MedChemComm

Accepted Manuscript

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/medchemcomm

MedChemComm

RSCPublishing

REVIEW

NMR and Molecular recognition. The application of ligand-based NMR methods to monitor molecular interactions

Cite this: DOI: 10.1039/x0xx00000x

Received 25thMarch 2014,
Accepted 00th January 2012

DOI: 10.1039/x0xx00000x

www.rsc.org/

Luca Unione,[a] Silvia Galante,[a] Dolores Díaz,[a] F. Javier Cañada,[a] Jesús Jiménez-Barbero[a]

NMR spectroscopy allows the monitoring of molecular recognition processes in solution. Nowadays, a plethora of NMR methods are available to deduce the key features of the interaction from both the ligand or the receptor points of view (receptor-observed or ligand-observed NMR methods, respectively). Herein, we present the most used ligand-observed NMR experiments for drug discovery making special emphasis on their experimental aspects and on their applications to the study of protein-carbohydrate interactions.

Introduction

Molecular recognition processes are involved in life regulation and progression. Due to the immense structure variability of the entities involved in these events, the code that controls ligand/receptor interactions is highly regulated. In fact, subtle differences at the atomic level may produce dramatic changes in the corresponding biological responses. Therefore, the exploration of these interactions at the maximum possible resolution has become a new frontier in molecular sciences.[1,2]

In this context, although underappreciated in the genomic and post-genomic eras, carbohydrates have now emerged as key players in the translation of information into functional activity, thanks to their large coding capacity and the increasing realization of their key role in receptor-mediated recognition events.[3] Indeed, many sugar-mediated processes are associated with cell–cell interactions, including cell proliferation and other life threatening or positive events.[4] Interactions involving saccharide molecules are also concerned in diverse disease-producing mechanisms, from bacterial adhesion to inflammatory processes.[5,6] Carbohydrate-protein interactions are also rising as medically relevant factors in, for instance, infection and tumour-related studies.[7,8] Thus, the investigation of the possibilities of controlling these events has become challenging in science, from chemistry and biochemistry to molecular and cell biology.

On this basis, the understanding, at high resolution, of the various mechanisms by which the binding sites of antibodies, lectins, and enzymes recognize other molecular entities is of evident importance.[9] From this point of view, it can be safely assumed that achieving a full conception of the energy, structural, conformational, and dynamic features of the molecular recognition processes between carbohydrates and their receptors can be exploited to explain the basic interaction features and hence to design new molecular probes and ultimately, novel sugar-based therapeutic agents.[10] In fact, the detailed study of the 3D structure of different ligand/receptor complexes has led to a deeper understanding of the molecular basis for drug–receptor interactions, and thus to better inhibitors of key enzymes.[11]

The power of X-ray crystallography to provide a detailed description of these features can be acknowledged since has become one of the major alternatives in this field.[12] Nevertheless, in this review article we have paid attention to NMR spectroscopy and its applications to explore molecular interactions.[13] In the last years, NMR has become a powerful tool to monitor molecular interactions and to deduce features of recognition processes at different levels of complexity,

Table 1. Summary of the ligand-based NMR methods cited herein.

Method		Generalities			
		Isotope labelled material	Receptor molecular weight	Throughput ligand screening	Range of applicability ($K_{d)}$
NOE type experiments	STD	no	>>10KDa	yes	nM-mM, high sensitivity
	Tr-NOESY	no	>>10KDa	very moderate	nM-mM, high sensitivity
	Water-LOGSY	no	>>10KDa	yes	nM-mM, high sensitivity
Diffusion based experiments	DOSY	no	>>10KDa	moderate	~100nM-mM, low sensitivity

both from the perspective of the receptor and of the ligand.[14] Nevertheless, there is still an open window for further development of new NMR methods and their applications to unravel molecular recognition problems. Thus, we have aimed to describe significant advances in the applications of ligand-based NMR techniques to understand molecular interactions, especially in the carbohydrate–protein recognition field.[10] It has not been our intention to provide an exhaustive review; we have instead attempted to focus on describing different technical and methodological aspects within this research area, also including examples of molecular interactions for other biomolecules.[15] We have therefore selected different recent examples that emphasize the application of different NMR protocols to circumvent the technical problems that can be often found in this type of investigations.

In contrast to the receptor-based NMR techniques,[16,17] which allow monitoring those changes that take place at the different residues of the interacting macromolecule, ligand-based NMR techniques can also be exploited to detect biomolecular interactions.[18] The advantages of these techniques are that they require very small amounts of the receptor (in the nmol scale) and, especially that labelling with stable isotopes (^{15}N, ^{13}C, ^{2}H) is not required. Nevertheless, there are also drawbacks in the application of a ligand-based NMR protocol. First, changes in the NMR parameters in the ligand only are those under scrutiny. Therefore, no straightforward indication on the protein site that is interacting may be directly extracted. In this case, it is necessary that additional NMR experiments are performed, especially competition experiments, in order to assess the binding site. Moreover, for certain relaxation-based techniques, as STD and trNOESY experiments, there is a strong requirement in the kinetics of the dissociation process.[19] For these experiments to be successful, the off-rate should be fast in the relaxation time scale. Since fast off-rates are usually associated to moderate or weak binding processes, these experiments are mostly used for investigating interaction processes in the mM or μM range. Their application to investigate tight binding events, around nM, can only be exploited in combination with other alternative approaches, including competition experiments. Table 1 gathers the most common ligand-observed NMR methods cited herein together with some of their relevant and practical properties.

Ligand-based NMR methods for drug discovery

Saturation transfer difference NMR (STD NMR)

The Saturation Transfer Difference (STD) experiment permits to deduce the existence of interaction between a small molecule and a target receptor and, in favourable cases, the identification of the ligand epitope.[20] STD has now been used for fifteen years to characterize ligand binding to receptors. The experiment derives from the previously used saturation transfer experiment, which permits to identify chemical exchange in a given molecule that is present in different states (i.e., free and bound), provided that the exchange rate is slow in the chemical shift time scale.[21] Thus, by saturating the NMR signals of the free ligand, the signals of the bound ligand can be identified, since they also receive the saturating radiofrequency field through the corresponding chemical exchange process.[22] The improvement of the method using difference spectroscopy permitted to detect ligand binding to a given receptor, even for mixtures of molecules (chemical libraries) and also for fast exchange.

Technically, the STD NMR experiment is the subtraction of two different ^{1}H-NMR spectra performed on the same ligand/receptor sample. In the first one (on-resonance spectrum), the selective irradiation of some protein protons is performed for a few seconds. The saturation frequency should be chosen so that no ligand protons are present at least within 1-2 ppm. The saturation is efficiently propagated across the entire protein through spin-diffusion, also

saturating those protons at the protein surface, including those located at the binding site and, therefore affecting the NMR active nuclei of any possible molecule interacting at the protein binding site. Thus, the signals from binding molecules display lower intensities in the associated spectrum.

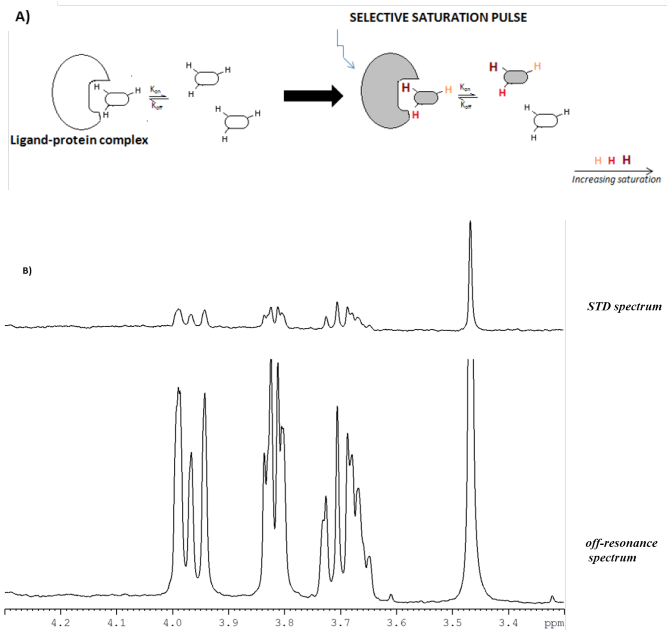

Figure 1. A) Schematic view of the STD experiment. The protein protons are selectively saturated and at a specific frequency (optimal frequency values are those between δ 0 and -2 ppm, since for globular proteins several aliphatic protons close to aromatic rings are present in this region. Alternatively, especially for natural carbohydrates devoid of aromatic protons, the δ 6.5-8.0 ppm region can also be employed). Then, the saturation is transferred to any ligand interacting with the protein, affecting the ligand protons signals differently, depending on their proximity to the protein protons (as indicated with the different colours and font sizes used for the ligand displayed in the figure). B) STD NMR experiment with concanavalin A and methyl α-mannoside at a 1:30 protein/ligand ratio.

Moreover, the intensity of the protons of the ligand will be differently affected, depending on their distance to the protein-binding site. Any putative ligand molecule that is not interacting with the receptor does not receive any saturation from the protein and, therefore, the saturation pulse does not affect its intensities. The second spectrum (off-resonance spectrum) is recorded as reference or blank spectrum. Herein, the chosen saturation region is free from ligand and protein signals (ca. δ 100 ppm). Therefore, the NMR signals of all molecules display their natural intensities (Fig. 1). Then, subtracting the on-resonance from the off-resonance spectra, the obtained STD spectrum only contains the ligand resonances whose intensities have been affected by the protein saturation. Thus, the STD technique permits to differentiate binders from non-binders in a complex mixture.

This experiment does not require a large amount of protein (actually 1:20 to 1:100 protein/ligand molar ratios are recommended, with a protein concentration less than 20 μM). As mentioned above, the suitable affinity values (Kd) are between 10^{-3} and 10^{-7} M. In fact, given the intrinsic properties of carbohydrate-protein interactions, they usually display moderate binding affinities, with optimal properties for the application of STD spectroscopy.

Different applications of STD to screen carbohydrates within chemical libraries or even plant extracts towards different receptors have been reported.[23] One of the key applications of STD relies on the possibility of deducing the binding epitope of the ligand. Since the protons closer to the protein are more effectively saturated, they display the largest intensities in the difference spectrum. This STD-based epitope mapping procedure provides essential information for drug lead optimization and its concomitant application in a structure-based drug design protocol.[24]

Remarkably, STD-based intensities can be employed for determining the inhibition constant (Ki) of a given ligand that competes with a reference ligand (with a known Kd). Competition experiments can be performed for designed mixtures of both molecules by evaluation of the gradual decay of the STD signals from the reference in the presence of increasing concentrations of the inhibitor.[25]

STD NMR has been employed to characterize the binding epitope of two glycomimetic platforms (thiodigalactoside and dithiodigalactoside). Of note, despite its apparent similarity, STD could demonstrate that the binding epitope of these pseudodisaccharides is different for human galectin-1, with medical relevance in tumor growth and methastasis, and for the plant toxin VAA (Viscum Album Agglutinin), a potent ribosomal inactivator with modulatory effect on the natural immune system.[26]

Although STD is commonly employed to monitor the binding of small molecules and fragments to the protein targets, recent studies have expanded the scope of the method to larger systems, much closer to the natural interacting partners.[27] For instance, N-glycans, which are ubiquitous in mammalian glycoproteins, may contain different epitopes in their structures, displaying multiple possibilities for recognition. In this case, the concatenation of different NMR experiments to the basic STD module helps to extract conclusions on the interacting epitope that cannot be accessed from simple 1D measurements. In particular, combinations of STD with TOCSY,[28] selective TOCSY,[29] or HSQC/HMQC [30] have been proposed, providing spectra with optimal dispersion of the STD signals, now devoid of significant overlap. Nevertheless, the acquisition time for these experiments may be significantly large. Using STD and STD-TOCSY experiments, it has been demonstrated that the interaction mode of large N-glycans towards a given receptor may differ from that existing for small model fragments of the same molecule. In fact, the presentation mode of the N-glycan is essential for the interaction to take place. Even new interacting epitopes may appear within a large structure, not possible for smaller systems. These results indicate that care should be taken when trying to extrapolate the obtained results for molecular fragments to the complete entity. Moreover, it should be emphasized that those conclusions arising from macroscopic measurements (i.e., SPR, ITC, glycan or lectin arrays) without detailed structural investigations may be misleading.[27]

Moreover, if the ligand contains NMR-active nuclei other than protons, filtering/editing schemes may also be employed to characterize binding. Examples for ^{13}C or ^{15}N-containing ligands have been reported.[31] For ^{19}F-containing molecules, an alternative STD approach focused on the detection of heteronuclear $^{1}H \rightarrow ^{19}F$ STD has been designed to study fluorinated carbohydrates as lectin ligands.[32] Obviously, the presence of NMR heteronuclei also open the avenue to perform NMR experiments in which the hetero atom acts as "spy" probe to monitor binding of other molecules present in the NMR tube by using competition experiments. Dalvit has exhaustively employed this approach in the drug design context, by using ^{19}F-containing molecules and T_1- and T_2-based experiments.[33] Obviously, the relaxation properties of the ligand nuclei dramatically change when passing from the free to the bound state, being the fluorine atom the proper sensor to detect binding of the ^{19}F-containing molecule or others, provided that they compete by the same protein-binding site. A similar approach has been used applying the $^{1}H \rightarrow ^{19}F$ STD experiment.[34]

In the case of ligand interactions with nucleic acids, STD not only provide evidences on binding epitopes, but also on the groove of DNA that may be involved in the interaction.[35] Due to the relatively few protons existing in DNA, the intermolecular saturation transfer may be restricted with a particular area of the DNA structure, thus permitting the assessment of local features of the recognition process.[36]

STD-NMR has also been applied to monitor interactions of ligands in the presence of more complex systems, including membrane preparations,[37] microtubule assemblies,[38] living cells,[39] or even bacteria,[40] and viruses.[41] These experiments open new avenues in the drug discovery context, also approaching the possibilities of studying ligand binding to G-protein coupled receptors (GPCRs)[42] and other membrane proteins that are very difficult to isolate and manipulate. Different examples expanding the application of solution based NMR spectroscopy have been presented. Alternatively, in the study of interactions with living cells,[43] it has been proposed that the method can also be safely combined with HR-MAS protocols, providing very good quality STD spectra.

Transferred NOE

As complement to the STD method, the exchange-transferred NOE experiment (tr-NOESY) allows the conformation of a particular ligand at the receptor's binding site to be deduced.[44] The tr-NOESY technique is based on the changes in the rotational motion properties of a ligand upon binding to a large macromolecular receptor.[45,46,47] For small molecules, whose rotational motion correlation time lies in the ps timescale, NOEs are positive. However, for large receptors, the correlation time is in the ns time scale and the associated NOEs are negative (Fig. 2).

Thus, when a small ligand binds to a given receptor, its correlation time dramatically changes. In consequence, the NOEs move from positive to negative values (Fig. 3). Therefore, the change in the NOEs cross peaks sign is a simple and clear evidence of the existence of binding. The major disadvantage of the tr-NOESY approach is the narrow accessible kinetic window. For the experiment to be successful, the off-rate should be fast in the relaxation time scale, (i.e., a few times per second). Otherwise, the NOE information acquired by the ligand at the receptor site will fade away by relaxation before it dissociates from the protein.

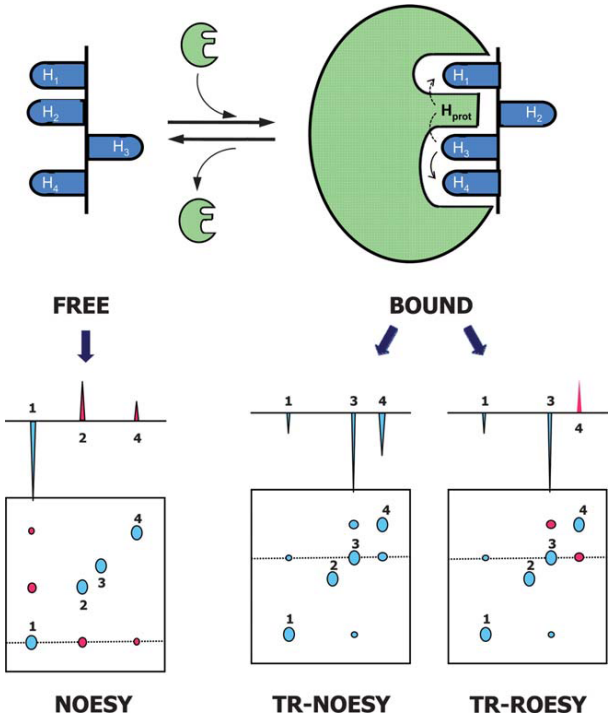

Figure 2. Schematic view of the tr-NOESY experiment for an exchanging ligand-receptor system. For a small molecule, the cross-peaks are positive and display opposite sign to the diagonal peaks. In the presence of the receptor, the cross-peaks for the complex change their signs and become negative, due to the increase of the effective rotational correlation time. The analysis of the negative cross-peaks provides information on the bound geometry of the ligand. The possible existence of spin diffusion effects can complicate tr-NOESY analysis. In this case, spin-diffusion-mediated cross-peaks can be detected by employing tr-ROESY experiments, due to the alternation of sign between direct and indirect effects.

The tr-NOESY approach can be also applied to discriminate ligand binders in complex mixtures, even when significant structural similarities could lead to extensive signal overlap.[48] In principle, tr-NOESY (as well as STD) have the advantage that separation of the components of a library prior to activity tests is not necessary. Small amounts of receptor are required, with protein/ligand molar ratios usually ranging from 1:5 to 1:50, depending on the affinity and kinetic parameters.

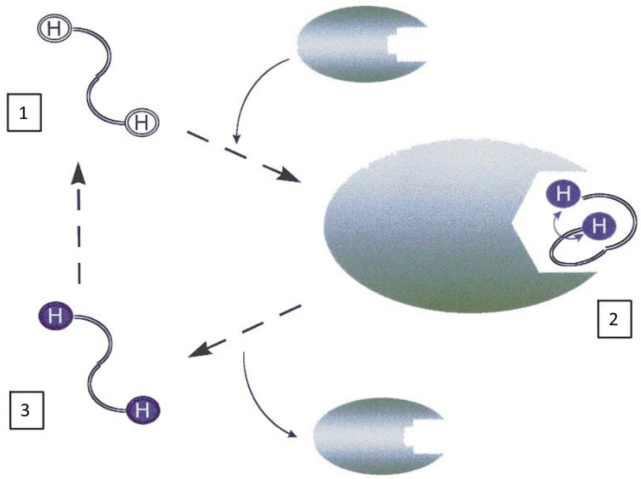

Figure 3. Schematic representation of the dynamic equilibrium between free and bound states. (1) Initial state: the free ligand displays fast rotational motion properties. (2) Formation of the complex: the bound ligand now rotates with the protein. (3) Dissociation of the complex. The free ligand retains the NOE features acquired in the bound state and produces negative NOEs, provided that the off-rate is fast in the relaxation timescale. After some time, the system reverts to the initial state.

For gaining quantitative information, it is essential to register different experiments with variations in the NOE build experimental parameters in a systematic manner, (i.e. mixing times) as well as employing several ligand/protein molar ratios.

Many examples have focused on the study of the structural events associated with molecular recognition processes. Indeed, different structural and conformational recognition patterns may take place. In many cases, the ligand NOE cross peak pattern in the free and bound state is very similar. In fact, for these systems, protein binding sites are well pre-organized to recognize one conformation of the ligand, which is close to its global minimum energy region. This condition considerably reduces the entropic cost for the binding. One example of this phenomenon has been described for the recognition of the microtubule-stabilizing agent discodermolide by tubulin dimers and by assembled tubulin.[49] This phenomenon has also been evidenced for other microtubule-stabilizing agents, whose conformations are fairly well defined in solution.[50]

However, there are several examples in which the rigid lock and key model fails. In this case, the plasticity of one or both partners involved in the interaction is the only possible explanation.

For instance, a case of protein-induced conformational selection has been reported when studying the interaction of human galectin-1 (hGal-1) with C-glycosyl compounds, as hydrolytically stable glycomimetics.[51] Although the C-glycosyl molecule exists as a complex conformation equilibrium with four distinct conformers in solution, the tr-NOESY experiment pointed out that hGal-1 only selects one of these geometries, which strikingly does not correspond to the global minimum structure.[52]

These tr-NOESY experiments also provide a robust manner to detect protein-induced conformational variations. The comparison of the NOESY acquired for the ligand in the isolated state with that in the presence of receptor provides a quick mean to detect conformational distortions. For instance, binding of rosmarinic acid to acetylcholinesterase (AChE) has been reported[53] to occur with a change in the conformation of the molecule, from an extended to a hairpin-like conformation. The tr-NOESY experiment showed a dramatic difference of the NOE cross peak patterns between the free and bound states. Noteworthy, the NMR study (STD and tr-NOESY) of the interaction of *salvia sclareoides* extracts with AChE permitted to identify this molecule as the active component in the plant.[53] In fact, the STD experiments of *S. sclareoides* crude extracts in the presence of AChE determined that rosmarinic acid was the only binder for AChE.

Depending on the architecture of the binding site of the receptor, different situations may take place, even for the same type of molecules. Different examples have focused on the study of heparin, heparin sulphate, and other glycosaminoglycans (GAGs), which regulate the activity of a wide variety of proteins. It is noteworthy to mention that these molecules are usually highly sulphated and that the presence of L-iduronic rings endows high flexibility to these entities. In particular, the conformational features of the interaction of the antithrombin-heparin pentasaccharide with the extracellular Ig2 domain of the fibroblast growth factor receptor (FGFR2) has been evaluated.[54] In the free state, the iduronate residue is in equilibrium between a skew boat ($^2S_0$) and a chair ($^1C_4$) conformation. However, when bound to the fibroblast growth factor receptor FGFR2, the NOE data, acquired thorough a ^{13}C filtered NOESY experiment, showed that the iduronate residue only displays the $^2S_0$ conformer. Therefore, a conformational selection process takes place. In contrast, the analysis of the binding of a similar heparin hexasaccharide to the acidic fibroblast growth factor FGF1 showed that this protein recognizes this ligand maintaining its intrinsic flexibility. Both chair and skew boat conformers are recognized in this case.[55] This event, involving the simultaneous recognition of different conformations has also been described for different dynamic ligands. For instance, for glycomimetics with thioglycoside chemical nature,[56] the energy barrier for rotation around the glycosidic torsions is low enough to access different geometries in the free state. The study of the interaction of these glycosidase inhibitors with a model glycosidase enzyme showed that the enzyme is able to accommodate different distorted geometries, which probably mimic the transition state for glycosidase-mediated hydrolysis.[56]

Ligand binding detection through bulk water: waterLOGSY

Although most of the ligand-observed NMR methods described above do not take explicitly into account the role that the surrounding solvent molecules may play during complex formation, the influence of protein solvation on ligand binding processes is evident.[57] The presence of water molecules at the binding site of both free and bound receptors has been confirmed by X-ray crystallography as well as by NMR spectroscopy, demonstrating that water molecules may mediate specific contacts between the ligand and the protein residues at the binding site.[58,59] As a matter of fact, the detection of water-ligand intermolecular NOEs have been taken

as a proof of the existence of water molecules squeezed at the interface of the protein-ligand complex and/or tightly bound through hydrogen bonds with the free ligand.[60,61,62] Based on these experimental observations, the so-called waterLOGSY experiment (water-ligand observed via gradient spectroscopy) has become one of the most popular NMR-based methods applied to target-directed drug discovery.[63,64] WaterLOGSY allows distinguishing between binder and non-binder molecules by transfer of magnetization from the bulk water, via the protein-ligand complex, to the free ligand.

The regular 1D waterLOGSYspectrum allows the easy and visual identification of active ligands from a ligand mixture in the presence of a potential receptor. In this experiment, the NMR resonances of free and protein-bound molecules exhibit different phase properties. While the signals of the interacting ligands are positive, the non-interacting ones are negative. This technique has become very popular in drug discovery, particularly as a fragment-based NMR screening experiment (Fig.4). The experiment is very well suited for medium-throughput screening of large libraries of compounds since its sensitivity is relatively high and a reduced amount of material (biomolecules and ligands) is needed.[65] Thus, the method has been employed for initial identification of possible hits from a fragment library of potential inhibitors targeting, for example, an alternative pocket on HIV-1 envelope glycoprotein gp41[66] or to evaluate the interaction of previously *in silico*-selected drugs with heparanase[67].

The main drawback of waterLOGSY NMR based experiments (as for STD) is that strongly binding ligands with slow dissociation rates cannot be directly detected. This problem can be easily overcome by using a reference compound with an intermediate binding affinity for the receptor and, as it happens with STD NMR, designing appropriate competition binding waterLOGSY experiments.[68] The effectiveness of this methodology was first demonstrated by screening a molecular library against human serum albumin (HSA) and using 6-methyl tryptophan as reference compound. In particular, the waterLOGSY signal intensity of the reference compound in the presence and absence of other ligands was compared. Changes in the sign or the intensity of the reference signal in the presence of the other molecules were related to the existence of a better binder in the mixture.[68]

Investigations on the blood coagulation process have also profited from this technique. Since most widely anticoagulants (i.e. heparin and warfarin) appear to exhibit adverse side effects and handling problems, pharmaceutical companies are looking for alternative drugs. Thus, thrombin has been envisaged as a potential target for direct inhibition. Indeed, water-LOGSY experiments have been employed as primary source of information to identify specific neutral fragments that bind to thrombin.[69]

Human serum albumin (HSA) is one of the most important plasma transport proteins. Thus, from a pharmacokinetic perspective, the complete understanding of how drugs interact with HSA is of importance for the development of new drugs. In this context, the interaction of a series of antimicrobial peptides with HSA has been investigated using competitive water-LOGSY experiments.[70]

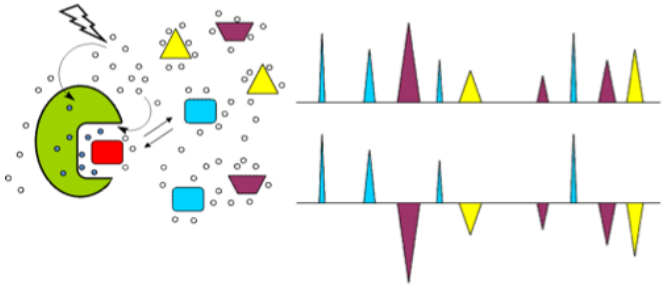

Figure 4. Schematic description of the waterLOGSY experiment. Left. The magnetization of the bulk water is selectively transferred, via the protein-ligand complex, to the ligand in the unbound state. Right. Representation of the regular NMR spectrum of the mixture (above) and of the resulting 1D waterLOGSY spectrum (below), where binder and non-binders ligands can be identified in a straightforward manner.

Other variants of the waterLOGSY experiment as the aroma-waterLOGSY experiment[71] or the polarization optimized waterLOGSY method (PO-waterLOGSY)[72] provide significant sensitivity improvement. In these cases, the fundamental waterLOGSY scheme is modified by including longer delays for the water signal to relax. Thus, aroma-waterLOGSY selectively identifies the aromatic signals from the ligand, while PO-waterLOGSY significantly improves the time efficiency.

The range of applicability of this technique is not restricted to the field of ligand screening. Other interesting applications have also been reported. The SALMON method (solvent accessibility, ligand binding, and mapping of ligand orientation by NMR spectroscopy) has been developed to detect not only ligand-receptor interactions, but also to investigate the orientation of the ligand with respect to the receptor. The SALMON experiments permits the deduction of the portions of the ligand that remain more exposed to the solvent upon binding (ligand solvent accessibility epitope).[73] This approach has been employed to deduce the relative orientation of a serial of antimicrobial peptides in detergent membranes. The detection of A variety of water-peptide contacts are detected for those peptide residues that remain solvent-exposed, in contrast with the absence of water-peptide interactions for those that are embedded into the membrane.[74]

An interesting application of this technique has shown to be very useful to detect structural water molecules,[75] previously identified by X-ray crystallography in the complex formed by a peptidic carbohydrate mimic and the anti-carbohydrate antibody SYA/J6. The combination of STD and waterLOGSY NMR experiments with molecular modelling protocols confirmed the presence of bound water molecules in the active site of the receptor. This study exemplifies how the combined use of NMR data and computational methods can be used for unravelling complex interaction processes.

Most often, ligand-observed NMR experiments employ ^{1}H detection. It is not unusual that the analysis of the resulting spectra results very difficult due to extreme overlapping of the ^{1}H NMR signals of the ligands present in the mixture. Therefore, as mentioned above, the use of fluorine atoms in the tested ligand molecules could be advisable. Thus, screening of large chemical mixtures may be

successfully achieved. In this context, a complete binding assay dubbed FAXS (fluorine chemical shift anisotropy and exchange for screening) has been developed to identify binders from a mixture of fluorinated molecules against the target biomolecule.[76] The method requires the presence of a fluorinated low affinity ligand in the mixture, the so-called "spy molecule".

NMR-based diffusion experiments

Although other ligand-based NMR experiments may be employed to monitor interactions, such as those involving paramagnetic agents[77] or residual dipolar couplings,[78] we will finish this review by focusing on the application of Diffusion Ordered spectroscopy (DOSY)[79] experiments in this field. DOSY data are presented as a 2D plot, as schematized in Fig.5. The vertical axis represents the diffusion coefficient values, while the abscissas' shows the regular 1D spectrum of the sample. The final plot somehow resembles the physical separation obtained by chromatography and the experiment is also known as NMR chromatography. Since the diffusion coefficient of a molecule depends on its size, molecular weight, and shape, this experiment has been found extremely useful in the field of molecular recognition and supramolecular chemistry to follow the formation of ligand-receptor and host-guest complexes, respectively. Obviously, the diffusion coefficients of the involved species are altered as consequence of the formation of the complex.[80,81]

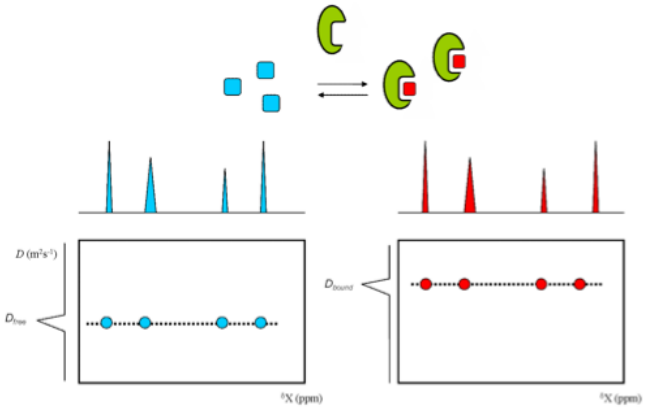

Figure 5. Schematic view of the DOSY experiment. The diffusion coefficient of the bound ligand is significantly larger than that of the free species.

In the molecular recognition field, NMR-based diffusion experiments have been used to describe protein aggregation[82] and folding,[83] as well as the formation of ligand-receptor complexes.[84] In the case of ligand binding, the hydrodynamic radius of any bound ligand molecule will apparently increase as consequence of the interaction with the receptor. Thus, ligand-binding processes can be monitored by tracing the observed changes in the diffusion coefficient of the ligand in the presence of the receptor. The formation of complexes between the antimicrobial peptide defensine and a variety of glycosaminoglycans has been demosntrated by detecting changes in the diffusion coefficient values of the species in both free and bound form.[85]

DOSY can simply be used to calculate ligand dissociation constants, as well as the stoichiometry of the process.[86] For instance, DOSY has been employed to evaluate the binding affinity of the anti-tumour agent topotecan (TPT) to DNA.[87] For screening purposes, DOSY can also be used to selectively detect the active ligands in a complex mixture that bind to a target protein. In contrast to the bound ligands, the diffusion coefficient of the non-interacting molecules in the mixture will remain unaffected. As mentioned above, this virtual separation of active and non-active compounds is conceptually comparable to affinity chromatography and this method has been therefore dubbed affinity NMR.[83,88]

For complex mixtures, signal overlapping may difficult the proper identification of the active species. It is then convenient to use the so-called DECODES method (Diffusion Encoded Spectroscopy-DECODES) that combines DOSY with a TOCSY experiment. Then, the resulting much less overcrowded spectrum allows the easy identification of the active species.[89,90]

Besides combining DOSY with other NMR methods, signal overlapping can be also overcome by employing other NMR active nuclei in the ligand molecule (e.g. ^{31}P, ^{19}F,...).[91,92]As example, in the study of the enzymatic reaction of a series of phosphorus-containing substrates with α-chymotrypsin, the resulting molecules that were covalently bound to the enzyme could be readily identified from their distinctive diffusion coefficient.[91]

A similar approach has been used that relies on the presence of fluorine tags in the ligand. The complexity of the acquired NMR spectra is easily reduced by simply monitoring the ^{19}F nuclei. Hence, identification of the bound ligands may be achieved by analyzing several NMR parameters that are affected upon binding (i.e. broadening of signals, waterLOGSY experiments,...), including the diffusion coefficients (DOSY) and the ^{19}F-^{1}H long-range connectivities (COSY). The absence of overlap in the ^{19}F spectra results in high quality DOSY spectra that are devoid of artefacts Following this approach, different ligands from a mixture of fluorinated molecules in the presence of bovine pancreatic trypsin have been identified.[92]

From a technical viewpoint, the overlapping of both ligand and receptor signals may lead to large errors in the calculation of the ligand diffusion coefficient. Thus, strategies to improve the precision of the measurement range have been proposed.[93] For instance, the suppression of the protein background signals has been implemented by simple spectral subtraction of the protein spectrum.[94]

DOSY can also be useful for epitope mapping, as complement to STD experiments. In fact, the experimental conditions employed to measure the ligand diffusion coefficient allow for a concomitant proton-proton magnetization transfer in the protein-ligand complex. Moreover, the intermolecular NOE transfers between specific parts of the ligand and the receptor produce slight changes in the loss of the intensities for each ligand resonance signal along the set of experiments that are needed to calculate the diffusion coefficient. These different intensity losses may permit to infer the ligand epitope, as exemplified in the study of the binding of trimethoprim to dihydrofolate reductase.[95] This protocol is particularly well suited for systems whose off-rate is much faster than cross relaxation.

Otherwise, spin diffusion prevails and no selective intermolecular NOE can be detected.[96]

Conclusions& Perspectives

In some cases, the NMR-based detection of the key features of molecular recognition processes may still be a complex problem. Nevertheless, new advances in new spectroscopic methods, advances in protein biochemistry, and the access to larger magnetic fields are taking place. For instance, the use of paramagnetic metals as tags, the employment of MAS or HR-MAS spectroscopy methods, together with new computational tools are examples of new developments. Hopefully, it will be early demonstrated for the successful application of these protocols that the only limitation is in our imagination.

Acknowledgements

We thank all the colleagues that have collaborated with us whose scientific works are reported here as examples. We thank the Ministery of Economy and Competitiveness of Spain for funding (CTQ2012-32025) and Comunidad de Madrid (MHit project). We also thank the Marie Curie Actions (DYNANO PITN-GA-2011-289033 and GLYCOPHARM PITN-GA-2012-317297 grants, respectively).

Notes and references

[a]Chemical and PhysicalBiology, Centro de Investigaciones Biológicas, CSIC, Ramiro de Maeztu 9, 28040 Madrid, Spain

1J. F. Rakus and L. K. Mahal, *AnalyticalChem. AnnualReview*, 2011, 4, 367-392.

2 M. F. García-Mayoral, A. Canales, D. Díaz, J. López-Prados, M. Moussaoui, J.L. de Paz, J. Angulo, P.M. Nieto, J. Jiménez-Barbero, E. Boix, E. Bruix, *ACS Chem. Biol*., 2013, **8**, 144-151.

3 H. J. Gabius, S. André, J. Jiménez-Barbero, A. Romero and D. Solís, *Cell Death Dis*., 2011, **36**, 298-313.

4 K. Fuchs, A. Hippe, A. Schmaus, B. Homey, J.P. Sleeman, V. Orian-Rousseau,*Cell Death Dis.*2013, **4**, e819.

5 C. Chemani, A. Imberty, S. de Bentzmann, M. Pierre, M. Wimmerová, B.P. Guery,K. Faure, *Infect Immun.*2009, **77**(**5**), 2065-75

6 A. De Soyza, F. Di Lorenzo, A. Silipo and A.Molinaro, *Carbohydr. Chem*., 2012, **38**, 13-39.

7 F.T.Liu, G.A.Rabinovich, *Nat Rev Cancer* 2005, **5**, 29-41.

8 D.H.Dube, C.R.Bertozzi, *Nat. Rev. Drug Discov*. 2005, **4**, 477-488.

9 J. Kopitz, Z.Fik, S.Andre,K. Smetana and H. J.Gabius, Am. Chem. Soc., 2013, **10**, 2054-61.

10 M. del Carmen Fernández-Alonso, D. Díaz, M. A.Berbis, F. Marcelo,J. Cañada, and J. Jiménez-Barbero, *Curr. ProteinPept. Sci.*2012;**13**, 816-830.

11 J. J. Lozano,M. Pastor,G.Cruciani,F.Sanzetal, *J. Comp.Aided Molecular Design*,2000,**14**, 341-353.

12 M.J. Borrok, L. L. Kiessling, K. T. Forest, *J. Prot. Sci*., 2007, **16**, 1032-1041.

13 G. Ortega,D. Castaño, T. Diercks, O. Millet, *J Am Chem Soc*., 2012, **134**, 19869-76.

14 M. Bartoloni,B. E. Domínguez, E. Dragoni, B. Richichi, M. Fragai, S. André,H. J. Gabius, A. Ardá, C.Luchinat, J. Jiménez-Barbero, C.Nativi, *Chem. Eur. J*., 2012, **19**, 1896-1902.

15 C.Coderch, Y. Tang, J. Klett, S.E. Zhang, Y.T. Ma, W. Shaorong, R. Matesanz, B. Pera, A. Canales, J. Jiménez-Barbero, A. Morreale, J.F. Díaz, W.S. Fang, F. Gago, *Org. Biomol. Chem*., 2013, **11**, 3046–3056.

16 H.Ippel,M. C. Miller,M. A. Berbís, D. Suylen, S. André, T.M. Hackeng, F.J. Cañada, C. Weber, H.J. Gabius, J. Jiménez-Barbero, K.H. Mayo, *Biomol. NMR Assign*., 2014, online.

17 J.L.Asensio, H. C. Siebert, C.W. von Der Lieth, J. Laynez, M. Bruix, U.M. Soedjanaamadja, J.J. Beintema, F.J. Cañada, J.Jiménez-Barbero, *PROTEINS: Struc.Funct. Gen*., 2000, **40**, 218–236.

18 H. Kogelberg, D. Solís, J. Jiménez-Barbero, *Curr. Opin.Struc.Biol*., 2003, **13**, 646-653.

19 L.Herfurth, B. Ernst,B. Wagner, D. Ricklin, D.S. Strasser, J.L. Magnani, A.J. Benie, T. Peters, *J. Med. Chem*. 2005, **48**, 6879-6886.

20 M. Mayer, B.Meyer, *Angew. Chem. Int. Ed.Engl*. 1999, **38**, 1784-1788.

21 P. J. Cayley, J. P. Albrand, J. Feeney, G. C. Roberts, E. A. Piper,A. S. Burgen, *Biochemistry* 1979,**18**, 3886–3895.

22 B. Meyer, T. Peters, *Angew. Chem. Int. Ed.Engl*. 2003, **42**, 864–890.

23 J.P. Ribeiro, S. André, F.J. Cañada, H.J. Gabius, A.P. Butera,R.J. Alves, J. Jiménez-Barbero, *ChemMedChem* 2010, **5**, 415–419.

24 C. Sandström, O. Berteau, E. Gemma, S. Oscarson, L. Kenne, A. M. Gronenborn, *Biochemistry* 2004, **43**, 13926-13931.

25 F. Marcelo, C. Dias, A. Martins, P.J. Madeira,T. Jorge, M.H. Florêncio, F.J. Cañada, E.J. Cabrita, J. Jiménez-Barbero, A.P. Rauter, *Chem. Eur. J*. 2013, **19**, 6641–6649.

26 S. Martín-Santamaría, S. André, E. Buzamet, R. Caraballo, G. Fernandéz-Cureses, M. Morando, J.P. Ribeiro, K. Ramirez-Gualito, B. de Pascual-Teresa, F.J. Cañada, M. Menéndez, O. Ramström, J. Jiménez-Barbero, D. Solis, H.J. Gabius, *Org. Biomol. Chem*. 2011, **9**, 5445-5455.

27 A. Arda, P. Blasco, D. Varon Silva, V. Schubert, S. André, M. Bruix, F.J. Cañada, H.J. Gabius, C. Unverzagt, J. Jiménez-Barbero, *J.Am.Chem.Soc*. 2013, **135**, 2667-2675.

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

29 M.A. Johnson, B.M. Pinto, *J. Am. Chem. Soc*. 2002, **124**, 15368-15374.

30 M. Vogther, T. Peters, *J. Am. Chem. Soc.* 2000, **122**, 6093–6099.

31 K. Feher, P. Groves, G. Batta, J. Jimenez-Barbero, C. Muhle-Goll, K.E. Kövér, *J. Am. Chem. Soc*. 2008, **130**, 17148–17153; J.L. Wagstaff, S.L. Taylor, M.J. Howard, *Mol. BioSyst*. 2013, **9**, 571—577.

32 L.P. Calle, F.J. Cañada, J. Jiménez-Barbero, *Nat. Prod. Rep*. 2011, **28**, 1118 – 1125.

33 B.J. Stockman, C. Dalvit , *Progress in Nuclear Magnetic Resonance Spectroscopy* 2002, **41**, 187–231.

34 T. Diercks, J.P. Ribeiro, F.J. Cañada, S. André, J. Jiménez-Barbero, H.J. Gabius, *Chem. Eur. J*. 2009, **15**, 5666-5668.

35 F. Souard, E. Muñoz, P. Peñalver, C. Badía, R. del Villar-Guerra, J.L. Asensio, J. Jiménez-Barbero, C. Vicent, *Chem. Eur. J*.2008, **14**, 2435–2442.

36 S. Di Micco, C. Bassarello, G. Bifulco, R. Riccio, L. Gomez-Paloma, *Angew.Chem*., 2006, **118**, 230–234.

37 F.M. Assadi-Porter, M. Tonelli, E. Maillet, K. Hallenga, O. Benard, M. Max, J.L. Markley, *J. Am. Chem. Soc*. 2008, **130**, 7212–7213.

38 A. Canales, J. Rodríguez-Salarichs, C. Trigili, L. Nieto, C. Coderch, J.M. Andreu, I. Paterson, J. Jiménez-Barbero, J.F. Díaz, *ACS Chem. Biol*. 2011, **6**, 789–799.

39 B. Claasen, M. Axmann, M. Meinecke, B. Meyer, *J. Am. Chem. Soc.* 2005, **127**, 916-919; S. Mari, D. Serrano-Gómez, F.J. Cañada, A.L. Corbí, J. Jiménez-Barbero, *Angew. Chem*. 2005, **117**, 300–302.

40 C.D.F. Milagre, L.F. Cabeça, L.G. Martins, A.J. Marsaioli, *J. Braz. Chem. Soc*., Vol. **22**, No.2, 286-291, 2011.

41 A.J. Benie, R. Moser, E. Baüml, D. Blaas, T. Peters, *J. Am. Chem. Soc.* 2003, **125**, 14-15.

42 S. Bartoschek, T. Klabunde, E. Defossa, V. Dietrich, S. Stengelin, C. Griesinger, T. Carlomagno, I. Focken, U. Wendt, *Angew. Chem*. 2010, **122**, 1468 –1471.

43 C. Airoldi, S. Giovannardi, B. La Ferla, J. Jiménez-Barbero, F. Nicotra, *Chem. Eur. J*. 2011, **17**, 13395–13399.

44 J.L. Asensio, F.J. Cañada, J. Jiménez-Barbero, *Eur. J. Biochem*. 1995, **233**, 618-630.

45 A. Poveda and J. Jiménez-Barbero,*Chem. Soc. Rev*. 1998, **27**, 133-143.

46 L. P. Calle, F. J. Cañada and J. Jiménez-Barbero, *Nat.Prod. Rep*., 2011, **28**, 1118-1125.

47 V. Roldós, F. J. Cañada and J. Jiménez-Barbero, *Chem. Bio. Chem*. 2011, **12**, 990-1005.

48 D. Henrichsen, B. Ernst, J.L. Magnani, W.T. Wang, B. Meyer, T. Peters, *Angew. Chem. Int.* Ed. 1999, **38**, 98-102.

49 A. Canales, R. Matesanz, N.M. Gardner, J.M. Andreu, I. Paterson, J.F. Díaz, J. Jiménez-Barbero, *Chemistry* 2008, **14**, 7557-69.

50 A. Canales, L. Nieto, J. Rodríguez-Salarichs, P.A. Sánchez-Murcia, C. Coderch, A. Cortés-Cabrera, I. Paterson, T. Carlomagno, F. Gago, J.M. Andreu, K.H. Altmann, J. Jiménez-Barbero, J.F. Díaz, *ACS Chem. Biol*. 2014**DOI:** 10.1021/cb400673h

51 J. Kopitz, C. von Reitzenstein, S. André, H. Kaltner, J. Uhl, V. Ehemann, M. Cantz, H.J Gabius, *J. Biol. Chem*., 2001,**276**, 35917–35923.

52 P. Vidal, V.Roldós, Y.Bleriot, M.delC. Fernández-Alonso, B. Vauzeilles, Y. Bleriot, F.J. Cañada, S. André, H.J. Gabius, J. Jiménez-Barbero, J.F. Espinosa, S. Martín-Santamaría, *Chem. Eur. J*., 2013, **19**, 14581-14590.

53 F. Marcelo, C.Dias, A. Martins, P.J. Madeira, T. Jorge, M.H. Florêncio, F.J Cañada, E. J. Cabrita, A.P. Rauter, *Chem. Eur.J*., 2013, **19**, 6641-6649.

54 L. Nieto, A. Canales, G. Giménez-Gallego, P. M. Nieto, and J. Jiménez-Barbero, *Chem. Eur. J*., 2011, **17**, 11204-11209.

55 A. Canales, J. Angulo, R. Ojeda, M. Bruix, R. Fayos, R.Lozano, G. Giménez-Gallego, M. Martín-Lomas, P.M. Nieto, J. Jiménez-Barbero, *J Am Chem Soc*., 2005, **127**, 5778-5779.

56 L. Calle, V. Roldós, F. J. Cañada, M.L. Uhrig, A.J. Cagnoni, V.E. Manzano, O. Varela, J. Jiménez-Barbero, *Chem. Eur. J*, 2013, **19**, 4262-4270.

57 P.X. Qi, J. L. Urbauer, E. J. Fuentes, M. F. Leopold and A. J. Wand, *Nat Struct Biol.*1994, **1**, 378-82.

58 M. Levitt and B. H.Park, *Structure* 1993, **1**, 223-226.

59 H. Kovacs, T. Agback and J. Isaksson, *J Biomol NMR* 2012, **53**, 85–92

60 G. Otting, E. Liepinsh, B. T. Farmer and K. Wuthrich, *J Biomol NMR* 1991, **1**, 209–215

61 G.M. Clore, A. Bax, P. T. Wingfield and A. M. Gronenborn, *Biochem*. 1990, **29**, 5671–5676

62 S. Grzesiek and A. Bax, *J Biomol NMR* 1993, **3**, 627–638.

63 C. Dalvit, P. Pevarello, M. Tato, M. Veronesi, A. Vulpetti and M. Sundström, *J. Biomol. NMR* 2000, **18**, 65–68.

64 C. Dalvit, G. P. Fliatto, A. Stewart, M. Veronesi and B.Stockman, *J. Biomol. NMR* 2001, **21**, 349–359.

65 C. Dalvit, D. Caronni, N. Mongelli, M. Veronesi and A. Vulpetti, *Curr. Drug Discov. Technol.* 2006, **3**, 115–124.

66 S. Chu and M. Gochin, *Bioorg. Med. Chem. Lett.* 2013, **23**, 5114–5118

67 R. Gozalbes, S. Mosulén, L. Ortí, J. Rodríguez-Díaz, R. J. Carbajo, P. Melnyk and A. Pineda-Lucena *Bioorg. Med. Chem*. 2013, **21**, 1944–1951

68 C. Dalvit, M. Fasolini, M. Flocco, S. Knapp, P. Pevarello and M. Veronesi, *J. Med. Chem*. 2002, **45**, 2610-4

69 M. Karle, W. Knecht and Y. Xue*Bioorg. Med. Chem. Lett.* 2012,**22**, 4839–4843

70 A. Sivertsen, J. Isaksson, H.-K. S Leiros, J. Svenson, J.-S. Svendsen and B. O. Brandsdal BMC Struc. Biol. 2014, 14:4

71 J. Hu, P.-O. Eriksson, and G. Kerna, *Magn. Reson. Chem*. 2010, **48**, 909–911.

72 A.D. Gossert, C. Henry, M. J. Blommers, W. Jahnke and C.Fernández, *J Biomol NMR*.2009, **43**, 211-7.

73 C. Ludwig, P. J. Michiels, X. Wu, K. L. Kavanagh, E. Pilka, A. Jansson, U. Oppermann and U. L. Günther, *J Med Chem.*2008, **51**, 1-3.

74 M. D. Díaz, M. Palomino-Schätzlein, F. Corzana, C. Andreu, R. J. Carbajo, M. del Olmo, A. Canales-Mayordomo, A. Pineda-Lucena, G. Asensio and J. Jiménez-Barbero, *J Biomol NMR*.2009, **43**, 211-7.

75 M. G. Szczepina, D. W. Bleile, J. Mullegger, A. R. Lewis and B. M. Pinto, *Chem. Eur. J*. 2011, **17**, 11438–11445.

76 C. Dalvit, P. E. Fagerness, D. T. A. Hadden, R. W. Sarver and B. J. Stockman, *J. Am. Chem. Soc*. 2003, **125**, 7696-7703.

77 T. Zhuang, H. S. Lee, B. Imperiali, J. H. Prestegard, *Protein Sci.* 2008,**17**, 1220-31.

78 R. D. Seidel, T. Zhuang and J. H. Prestegard, *J Am Chem Soc*. 2007, **129**, 4834-9.

79 C.S. Johnson Jr., *Progress in NMR* 1999, **34**, 203-256.

80 Y. Cohen, L. Avram and L. Frish, *Angew. Chem. Int.* Ed. 2005, **44**, 520–554.

81 K.S. Cameron and L. Fielding, *Magn. Reson.Chem*. 2002, **40**, S106-S109.

82 A. Buevich and J. Baum, *J. Am. Chem. Soc*. 2000, **124**, 7156-62.

83 J. A. Jones, D. K. Wilkins, L. J. Smith and C. M. Dobson, *J. Biomol. NMR* 1997, **10**, 199-203.

84 L. H. Lucas and C. K. Larive, *Concepts MagnReson Part A* 2004, **20A**, 24-41.

85 E E. S. Seo, B. S. Blaum, T. Vargues, M. De Cecco, J. A. Deakin, M. Lyon, P. E. Barran, D. J. Campopiano and D. Uhrín , *Biochem.* 2010, **49**, 10486–10495

86 L. Fielding, *Curr Top Med Chem.* 2003;**3**, 39-53.

87 W. Bocian, R. Kawecki, E. Bednarek, J. Sitkowski, M. P. Williamson, P. E. Hansen and L. Kozerski, *Chem. Eur. J.* 2008, **14**, 2788–2794.

88 M. Lin, M. J. Shapiro and J. R. Wareing, J. *Org. Chem*. 1997, **62**, 8930-8931.

89 M. Lin, M. J. Shapiro and J. R. Wareing, *J. Am. Chem. Soc*. 1997, **119**, 5249-5250.

90 R. C. Anderson, M. Lin and M. J. Shapiro, *J. Comb. Chem*. 1999, **1**, 69-72.

91 O. Segev, I. Columbus, Y. Ashani and Y. Cohen, *J. Org. Chem.* 2005, **70**, 309-314.
92 C. Dalvit and A. Vulpetti, *Magn.Reson. Chem.* 2012, **50**, 592–597.
93 B. A. Becker, K. F. Morris and C. K. Larive, *J. Magn. Reson*.2006, **181**, 327–330.
94 T. S. Derrick, E. F. McCord and C. K. Larive, *J. Magn.Reson*. 2002, **155**, 217–225.
95 J. Yan, A. D. Kline, H. Mo, E.R. Zartler and M. J. Shapiro, *J. Am. Chem. Soc*. 2002, **124**, 9984-9985.
96 B. A. Becker and C. K. Larive, J. Phys. Chem. B, 2008, **112**, 13581-7.