The Interaction of La^{3+} Complexes of DOTA/DTPA-Glycoconjugates with the RCA_{120} lectin: A Saturation Transfer Difference (STD) NMR Spectroscopic Study

João M. C. Teixeira, [a] David M. Dias,[a] F. Javier Cañada,[b] José A. Martins,[c] João P. André,[c] Jesús Jiménez-Barbero,[b] and Carlos F. G. C. Geraldes * [a]

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 DOTA mono(amide) and DTPA bis(amide) 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 ASGPR (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 pathologies. The *in vitro* affinity constants of the divalent La(III)- glycoconjugate complexes to 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 binding mode of mono and divalent forms of these CAs to the galactose 1α binding sites of RCA_{120}. The atomic details of the molecular interactions allowed corroborating and supporting the interaction of both the 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 • STD-NMR spectroscopy • MRI contrast agents • protein-ligand interaction • compound docking

[a] João M. C. Teixeira, D.M. Dias, C.F.G.C. Geraldes (✉)
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
Tel: (+) 351239853608; Fax: (+)351239853607
e-mail: geraldes@bioq.uc.pt

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

[c] 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, ELISA (Enzyme Linked Immunosorbent Assay(1), Immunoblotting, RIA (Radioimmuno-assay),(2) affinity chromatography(3) or 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 NOESY,(6) NOE pumping (7) and 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 characterizing ligand binding through intermolecular saturation transfer and, moreover, allows screening of ligand libraries,(11) as well as

calculating 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 pathologies 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 therapy. Based on this knowledge, a new class of CAs has been recently developed with the intent to be selectively taken up by ASGPR hepatic receptor.(26-28) DOTA-like chelators [DOTA = 1,4,7,10-tetrakis(carboxymethyl)-1,4,7,10-tetraazactclododecane] were attached to sugar moieties, either galactosyl, glucosyl or lactosyl residues, by pendant arms containing aliphatic chains and amide bonds, resulting in mono or multivalent glycoconjugate derivative agents. After the development of the DOTA-based glycoconjugates,(26) DTPA bis(amide)-based glycoconjugates [DTPA = diethylenetriaminepentaacetic acid] were also devised and studied.(27) In both types of CAs, DOTA- and DPTA-based chelates, the structural peculiarities are similar: a central reporter group complexing a paramagnetic metal center (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 (Figure 1).

(Insert Figure 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 binding to proteins (*i.e.*, tetra > tri > di > mono).(33,34) However, the way how 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: *intra-* or *inter-molecular* 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 as chelate-type binding, as the glycoside simulates a

chelate motif imprisoning the *metal* atom, in this case, the lectin. Moreover, in order 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 an 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 herein 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 to 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$_2$) (Figure 1). Thus, 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 asialoglycoprotein-receptor (ASGPR). Although the carbohydrate recognition domain (CRD) 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 the present proof-of-principle study of the employed methodology 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, while 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 revealed by site mutations.(45-48).

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

Materials and Methods

Samples: The DOTA and DTPA glycoconjugate derivatives and their La(III) complexes were synthesised and characterized as described previously.(26, 27) The La(III) complexes were dissolved as 99.9% D_2O / 10% PBS buffered solutions. *Ricinus communis* agglutinin (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. Proteins concentration ranged from 10-25 μM depending on the compound studied and the expected affinity to the protein, in order to achieve a large range of ligand excess. The concentration of the ligands were selected in order to obtain the majority 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 concentration of the protein and ligand was not constant for every compound, and were defined according to the desired ligand excess ratio and the available quantities.

NMR studies: All ^1H NMR spectra were acquired using a 5 mm pulse field gradient (PFG) triple resonance inverse probe on a Varian NMRS-600 NMR Spectrometer working at 599.72 MHz. For each sample an one-dimensional (1D) ^1H spectrum was obtained, and the spectral assignments from the literature (12,26,27) were used after being confirmed by two dimensional (2D) gCOSY spectra (data not shown). Saturation transfer difference (STD) NMR spectra were then acquired, where the Double Pulse Field Gradient 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 ^1H 1D 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 ^1H spectra and between 1024 and 2048 in 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 second and the repetition time was 3.5 seconds. STD experiments were performed using a saturation delay of 2.5 seconds. In order to compare the reference spectra with the STD NMR spectra, the different number of acquisitions was normalized according to equation (1),

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

where I_{STD} is the peak intensity of the STD NMR spectra, I0 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),

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

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 the one-site binding model, (13, 16) equation (3)

where α_{STD} is the maximum A_{STD} and [L] is the total ligand concentration. Plots and fits were obtained using GnuPlot v4.2-3.

$$A_{STD} = \frac{\alpha STD \times [L]}{[L] + K_D} \qquad (3)$$

Docking Calculations: Automated docking was performed using Autodock 4.1(51) and the Lamarckian Genetic Algorithm (52) as searching procedure. The PDB file corresponding to the protein used, RCA_{120}, was 1RZO. The protein model was kept rigid, torsions were allowed only at the ligand level. Due to the large number of torsion and the size of the ligands, only the sugar moiety and the adjacent arms were docked. Thus, 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 three-dimensionally designed using Maestro Schrödinger Software.(53) For glycoconjugate derivatives grid maps were constructed using 50 x 50 x 50 points, with 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 27000, the elitism was 1, the probability of a gene to 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 evaluations was kept around 1.1 x 10^7. For galactose calculations, grid maps were constructed using 40 x 40 x 40 points, with 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 x 10^7, the maximum number of generations was 40000, 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 r.m.s.d. cutoff of 0.5 Å.

Results and Discussion

Figure 2 represents both the ^{1}H 1D 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 based on previous publications (12,26,27) and on gCOSY analysis. The resonances are identified in Figure 2, following the proton numbering schemes shown in Figure 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. Due to the nature of the STD experiment, it is possible to characterize the binding epitope of the ligand for a particular interaction. Table 1 summarises the determined saturation profiles (relative STD%) for the sugar and linker protons of the four compounds studied. *H* protons refer to the protons from the galactosyl residue, while *H'* protons refer to those of the glucosyl residues, in the case of lactosyl derivatives. The evaluation of the saturation profile of the six (twelve) groups of protons from the

(Insert Figure 2)

galactosyl (lactosyl) residues proves that the sugar protons that remain closer to the protein are always H3 and H2 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$^{a'}$ and H3/3' in the case of lactosyl derivatives was not considered because these peaks are superimposed with 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 interactions.(12,29) The anomeric protons were not considered in the group epitope mapping evaluations, because the DPFGSE water suppression scheme dramatically affected its intensity. Noteworthy, additional STD effects were observed for the linker protons (see Figure 2), which will be discussed later.

(Insert Table 1)

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 from the galactosyl and lactosyl derivatives, each one representing both monovalent and divalent forms. Table 2 summarises 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, due to 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) the K_D determinations presented here have to be considered 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)

(Insert Table 2 and Figure 3)

The stronger binding of the divalent compounds to the lectin clearly observed in the present experiments, is in agreement with previous *in vivo* studies, where the higher affinity of the multivalent forms of this type of glycoconjugate derivatives 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 mono and divalent molecules with RCA_{120}. For the monovalent compounds, La(DOTAGal) and La(DOTALac) (Figure 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 GEM 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) *l* was measured to have 29% of saturation and *gj* and *hi*, 37% and 28%, respectively, normalized to H3 (Table 1). Although we have to consider that the *gj* and *hi* signal contribution comes from 8 protons, we cannot neglect that the linkers of these divalent compounds interact with the lectin surface.

An 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* to *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 an interaction of the protons *g-j* of La(DOTALac_2) with the protein.

The presence of hydrophobic interactions between a given ligand and the neighboring regions of the sugar binding site in the lectin surface has already been reported.(31,35) Although very frequently this type of binding features has been associated to intramolecular, or chelate-type, binding modes,(55,56) the 3D structure of the RCA_{120} protein does not allow such 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 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, of more than 50 Å. Thus, the ten-fold increased affinity of the divalent compounds to 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.

(Insert Figure 4 and 5)

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 the RCA$_{120}$ binding site 1α, considering only one of the sugar arms, following the protocol stated in the Experimental Section. The highest rank cluster encompassed eight possible binding conformations, with the output geometries clustered using a root-mean-square deviation (r.m.s.d.) of 2 Å. Three out 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 obtained binding mode 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, while Figure 4B represents three superimposed structures of 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 take place for the different solutions for this glycoconjugate. With respect to the interactions of the linker with the protein surface, the obtained models set the long hydrophobic linker of the La(DTPAGal$_2$) chelate close to a hydrophobic region of the protein surface, thus interacting with the side chains of different aminoacids. Due to the size of the docked ligand, the obtained structures can be considered as a good approximation of the interaction mode, which can not be seen within the concept of a rigid, static representation. Very probably, different orientations of the linker may be adopted to properly interact with the lectin, as suggested by the docking calculations.

(Insert Table 3)

Docking studies of the La(DOTAGal) single-arm molecule with the RCA$_{120}$ binding site 1α were also performed, as described in the Experimental Section. Only the structures, which fit the STD data, were selected for further analysis (Figure 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 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 of La(DTPAGal2) and La(DOTAGal) and the RCA_{120} binding site 1α surface were measured using Autodock 4 (Table 3). In average, the protons *f-k* of the hydrophobic patch at La(DTPAGal$_2$) linker are closer to the protein surface protons (calculated average distance of 2.9 Å) than the 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.

(Insert Table 4)

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 to 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). The values should be regarded as merely qualitative, given the simplification of the employed model. The affinity of galactose to RCA_{120} is 2.2 x 10^3 M^{-1},(46) a value which lies between the calculated values for the mono- 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 to hepatocyte cells with the corresponding $^{153}Sm^{3+}$ chelates.(54) The so-called cluster glycoside effect may be invoked to explain the observations. Our present 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/or 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, specially 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 to 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 the first time, can be very useful in the design of novel targeted MRI contrast agents. In particular, novel design and production of high affinity glycoconjugates, such these ones, 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 approach will aim at increasing entropic and enthalpic savings that derive from the linker. A good knowledge of the structure of the target receptor is also of extreme importance in order to design specific and protein-directed ligands.

Acknowledgements

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 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 RNRMN (Rede Nacional de RMN). This work was carried out in the framework of the COST D38 Action. The group at Madrid thanks the Ministery of Science and Innovation of Spain for financial support (Grant CTQ2009-08536). We also thank Dr. Eurico Cabrita for useful discussions.

References:

1. E. Engvall, P. Perlmann, *Immunochemistry* **1971**, *8*, 871-874.
2. B. R. S. Yalow, S. A. Berson, *J Clin Invest.* **1960**, *39*, 11-13.
3. P. Cuatrecasas, M. Wilchek, C. B. Anfinsen, *Biochemistry* **1968**, 61, 636-643.
4. J. Homola, S. S. Yee, G. Gauglitz, *Sensors and Actuators B: Chemical* **1999**, *54*, 3-15.
5. B. Meyer, T. Peters, *Angew. Chem. Int. Ed.* **2003**, *42*, 864-90.
6. F. Ni, *Progr. Nucl. Magn. Reson. Spectroscopy* **1994**, *26*, 517-606.
7. A. Chen, M. J. Shapiro, *J. Am. Chem. Soc.* **1998**, *120*, 10258-10259.
8. C. Dalvit, P. Pevarello, M. Tatò, M. Veronesi, A. Vulpetti, M. Sundström, *J. Biomolecular NMR.* **2000**, *18*, 65-8.

9. C. Dalvit, G. Fogliatto, A. Stewart, M. Veronesi, B. Stockman, *J. Biomolecular NMR*. **2001**, *21*, 349-59.
10. M. Mayer, B. Meyer, *Angew. Chem. Int. Ed.* **1999**, *38*, 1784-1788.
11. M. Vogtherr, T. Peters, *J. Am. Chem. Soc.*. **2000**, *122*, 6093-6099.
12. M. Mayer, B. Meyer, *J. Am. Chem. Soc.*. **2001**, *123*, 6108-17.
13. C. Lepre, J. M. Moore, J. W. Peng, *Chem. Rev.*. **2004**, *104*, 3641-76.
14. L. Fielding, *Progr. Nucl. Magn. Reson. Spectroscopy*. **2007**, *51*, 219-242.
15. R. Meinecke, B. Meyer, *J. Med. Chem.* **2001**, *44*, 3059-65.
16. A. T. Neffe, M. Bilang, I. Grüneberg, Meyer, B. *J. Med. Chem.* **2007**, *50*, 3482-8.
17. J. Angulo, P. M. Enríquez-Navas, P. M. Nieto, *Chem. Eur. J.*. **2010**, *16*, 7803-12.
18. P. J. Hajduk, J. C. Mack, E. T. Olejniczak, C. Park, P. J. Dandliker, B. Beutel, *J. Am. Chem. Soc.*. **2004**, *126*, 2390-8.
19. J. Klein, R. Meinecke, M. Mayer, B. Meyer, *J. Am. Chem. Soc.*. **1999**, *121*, 5336-5337.
20. M. Kolympadi, M. Fontanella, C. Venturi, S. André, H-J Gabius, J. Jiménez-Barbero, P. Vogel, *Chem. Eur. J*. **2009**, *15*, 2861-73.
21. J. Jiménez-Barbero, E. Dragoni, C. Venturi, F. Nannucci, A. Ardá, M. Fontanella, S. André, F. J. Cañada, H.-J. Gabius, C. Nativi, *Chem. Eur. J.* **2009**, *15*, 10423-31.
22. H. Leffler, S. Carlsson, M. Hedlund, Y. Qian, F. Poirier, *Glycoconjugate Journal* **2004**, *19*, 433-40.
23. J. Almkvist, A. Karlsson, *Glycoconjugate Journal*, **2004**, *19*, 575-81.
24. P. Critchley, M. N. Willand, A. K. Rullay, D.H. Crout, *Org. Biomol. Chem.* **2003**, *1*, 928-38.
25. G. Ashwell, J. Harford, *Ann. Rev. Biochem*. **1982**, *51*, 531-544.
26. J.P. André, C. F. G. C. Geraldes, J. A. Martins, A. E. Merbach, M. I. M. Prata, A. C. Santos, J. J. P. de Lima, J. P., E. Tóth, *Chem. Eur. J.*. **2004**, *10*, 5804-16.
27. P. Baía, J. P. André, C. F. G. C. Geraldes, J. A. Martins, A. E. Merbach, E. Tóth, *Eur. J. Inorg. Chem.* **2005**, 2110-2119.
28. D.A. Fulton, E.M. Elemento, S. Aime, L. Chaabane, M. Botta, D. Parker, *Chem. Commun.* **2006**, 1064-6.
29. Y. C. Lee, *FASEB J*. **1992**, *6*, 3193-3200.
30. R. T. Lee, Y. C. Lee, *Glycoconjugate Journal*. **2001**, *17*, 543-51.
31. J. Lundquist, E. J. Toone, *Chem. Rev.*. **2002**, *102*, 555-78.
32. Y. C. Lee, R. T. Lee, *Acc. Chem. Res.* **1995**, *28*, 321-327.
33. Y. C. Lee, R. R. Townsend, M. R. Hardy, J. Lönngren, J. Arnarp, M. Haraldsson, H. Lönn, *J. Biol. Chem.*. **1983**, *258*, 199-202.
34. E. A. Biessen, H. Broxterman, J. H. van Boom, T. J. van Berkel, *J. Med. Chem.* **1995**, *38*, 1846-

52.

35. J. J. Lundquist, S. D. Debenham, E. J. Toone, *J. Org. Chem*. **2000**, *65*, 8245-50.

36. J. Corbell, *Tetrahedron Asymmetry*. **2000**, *11*, 95-111.

37. M. Meier, M.D. Bider, V.N. Malashkevich, M. Spiess, P. Burkhard, *J. Mol. Biol*. **2000**, *300*, 857-65.

38. J. Nahálková, J. Svitel, P. Gemeiner, P. Danielsson, B. Pribulová, L. Petrus, *J. Biochem. Biophys. Methods* **2002**, *52*, 11-8.

39. R. D'Agata, G. Grasso, G. Iacono, G. Spoto, G. Vecchio, *Org. Biomol. Chem*. **2006**, *4*, 610-2.

40. M. Lee, S. Park, I. Shin, *Bioorg. Med. Chem. Lett.* **2006**, *16*, 5132-5.

41. G. L. Nicolson, J. Blaustein, M. E. Etzler, *Biochemistry*. **1974**, *13*, 196-204.

42. S. Olsnes, E. Saltvedt, A. Pihl, *J. Biol. Chem.*. **1974**, *249*, 803-10.

43. Y. Endo, K. Tsurugi, *J. Biol. Chem*. **1987**, *262*, 8128-30.

44. F. I. Lamb, L. M. Roberts, J. M. Lord, *Eur. J. Biochem*. **1985**, *148*, 265-70.

45. L. L. Houston, T. P. Dooley, *J. Biol. Chem.*. **1982**, *257*, 4147-4151.

46. S. Sharma, S. Bharadwaj, A. Surolia, S. K. Podder, *Biochem. J.* **1998**, *333 [Pt 3*, 539-42.

47. T. K. Dam, C. F. Brewer, *Chem. Rev.*. **2002**, *102*, 387-429.

48. N. Sphyris, J. M. Lord, R. Wales, L. M. Roberts, *J. Biol. Chem.*. **1995**, *270*, 20292-7.

49. A. Rivera-Sagredo, J. Jiménez-Barbero, M. Martín-Lomas, D. Solís, T. Díaz-Mauriño, *Carbohyd. Res.* **1992**, *232*, 207-226.

50. T.-L. Hwang, A. J. Shaka, *J. Magn. Reson. A.* **1995**, *112*, 275-279.

51. D. S. Goodsell, G. M. Morris, A. J. Olson, *J. Mol. Rec.* **1996**, *9*, 1-5.

52. G. M. Morris, D. S. Goodsell, R. S. Halliday, R. Huey, W. E. Hart, R. K. Belew, A. J. Olson, *J. Computational Chem*. **1998**, *19*, 1639-1662.

53. Maestro, version 8.5, Schrödinger, LLC, New York, NY **2008**.

54. M. I. M. Prata, A. C. Santos, S. Torres, J. P. André, J. A. Martins, M. Neves, M. L. García-Martín, T. B. Rodrigues, P. López-Larrubia, S. Cerdán, C. F. G. C. Geraldes, *Contrast Med. Mol. Imaging*. **2006**, *258*, 246-258.

55. M. Kanai, K. H. Mortell, L. L. Kiessling, *J. Am. Chem. Soc.*. **1997**, *119*, 9931-9932.

56. R. H. Kramer, J. W. Karpen, *Nature*. **1998**, *395*, 710-3.

57. A. Bernardi, D. Arosio, D. Potenza, I. Sánchez-Medina, S. Mari, F. J. Cañada, J. Jiménez-Barbero, *Chem. Eur. J.*. **2004**, *10*, 4395.

58. G. Terraneo, D. Potenza, A. Canales, J. Jiménez-Barbero, K. K. Baldridge, A. Bernardi, *J. Am. Chem. Soc.*. **2007**, *129*, 2890-900.

Figures and tables

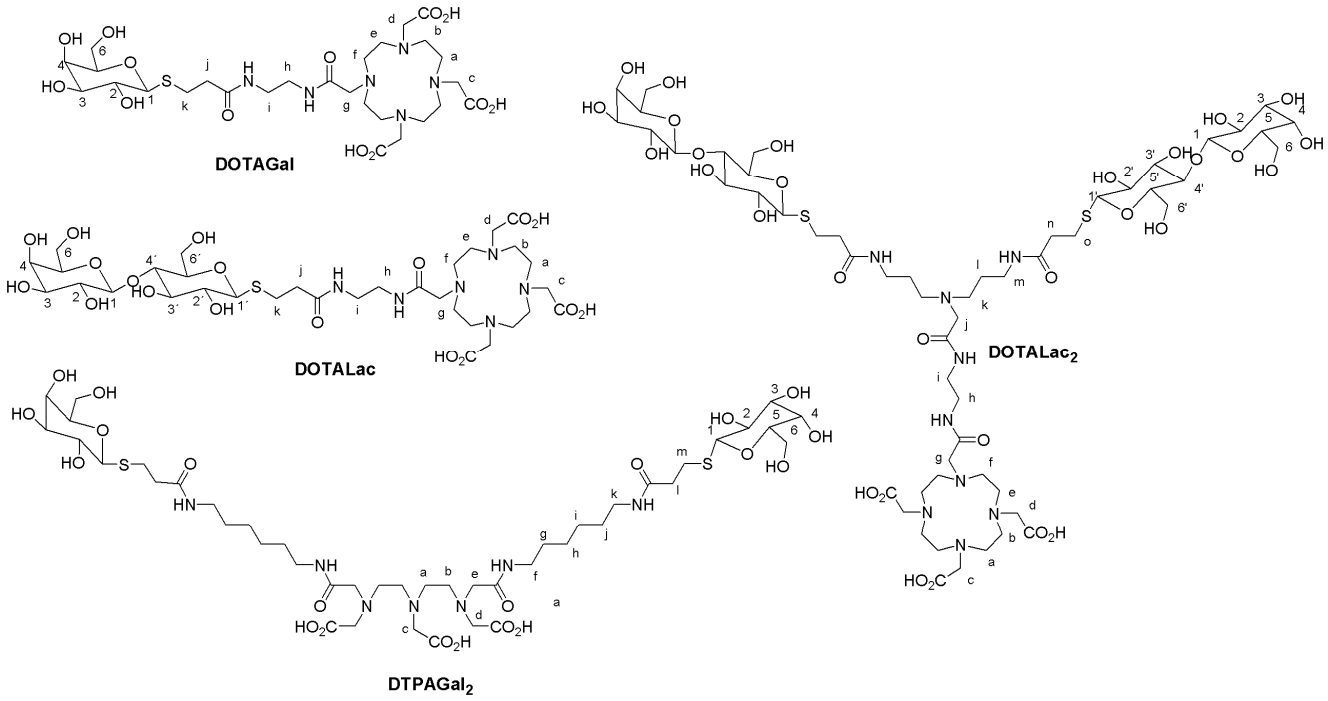


Figure 1. Chemical structures and proton numbering scheme of the DOTA and DTPA glycoconjugates.

Table 1. Summary of the characterised binding epitopes with STD values relative to H2 or H3 in 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

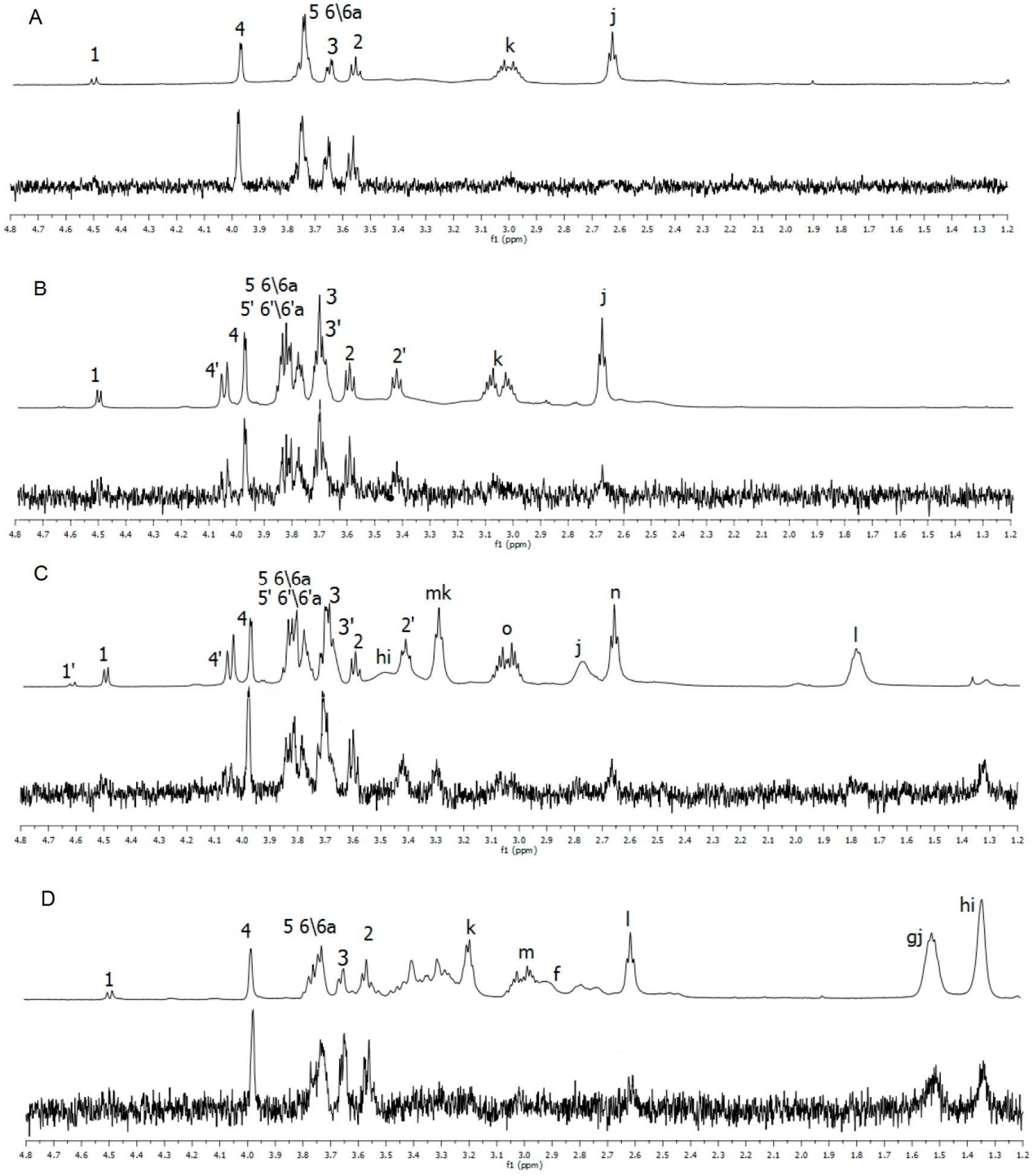
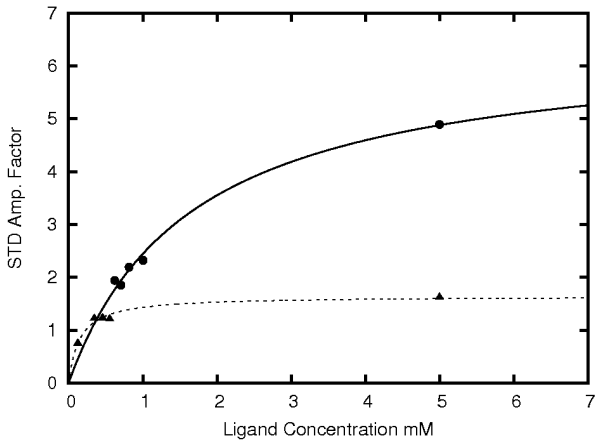


Figure 2. ^{1}H 1D (top) and STD NMR (bottom) of (A) La(DOTAGal) 0.82 mM (B) La(DOTALac) 6.2 mM (C) La(DOTALac$_2$) 1.1 mM (D) La(DTPAGal$_2$) 0.58 mM in the presence of RCA$_{120}$ with the following concentrations (A) 25 μM (B) 10 μM (C) 15 μM and (D) 15 μM. Selective saturation (2.5 s) was performed at the aromatic region of the protein. The spin-lock pulse was calibrated to avoid unwanted protein resonances.

(A)



(B)

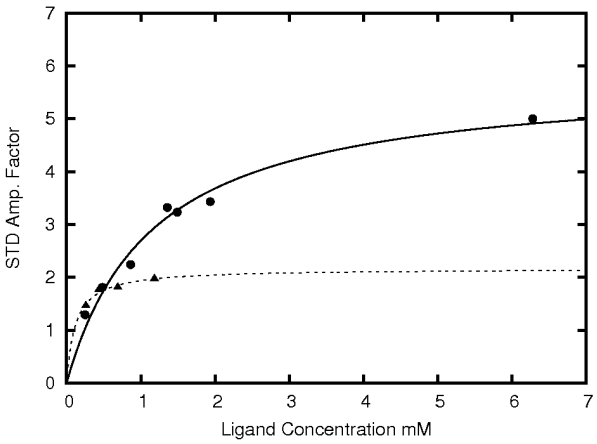


Figure 3. Direct determination of the K_D values of La(III) complexes of (A) DOTAGal (●) and $DTPAGal_2$ (▲) glycoconjugates and (B) DOTALac (●) and $DOTALac_2$ (▲) binding to RCA_{120} by fitting the acquired data points to equation 2.

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

Proton	Individual K_D value [mM]	Max. 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

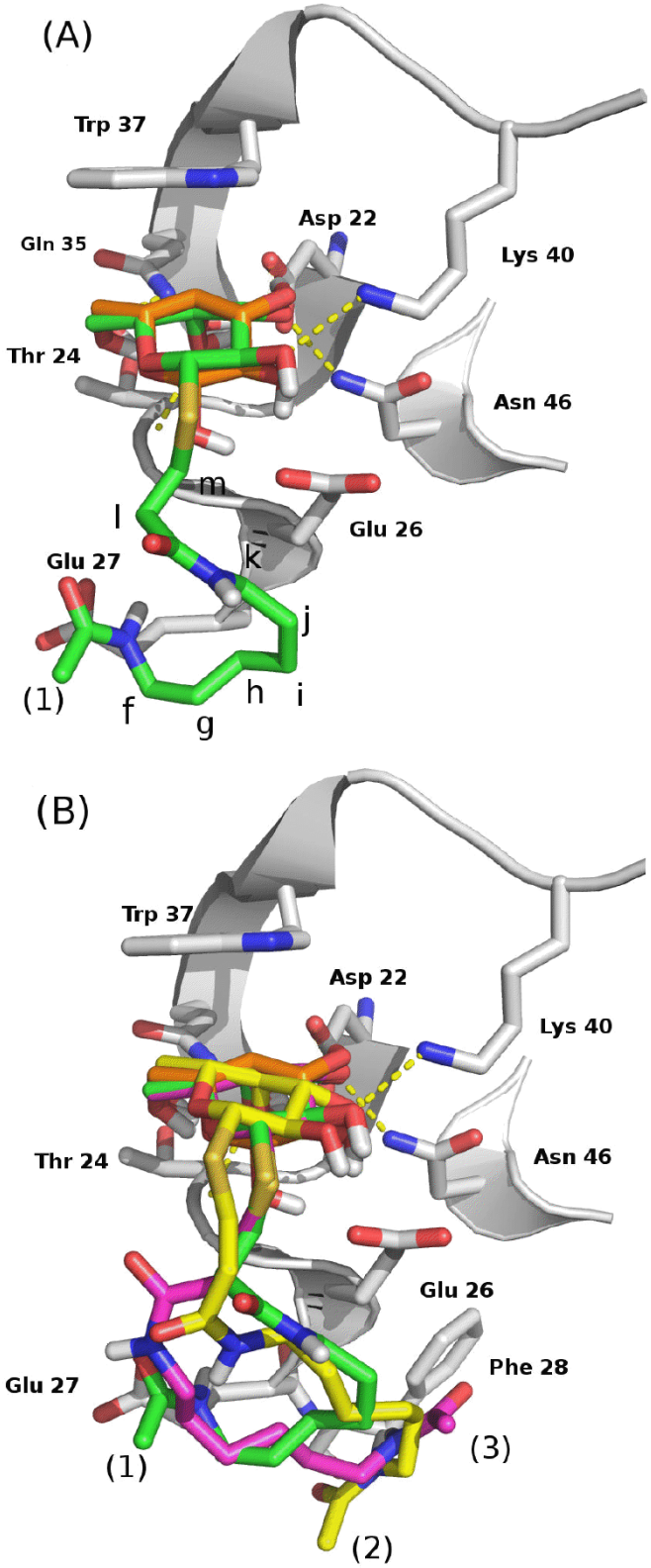


Figure 4. Automated docking structures of one arm of La(DTPAGal$_2$) in the binding site 1α of the Ricinus communis Agglutinin. (A) shows in green one of the resulted runs (B) the three most reliable runs. The single docked galactose molecule was also superimposed to allow a better comparison, and it 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, OH6 and Gln35. CH-π stacking interactions with Trp37 also take place. [52]

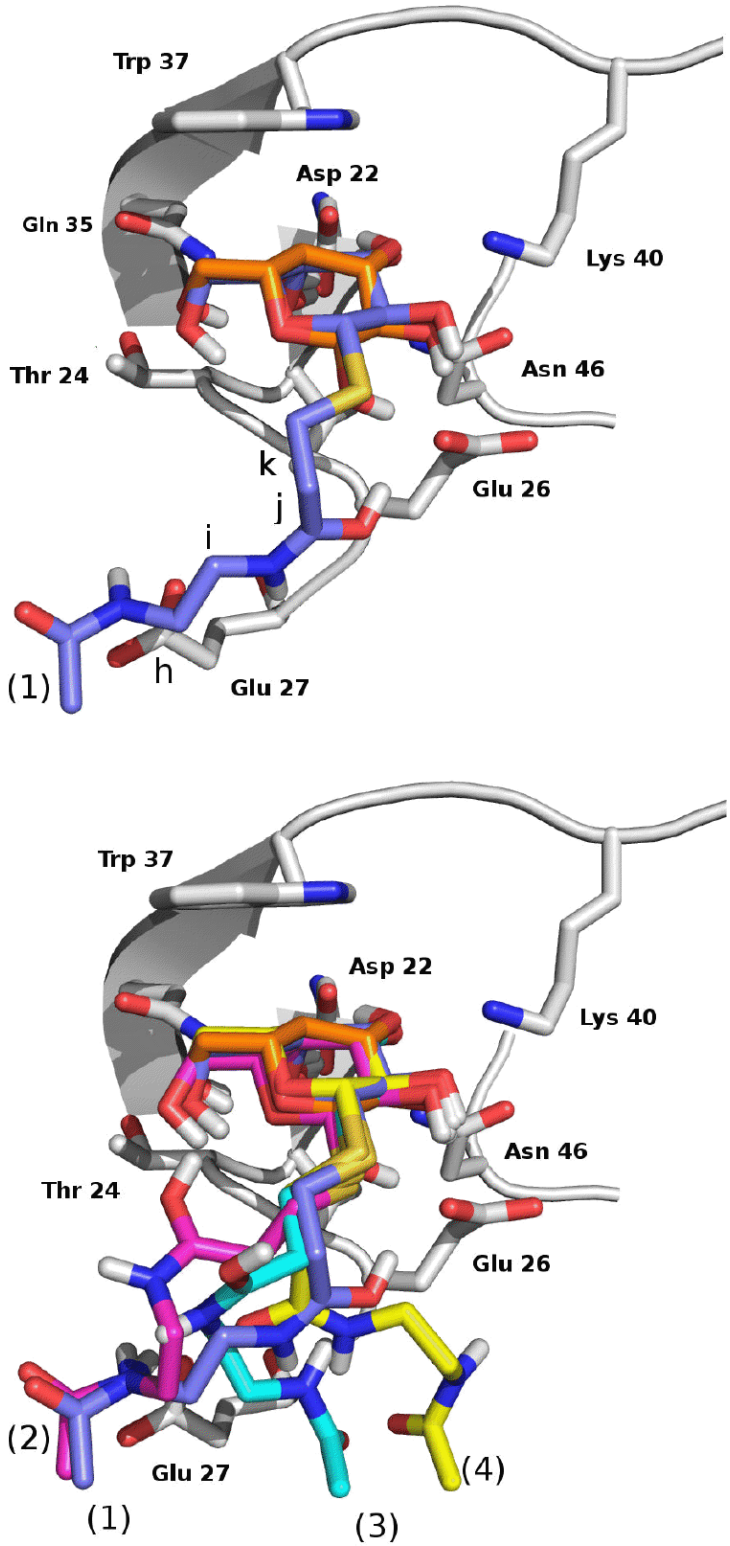


Figure 5. Docking results of the sugar-linker moiety of La(DOTAGal). (A) represents one of the docked results and (B) represents the superimposed view of all the selected possible structures. The number of torsional bonds was 13.

Table 3. H-H linker-protein distance of docked arms of (La)DOTAGal and (La)DTPAGal2. Distances were measured relative to Cα G25, Cβ E26 and Cβ E27 and are presented as average values.

(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	i-h	2.9
h	6	g-j	3.1

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