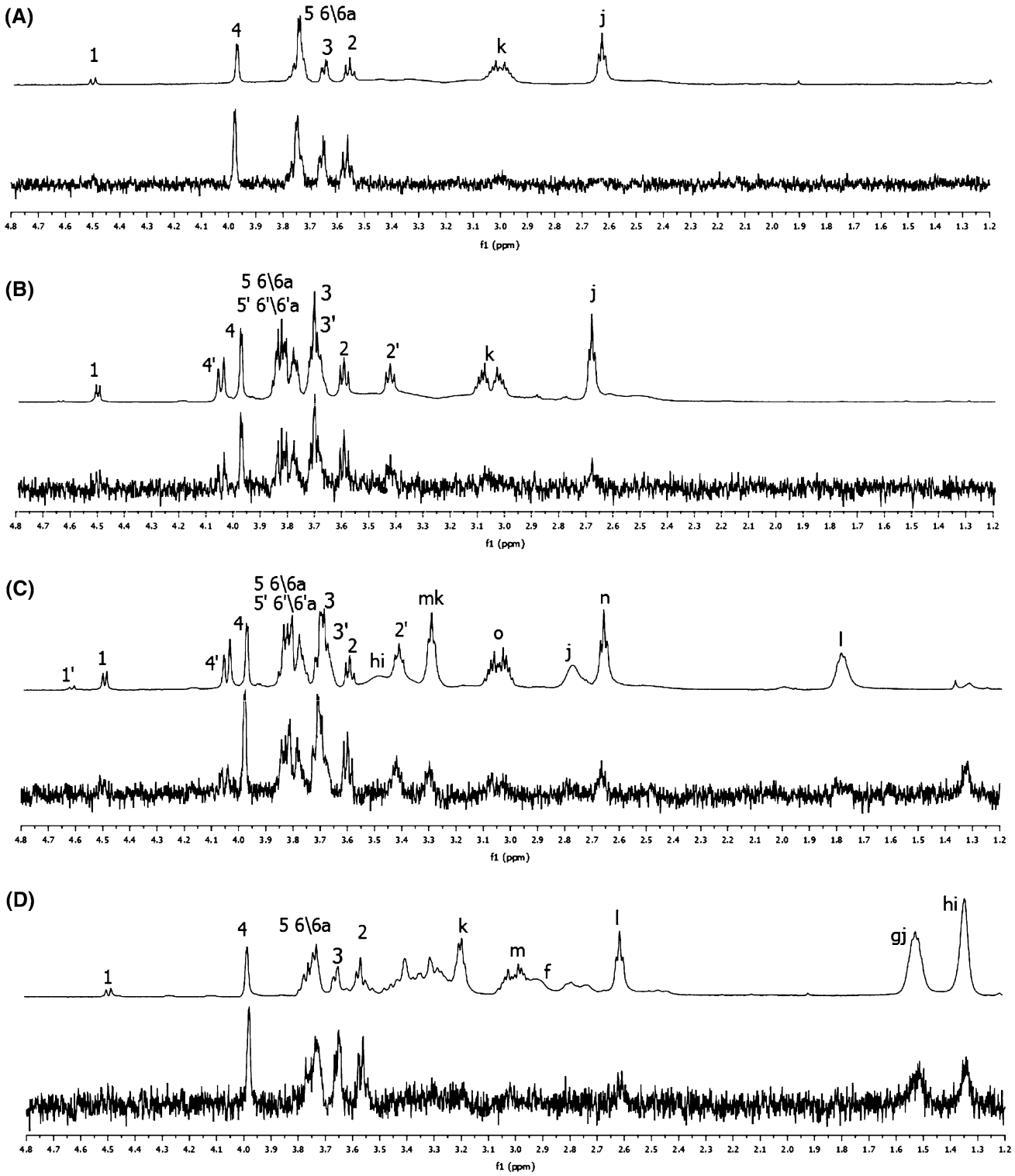


Fig. 2 One-dimensional ^{1}H NMR (*top*) and saturation transfer difference (STD) NMR (*bottom*) spectra of **a** 0.82 mM La(DOTA-Gal), **b** 6.2 mM La(DOTALac), **c** 1.1 mM La(DOTALac$_2$) and **d** 0.58 mM La(DTPAGal$_2$) in the presence of *Ricinus communis* agglutinin (RCA$_{120}$) with the following concentrations: **a** 25 μM, **b** 10 μM, **c** 15 μM and **d** 15 μM. Selective saturation (2.5 s) was performed at the aromatic region of the protein. The spin-lock pulse was calibrated to avoid unwanted protein resonances

Table 1 Summary of the characterized binding epitopes with saturation transfer difference (STD) values relative to those of H2 or H3 as a percentage

Protons	La(DOTAGal)	La(DPTAGal)$_2$	La(DOTALac)	La(DOTALac$_2$)
H2	94	95	100	100
H3	100	100	–	–
H4	93	64	68	94
H5 H6[a]/6[b]	64	65	–	–
H2′			72	80
H4′			37	34
k	11			
j	10			
hi		28		
gj		37		
l		29		
n				37
o				44

See Fig. 1 for the structures of the epitopes

from the galactosyl and lactosyl derivatives, each one representing both monovalent and divalent forms. Table 2 summarizes the values of the affinity constants estimated for the corresponding interactions with the lectin. To estimate the affinity constants for the different compounds, the data points obtained from the H3 protons of the galactosyl derivatives were used for the calculations since they showed the highest degree of saturation, thus allowing a more accurate K_D estimation [16, 17]. Also, because of degeneration of different proton resonances in the spectra of the lactosyl derivatives, the affinity constant for these compounds was estimated according to the data obtained from the galactosyl H4 proton, which remained fairly isolated. However, taking into account the recently published detailed analysis of the factors affecting the determination of ligand–receptor dissociation constants by STD NMR titration experiments [17], one has to consider the K_D values presented here as approximations of the real values. The calculated K_D values (Table 2) reflect an increase in binding affinity for both divalent compounds relative to the monovalent ones, also evidenced by the lower A_{STD} values output by the fitting curves, with a substantial decrease of the dissociation rate (k_{off}) with respect to that of the monovalent analogues, as it is rather unlikely that, when completely bound, both sugar residues of the divalent compounds simultaneously dissociate from the two binding sites [5].

The stronger binding of the divalent compounds to the lectin clearly observed in the present experiments is in agreement with findings of previous in vivo studies, where the higher affinity of the multivalent forms of glycoconjugate derivatives of this type showed a more rapid incorporation by the liver than the monovalent forms [54].

The STD NMR spectra provided additional information on the slightly different interaction features for the monovalent and divalent molecules with RCA$_{120}$. For the monovalent compounds, La(DOTAGal) and La(DOTALac) (Fig. 2a, b), the resonances from protons j and k in the linker arms are barely visible. On the other hand, several proton resonances are clearly visible in the STD NMR spectra of the divalent compounds, and it is possible to measure the saturation profiles of resonances n and o for La(DOTALac$_2$) and resonances l, gj and hi for La(DTPAGal$_2$). The calculated STD values for protons n and o of La(DOTALac$_2$) were 37 and 44%, respectively, normalized to H2. In the case of La(DTPAGal$_2$) resonance l was measured to have 29% saturation and resonances gj and hi had 37 and 28%, respectively, normalized to H3 (Table 1). Although we have to consider that the gj and hi signal contribution comes from eight protons, we cannot ignore that the linkers of these divalent compounds interact with the lectin surface.

A hydrophobic linker is more prone to establish interactions with a lectin protein surface than a hydrophilic one [55]. In fact, the La(DTPAGal$_2$) linker is longer and more hydrophobic, when compared with the linker (protons g–j) of La(DOTALac$_2$). It also displays more torsional degrees of freedom in solution, thus facilitating the interaction with the surface of the protein. That might explain why we do not observe interaction of protons g–j of La(DOTALac$_2$) with the protein.

The presence of hydrophobic interactions between a given ligand and the neighbouring regions of the sugar binding site in the lectin surface has already been reported [31, 35]. Although very frequently binding features of this type have been associated with intramolecular, or chelate-type, binding modes [55, 56], the 3D structure of the RCA$_{120}$ protein does not allow such a type of binding. The X-ray diffraction structure available for RCA$_{120}$ (PDB file 1RZO) shows a dimer of AB ricin heterodimers in the crystal. In fact, the distance between the two galactose binding sites within one heterodimer is 36 Å, too far for the

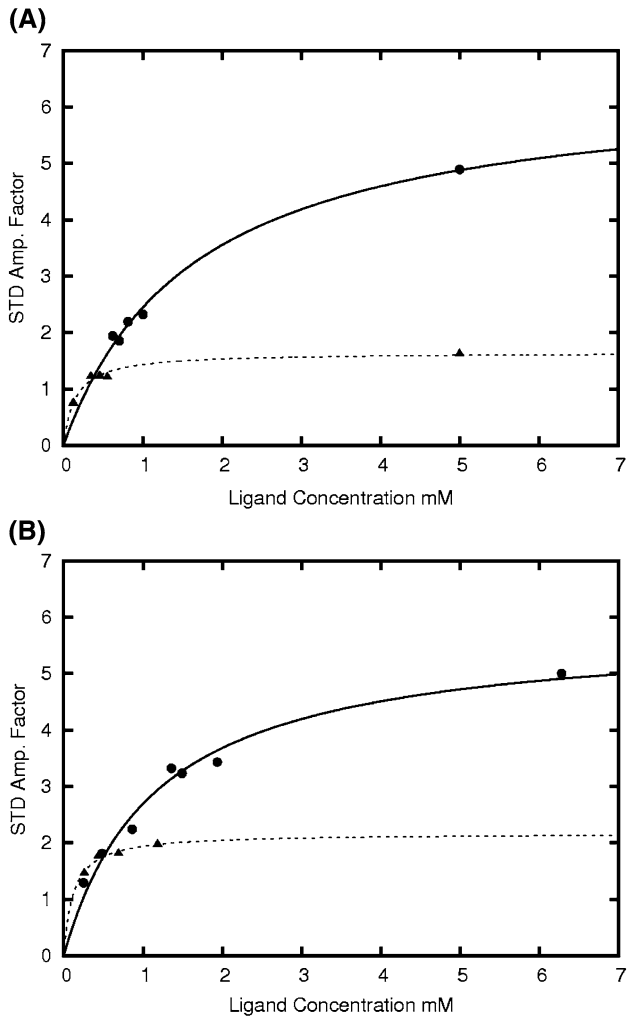


Fig. 3 Direct determination of the K_D values of La(III) complexes of **a** DOTAGal (*circles*) and $DTPAGal_2$ (*triangles*) and **b** DOTALac (*circles*) and $DOTALac_2$ (*triangles*) binding to RCA_{120} by fitting the acquired data points to Eq. 2

Table 2 Individual K_D values for the protons of the glycoconjugate compounds obtained from the STD NMR experiments

Protons	Individual K_D value (mM)	Maximum STD amplification factor
(La)DOTAGal–H3	1.66	6.5
(La)$DTPAGal_2$–H3	0.15	1.6
(La)DOTALac–H4	1.16	5.8
(La)$DOTALac_2$–H4	0.12	2.2

two arms of the divalent ligands to span them and account for a possible simultaneous intramolecular effect. Moreover, the distance between the two closest galactose binding sites, each one on each B-chain subunit of the two dimers, is even larger, more than 50 Å. Thus, the tenfold increased affinity of the divalent compounds for the lectin relative to the monovalent compounds cannot be explained by an intramolecular mechanism of a cluster glycoside effect. In the case of an intermolecular mechanism, lectin–lectin interactions and finally precipitation should occur to produce the increased affinity, which was not observed in the present case. Therefore, with the available data, the statistical effect of the multiple carbohydrate epitopes present together with the interaction of the linkers with the protein surface (see later) is considered to be responsible for the observed increased affinity.

The empirical results of the STD experiments were substantiated with a 3D model of the complex by using molecular modelling calculations based on Autodock 4.1 [51]. Docking calculations were then performed for La($DTPAGal_2$) with RCA_{120} binding site 1α, considering only one of the sugar arms, following the protocol described in "Docking calculations". The highest ranking cluster encompassed eight possible binding conformations, with the output geometries clustered using a RMSD of 2 Å. Three of the eight calculated conformers were selected according to the orientation of the sugar moiety inside the binding pocket, and considering they keep conformity with the STD NMR results and with the chemical nature of the molecule itself. Indeed, the binding mode obtained was completely in agreement with that obtained for an isolated galactosyl moiety, thus validating the orientation of the saccharide residue of La($DTPAGal_2$) within the binding site. Figure 4a shows one representative structure of the selected cluster, and Fig. 4b represents three superimposed structures from the above-mentioned calculations. The galactosyl residues from the different runs, including that for a single galactose moiety, are oriented in a similar manner, although they are not perfectly superimposed. Nonetheless, all the intermolecular hydrogen bonds that occur for galactose binding also occur for the different solutions for this glycoconjugate. With respect to the interactions of the linker with the protein surface, the models obtained set the long hydrophobic linker of the La($DTPAGal_2$) chelate close to a hydrophobic region of the protein surface, and thus it interacts with the side chains of different amino acids. Owing to the size of the docked ligand, the structures obtained can be considered as a good approximation of the interaction mode, which cannot be seen within the concept of a rigid, static representation. Very probably, different orientations of the linker may be adopted for it to properly interact with the lectin, as suggested by the docking calculations.

Docking studies of the La(DOTAGal) single-arm molecule with RCA_{120} binding site 1α were also performed, as described in "Docking calculations". Only the structures which fit the STD data were selected for further analysis (Fig. 5). In this case, the STD data suggested very weak interactions between the linker of these monovalent derivatives and the protein surface. Again, the same region

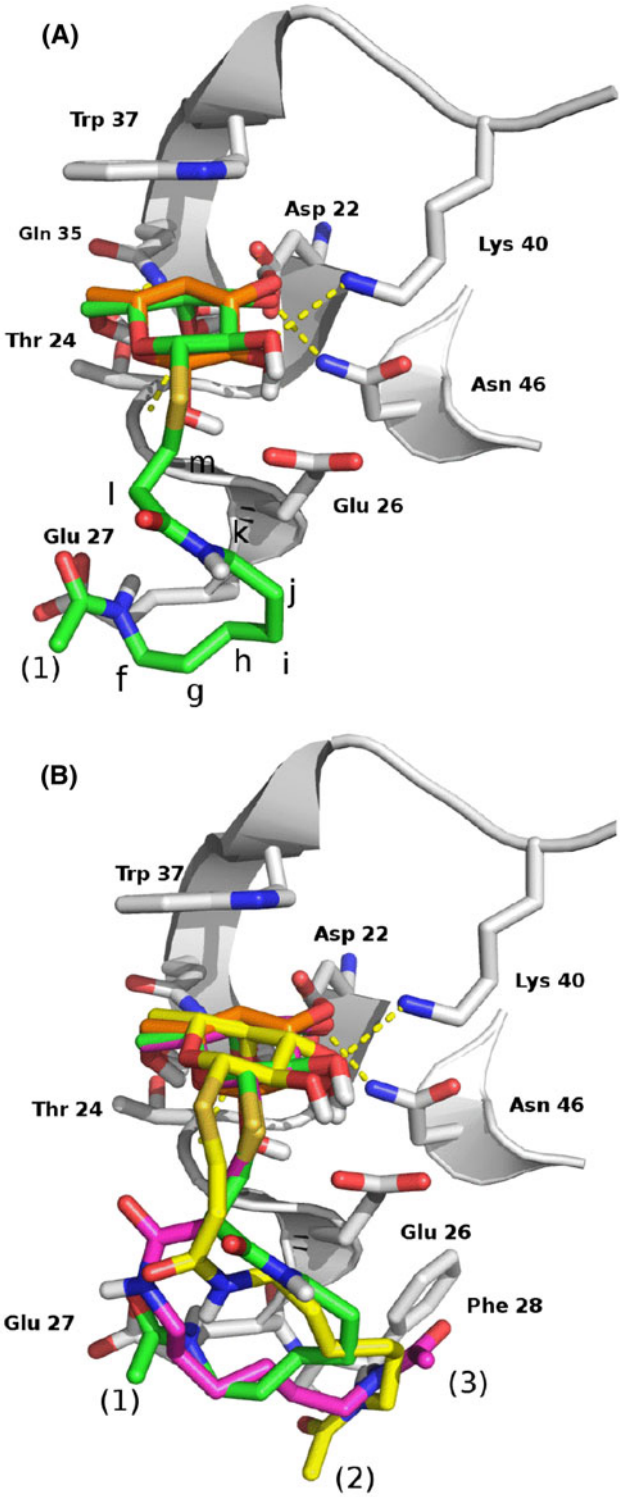


Fig. 4 Automated docking structures of one arm of La(DTPAGal$_2$) in binding site 1α of RCA$_{120}$. **a** One of resulting runs is shown in *green* and **b** the three most reliable runs. The single docked galactose molecule was also superimposed to allow a better comparison, and it is displayed in *orange*. Marked in *yellow* are the hydrogen bonds considered between the ligand and the protein, involving OH2 and Lys40, OH3 and Asn46, OH4 and Gly25, and OH6 and Gln35. CH–π stacking interactions with Trp37 also occur [52]

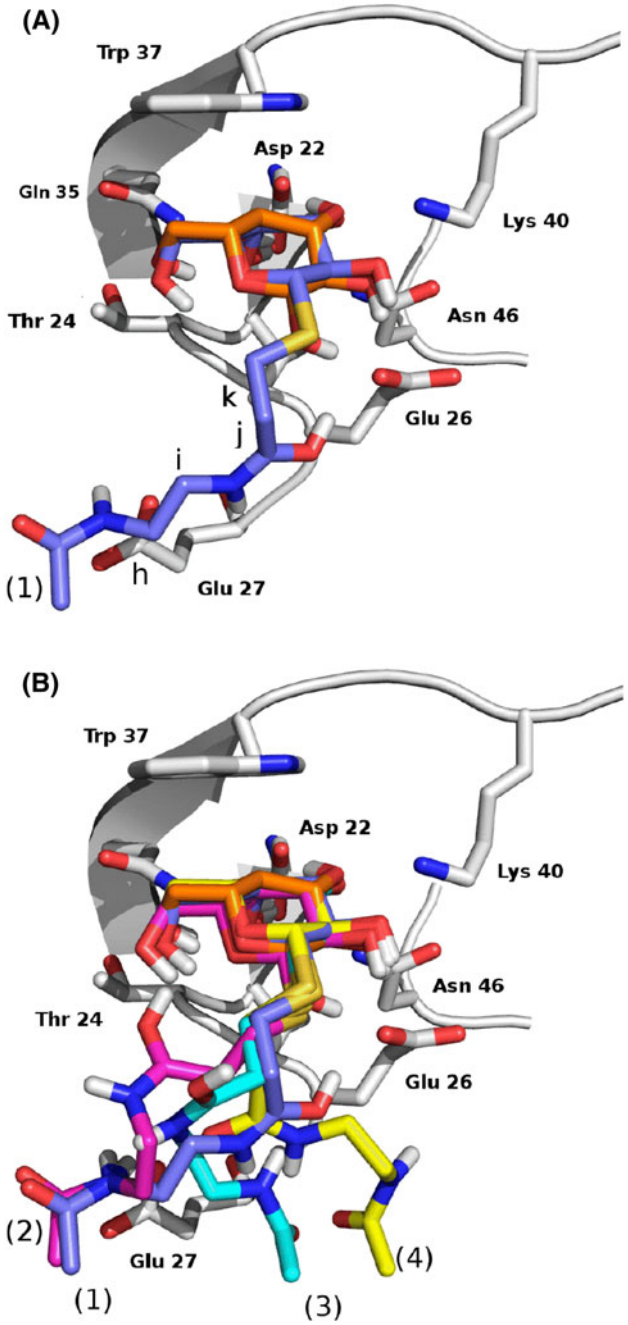


Fig. 5 Docking results of the sugar-linker moiety of La(DOTAGal). **a** One of the docked results and **b** the superimposed view of all the selected possible structures. The number of torsional bonds was 13

of the protein was targeted by the linker, as for La(DTPAGal$_2$). It can be observed that the charged regions of the linker were placed near the charged atoms of the surface amino acids, which could stabilize the conformation by polar interactions. H–H distances between the linkers of both docking results for La(DTPAGal2) and La(DOTAGal) and the surface of RCA$_{120}$ binding site 1α

Table 3 H–H linker–protein distance of docked arms of La(DOTAGal) and La(DTPAGal$_2$)

La(DOTAGal)		La(DTPAGal$_2$)	
Protons	Distance (Å)	Protons	Distance (Å)
k	3.2	m	2.8
j	4.1	l	3
i	5.2	hi	2.9
h	6	gj	3.1

Distances were measured relative to Cα G25, Cβ E26 and Cβ E27 and are presented as average values

were measured using Autodock 4 (Table 3). On average, protons f–k of the hydrophobic patch of the La(DTPAGal$_2$) linker are closer to the protein surface protons (calculated average distance of 2.9 Å) than protons i–h of the polar linker of La(DOTAGal) (calculated average distance of 5.8 Å). In fact, as one moves from the sugar moiety of the molecule, the polar linker of La(DOTAGal) tends to move away from the protein surface, whereas the hydrophobic parts of the La(DTPAGal$_2$) linker far from the sugar moiety stay quite close to the protein surface.

Finally, Table 4 shows the output values for the calculated binding structures of La(DOTAGal) and La(DTPAGal$_2$). The intermolecular energy is lower in La(DOTAGal) and La(DTPAGal$_2$) runs when compared with the single galactose molecule. As expected from the higher number of torsions, the average internal energy upon binding is lower in the La(DTPAGal$_2$) runs than in the La(DOTAGal) runs. The values should be regarded as merely qualitative, given the simplification of the model employed.The affinity of galactose for RCA$_{120}$ is 2.2×10^3 M^{-1} [46], a value which lies between the calculated values for the monovalent and divalent compounds, in qualitative agreement with the estimated free energies obtained by docking calculations.

Conclusion

An STD NMR analysis has shown that the divalent La(DTPAGal$_2$) and La(DOTALac$_2$) glycoconjugate derivatives have higher affinity for the RCA$_{120}$ lectin than their monovalent La(DOTAGal) and La(DOTALac) analogues. This effect is therefore concordant with the results observed in in vivo binding studies with hepatocyte cells and the corresponding $^{153}Sm^{3+}$ chelates [54]. The so-called cluster glycoside effect may be invoked to explain the observations. Our studies have tried to clarify the binding mode of this new class of potential liver imaging agents, using the RCA$_{120}$ lectin as a simple model receptor, in order to provide new insights into the development of lead compounds and optimization of those already developed. The STD NMR data, assisted by docking calculations, suggest the existence of interactions between the linkers of the divalent compounds and the protein surface.

The structural features of RCA$_{120}$ and the glycoconjugate imaging agents used in this work preclude the existence of an intramolecular binding process. An intermolecular type of binding cannot be considered, as it would imply protein clustering and precipitation, which did not occur in the experimental conditions used. Taking into consideration the STD NMR data and the docking results obtained, we can conclude that the main interaction between these ligands and the lectin protein occurs through the sugar residues, through a combination of hydrogen bonds, van der Waals forces and CH–π stacking interactions [57, 58], but the hydrophobic linker arms also interact with the protein surface, especially for the divalent agents. These interactions, together with a statistical effect of the presence of multiple carbohydrate epitopes, are considered to be responsible for the increased affinity of the divalent compounds for the lectin. We believe that the approach to study CA–target protein interactions combining NMR and modelling tools, proposed and exemplified in this work for

Table 4 Calculated energies for the "single arm" of La(DOTAGal) and La(DTPAGal$_2$) (kcal mol^{-1})

Run La(DOTAGal)/ La(DTPAGal$_2$)	Intermolecular energy	Internal energy	Torsional free energy	Unbound system energy	Estimated free energy
Run 1	−8.62	−1.55	+3.57	−0.46	−6.14
Run 2	−7.01	−2.48	+3.57	−0.46	−5.47
Run 3	−8.16	−1.45	+3.57	−0.46	−5.58
Run 4	−8.73	−1.56	+3.57	−0.46	−6.25
Run 1	−8.60	−1.82	+4.39	−0.60	−5.43
Run 2	−9.13	−1.62	+4.39	−0.60	−5.76
Run 3	−7.86	−1.78	+4.39	−0.60	−4.65
Galactose	−6.69	−1.49	1.65	−0.37	−6.16

The first set of runs 1–4 and second set of runs 1–3 refer to La(DOTAGal) and La(DTPAGal$_2$), respectively

the first time, can be very useful in the design of novel targeted MRI CAs. In particular, novel design and production of high-affinity glycoconjugates, such as ones investigated here, to interact with lectins should focus on the optimization of the linker arms as a protein binding complement to the sugar residues, regarding their length, flexibility and chemical nature. Such an approach will aim at increasing entropic and enthalpic savings that derive from the linker. Good knowledge of the structure of the target receptor is also of extreme importance in order to design specific and protein-directed ligands.

Acknowledgments This work was supported by the Fundação para a Ciência e a Tecnologia (FCT), Portugal (project PTDC/QUI/70063/2006) and FEDER. The Varian VNMRS 600 MHz NMR spectrometer in Coimbra was acquired with the support of the Programa Nacional de Reequipamento Científico of FCT, Portugal, contract REDE/1517/RMN/2005—as part of Rede Nacional de RMN (RNRMN). This work was carried out in the framework of the COST D38 Action. The group in Madrid thanks the Ministery of Science and Innovation of Spain for financial support (grant CTQ2009-08536). We also thank Eurico Cabrita for useful discussions.

References

1. Engvall E, Perlmann P (1971) Immunochemistry 8:871–874
2. Yalow BRS, Berson SA (1960) J Clin Invest 39:11–13
3. Cuatrecasas P, Wilchek M, Anfinsen CB (1968) Biochemistry 61:636–643
4. Homola J, Yee SS, Gauglitz G (1999) Sens Actuators B Chem 54:3–15
5. Meyer B, Peters T (2003) Angew Chem Int Ed 42:864–890
6. Ni F (1994) Progr Nucl Magn Reson Spectrosc 26:517–606
7. Chen A, Shapiro MJ (1998) J Am Chem Soc 120:10258–10259
8. Dalvit C, Pevarello P, Tatò M, Veronesi M, Vulpetti A, Sundström M (2000) J Biomol NMR 18:65–68
9. Dalvit C, Fogliatto G, Stewart A, Veronesi M, Stockman B (2001) J Biomol NMR 21:349–359
10. Mayer M, Meyer B (1999) Angew Chem Int Ed 38:1784–1788
11. Vogtherr M, Peters T (2000) J Am Chem Soc 122:6093–6099
12. Mayer M, Meyer B (2001) J Am Chem Soc 123:6108–6117
13. Lepre C, Moore JM, Peng JW (2004) Chem Rev 104:3641–3676
14. Fielding L (2007) Progr Nucl Magn Reson Spectrosc 51:219–242
15. Meinecke R, Meyer B (2001) J Med Chem 44:3059–3065
16. Neffe AT, Bilang M, Grüneberg I, Meyer B (2007) J Med Chem 50:3482–3488
17. Angulo J, Enríquez-Navas PM, Nieto PM (2010) Chem Eur J 16:7803–7812
18. Hajduk PJ, Mack JC, Olejniczak ET, Park C, Dandliker PJ, Beutel B (2004) J Am Chem Soc 126:2390–2398
19. Klein J, Meinecke R, Mayer M, Meyer B (1999) J Am Chem Soc 121:5336–5337
20. Kolympadi M, Fontanella M, Venturi C, André S, Gabius H-J, Jiménez-Barbero J, Vogel P (2009) Chem Eur J 15:2861–2873
21. Jiménez-Barbero J, Dragoni E, Venturi C, Nannucci F, Ardá A, Fontanella M, André S, Cañada FJ, Gabius H-J, Nativi C (2009) Chem Eur J 15:10423–10431
22. Leffler H, Carlsson S, Hedlund M, Qian Y, Poirier F (2004) Glycoconj J 19:433–440
23. Almkvist J, Karlsson A (2004) Glycoconj J 19:575–581
24. Critchley P, Willand MN, Rullay AK, Crout DH (2003) Org Biomol Chem 1:928–938
25. Ashwell G, Harford J (1982) Ann Rev Biochem 51:531–544
26. André JP, Geraldes CFGC, Martins JA, Merbach AE, Prata MIM, Santos AC, de Lima JJP, Tóth E (2004) Chem Eur J 10:5804–5816
27. Baía P, André JP, Geraldes CFGC, Martins JA, Merbach AE, Tóth E (2005) Eur J Inorg Chem 2110–2119
28. Fulton DA, Elemento EM, Aime S, Chaabane L, Botta M, Parker D (2006) Chem Commun 1064–1066
29. Lee YC (1992) FASEB J 6:3193–3200
30. Lee RT, Lee YC (2001) Glycoconj J 17:543–551
31. Lundquist J, Toone EJ (2002) Chem Rev 102:555–578
32. Lee YC, Lee RT (1995) Acc Chem Res 28:321–327
33. Lee YC, Townsend RR, Hardy MR, Lönngren J, Arnarp J, Haraldsson M, Lönn H (1983) J Biol Chem 258:199–202
34. Biessen EA, Broxterman H, van Boom JH, van Berkel TJ (1995) J Med Chem 38:1846–1852
35. Lundquist JJ, Debenham SD, Toone EJ (2000) J Org Chem 65:8245–8250
36. Corbell J (2000) Tetrahedron Asymmetry 11:95–111
37. Meier M, Bider MD, Malashkevich VN, Spiess M, Burkhard P (2000) J Mol Biol 300:857–865
38. Nahálková J, Svitel J, Gemeiner P, Danielsson P, Pribulová B, Petrus L (2002) J Biochem Biophys Methods 52:11–18
39. D'Agata R, Grasso G, Iacono G, Spoto G, Vecchio G (2006) Org Biomol Chem 4:610–612
40. Lee M, Park S, Shin I (2006) Bioorg Med Chem Lett 16:5132–5135
41. Nicolson GL, Blaustein J, Etzler ME (1974) Biochemistry 13:196–204
42. Olsnes S, Saltvedt E, Pihl A (1974) J Biol Chem 249:803–810
43. Endo Y, Tsurugi K (1987) J Biol Chem 262:8128–8130
44. Lamb FI, Roberts LM, Lord JM (1985) Eur J Biochem 148:265–270
45. Houston LL, Dooley TP (1982) J Biol Chem 257:4147–4151
46. Sharma S, Bharadwaj S, Surolia A, Podder SK (1998) Biochem J 333(3):539–542
47. Dam TK, Brewer CF (2002) Chem Rev 102:387–429
48. Sphyris N, Lord JM, Wales R, Roberts LM (1995) J Biol Chem 270:20292–20297
49. Rivera-Sagredo A, Jiménez-Barbero J, Martín-Lomas M, Solís D, Díaz-Mauriño T (1992) Carbohydr Res 232:207–226
50. Hwang T-L, Shaka AJ (1995) J Magn Reson A 112:275–279
51. Goodsell DS, Morris GM, Olson AJ (1996) J Mol Recognit 9:1–5
52. Morris GM, Goodsell DS, Halliday RS, Huey R, Hart WE, Belew RK, Olson AJ (1998) J Comput Chem 19:1639–1662
53. Schrödinger (2008) Maestro, version 8.5. Schrödinger, New York
54. Prata MIM, Santos AC, Torres S, André JP, Martins JA, Neves M, García-Martín ML, Rodrigues TB, López-Larrubia P, Cerdán S, Geraldes CFGC (2006) Contrast Med Mol Imaging 258:246–258
55. Kanai M, Mortell KH, Kiessling LL (1997) J Am Chem Soc 119:9931–9932
56. Kramer RH, Karpen JW (1998) Nature 395:710–713
57. Bernardi A, Arosio D, Potenza D, Sánchez-Medina I, Mari S, Cañada FJ, Jiménez-Barbero J (2004) Chem Eur J 10:4395
58. Terraneo G, Potenza D, Canales A, Jiménez-Barbero J, Baldridge KK, Bernardi A (2007) J Am Chem Soc 129:2890–2900

RSC Publishing Organic & Biomolecular Chemistry

Enantioselective binding of a lanthanide(III) complex to human serum albumin studied by 1H STD NMR techniques

Journal:	*Organic & Biomolecular Chemistry*
Manuscript ID:	OB-COM-04-2011-005524.R1
Article Type:	Communication
Date Submitted by the Author:	16-May-2011
Complete List of Authors:	Dias, David; University of Coimbra, Life Sciences and Center of Neurosciences and Cell Biology, Faculty of Science and Technology Teixeira, João; University of Coimbra, Life Sciences and Center of Neurosciences and Cell Biology, Faculty of Science and Technology Kuprov, Ilya; Durham University, Chemistry New, Elizabeth; Durham University, Chemistry Parker, David; Durham University, Chemistry Geraldes, Carlos; University of Coimbra, Life Sciences and Center of Neurosciences and Cell Biology, Faculty of Science and Technology

Enantioselective binding of a lanthanide(III) complex to human serum albumin studied by ^{1}H STD NMR techniques

David M. Dias,[a] João M.C. Teixeira,[a] Ilya Kuprov[b], Elizabeth J. New [b], David Parker [b] and Carlos F.G.C. Geraldes[a]

[a]*Department of Life Sciences and Center of Neurosciences and Cell Biology, Faculty of Science and Technology, University of Coimbra, P.O. Box 3046, 3001-401 Coimbra, Portugal; E-mail*: *geraldes@bioq.uc.pt*
[b]*Department of Chemistry, Durham University, South Road, Durham, UK DH1 3LE. E-mail: david.parker@durham.ac.uk*

The enantioselective binding of the (*SSS*)-Δ isomer of an yttrium (III) tetraazatriphenylene complex to 'drug-site II' of human serum albumin (HSA) was detected by the intensity differences of its STD ^{1}H NMR spectrum relative to the (*RRR*)-Λ isomer, by the effect of the competitive binder to that site, *N*-dansyl sarcosine, upon the STD spectrum of each isomer, in the presence of HSA and by 3D docking simulations.

Highly emissive complexes of the lanthanide (III) ions show great promise as chemosensors for potential application in studying biological systems.[1-9] Much current research focuses on the relationship of structure to various aspects of chemical and biological behaviour, in order to improve promising complexes and to design new complexes for specific purposes.[10-14]

In order for a lanthanide (III) complex to have potential utility as a biological probe, it must retain its selectivity and response in the cellular environment, notably in the presence of the high intracellular concentrations of proteins. Recent studies of potential luminescent lanthanide probes have investigated the effect of protein binding on their *in vitro* behaviour.[10,12,14-16] One such investigation yielded a novel example of dynamic helicity inversion following enantioselective protein binding: the (*SSS*)-Δ enantiomer of $[Ln.\mathbf{L}]^{3+}$ (Scheme 1, Ln = Tb, Eu) binds to human serum albumin (HSA) with a resulting helicity change as measured by circularly polarised luminescence (CPL).[17] In contrast, the (*RRR*)-Λ enantiomer associates more weakly with the protein and does not exhibit this inversion. This behaviour was unique to the complex structure, and was not replicated by other complexes bearing either the same chromophore or pendant arms. Here, we report efforts to define the regions of the complex which bind to the protein by employing saturation transfer difference (STD) NMR techniques.

STD NMR is commonly used to characterise the binding in ligand-receptor complexes.[18-21] In this method, selective irradiation of the protein NMR spectrum results in saturation of the protein signals and of any ligand protons interacting with the protein. Subtraction of this spectrum from the one collected in non-saturated protein and ligand conditions gives the saturation transfer difference spectrum, in which only protein-interacting protons are visible and the epitope is mapped by analysing the signal intensities of these protons in the STD spectrum.[19]

STD NMR requires the use of a diamagnetic metal ion, which precludes the direct study of Eu(III) or Tb(III) complexes. The diamagnetic yttrium (III) is a good model for Eu(III) and Tb(III), with an ionic radius that is only 0.02 Å smaller than Tb(III) in nine-coordinate systems.[22] Y(III) complexes were synthesized from the (*RRR*) and (*SSS*) isomers of **L** from yttrium acetate, using analogous procedures to those described previously for Eu(III), Tb(III) and Gd(III) complexes of **L**.[4]

While complete assignment of the ^{1}H NMR spectra of **L** or $[Eu.\mathbf{L}]^{3+}$ was hampered by exchange broadening associated with the conformational flexibility of the ligand and the broad peaks observed for the paramagnetic Eu(III) complex, the $[Y.\mathbf{L}]^{3+}$ complexes were found to exist primarily as a single enantiomer in solution, giving rise to sharp NMR resonances. This enabled assignment of most protons in the complex from the well-defined COSY spectrum (Figures S1, S2). Two of the three phenyl amide pendant arms exhibit very similar chemical shifts, while resonances of the third lie about 1 ppm to lower frequency. This shift is consistent with a deshielding ring current effect associated with a through-space interaction involving another aromatic system. The molecular geometry of the two isomers of $[Y.\mathbf{L}]^{3+}$ was calculated using density function theory (DFT), which indicated that one of the pendant arms adjacent to the chromophore lies directly above the aromatic portion of the chromophore, giving rise to this deshielding effect (Figure 1).

Following assignment of the $[Y.\mathbf{L}]^{3+}$ spectrum, STD NMR spectra were acquired for the two enantiomers in the presence of HSA, with pre-saturation of the protein at -0.5 ppm (Figure 2). The two complexes were compared by calculation of the STD amplification factors, A_{STD}, the relative intensity of the STD signal compared to that of the most shifted ligand proton (Table 1). The A_{STD} values for all measured protons of the (*SSS*)-Δ isomer are larger than for the corresponding protons of the (*RRR*)-Λ; for almost every proton, the difference between the two isomers is greater than 10%. Given the HSA binding isotherm for the (*SSS*)-Δ isomer (log K = 5.1 for a 1:1 binding model) and the observed multi-site binding of the (*RRR*)-Λ isomer consistent with stepwise formation of

various adducts of lower affinity,[17] we attribute the higher A_{STD} values of the (*SSS*)-Δ isomer protons to its stronger protein interaction, as found previously.

Analysis of the STD NMR spectrum of (*SSS*)-[Y.**L**]$^{3+}$ allows identification of the region of the complex that interacts with the protein. The protons that appear in the standard NMR spectrum, but not in the STD spectrum, correspond to the cyclen ring and the amide methylene groups, which do not interact with the protein, consistent with their location in the centre of the complex, shielded by the pendant arms and chromophore. In addition, protons from the pendant arm that interacts with the chromophore and the cyclohexyl ring of the chromophore do not appear in the STD spectrum, or have very similar A_{STD} values to the (*RRR*)-Λ enantiomer, indicating that this moiety is not in the main core of the interaction with HSA.

Previously-reported proton relaxometric studies of the corresponding Gd(III) complex demonstrated the selectivity of binding of the complex to drug site II of HSA, identified by competition assays with the site-selective ligand *N*-dansyl sarcosine.[17] This interaction was also studied by an STD NMR experiment, in which *N*-dansyl sarcosine was added to the mixtures of protein and complex (Figures S3 and S4). STD amplification factors for the (*SSS*)-Δ isomer in the presence of *N*-dansyl sarcosine are vastly different from those measured in the absence of the competitor, while there is less difference for the (*RRR*)-Λ enantiomer (Table 1). These results corroborate the findings of the relaxivity studies.

Since the X-ray structure of N-dansyl sarcosine-bound HSA is available, a molecular docking simulation of the interaction of HSA binding site II with the most interacting (*SSS*)-Δ enantiomer was performed using Autodock Vina[23] in order to provide a 3D model of the interaction. Among the population of possible orientations for the ligand at the HSA binding site, the one with the lowest docked energy was selected (Fig. 3). This software, however, does not support Yttrium, and thus it was removed, remaining its structurally unchanged cage. Since, due to this limitation, three units of positive charge and its subsequent energy interaction are lost, these simulations are considered a qualitative three-dimensional search and not a quantitative model. The simulated 3D orientation of N-dansyl sarcosine at the HSA binding site II and the 2D interaction model coincide with the X-ray structure (Figure 3A,B). The 3D scheme (Figure 3C,D) shows that the (*SSS*)-Δ isomer interacts directly with the drug site II, mainly through the phenylmethyl pendant arms and the aromatic region of the chromophore, as suggested by the STD NMR data. This provides an explanation for the observation that this binding behaviour requires conservation of both the chromophore structure and the pendant arms. A 2D model (Figure 3E) was also established based on the docking results and using Poseview,[24,25] which highlights the direct competition of the complex with N-dansyl sarcosine for interaction with the same residues, mainly Tyr 411, Asn 391 and Leu 387. This 2D view depicts the hydrophobic contact of the phenylmethyl pendant arms and the chromophore with HSA, which is further stabilized by an hydrogen bond between Asn 391 and a chromophore nitrogen.

In summary, this work has demonstrated the utility of STD NMR experiments in providing detailed structural information about the protein binding interaction. The study confirms that (*SSS*)-Δ [Ln.**L**]$^{3+}$ associates selectively with HSA compared to the (*RRR*)-Λ enantiomer, and identifies the regions of the complex which directly interact with drug site II of the protein.

We thank the Association of Commonwealth Universities (EJN), the Fundação para a Ciência e Tecnologia (FCT), Portugal (project PTDC/QUI/70063/2006) FEDER and ESF-COST D38 for support. The Varian VNMRS 600 NMR spectrometer in Coimbra was acquired with the support of the FCT Programa Nacional de Reequipamento Científico, contract REDE/1517/RMN/2005, as part of RNRMN (Rede Nacional de RMN).

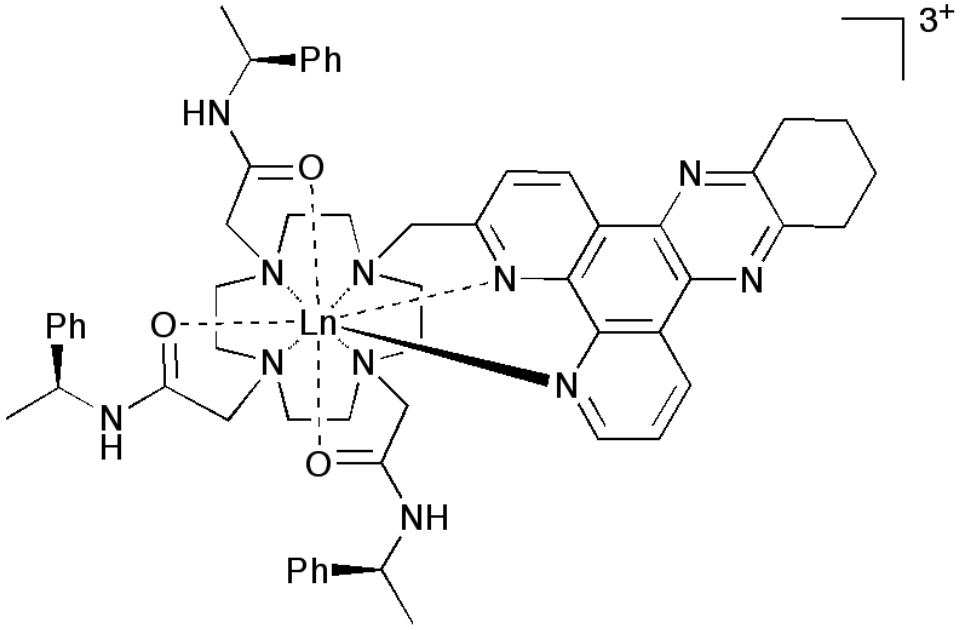
Notes and references

† Electronic Supplementary Information (ESI) available: 600 MHz COSY NMR spectrum of (*SSS*)- Δ[Y.**L**]$^{3+}$ (Fig. S1) and 1D spectrum with partial assignments (Fig. S2); ^{1}H STD NMR spectra of 30 μM HAS and 900 μM (*SSS*)-Δ [Y.**L**]$^{3+}$(Fig. S3) or 900 μM (*RRR*)-Δ[Y.**L**]$^{3+}$ (Fig. S4) in the absence and in the presence of 2mM *N*-dansyl sarcosine. See DOI: 10.1039/b000000x/

1 Bretonniere, Y.; Cann, M. J.; Parker, D.; Slater, R. *Chem. Commun.* **2002**, 1930-1931.
2 Bünzli, J.-C. G. *In Metal Ions in Biological Systems - The Lanthanides and Their Interrelations with Biosystems*; Sigel, A., Sigel, H., Eds.; Marcel Dekker: New York, 2003.
3 Hanaoka, K.; Kikuchi, K.; Kojima, H.; Urano, Y.; Nagano, T. J. *Am. Chem. Soc.* **2004**, 126, 12470-12476.
4 Poole, R. A.; Bobba, G.; Cann, M. J.; Frias, J.-C.; Parker, D.; Peacock, R. D. *Org. Biomol. Chem.* **2005**, 3, 1013-1024.
5 Pandya, S.; Yu, J.; Parker, D. Dalton Trans. **2006**, 2757-2766.
6 Yu, J.; Parker, D.; Pal, R.; Poole, R. A.; Cann, M. J. *J. Am. Chem. Soc.* **2006**, 128, 2294-2299.
7 Vandevyver, C. D.; Chauvin, A. S.; Comby, S.; Bünzli, J.-C. G. *Chem. Commun.* **2007**, 1716-1718.
8 Montgomery, C. P.; Murray, B. S.; New, E. J.; Pal, R.; Parker, D. *Acc. Chem. Res.* **2009**, 42, 925-937.
9 New, E. J.; Parker, D.; Smith, D. G.; Walton, J. W. *Curr. Opin. Chem. Biol.* **2010**, 14, 238-246.
10 Poole, R. A.; Montgomery, C. P.; New, E. J.; Congreve, A.; Parker, D.; Botta, M. *Org. Biomol. Chem.* **2007**, 5, 2055-2062.
11 Kielar, F.; Law, G.-L.; New, E. J.; Parker, D. *Org. Biomol. Chem.* **2008**, 6, 2256-2258.
12 Murray, B. S.; New, E. J.; Pal, R.; Parker, D. *Org. Biomol. Chem.* **2008**, 6, 2085-2094.
13 New, E. J.; Parker, D. *Org. Biomol. Chem.* **2009**, 7, 851-855.
14 New, E. J.; Parker, D.; Peacock, R. D. *Dalton Trans.* **2009**, 672-679.
15 Pal, R.; Parker, D.; Costello, L. C. *Org. Biomol. Chem.* **2009**, 7, 1525-1528.

16 Montgomery, C. P.; New, E. J.; Palsson, L. O.; Parker, D.; Batsanov, A. S.; Lamarque, L. *Helv. Chim. Acta*. **2009**, 92, 2186-2213.
17 Montgomery, C. P.; New, E. J.; Parker, D.; Peacock, R. D. *Chem. Commun.* **2008**, 4261-4263.
18 Mayer, M.; Meyer, B. *Angew. Chem. Int. Ed*. **1999**, 38, 1784-1788.
19 Mayer, M.; Meyer, B. *J. Am. Chem. Soc*. **2001**, 123, 6108-6117.
20 Meyer, B.; Peters, T. *Angew. Chem. Int. Ed*. **2003**, 42, 864-890.
21 Angulo, J.; Enríquez-Navas, P. M.; Nieto, P. M. *Chem-Eur. J*. **2010**, 16, 7803-7812.
22 Shannon, R. D. *Acta Cryst.* **1976**, A32, 751-767.
23 O. Trott, A. J. Olson, *Journal of Computational Chemistry*, **2010**, 31, 455-461.
24 K. Stierand, P.C. Maass, M. Rarey, *Bioinformatics*, **2006**, 22, 1710-1716.
25 K. Stierand and M. Rarey, *ChemMedChem*, **2007**, 2, 853 – 860.

Tables, Schemes and Figures



Scheme 1. Chemical Structure of the (*SSS*)-Δ enantiomer of $[Ln.\mathbf{L}]^{3+}$ (Ln = Tb, Eu, Y).

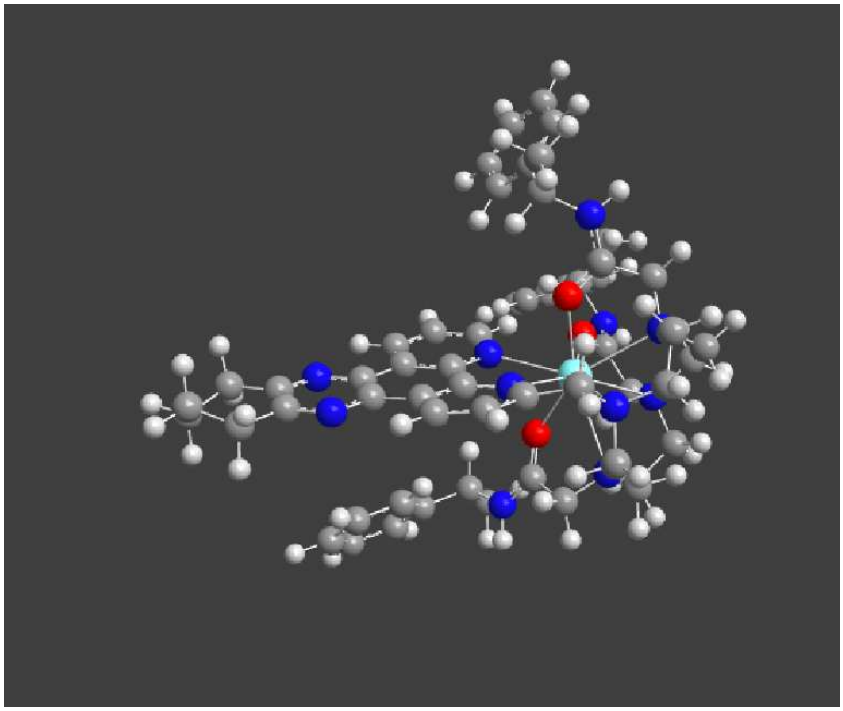


Figure 1. Molecular geometry of *(SSS) -Δ* $[Y.\mathbf{L}]^{3+}$, optimized in Gaussian03 using DFT with PBE exchange-correlation functional, cc-pVDZ basis set on the light atoms and Stuttgart RSC 1997 ECP basis set on yttrium.

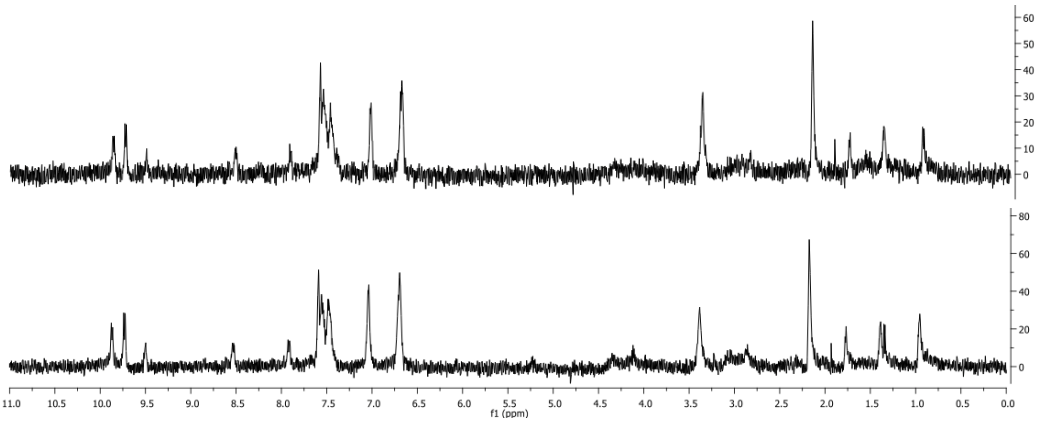


Figure 2. ^{1}H STD NMR of *(SSS)*- Δ (upper) and *(RRR)*- Λ $[Y.\mathbf{L}]^{3+}$ (lower) in the presence of 30 μM HSA (1mM complex, 25ºC, D_2O, 600 MHz, -0.5 ppm saturation)

Table 1. STD amplification factor (A_{STD}) and relative STD (to the most shifted proton) for 900 μM *(SSS)*-Δ and *(RRR)*- *Λ* $[Y.\mathbf{L}]^{3+}$ in the presence of 30 μM HSA and changes in the relative STD for (*SSS*)-Δ and (*RRR*)-Λ

Proton - ppm	(*RRR*)- *Λ* $[Y.L]^{3+}$		(*SSS*)- *Δ* $[Y.L]^{3+}$		% (SSS) – %(RRR)	% change after N-dansyl-sarcosine	
	A_{STD}	Relative STD (%)	A_{STD}	Relative STD (%)		(***RRR***)-Λ	(***SSS***)-Δ
1' – 9.6	3.51	100	4.08	100	-	-	-
3 – 9.5	2.50	71.2	3.86	94.6	23.4	8.8	-30.5
3' – 9.3	1.46	41.5	2.41	59.0	17.5	12.0	-36.8
2' – 8.35	2.63	75.0	3.24	79.5	4.5	23.9	24.7
f – 7.04	1.88	53.6	3.19	78.1	24.5	0.5	-62.8
2(e.d) – 6.4	2.48	70.8	3.75	91.8	21.1	15.4	-37.7
4.4' – 3.2	1.65	47.1	2.67	65.4	18.2	-16.5	-41.1
5.5' – 2.0	1.74	49.6	2.86	70.0	18.5	-13.1	-33.8
Me.2 – 1.6	0.62	17.6	1.03	25.2	7.6	-4.9	-14.6
Me.1 – 1.1	1.34	38.2	2.07	50.8	12.5	-10.8	-30.6
Me.3 – 0.8	1.21	34.6	2.11	51.7	17.1	-0.5	-15.9

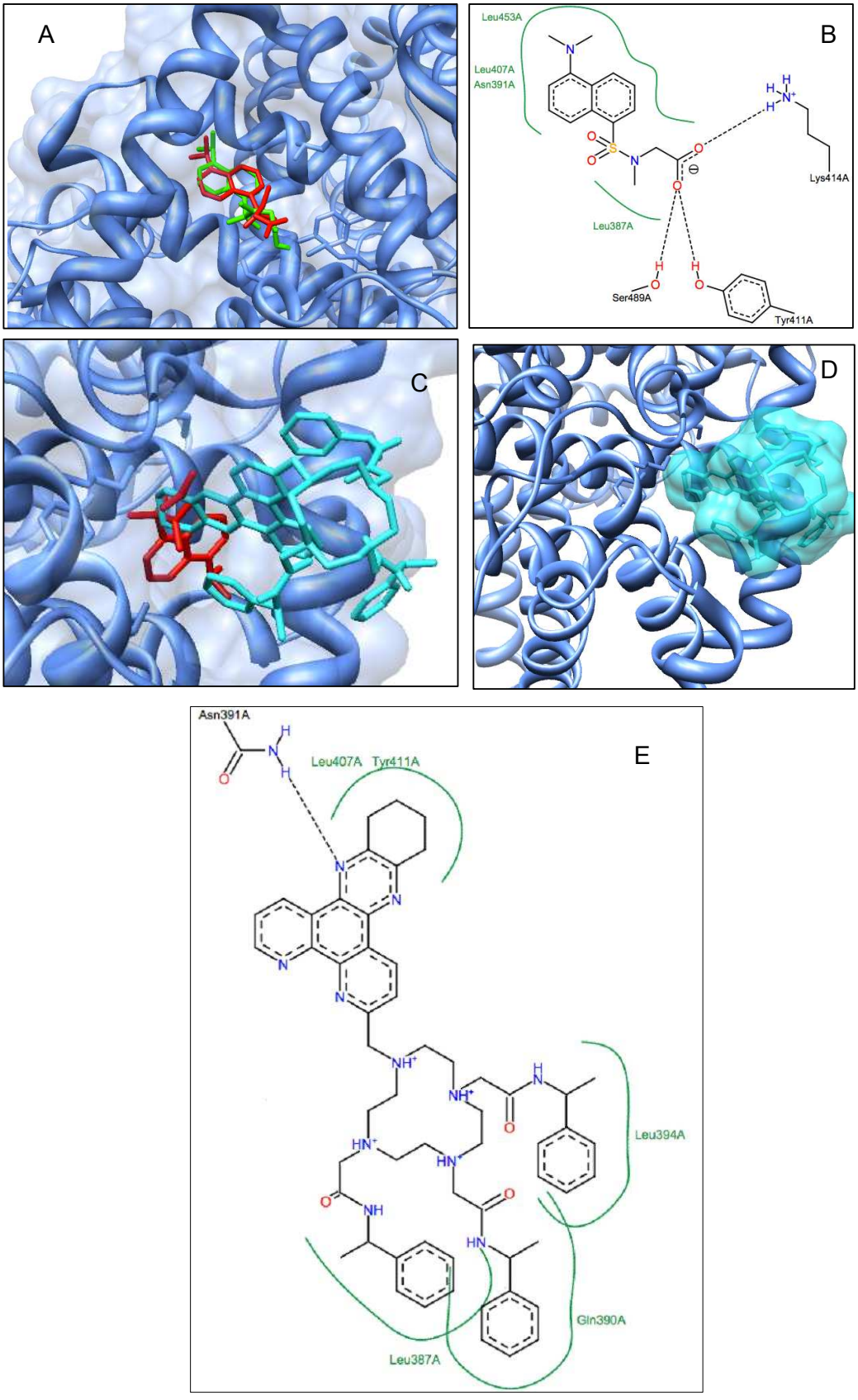

A
B
Leu453A
Leu407A
Asn391A
Lys414A
Leu387A
Ser489A
Tyr411A
C
D
E
Asn391A
Leu407A
Tyr411A
Leu394A
Gln390A
Leu387A

Figure 3. A) N-dansyl sarcosine, a known a binder of HSA drug site II. Overlay of structures from X-ray structure 2XVQ.pdb, (red) and docked with Autodock Vina (green); B) 2D scheme for interactions between N-dansyl sarcosine and site II of HSA, (black dashed lines – hydrogen bonds; green solid lines – hydrophobic interactions); C and D) Docked structure for (*SSS*)-Δ isomer (cyan) overlaid with and without N-dansyl sarcosine (red), respectively. Indicating a specific competition for this site and a lodging inside the binding site. E) 2D scheme for interactions between (*SSS*)-Δ isomer and site II of HSA (black dashed lines –hydrogen bonds, green solid lines – hydrophobic interactions; green dashed lines – Pi-cation interactions).

Electronic Supplementary Information for Organic and Biomolecular Chemistry

Supporting information

Enantioselective binding of a lanthanide(III) complex to human serum albumin studied by ^{1}H STD NMR techniques

David M. Dias,[a] João M. C. Teixeira,[a] Ilya Kuprov[b], Elizabeth J. New,[b] David Parker[b] and Carlos F.G.C. Geraldes[a]

[a] Department of Life Sciences and Centre of Neurosciences and Cell Biology, Faculty of Science and Technology, University of Coimbra, P.O. Box 3046, 3001-401 Coimbra, Portugal; Tel: +351239853608; Fax: +351239853607; E-mail: geraldes@bioq.uc.pt

[b]Department of Chemistry, Durham University, South Road, Durham, UK DH1 3LE; Tel +1913342033.; Fax +1913844737 ; E-mail: david.parker@durham.ac.uk

Experimental Section

Materials. Chemicals were purchased from Sigma-Aldrich and used without further purification. *(SSS)*-**L** and *(RRR)*-**L** were synthesised as previously described.[1]

Synthesis of *(SSS)*-[Y.**L**]Cl_3: 1-(6-Methyl-10,11,12,13-tetrahydro-4,5,9,14-tetraazabenzo[b]triphenyl ene)-4,7,10-tris[*(S)*-1-(1-phenyl)ethylcarbamoylmethyl]-1,4,7,10-tetraazacyclododecane, ligand **L** (26 mg, 0.027 mmol) and yttrium (III) acetate (6.7 mg, 0.025 mmol) were dissolved in methanol:water (1:1, 5 mL) and the resulting solution was heated to reflux under argon overnight. The solution was then added dropwise to ether (20 mL), the precipitate centrifuged and the solvent decanted. The solid was redissolved in acetonitrile and the process repeated to yield a pale yellow solid. This solid was then converted to the chloride salt by stirring with Dowex 1×8 200-400 mesh Cl which had previously been washed with 1 M hydrochloric acid and neutralised with water. The solid Dowex was removed by filtration and the solvent removed by lyophilisation to yield the title compound as a yellow solid (23 mg, 0.019 mmol, 75%).

^{1}H NMR (700 MHz, D_2O): 0.79 (3H, m, Me.1), 1.11 (3H, m, Me.2), 1.25 (1H, m, cyclen CH_2), 1.39 (1H, m, cyclen CH_2), 1.61 (3H, m, Me.3), 1.84 (2H, m, amide CH_2), 2.01 (5H, m, 5, 5', cyclen CH_2),

2.17 (3H, m, cyclen CH_2), 2.26 (1H, m, cyclen CH_2), 2.34 (1H, m, amide CH_2), 2.52 (2H, m, amide CH_2), 2.57 (1H, m, 1c), 2.75 (2H, m, cyclen CH_2), 2.97 (1H, m, cyclen CH_2), 3.20 (3H, m, cyclen CH_2), 3.46 (4H, m, 4, 4'), 3.33 (1H, m, amide CH_2), 3.68 (1H, m, cyclen CH_2), 4.17 (1H, m, 2c), 5.14 (1H, m, 3c), 6.37 (2H, br, amide NH), 6.51 (4H, m, 3e, 3d), 6.76 (1H, m, amide NH), 6.87 (3H, m, 1f, 2f, 3f), 7.33-7.50 (10H, m, 1e, 1d, 2e, 2d, aryl CH), 7.75 (1H, m, 2), 8.35 (1H, m, 2'), 9.30 (1H m, 1'), 9.53 (1H, m, 3), 9.59 (1H, m, 3').

HRMS (+ m/z): $[M]^{3+}$ calculated for $C_{57}H_{65}N_{11}O_3Y$, 347.4824; found 347.4816.

(RRR)-[Y.**L**]Cl_3 was synthesised using an analogous procedure, using *(RRR)*-**L**. Spectroscopic characterisation was identical to that observed for *(SSS)*-[Y.**L**]Cl_3.

NMR studies. Solutions for NMR studies were prepared by mixing 99.6% D_2O (purchased from Sigma-Aldrich) solutions of each complex, HSA (defatted, from Fluka Biochemika, Switzerland) and *N*-dansyl sarcosine, in appropriate concentrations, to a total volume of 200 μL and then transferred to 3 mm NMR tubes. No buffer was used.

All ^{1}H NMR spectra were acquired on a Varian VNMRS 600 (14.09 T, 599.72 MHz) NMR spectrometer (Coimbra) using a 3 mm pulse field gradient (PFG) inverse probe and on Varian VNMRS-700 (16.44 T, 699.731 MHz) NMR spectrometer (Durham). ^{1}H chemical shifts (δ) are given in ppm relative to TSP as internal reference (^{1}H, δ 0.0). For each sample, one-dimensional (1D) ^{1}H spectra were obtained using a spectral window of 10 kHz, an acquisition time of 2 seconds, a repetition time of 5 seconds, 128 scans and a line broadening of 0.6 Hz. The proton 1D spectra were assigned using 2D g-COSY spectra. The ^{1}H saturation transfer difference (STD) NMR spectra were acquired on the Varian VNMRS 600 (Coimbra) spectrometer using the Double Pulse Field Gradient Spin Echo (DPFGSE) sequence[2] for water suppression. Since in this NMR system the STD NMR spectra are acquired directly from phase cycling, the ^{1}H 1D NMR spectra were used as off-resonance references in order to calculate the STD amplification factor.[3] All STD spectra were acquired using the same parameters: equal spectrometer gain value, the same acquisition parameters as for 1D spectra, except that 2048 scans were used. A selective Gaussian saturation pulse of 2.5 seconds was applied at protein resonances. A previously calibrated spin-lock filter ($T_{1\rho}$) was used to remove protein resonances. All spectra were analyzed using Mestre Nova Software v5.3.1-4825. In order to compare the reference spectra with the STD NMR spectra, the different number of acquisitions was normalized according to equation (1),

$$\text{Rel.STD}\,\% = \frac{I_{STD} \times 2 \times scans_{reference}}{I_0 \times scans_{STD}} \quad (1)$$

where I_{STD} is the peak intensity of the STD NMR spectra, I_0 is the intensity of the peaks in the ^1H reference spectra. Then, the peak intensities were normalized to the amplification factor STD (A_{STD}), equation (2),

$$A_{STD} = \text{Rel.STD}\,\% \times \text{Lig.Exc.} \quad (2)$$

DFT. The DFT calculations were performed by using the Gaussian 03[4] package Y^{3+} complexes; the structures of complexes of other lanthanides are likely to be nearly identical. Importantly, however, the use of diamagnetic La^{3+} and Y^{3+} for DFT calculations avoids a host of largely unresolved theoretical issues with spin-orbit coupling and zero-field splitting in open-shell lanthanides. The Y^{3+} results were found to be very similar to those obtained with La^{3+}, and given the similarity of the ionic radii of the central lanthanide ions, only Y^{3+} complex computations were undertaken. Gaussian 03[4] logs and checkpoints are available upon request. Molecular geometries were optimised *in vacuo* by using spin-restricted B3LYP exchange-correlation functional with a compound basis set (ccpVDZ for CHNOFS, Stuttgart ECP28MWB for La and WGBS for Y).

Docking. The molecular docking simulations were performed using Autodock Vina 1.1.1. [5] from the Scripps Research Institute. Autodock Tools 1.5.4 revision 29 was used to prepare all the pdbqt files for Autodock Vina.

The HSA structure used was the one with the entry 2XVQ.pdb, because this X-ray structure contains the Dansyl Sarcosine, which is known to interact with drug site II, and therefore, the adequate model to study the possible binding site to our ligands.

From the structure contained in 2XVQ.pdb, the chain A was selected and polar hydrogen atoms were added to the HSA and its nonpolar hydrogen atoms were merged.

The ligands, obtained from DFT calculations performed previously from Gaussian 03. With Autodock Tools automatically add gasteiger charges, merge non-polar hydrogens, detects the aromatic carbons and the rotatable bounds.rotatable according to the stereochemistry of the molecule.

All calculations for protein-fixed ligand-flexible docking were done using the Lamarckian Genetic Algorithm (LGA) method. The GRID box parameters used were: center x = 0, center Y = 5, center Z = -24, dimension 30 x 30 x30 and spacing of 1 Angstrom, with a total of 29791 total grid points per map.

The grid box was set around dansyl sarcosine binding site to fully cover the entire binding site and accommodate ligands to move freely. The best pose was chosen with the lowest docked energy, after the docking search was completed.

USFC Chimera (version 1.5.2 build 32411) was used to read the Autodock data and to obtain the images /views presented in this work. For the 2D interactions PoseviewWeb 1.97.0[6,7] was used to represent all possible interactions al fully characterize them. It was used 2XVQ.pbd chain A and the best poses for all ligands in mol2 files, that retains all the Cartesian coordinates like the pdb files.

References

1. R. A. Poole, G. Bobba, M. J. Cann, J.-C. Frias, D. Parker and R. D. Peacock, *Org. Biomol. Chem.,* **2005**, *3*, 1013-1024
2. T.-L. Hwang, A. J. Shaka, *J. Magn. Reson. A*. **1995**, *112*, 275-279.
3. M. Mayer, B. Meyer, *J. Am. Chem. Soc.*. **2001**, *123*, 6108-6117.
4. Gaussian 03, Revision E.01, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery, Jr., T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S.Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari,J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez, J. A. Pople, Gaussian, Inc., Wallingford CT, 2004.
5. O. Trott, A. J. Olson, *Journal of Computational Chemistry*, **2010**, 31, 455-461.
6. K. Stierand, P.C. Maass, M. Rarey, *Bioinformatics*,**2006**, 15;22(14):1710-1716
7. K. Stierand and M. Rarey, *ChemMedChem*, **2007**, 2, 853 – 860

Supplementary Results

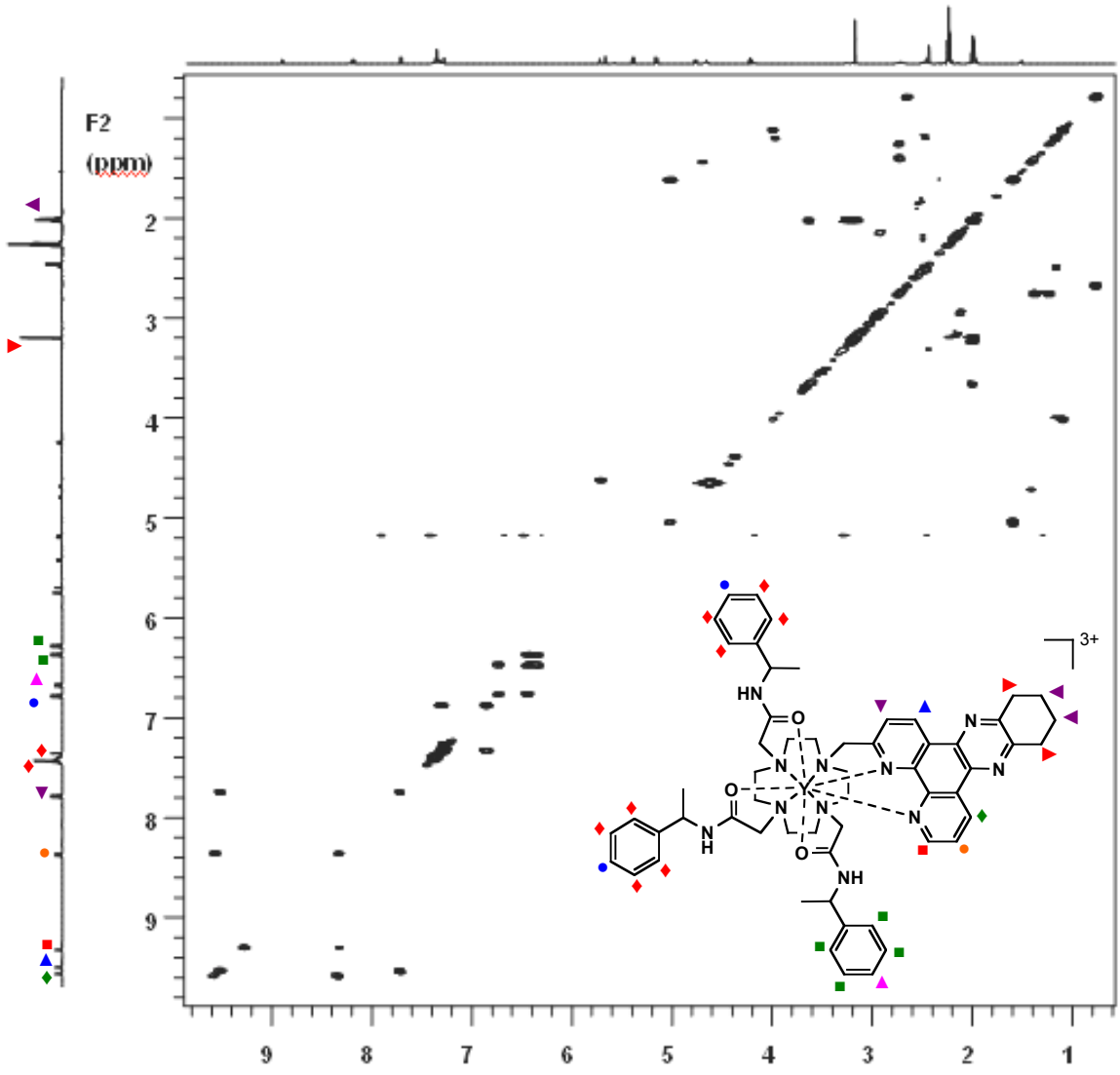


Figure S1. COSY NMR spectrum of *(SSS)*- Δ [Y.**L**] $^{3+}$ (600 MHz, pH 7.0, 295 K) showing partial assignments.

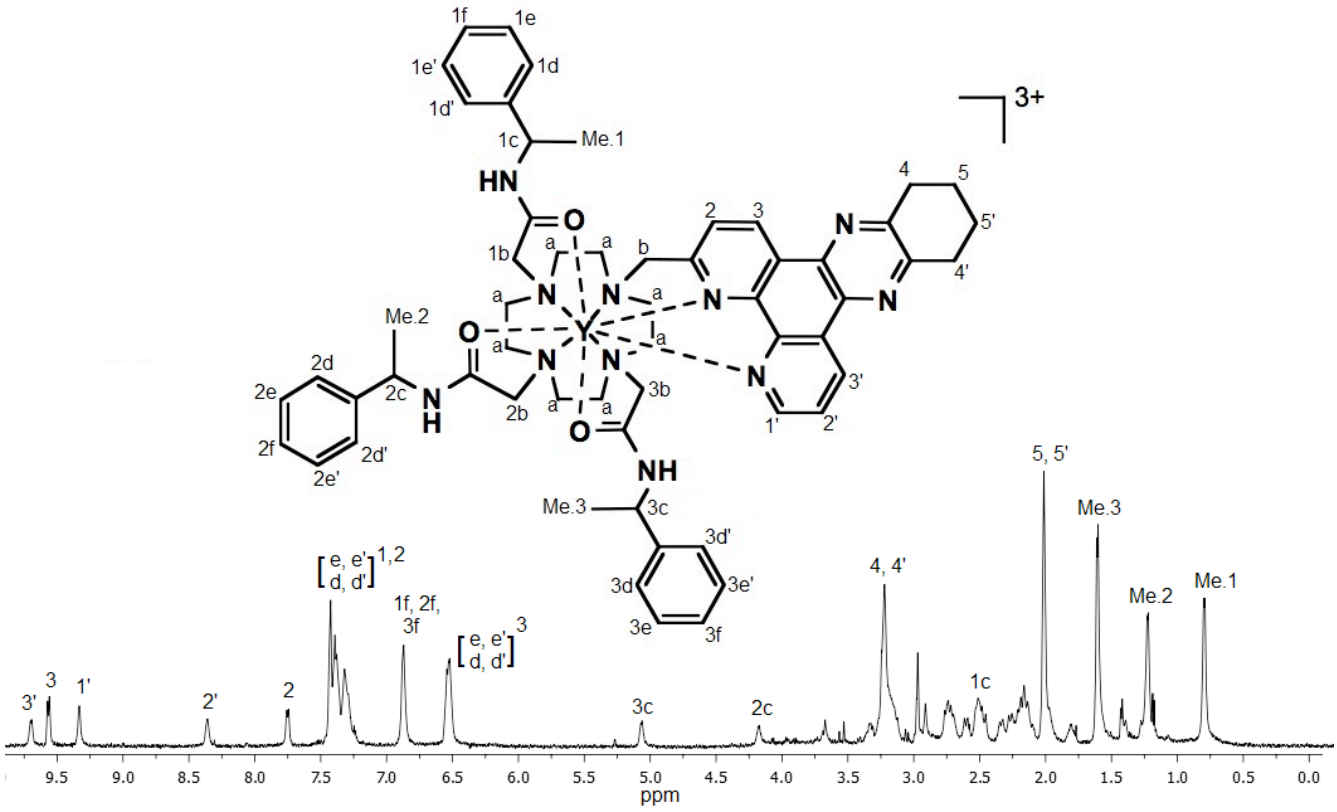


Figure S2. ^{1}H NMR assignments for the *(SSS)*- Δ [Y.**L**]$^{3+}$ complex.

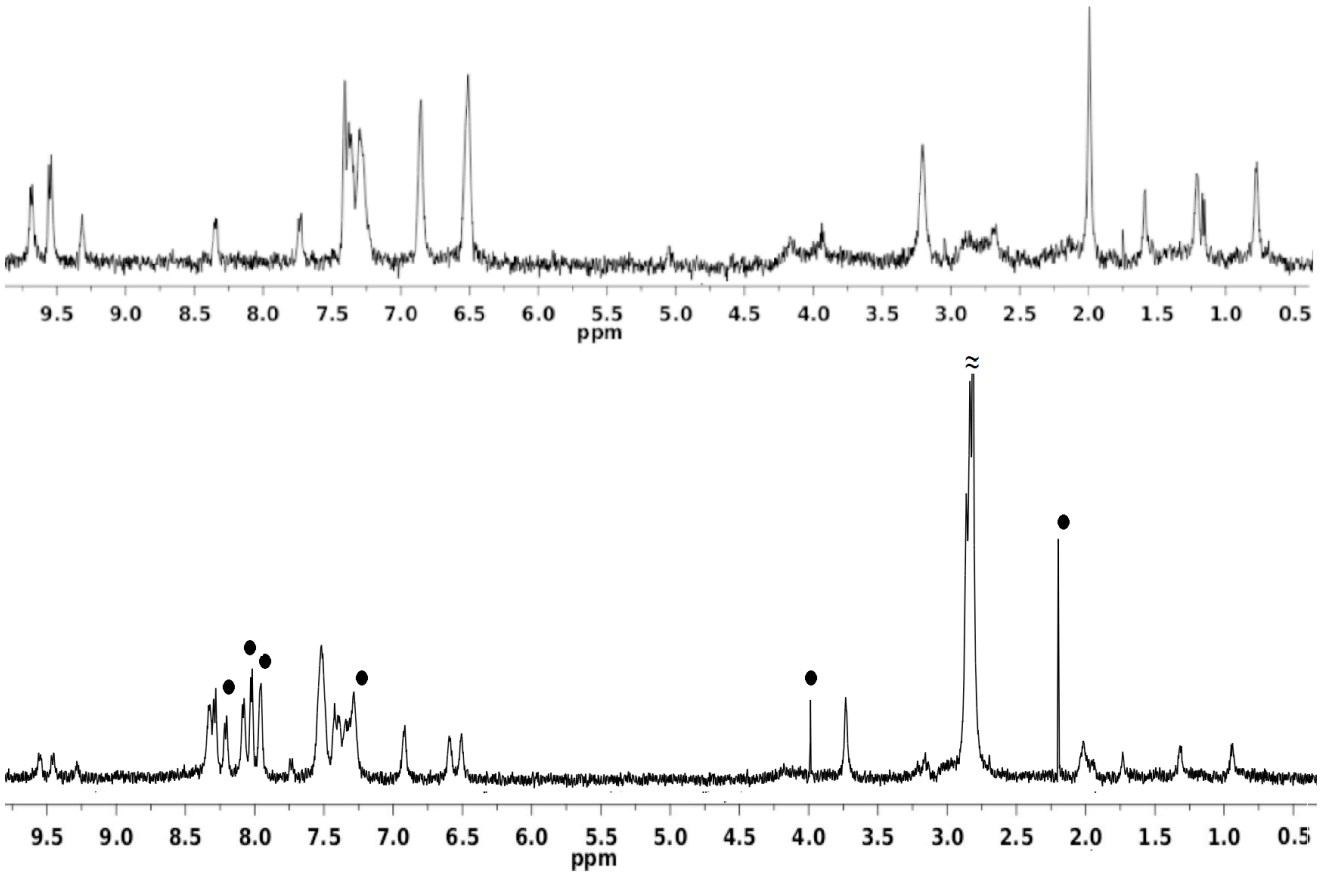


Figure S3. ^{1}H STD NMR of 30 μM HSA, 900 μM (*SSS*)- Δ [Y.**L**]$^{3+}$ in the absence (upper) and in the presence of 2mM *N*-dansyl sarcosine (● – N-Dansyl Sarcosine resonances) at 25ºC, D2O, 600 MHz, -0.5 ppm saturation.

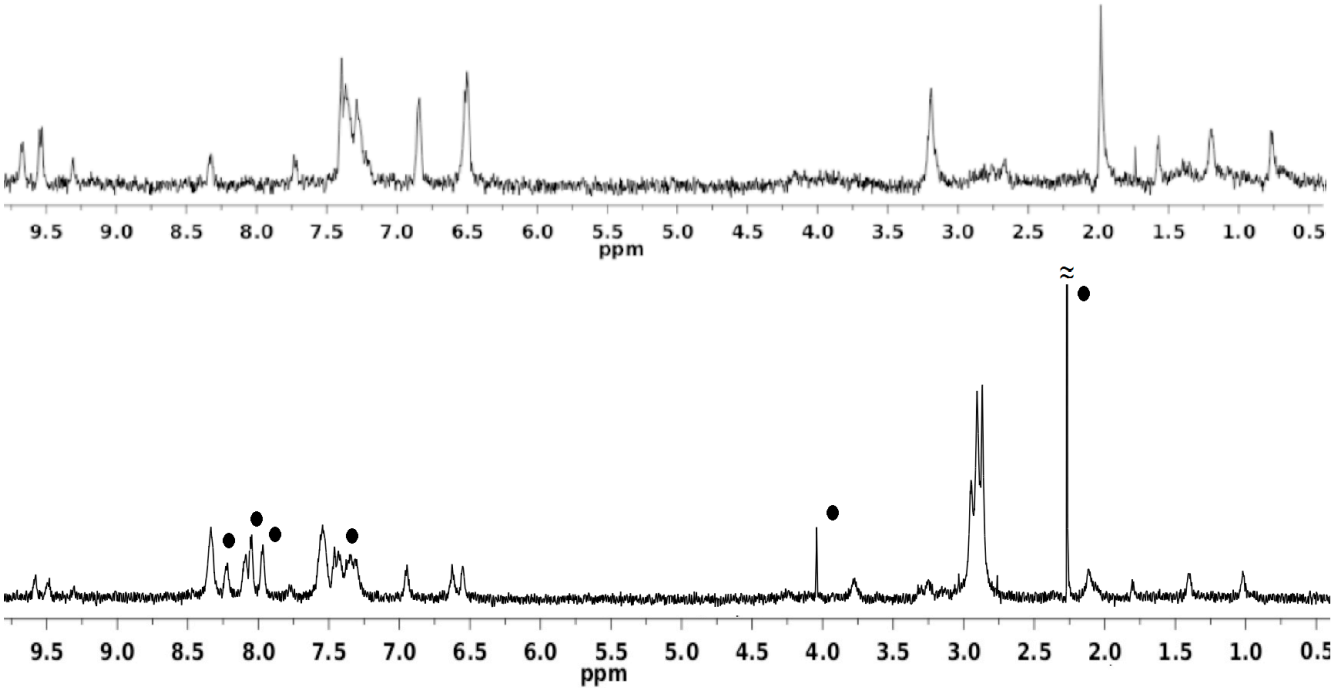


Figure S4. ^{1}H STD NMR of 30 μM HSA, 900 μM (*RRR*)- Λ [Y.**L**]$^{3+}$ in the absence (upper) and in the presence of 2mM *N*-dansyl sarcosine (● – *N*-dansyl Sarcosine resonances) at 25ºC, D_2O, 600 MHz, -0.5 ppm saturation.