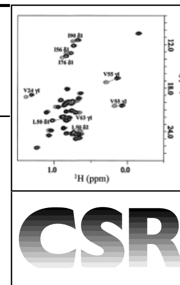
NMR-based methods and strategies for drug discovery



Xavier Salvatella†[a] and Ernest Giralt*[ab]

[a] *Institut de Recerca Biomèdica, Parc Científic de Barcelona, Universitat de Barcelona, Josep Samitier 1, E-08028 Barcelona, SPAIN. E-mail: egiralt@pcb.ub.es; Fax: +34-934037126; Tel: +34-934037125*

[b] *Departament de Química Orgànica, Universitat de Barcelona, Martí i Franqués 1-11, E-08028 Barcelona, SPAIN*

Received 17th July 2003
First published as an Advance Article on the web 20th October 2003

Nuclear Magnetic Resonance (NMR) spectroscopy has long been a favourite tool of chemists interested in host–guest systems because it permits access to a wealth of information about the molecular recognition reaction. NMR has evolved dramatically in the last 15 years and, in parallel with the development of NMR methods for the determination of protein structure, a variety of tools aimed at detecting protein ligand interactions have been proposed and are being now used both in industrial and academic laboratories as valuable tools for structure-based drug discovery. Very recent developments have considerably increased the fraction of therapeutic targets that can be tackled by NMR and significantly reduced the amount of sample required for analysis; in this tutorial review we outline the essential NMR-based techniques and describe some examples of their implementation as part of drug discovery programmes.

1 Introduction

Nuclear magnetic resonance (NMR) spectroscopy is a very powerful analytical technique for studying intermolecular interactions and is unique in its ability to provide information on the structural, thermodynamic and kinetic aspects of the binding reaction. It has thus found widespread use in the field of molecular recognition and supramolecular chemistry,[1] a natural consequence of this being its utilization in structure-based drug discovery.[2]

The great methodological and technical advances of recombinant DNA technology and of NMR spectroscopy,[3] together with the increase in the awareness that structure-based strategies will be key in the search for new drugs,[4] have placed NMR in the center of a silent revolution in the field of drug discovery. This is especially relevant in the context of structural genomics, the global research program which aims at determining the 3D atomic structures of all important therapeutic targets and structurally interesting proteins by a combination of X-ray crystallography, NMR spectroscopy and structural bioinformatics.

The NMR techniques initially developed by spectroscopists to detect interaction between the host and the guest have evolved into methodologies for lead generation and optimization in which NMR experiments and careful ligand design work in synergy towards the development of potent *in vitro* drug candidates. It is thus timely to review here the techniques that NMR spectroscopy has to offer to detect binding and their implementation in the overall strategy of drug discovery research programs.

† Current address Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, UK.

Xavier Salvatella obtained a BSc in Chemistry from the University of Barcelona in 1996 and, after working in organic electrochemistry under the supervision of Professor J. H. P. Utley towards a MSc in Chemical Research at the Queen Mary and Westfield College (University of London), he returned to Barcelona as a graduate student to work in the lab of Ernest Giralt. In there he studied, mainly using Nuclear Magnetic Resonance, peptide and protein surface recognition by synthetic ligands and the solution structure of marine cyclic peptides. He obtained his PhD from the University of Barcelona in 2002 and is currently a Marie Curie Fellow at the Department of Chemistry of the University of Cambridge, where he is developing methods for the determination of non-native protein structures using NMR and novel computational methods with Michele Vendruscolo and Chris Dobson.

Ernest Giralt received his first degree in 1970 and his PhD in 1974 from the University of Barcelona. After postdoctoral work at the University of Montpellier, France, he returned to Barcelona as Assistant Professor. He was subsequently promoted to Associate Professor in 1977 and to Full Professor in 1986. He was Visiting Professor at the University of California, San Diego, and Research Associate at the Scripps Research Institute, USA, in 1991. Professor Giralt's main research interests lie in the fields of molecular recognition, peptide synthesis and structure determination, in particular using Nuclear Magnetic Resonance spectroscopy. He was one of the founding members of the European Peptide Society and received the Narcis Monturiol prize in 1992, the Leonidas Zervas award in 1994 and the National Prize from the Royal Spanish Chemical Society in 2003.

DOI: 10.1039/b210047a

This journal is © The Royal Society of Chemistry 2003

2 Techniques to detect target–ligand binding

The reasons that have made NMR the essential tool to characterize molecular recognition reactions are the same that made it the most important technique for the structural characterization of chemical compounds. The NMR parameters which one can obtain from a single NMR experiment (chemical shifts, coupling constants, signal intensities and linewidths) are intimately dependent on the very precise chemical environment of each nucleus of the molecule. With careful experimental control (temperature, concentration, solvent, pH and ionic strength if applicable), the modification of the NMR spectrum by addition of a second compound to the first sample clearly indicates the formation of a complex.

The aim of the NMR techniques developed by spectroscopists interested in molecular recognition is to measure one or more of these parameters and obtain information about the binding reaction: which atoms, functional groups, residues or secondary structure elements are involved in establishing the non-covalent interactions, what is the affinity between the two compounds, what is the rate at which the complex is formed and how quickly it dissociates. Since all this information is relevant to the structure-based approach to drug discovery NMR has become a very attractive tool in this field.

The techniques to detect and investigate target–ligand binding can, from a very practical perspective, be divided in two classes: those that measure the NMR properties of the target (usually a macromolecule such as a protein) and those which detect binding by measuring the NMR spectrum of the ligand.

2.1 Monitoring the signals of the target

The most commonly employed method to detect interaction between a drug candidate and a target is the chemical shift perturbation method,[5,6] which, as its names suggests, analyzes the chemical shift changes observed in the target when a ligand interacts with its surface. To relate the changes observed with the primary structure of the protein it requires the chemical shifts of the protein to be assigned and for the method to yield the maximal amount of information the three-dimensional structure of the target should be known.

The NMR experiments more commonly used for this purpose are the 2D [^{1}H–^{15}N] or [^{1}H–^{13}C]-HSQC[7] in the absence and presence of ligand. (For a list of acronyms see Table 1.) While the former allows detection of changes in the amide protons and nitrogen nuclei of the backbone and Asn and Gln side chains and requires the protein sample to be enriched in ^{15}N, the latter requires ^{13}C enrichment but yields information on chemical shift changes in all side chains. Although ^{13}C labeling allows the Chemical Shift Perturbation approach to sample hydrophobic patches on the surface of the protein, the ^{15}N experiment is normally preferred because it requires neither the relatively costly ^{13}C enrichment nor the often lengthy process of side chain assignment. In both cases, by measuring the chemical shift changes as a function of ligand concentration the affinity constant between the ligand and the target can be accurately measured. A typical chemical shift perturbation experiment is shown in Fig. 1. The most important feature of this method is the structural information that it delivers; the binding site can be

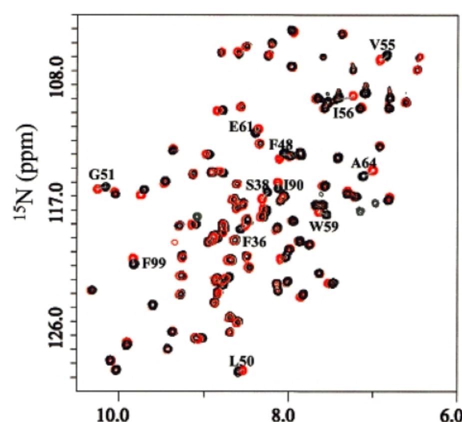


Fig. 1 Detection of ligand binding using chemical shift perturbation using 2D [^{1}H–^{15}N]-HSQC spectroscopy.[13] The black contours correspond to FKBP, the macromolecular target, whereas the red contours correspond to the complex formed by FKBP and phenylimidazole. Reproduced with permission from ref. 13. Copyright 2000, American Chemical Society.

Table 1 Acronyms commonly used in NMR-based drug discovery

Acronyms	Full Name	Brief Definition
HSQC/HMQC	Heteronuclear Single Quantum Correlation/Heteronuclear Multiple Quantum Correlation	2D experiments which correlate proton and heteronuclear resonances. Very useful for protein binding studies and central to the chemical shift perturbation method.
TROSY	Transverse Relaxation-Optimised Spectroscopy	Technique that, by taking advantage of the interference of different relaxation mechanisms, allows for a significant increase in the molecular weight limit of biomolecular NMR.
SEA-TROSY	Solvent-Exposed Amides Transverse Relaxation-Optimised Spectroscopy	TROSY-based experiment that, by detecting only solvent-exposed amides, greatly simplifies the spectrum of high molecular weight proteins. This facilitates the use of chemical shift perturbation methods for screening.
INEPT	Insensitive Nuclei Enhanced by Polarization Transfer	Technique which uses heteronuclear coupling constants to transfer magnetization to insensitive nuclei, allowing their detection.
CRIPT	Cross Relaxation-Induced Polarization Transfer	Alternative to the INEPT technique where magnetization is transferred using cross-correlated relaxation.
NOE/NOESY	Nuclear Overhauser Effect/NOE Spectroscopy	Effect that can be used to measure approximate through-space proton to proton distances. NOEs are the main NMR parameter used for conformational analysis and protein structure determination by NMR. NOEs are in general measured using a 2-D NMR experiment termed NOESY.
STD	Saturation Transfer Difference	Technique that allows the identification of ligands from a mixture of low molecular weight compounds by transferring saturation from the macromolecular target to the ligands.
Water-LOGSY	Water-ligand observed *via* gradient spectroscopy	Technique which uses water molecules to mediate the transfer of magnetization from the macromolecular target to the ligand.
SAR by NMR	Structure Activity Relationships by NMR	Structure-based NMR approach for the discovery of high affinity protein ligands based on chemical shift perturbation.
NMR-DOC	NMR docking of compounds	Structure-based NMR approach well-suited for very high molecular weight proteins which relies on selective isotope enrichment and requires no previous knowledge of the chemical shift assignments of the protein.

mapped on the surface of the protein. As will be seen in later sections, this information can be used in the refined design of further drug candidates.[8]

The chemical shift perturbation method cannot differentiate by itself between direct chemical shift changes due the proximity of the ligand and indirect chemical shift changes which have a conformational origin; indeed binding of a ligand can cause chemical shift changes to be observed not in the binding site but in structurally related residues. In order to prevent mis-assignment of the binding site a method termed secondary chemical shift mapping has been proposed,[9] whereby direct and indirect chemical shift changes are differentiated through the use of a chemically modified ligand, enabling unambiguous assignment of the binding site. Very often the targets of drug-discovery programs are very large proteins. Unfortunately, the very long correlation time of proteins with a molecular weight higher than 30 kDa causes their NMR resonances to be too wide to be detected by the traditional methods and would produce, if detected, very crowded 2D spectra in which it would be difficult to resolve the individual peaks. Recent advances have however helped reduce the impact of these limitations.

Novel techniques for high molecular weight targets. New developments such as ^{2}H labeling,[10] Transverse Relaxation Optimized Spectroscopy (TROSY)[11] and Cross Relaxation Induced Polarization Transfer (CRIPT)[12] have increased the molecular weight limit of NMR spectroscopy to values greater than 100 kDa. Whereas perdeuteration of a protein eliminates the most important relaxation mechanism of the $^{13}C^{\alpha}$ nuclei, namely the dipolar coupling to the directly bound $^{1}H^{\alpha}$, the TROSY[11] method takes advantage of the partial cancellation of the two most important transverse relaxation mechanisms in large proteins, dipolar couplings and chemical shift anisotropy, that takes place at very high magnetic fields. The combination of both strategies has resulted in a significant increase in the molecular wight limit of biomolecular NMR. The CRIPT[12] method, on the other hand, is a new approach to polarization transfer in heteronuclear NMR that is much less sensitive to relaxation than the traditionally used INEPT. A particularly striking example of the combined effect of the TROSY and CRIPT techniques is shown in Fig. 2, where the interaction surface between GroEL and GroES (with a combined molecular weight in excess of 800 kDa) has been mapped on the surface of GroES using chemical shift perturbation.

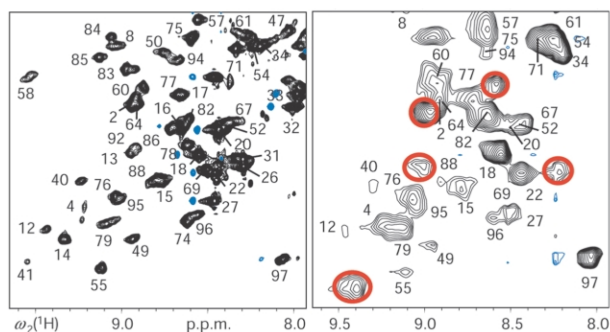


Fig. 2 Detection of the formation of a 800 kDa protein–protein complex by chemical shift perturbation using a heteronuclear 2D experiments implementing the TROSY (GroES reference spectrum, left hand side) and both the TROSY and CRIPT techniques (GroES-GroEL complex, right-hand side).[12] Reproduced with permission from ref. 12. Copyright 2002, Nature Publishing Group (http://www.nature.com/).

Approaches for the minimization of spectral overlap. While the use of deuterated targets and of the TROSY and CRIPT techniques permits the study of complexes of very high molecular weight the spectra that are obtained can be of extremely high complexity. Several methods which aim at simplifying the spectra in these cases have been proposed and will be discussed below.

The most general approach is the use of novel methods of protein expression which allow the selective labelling of a given residue type. For example, addition of the selectively labeled ^{13}C biosynthetic precursors of Leu, Val and Ile to the culture medium during the expression of the target yields a protein in which only the carbon nuclei of the methyl groups of Leu, Val and Ile are ^{13}C enriched. This greatly simplifies the [^{1}H–^{13}C]-HSQC spectrum of the target and makes it amenable to implementation in high throughput protocols[13] as shown in Fig. 3. An additional advantage of this method is the high signal to

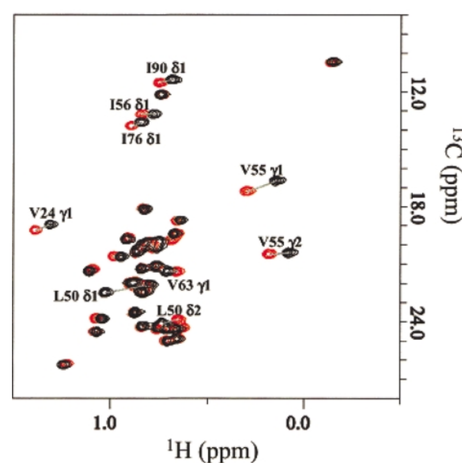


Fig. 3 Simplified 2D [^{1}H–^{13}C]-HSQC spectrum obtained using selective labelling schemes: selective incorporation of ^{13}C labelled Val, Ile and Leu. The system studied here is the same as the one shown in Fig. 1.[13] Reproduced with permission from ref. 13. Copyright 2000, American Chemical Society.

noise ratio of the [^{1}H–^{13}C] cross peaks of methyl groups as compared to those of methylene or methyne. Combination of these two factors greatly shortens the acquisition time to record a good quality [^{1}H–^{13}C]-HSQC spectrum and therefore the number of compounds that can be screened per unit time. Very recently a similar method has recently been described for the incorporation of ^{13}C or ^{19}F labeled Trp residues.[14,15]

An alternative and simpler method of spectral simplification is based on detection of only those NH protons which are exposed to the solvent. This new experiment, termed Solvent-Exposed Amides with TROSY (SEA-TROSY),[16] is a modification of the TROSY-[^{1}H,^{15}N]-HSQC pulse sequence.[11] In spite of a small signal loss as compared to a regular TROSY-[^{1}H,^{15}N]-HSQC spectrum the large increase in spectral simplicity allows chemical shift perturbation experiments to be carried out with proteins of *ca.* 300 residues without the need to selectively label with ^{13}C, as can be seen in Fig. 4.

2.2 Monitoring the signals of the ligand

A series of alternative experiments have been designed that monitor the resonances of the ligand. This is an important technique in cases where isotopically labelling is not possible, for targets of very high molecular weight or when the protein spectrum is too complex.

Principle of the transfer experiments. Rather than monitoring chemical shift changes, which require the use of stoichiometric amounts of target, these transfer methods use pulse sequences that transfer magnetization or coherence (or the lack of it) from the protein to the ligand (or *vice-versa*) when the

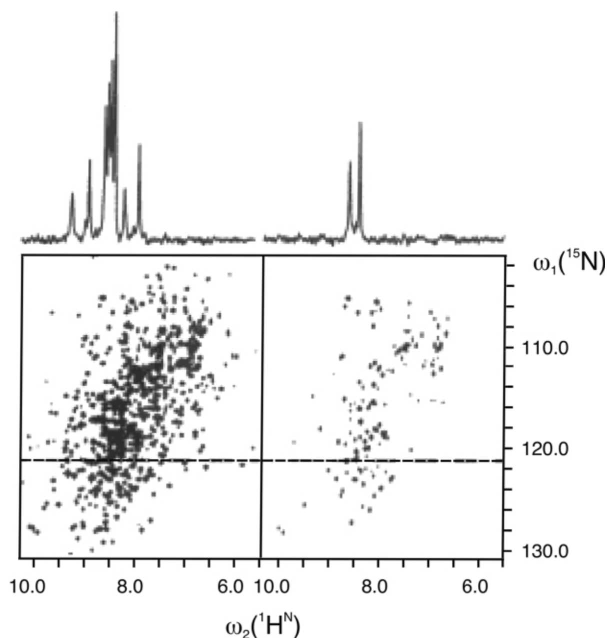


Fig. 4 The left hand side of the figure shows the TROSY spectrum of a ^{2}H and ^{15}N labelled sample of P450 whereas the right hand side shows the corresponding SEA-TROSY simplified spectrum suitable for screening.[16] The 1D spectra shown correspond to traces taken at the position shown with a dashed line. Reproduced with permission from ref. 16. Copyright 2001, American Chemical Society.

ligand is bound to the target. This transferred property affects the intensity of the resonances of the ligand and after subtraction of a reference spectrum affords the *transfer* spectrum of the ligand.

Typical concentrations used for this type of experiments are 1mM ligand and 50 μM protein. For this stoichiometry, assuming that the binding constant is sufficiently high for effective saturation of the binding site, each ligand molecule spends 1/20 of its time as part of a target–ligand complex with relaxation and hydrodynamic properties which are very different to those of the free ligand. If the exchange between the free and bound states is sufficiently fast so as to take place many times during the mixing (or saturation) time of the experiment the property of choice will be *transferred* to the free ligand.

The transfer-NOE experiment. Possibly the most classical and informative of the experiments which fall in this category is the transfer-NOE.[17] The 2D-NOESY experiment, when carried out in a concentrated ligand sample in the presence of sub-stoichiometric amounts of protein, can yield cross peaks which are not the consequence of the conformation of the ligand free in solution but that of the conformation of the ligand when it is bound to the binding site of the protein, *i.e.* the NOEs of the bound state are therefore *transferred* to the free state as shown in Fig. 5. The fact that the ^{1}H–^{1}H NOE is much stronger and of opposite sign in molecules of large molecular weight such as target–ligand complexes is the driving force for the transfer of the NOE. Although in principle it is necessary to record a reference spectrum in which no target is present (which needs to be subtracted from the transfer-NOE spectrum) this can be avoided by careful tuning of the experimental conditions, especially mixing time and temperature, so that no NOEs from the free state are observed.

The transfer NOE method yields information on the conformation of the ligand when bound to the protein;[18] from a structure-based drug design perspective this information is very valuable for the further design of compounds of improved affinity. The transfer NOE experiment is, however, not optimally suited for the high throughput screening of libraries because the 2D-NOESY experiment on which it is based is a

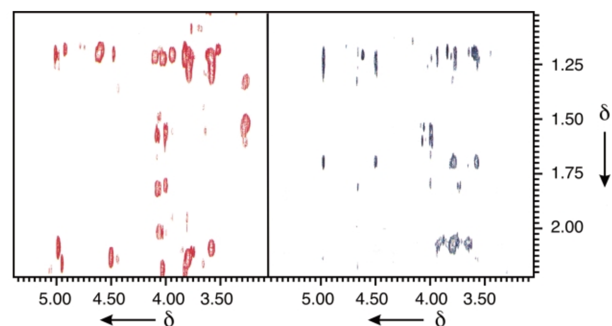


Fig. 5 Example of application of the transfer NOE experiment. The cross peaks (all positive) on the right-hand side spectrum are due to the bound conformation of the ligand when bound to its target, E-selectin.[18] The spectrum on the left-hand side is that of the free ligand and is shown for reference (negative cross peaks). Reproduced with permission from ref. 18. Copyright 1999, Wiley-VCH.

relatively insensitive 2D proton-homonuclear experiment which requires relatively long acquisition times to provide spectra of sufficient quality.

NOE pumping experiments. A very interesting approach to the detection of transfer NOE has been implemented in the *NOE pumping* sequence.[19] This experiment affords the 1D spectrum of the protein and all strongly interacting compounds. A reference spectrum is not required since all signals arising from unbound compounds are removed by a diffusion filter.

A pulse sequence based on the same principle but with an opposite magnetization flow has been termed *reverse NOE pumping* by the same authors.[20] Although this experiment requires the measurement of a reference spectrum it has a significant advantage in avoiding T_2 relaxation during the spin-echoes of the NOE pumping sequence, which limited to a great extent the sensitivity of the original *NOE pumping* experiment. The reverse pumping experiment in shown in Figure 6.

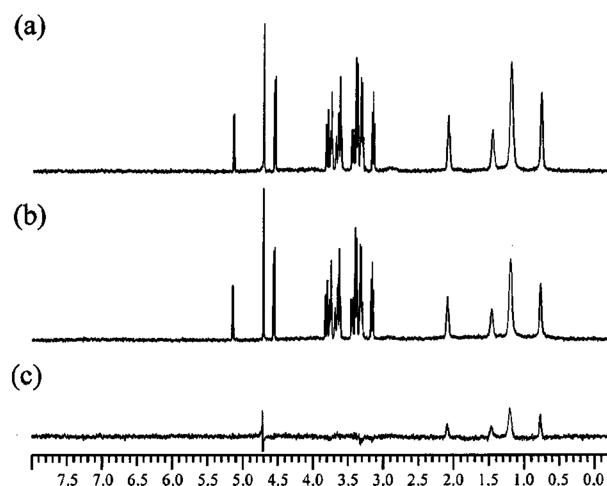


Fig. 6 The reverse NOE pumping experiment.[20] a) Reference spectrum of the mixture (1 mM octanoic acid 1 mM and glucose) in the presence of the target (20 μM HSA), b) reverse pumping NOE spectrum of the same sample, c) difference spectrum showing the resonances of the actual ligand, octanoic acid. Reproduced with permission from ref. 20. Copyright 2000, American Chemical Society.

Saturation Transfer Difference experiments. Possibly the most successful experiment of those described for transferring magnetization from the macromolecular target to the ligand (or *vice-versa*) is the Saturation Transfer Difference (STD) experiment which was first described in 1999[21] and has since found widespread use in the drug discovery industry.

The pulse sequence is equivalent to that of the steady state NOE, and the experiment is carried out by subtracting the spectrum obtained when irradiation is placed on a protein resonance from the spectrum obtained when irradiation is off resonance. On resonance irradiation causes saturation of the protein resonance and spread of the saturation to the rest of protein and to any interacting ligand whereas off resonance irradiation should yield the 1D spectrum of the sample.

The difference spectrum yields the ^{1}H 1D spectrum of the protein and all interacting low molecular weight compounds as seen in Fig. 7. In this case no coherence is transferred from the complex to the ligand; it is indeed the lack of coherence due to saturation which is very quickly and efficiently transferred in high molecular weight species. In order to prevent the appearance of false positives by direct saturation it is important in STD spectroscopy to carefully select the irradiation frequency. One important feature of this method is the possibility of obtaining information about the binding epitope of the ligand[22,23] through the analysis of the relative amount of saturation transferred to each of its atoms; those atoms which show a higher intensity in the transfer spectrum are those which lie at the target–ligand interface. The versatility of the STD method is remarkable since it can be applied to systems such as proteins immobilized onto a solid support[24] or to membrane proteins embedded onto liposomes.[25]

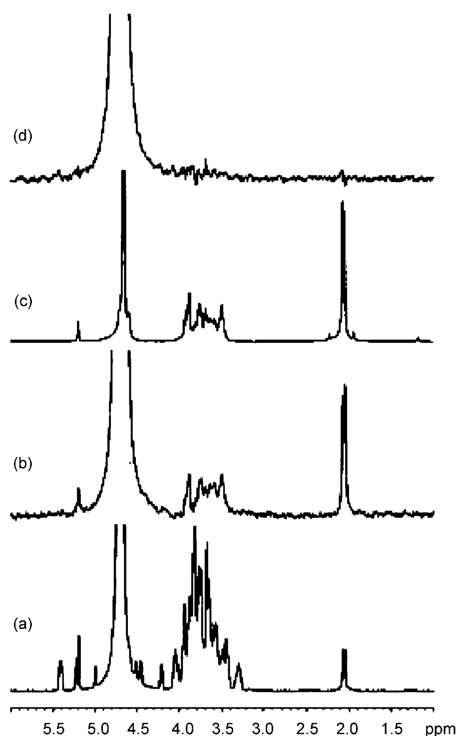


Fig. 7 The STD experiment: a) spectrum of a mixture of 7 oligosaccharides in the presence of immobilized WGA, b) STD spectrum, c) spectrum of the mixture when only the actual ligand is present, d) STD spectrum of the mixture when the actual ligand is absent.[24] Reproduced with permission from ref. 24. Copyright 1999, American Chemical Society.

Water-LOGSY. A related method takes advantage of the fact that water molecules are often found in protein–ligand interfaces to devise a very sensitive technique for ligand screening by NMR termed water-LOGSY.[26,27] The method is based on cross relaxation between the solvent and the ligand (or on transfer of saturation) using the interfacial water molecules as mediators. One disadvantage of this method is that the interpretation of the difference spectra is less straightforward than in the NOE pumping and STD approaches: the fact that binding and non-binding compounds give signals of opposite phase can give rise to complications in the case of crowded spectra so common to combinatorial libraries. However, one important advantage of this approach is its sensitivity. Since the magnetization transfer takes place with water molecules instead of protein molecules this increases the intensity change detected in the difference spectrum.

Relaxation-based methods. The change in the relaxation time ($T1$, $T2$ and $T1\rho$) of the ligand by interaction with a high molecular weight target has been used extensively for screening.[28] Complications due to signal overlap render these techniques not optimal for the analysis of mixtures but a new approach based on double quantum coherence has been shown to solve this problem to a great extent. Identification of the ligands is carried out by comparison of the 2D-double quantum re-focused spectrum of the mixture in the presence and in the absence of the target.[29] The cross peaks of the ligand are not present in the former due to the combination of the very fast relaxation of double quantum coherence of large macromolecules and the anti-phase character of the cross peaks in the indirect dimension.

Competition experiments. One important drawback of all ligand-based techniques discussed so far is their inability to screen for ligands of very high affinity. When the affinity between the target and the ligand approaches the nanomolar range the residence time of the ligand on the surface of the target is too long for the transfer experiments to yield sufficient signal. This problem is especially relevant at the later stages of drug discovery, where it is important to discriminate between ligands of very high affinity. One particularly simple approach to the solution of this problem is the use of competition experiments. A library of compounds can be screened by monitoring the ability of each compound to displace a low to medium affinity compound from the surface of the target. NMR is therefore used for the detection of the dissociation of the low stability complex rather than for the formation of the high stability complex.[30,31]

3 NMR-based strategies used for drug discovery

Structure-based drug discovery strategies are increasingly applying NMR techniques to several of the steps involved in drug development. NMR, which in its early stages in the drug discovery field was merely used as an alternative way of screening, can now be employed for lead generation, lead optimization and even, in certain cases, High Throughput Screening. The following section will review the integrated approaches to drug discovery reported by several research groups, mainly from pharmaceutical or drug discovery companies.

3.1 SAR by NMR

The first NMR-based method for structure-based drug discovery was reported in 1996 by Fesik and co-workers at Abbot.[8] The method, termed SAR by NMR, uses chemical shift perturbation for detecting low affinity lead compounds which are then optimized in a structure-based fashion.

The first step of the method is the screening of libraries of low molecular weight compounds using the [^{1}H,^{15}N]-HSQC spectrum of the target as the probe for chemical shift perturbation analysis. In order to screen libraries of compounds in a time-efficient manner each screening experiment is carried out with 10 different compounds simultaneously. Once a hit is identified the binding site can be discovered by deconvolution of the mixture, identification of the actual binding compound and

mapping the chemical shift changes on the surface of the target, and this information is used to guide the combinatorial search for modified ligands of higher affinity.

The following step requires the screening of a different library in search of compounds that will cause chemical shift perturbation in a different second site at the surface of the protein. This step requires the screening to be carried out on a protein solution in the presence of saturating amounts of the first ligand so that the first binding site is fully occupied. When a hit for the second binding site is obtained the chemical shift changes are again mapped on the surface of the protein in order to ascertain the relative positions of the two binding sites. After optimization of the affinity of the second hit the two compounds can be covalently linked, yielding a lead compound of high affinity due to the chelate effect.

An interesting example of the application of this methodology is the development of non-peptidic inhibitors of Stromelysin at Abbot and can be followed in Fig. 8. Stromelysin is a

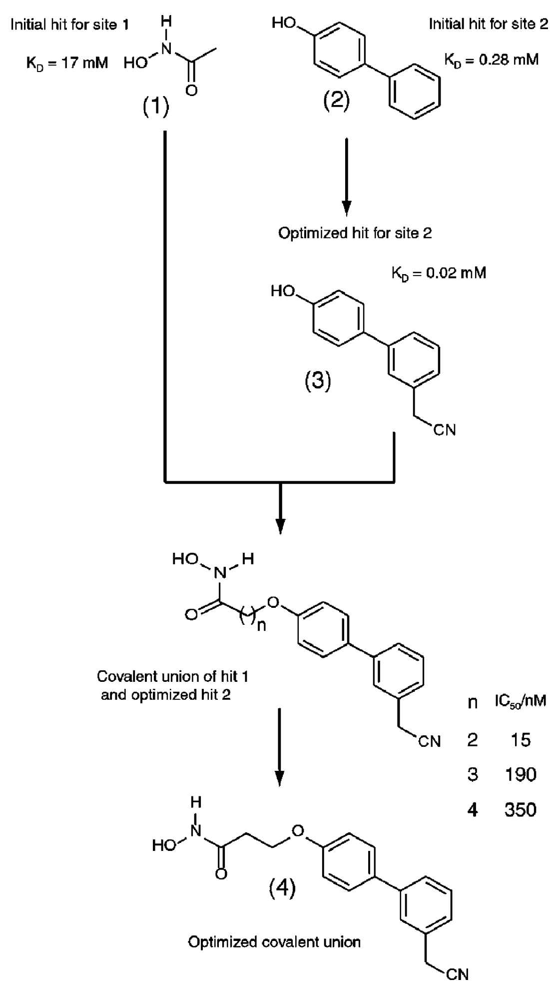


Fig. 8 Example of use of the SAR by NMR approach for the discovery of inhibitors of Stromelysin.[8,32]

zinc-dependent endoproteinase which is a target in arthritis and tumor metastases for which there had been difficulties in finding non-peptide inhibitors. The SAR by NMR method was used in a slightly modified fashion because there was already a wealth of information from poorly bioavailable peptide inhibitors and other non-peptide inhibitors of low activity.

Acetohydroxamic acid (**1**), with a very low (17 mM) affinity for the active site was the compound that was used for the first binding site in the protein since most previously known inhibitors were known to contain a hydroxamate moiety. The NMR-based screening of the second site was carried at very high concentrations of hydroxamate acid in order to occupy the first binding site and prevent the autolytic degradation of the protein. The chemical shift perturbation method showed the biphenyl scaffold (**2**) to have a certain affinity for the hydrophobic S_1' binding site of the endoproteinase. After a series of optimization steps compound **3** was shown to bind the S_1' active site of the enzyme with low affinity (0.02 mM). Structure-based design of a compound that would combine a hydroxamate moiety and the biphenyl scaffold of compound 30 yielded **4**, a potent (15 nM) non-peptide inhibitor of Stromelysin.[32]

The success of this method in a number of cases,[33] combined with the very recent increases in protein size that can be addressed by 2D-[^{1}H,^{15}N] heteronuclear spectroscopy highlights the impact that chemical shift perturbation-based methods can have in structure-based drug discovery.

3.2 The SHAPES strategy

An alternative and in some ways complementary methodology, termed SHAPES,[34] was developed at Vertex by Moore and coworkers. The SHAPES method uses the signals of the ligand as a means of screening libraries of compounds for lead generation, which can later be optimized on the basis of structural information derived from other NMR experiments. The most important aspect of the SHAPES strategy is the design of the library,[35] which has to incorporate, in a size amenable to NMR screening, as many scaffolds and chemical functionalities as possible. Indeed, in order to maximize the probability that a given hit obtained from the library leads, after optimization, to a drug candidate, it is important that the compounds of the library present drug-like properties, good solubility and synthetic accessibility.

Signal broadening in 1D ^{1}H spectra, the transfer NOE and more recently the STD method are the specific NMR techniques which are used for screening under the SHAPES scheme. Although NMR can be used in further stages of the drug discovery process for obtaining information about the structure of the bound ligand (using transfer NOE data) or on the detailed architecture of the non-covalent interactions between ligand and active site (based on the determination of the solution structure of the complex) the low affinity ligands provided by the SHAPES method are typically used as seeds in the design of libraries to be screened by methods more amenable to high throughput screening (HTS) approaches.

An illustrative example of the application of the SHAPES method was described by Moore and co-workers at Vertex and has been outlined in Fig. 9.[36] A library of drug-like compounds was screened by NMR for interactions with the active site of protein p38. The first search yielded the imidazole derivative **5**. Modification of the core derivative by addition of other aromatic moieties allowed the discovery of compounds **6** and **7**, which showed an enhanced affinity for the target. Using these ligands as lead compound the high affinity ligand **8** was discovered using a HTS screening method.

Hence the SHAPES approach to NMR-based drug discovery uses NMR for the initial screening of a small but very carefully designed library, allowing sampling of the most commonly found functional groups in current drugs. It is important to mention that NMR is especially powerful in this context because it can be used with virtually any organic compound and does not require any specific assay to be developed and validated and because it is ideally suited for the detection of low affinity compounds.

3.3 Alternative approaches

An alternative use of NMR spectroscopy for the screening of mixtures has been proposed very recently by Pellecchia and coworkers at TRIAD Therapeutics. The most important feature of

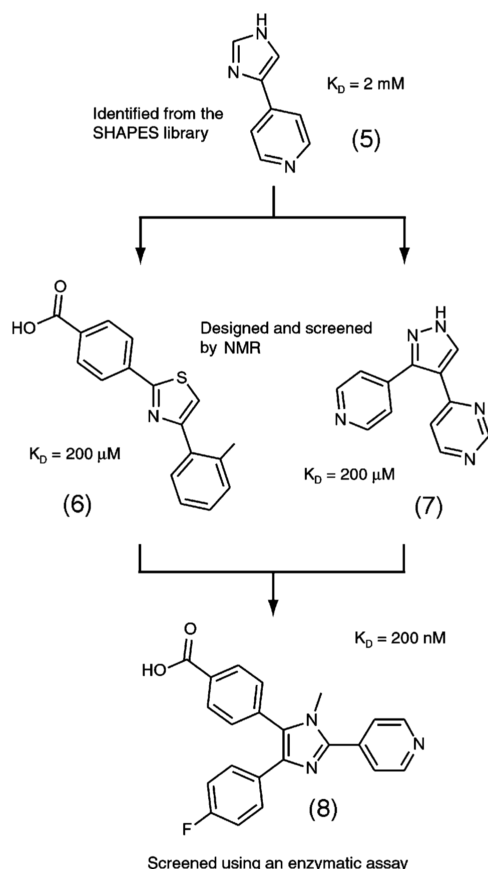


Fig. 9 Use of the SHAPES method for the discovery of ligands for the active site of protein p38.[36]

their NMR-DOC[37] approach is that it does not require any previous knowledge of the chemical shifts of the target and thus can be applied to proteins of very large molecular weight. The technique requires ^{2}H-enrichment of the protein and selective ^{13}C labeling of a given type of residue which is known to be present in or at least very close to the active site (*e.g.* Met, Ile and Thr in ref. 37). The chemical shifts of the labelled residue can be assigned either by chemical shift perturbation or by the detection of intermolecular NOEs after addition of a ligand. After identification of its cross peaks in the already very simple [^{13}C–^{1}H]-HMQC spectrum of the protein this experiment can be used to screen libraries and measure the relative affinities of the ligands. This technique can be applied to proteins of unknown structure when the protein presents two adjacent binding pockets (*e.g.* in dehydrogenases). Knowledge of a ligand for the first site can be used for the assignment of the cross peak to be used as a probe for binding and to direct a library of candidates for the second site by covalent linkage.

4 Conclusions and outlook

The number of methods and strategies for drug discovery based on NMR is increasing steadily due to the several interesting advantages of the NMR approach: the ability to provide structural information about the binding mode in solution, the possibility to screen for target binding in the presence or absence of cofactors and in different environmental conditions, the low tendency to produce false positives (in contrast to HTS) and, importantly, the possibility to screen for ligands of mediocre affinity which can be used as scaffolds for combinatorial approaches.

The methods based on monitoring the resonances of the target offer a unique opportunity to identify the binding site on the protein surface. These methods require the production of relatively large amounts of isotopically labeled protein and are most potent when the chemical shifts and structure of the target are known; they are therefore very well suited to those drug discovery research programs where the target is of low to medium molecular weight and has already been studied by NMR or where it is essential to obtain structural information about the binding site.

The methods that monitor the signals of the ligand have a series of advantages that render them more applicable to the most common targets of drug discovery, *i.e.* proteins of very high molecular weight, of unknown high resolution structure and which have never been studied using NMR spectroscopy. In these cases the ligand-based 1D and 2D techniques herein presented can be considered as very general screening methodologies which are especially adequate for those projects where no knowledge about the architecture of the binding site is required.

What approach is the most appropriate for a given target and drug candidate (or library of candidates) depends therefore on the size of the target, the availability of its structure and chemical shift assignments and the number of drug candidates to screen. The two main approaches described in this review (chemical shift perturbation as used in SAR by NMR and 1D-transfer experiments as used in the SHAPES approach) are complementary both in their molecular weight limitations and in the structural information they deliver about the system. Indeed a very thorough characterization of the interactions that take place between the target and the ligand, which is the basis of structure-based drug discovery, requires the use of techniques which fall in both categories.

The developments described in this review, together with the very promising advances in hardware which are being presented by manufacturers suggest that NMR will definitely have a role to play in future structure-based drug discovery research initiatives.

Acknowledgements

The authors wish to thank Drs Miguel Feliz and John Christodoulou for helpful discussions and suggestions. X. S. acknowledges the Generalitat de Catalunya for a predoctoral fellowship. This work was partially supported by grants from MCYT-FEDER (BIO2002-2301) and Generalitat de Catalunya (Grup Consolidat 1999SGR0042 and Centre de Referència en Biotecnologia).

References

1 M. Pons and O. Millet, *Prog. Nucl. Magn. Reson. Spectrosc.*, 2001, **38**, 267–324.
2 J. W. Peng, C. A. Lepre, J. Fejzo, N. Abdul-Manan and J. M. Moore, *Methods Enzymol.*, 2001, **338**, 202–230.
3 K. Wüthrich, *Nature Struct. Biol.*, 1998, **5**, 492–495.
4 M. von Dongen, J. Weigelt, J. Uppenberg, J. Schultz and M. Wikström, *Drug Discovery Today*, 2002, **7**, 471–478.
5 J. M. Johnson, E. M. Meiering, J. E. Wright, J. Pardo, A. Rosowsky and G. Wagner, *Biochemistry*, 1997, **36**, 4399–4411.
6 J. Kallen, C. Spitzfaden, M. G. M. Zurini, G. Wider, H. Widmer, K. Wüthrich and M. D. Walkinshaw, *Nature*, 1991, **353**, 276–279.
7 G. Bodenhausen and J. Ruben, *Chem. Phys. Lett.*, 1980, **69**, 185–189.
8 S. B. Shuker, P. J. Hajduk, R. P. Meadows and S. W. Fesik, *Science*, 1996, **274**, 1531–1534.
9 A. Medek, P. J. Hajduk, J. Mack and S. W. Fesik, *J. Am. Chem. Soc.*, 2000, **122**, 1241–1242.
10 K. H. Gardner and L. E. Kay, *Annu. Rev. Biophys. Biomol. Struct.*, 1998, **27**, 357–406.
11 K. Pervushin, R. Riek, G. Wider and K. Wüthrich, *Proc. Natl. Acad. Sci. USA*, 1997, **94**, 12366–12371.

12 J. Fiaux, E. B. Bertelsen, A. L. Horwich and K. Wüthrich, *Nature*, 2002, **418**, 207–211.
13 P. J. Hajduk, D. J. Augeri, J. Mack, R. Mendoza, J. G. Yang, S. F. Betz and S. W. Fesik, *J. Am. Chem. Soc.*, 2000, **122**, 7898–7904.
14 R. A. Rodríguez-Mias and M. Pellecchia, *J. Am. Chem. Soc.*, 2003, **125**, 2892–2893.
15 M. Leone, R. A. Rodríguez-Mias and M. Pellecchia, *ChemBioChem*, 2003, **4**, 649–650.
16 M. Pellecchia, D. Meininger, A. L. Schen, R. Jack, C. B. Kasper and D. S. Sem, *J. Am. Chem. Soc.*, 2001, **123**, 4633–4634.
17 F. Ni, *Prog. Nucl. Magn. Res. Spectrosc.*, 1994, **26**, 517–606.
18 D. Henrichsen, B. Ernst, J. L. Magnani, W.-T. Wang, B. Meyer and T. Peters, *Angew. Chem. Int. Ed.*, 1999, **38**, 98–102.
19 A. Chen and M. J. Shapiro, *J. Am. Chem. Soc.*, 1998, **120**, 10258–10259.
20 A. D. Chen and M. J. Shapiro, *J. Am. Chem. Soc.*, 2000, **122**, 414–415.
21 M. Mayer and B. Meyer, *Angew. Chem. Int. Ed.*, 1999, **38**, 1784–1788.
22 H. Moller, N. Serttas, H. Paulsen, J. M. Burcell, J. Taylor-Papadimitriou and Bernd Meyer, *Eur. J. Biochem.*, 2002, **269**, 1444–1455.
23 M. Mayer and B. Meyer, *J. Am. Chem. Soc.*, 2001, **123**, 6108–6117.
24 J. Klein, R. Meinecke, M. Mayer and B. Meyer, *J. Am. Chem. Soc.*, 1999, **121**, 5336–5337.
25 R. Meinecke and B. Meyer, *J. Med. Chem.*, 2001, **44**, 3059–3065.
26 C. Dalvit, P. Pevarello, M. Tatò, M. Veronesi, A. Vulpetti and M. Sundström, *J. Biomol. NMR*, 2000, **18**, 65–68.
27 C. Dalvit, G. P. Fogliatto, A. Stewart, M. Veronesi and B. J. Stockman, *J. Biomol. NMR*, 2001, **21**, 349–359.
28 P. J. Hajduk, E. T. Olejniczak and S. W. Fesik, *J. Am. Chem. Soc.*, 1997, **119**, 12257–12261.
29 (C. Dalvit, C. Batistini, P. Caccia, P. Giordano, P. Pevarello, M. Sundström and M. Tato), Pharmacia and Upjohn S.P.A, WO-0133243, 1999.
30 C. Dalvit, M. Fasolini, M. Flocco, S. Knapp, P. Pevarello and M. Veronesi, *J. Med. Chem.*, 2002, **45**, 2610–2614.
31 C. Dalvit, M. Flocco, S. Knapp, M. Mostardini, R. Perego, B. J. Stockman, M. Veronesi and M. Varasi, *J. Am. Chem. Soc.*, 2002, **124**, 7702–7709.
32 P. J. Hajduk, G. Sheppard, D. G. Nettesheim, E. T. Olejniczak, S. B. Shuker, R. P. Meadows, D. H. Steinman, G. M. Carrera Jr., P. A. Marcotte, J. Severin, K. Walter, H. Smith, E. Gubbins, R. Simmer, T. F. Holzman, D. W. Morgan, S. K. Davidsen, J. B. Summers and S. W. Fesik, *J. Am. Chem. Soc.*, 1997, **119**, 5818–5827.
33 P. J. Hajduk, J. Dinges, G. F. Miknis, M. Merlock, T. Middleton, D. J. Kempf, D. A. Egan, K. A. Walter, T. S. Robins, S. B. Shuker, T. F. Holzman and S. W. Fesik, *J. Med. Chem.*, 1997, **40**, 3144–3150.
34 J. Fejzo, C. A. Lepre, J. W. Peng, G. W. Bemis, Ajay, M. A. Murcko and J. M. Moore, *Chem. Biol.*, 1999, **6**, 755–769.
35 C. A. Lepre, *Drug Discovery Today*, 2001, **6**, 133–140.
36 J. M. Moore, *Curr. Opin. Biotechnol.*, 1999, **10**, 54–58.
37 M. Pellecchia, D. Meininger, Q. Dong, E. Chang, R. Jack and D. S. Sem, *J. Biomol. NMR*, 2002, **22**, 165–173.