CMLS, Cell. Mol. Life Sci. 61 (2004) 580–599
1420-682X/04/050580-20
DOI 10.1007/s00018-003-3382-3

CMLS Cellular and Molecular Life Sciences

Review

Protein NMR in biomedical research

W. Jahnke* and H. Widmer

Novartis Institutes for BioMedical Research, Novartis Pharma AG, WSJ-88.904, Lichtstrasse 4002 Basel (Switzerland), Fax: +41 61 324 2686, e-mail: wolfgang.jahnke@pharma.novartis.com

Received 7 October 2003; received after revision 11 November 2003; accepted 27 November 2003

Abstract. Nuclear magnetic resonance (NMR) spectroscopy is a versatile biophysical technique with wide applicability in drug discovery research, particularly for the detection and characterization of molecular interactions. This review highlights in a comprehensive manner the aspects of biomolecular NMR which are most beneficial for pharmaceutical research and presents them as contributions to the different stages of a drug discovery program: target selection, assay development, lead generation and lead optimization. Emphasis is put on the concept of the particular NMR application, rather than on technical details, and on recent examples. Finally, an appendix of frequently asked questions is given.

Key words. Fragment-based screening; drug discovery; NMR screening; SAR-by-NMR; structure determination; high throughput screening; hit validation.

Introduction

Drug discovery is a process of continuous innovation. Recent successful innovative steps include fully automated high-throughput screening (HTS), combinatorial chemistry and structure-guided drug design. They depend on the application and further development of underlying techniques; in the case of structure-guided drug design, these are mainly parallelized protein production, X-ray crystallography, nuclear magnetic resonance (NMR) spectroscopy and in silico docking and scoring. This review article focuses on the role of NMR in the drug discovery process. NMR can be used as a tool for the determination of structure and dynamics of proteins and protein-ligand complexes, but can also be applied in a broader scope for the detection and characterization of molecular interactions. For example, it can be used to (in)validate compound hits from HTS and thus to separate the wheat from the chaff. It can also generate lead compounds de novo by a fragment-based approach. Thus, NMR is not only a technique for structure-guided drug design, but it can also be integrated with HTS and combinatorial chemistry. Being so versatile, NMR is applied in all phases of a drug discovery program (fig. 1), including target selection, assay development, lead generation and lead optimization. This review will describe the potential and the contributions of NMR in these drug discovery phases, with emphasis on lead generation (including hit identification and hit validation). In the last part, a list of 'frequently asked questions' and their answers is compiled.

NMR spectroscopy in this article refers to biomolecular NMR to study biomolecules such as proteins or nucleic acids and their interactions with ligands in aqueous solution. Other applications of NMR, such as NMR used for analytic purposes to investigate covalent structures of small molecules, as well as NMR used as a tool for imaging [1] or metabonomics [2, 3], are not discussed here. The reader is also referred to previous reviews on the role of biomolecular NMR in drug discovery [4–20]. Many of these reviews focus on NMR screening techniques. This review covers all aspects of NMR that we consider most useful for the drug discovery process.

* Corresponding author.

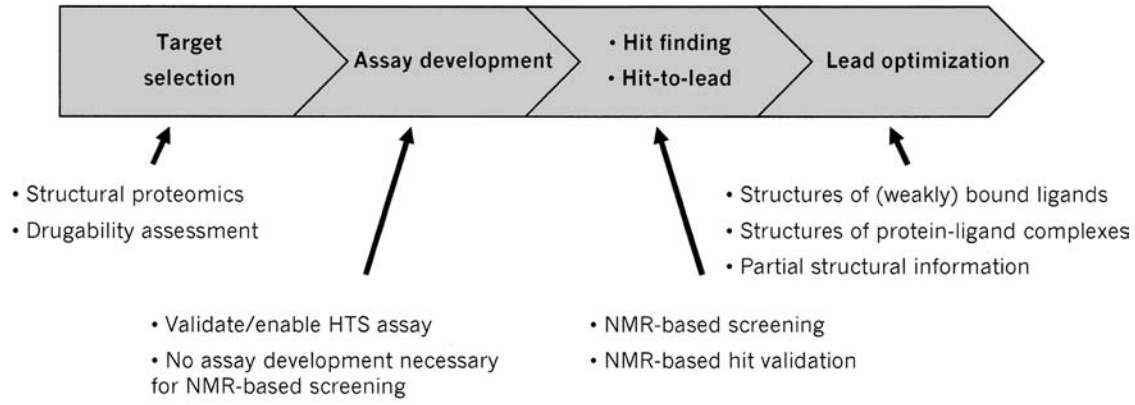

Figure 1. Overview of the contributions of biomolecular NMR to the drug discovery phases.

NMR in the different phases of drug discovery

Figure 1 displays the classical stages of a drug discovery program. In the exploratory phase, basic research is carried out to identify and preliminarily validate and select potential targets, often protein targets, for a given disease. Once the target is selected, an assay needs to be developed in order to identify compounds that bind to the target protein and modulate its activity. This assay should be sensitive and robust and preferably suitable for HTS. The lead generation phase is divided into the actual (high-throughput) screening phase, in which large compound libraries are screened to produce 'hits', and a hit-to-lead phase, in which hits are validated and characterized with respect to their physicochemical profile and initial structure-activity relationship (SAR). The result of this phase should be a 'lead' compound, which is then optimized by medicinal chemistry to hopefully yield a development candidate. There are trends to modify this classical concept of drug discovery by applying lead finding and lead optimization efforts even before target validation. In the following sections, the potential of NMR to support the different drug discovery phases will be discussed. Special attention will be given to explain the underlying principles, and recent examples will be presented to highlight techniques and applications.

Target selection

This is a phase of basic research which can be carried out at academic research institutes, biotech or big pharma companies. A potential drug target is defined as a protein with a propensity to bind small molecules that modulate its activity ('drugable target') and with the potential to be disease modifying. The sequencing of the human genome suggests that there are between 600 and 3000 genes that code for potential drug targets for human diseases [21]. Since the human proteome is much larger than the genome, many more possibilities for drug intervention result [22]. About 400 of these gene products are currently in the research portfolio of pharmaceutical companies, and only 120 of them are targeted by marketed drugs. New targets are therefore of high interest for pharmaceutical research. As a complement to X-ray crystallography, NMR can contribute by determining the solution structure of a potential new target, particularly for monomeric proteins < 40 kDa or oligomeric proteins < 60 kDa [23–25]. Knowledge of its structure can hint to the biological function of a protein [26, 27]. Furthermore, NMR can not only determine the three-dimensional (3D) structure of a protein, but can also detect proteins that do not have a defined 3D structure in vitro. An increasing number of such intrinsically unstructured proteins have been identified in the human proteome [28]. Although at least some of them adopt a defined fold when interacting with binding partners to carry out their biological functions [29–31], intrinsically unstructured proteins should not be considered ideal targets since the entropic cost of an intrinsically unstructured protein to bind a small molecule is probably too high.

In addition to characterization and determination of the 3D structure of a potential target, its drugability can be estimated not only by inspection of the binding pocket after 3D structure determination or chemical shift mapping of a protein-protein interaction, but also by a small-scale NMR screening approach (see below) in which the hit rate allows conclusions about its general drugability. The chemical nature of hits hints to the nature of the binding site, and to the general function of the protein (see below).

Assay development

One of the strengths of NMR in lead generation is that the NMR assay is a universal binding assay which can be generally applied to essentially all proteins and potential ligands. Development of a customized NMR assay is therefore not necessary. The lack of assay development typically gives NMR screening a head start of about 6

months if structural biology is involved early in the program.

The main assay to be carried out for a new target is often not NMR screening, but high-throughput screening by other methods. For novel target proteins, especially for those with unknown function, reagents for HTS are often not available. For example, a fluorescence-based competitive binding assay needs a fluorescent ligand whose displacement can be monitored. If no fluorescent ligands are known, NMR screening can identify a scaffold and an appropriate site to attach a fluorescent moiety. If fluorescent ligands are known but not sufficiently well characterized, their binding site on the protein can be determined for small protein targets, to ensure that the binding site is indeed the active site. This is not always the case. As recently reported by the Vertex group [32], fluorescently tagged fatty acid substrates did not bind as expected to the active site of adipocyte fatty acid-binding protein, a possible target for the treatment of type II diabetes. Instead, it was found to bind at multiple locations on the protein surface, so that it is useless as a HTS tool, and a high-throughput screen using this substrate would have yielded artifacts. As a consequence, NMR screening was used to identify active-site ligands with potencies up to 300 nM. Next, commercially available fluorescent analogs were selected and screened, and a suitable fluorescent probe, 1-anilinonaphthalene-8-sulfonic acid, was identified as active-site binder that was suited for the development of a fluorescence-based competition assay for adipocyte fatty acid-binding protein [32].

A strategy for the development of a HTS assay based on NMR-derived ligands has recently been reported [33]. It focuses on novel potential targets for infectious diseases that have been identified on the basis of their genomic sequences. Since generally neither the function or biological activity nor any substrate or ligand is known for such targets, a functional assay or a competitive binding assay cannot be developed. To support assay development for these cases, a library of common natural ligands has been assembled. This 'functional genomics library' contains about 160 compounds, comprising natural substrates, cofactors, metals, enzyme inhibitors, amino acids and nucleotides as 'usual suspects' of potential ligands [33, 34]. Screening of this library against HI-0033, a conserved bacterial protein from *Haemophilus influenza* with unknown biochemical activity, yielded three ligands with K_D < 10 µM, all of which were adenosine analogs and bound to the same binding site on HI-0033. Based upon limited SAR information, fluorescent analogs were tested, and a fluorescein-containing analog which binds with a K_D of 700 nM was selected for development of an FPA-based HTS assay [33]. In addition to assay development, knowledge about adenosine binding to HI-0033 may give valuable clues to the function of HI-0033.

Taken together, applications of NMR have the potential to become an important tool for quality assurance or even the design of a HTS assay.

Lead generation

Lead generation can be a time-consuming process and is generally divided into two parts: the actual screening process, the result of which is a 'hit', and a hit-to-lead process, in which hits are further characterized and perhaps chemically derivatized, the result of which is a 'lead'.

In cases where a high-throughput screen by other methods has been carried out and hits have been identified, these hits need to be validated and characterized. First of all, their binding to the target has to be confirmed. Identification of the ligand binding site on the target, as well as the binding epitope of the ligand, are beneficial, and competition with known ligands should be investigated. These aspects will be discussed below.

Screening by NMR can complement and in some special cases even substitute HTS by other methods. It is often applied in cases where HTS has failed to identify viable hits that can be transformed into leads. It can also be applied in addition and parallel to HTS, particularly for high-value targets. The NMR screening approach will also be discussed below.

For both hit validation and NMR screening, the essential feature of NMR is its ability to detect and characterize even weak protein-ligand interactions. This section will start with the description of techniques which are currently used for this purpose.

Principles of the NMR-binding assay

NMR has long been recognized as a valuable tool for the identification and characterization of interactions between small molecules and proteins or DNA. Back in 1965, Jardetzky et al. [35] detected and characterized penicillin binding to serum albumin using methods that are similar to the ones currently employed. Numerous investigations have followed in which protein-ligand interactions were identified and characterized [36–43], and several textbooks and reviews have summarized this important feature of biomolecular NMR [44–47].

All of these investigations have relied, and still rely, mainly on two aspects of a binding event: when protein is complexed by a ligand, the protein resonances experience chemical shift perturbations caused by the binding event [38–43]. These can be due to direct contact with ligand or to secondary effects such as conformational changes. Titration of the protein with ligand has been used to determine the dissociation constant, K_D [40–42]. Since 1D protein NMR spectra are often too crowded, 2D spectra such as COSY or ^{15}N, ^{1}H-heteronuclear single quantum coherence (HSQC) have been used to reduce spectral

overlap [40, 41, 43]. On the ligand side, a bound ligand temporarily behaves like a large molecule, and adopts properties of a large molecule, such as broad resonances and large negative nuclear Overhauser effects (NOEs) [37, 38, 40, 41, 44, 46]. Measurements of protein or ligand chemical shift changes, ligand relaxation or ligand NOE properties have therefore long been used to identify and characterize protein-ligand interactions. This is described in more detail below, together with a new technique in which the resonances of a reporter ligand are observed.

Protein observation

Protein observation [41, 44, 48, 49] relies on the fact that protein chemical shifts depend sensitively on the chemical environment of the respective residue. Binding of a ligand gives rise to chemical shift changes for the resonances of the protein atoms which are either in direct contact with the ligand, or which experience secondary effects (conformational changes or dynamic changes) upon ligand binding. Since a 1D protein spectrum is too crowded to be analyzed, heteronuclear 2D methods such as HSQC are usually employed to reduce signal overlap. As a consequence, ^{15}N-isotopically labeled protein is required for these experiments. An example is shown in figure 2 where part of the 2D ^{15}N, ^{1}H-HSQC of the anti-apoptotic protein, Bcl-xL, is displayed. In the ^{15}N, ^{1}H-HSQC spectrum, each amino acid residue (except proline) gives rise to one resonance at the position of the ^{15}N and ^{1}H frequencies of its amide group. For ligand binding studies, HSQC spectra are usually recorded in the absence (black) and presence (red) of a ligand. Ligand resonances are not observed since the ligand is not ^{15}N-labeled. Protein chemical shift changes indicate binding, and if resonance assignments and a 3D structure are available, the signals with the largest chemical shift changes can be mapped to the protein structure to yield the ligand binding site on the protein (fig. 2). As an alternative to ^{15}N, ^{1}H-HSQCs using ^{15}N-labeled protein, it has been suggested to record ^{13}C-^{1}H correlation spectra with ^{13}C-methyl-labeled protein [50].

If the target is a nucleic acid rather than a protein, the nucleic acid would correspondingly be observed, preferably the imino protons at or between the bases. The imino protons resonances are highly sensitive to ligand binding or structural changes and are often readily assigned. Since they are also generally well resolved, isotopic labeling is not required for screening of medium-sized DNA or RNA [51]. The advantage of protein detection is the high spectral information content and the robustness of the method. Drawbacks are the need for very large amounts (about a milligram of protein per sample) of isotopically labeled protein, and the need for high ligand solubility. Poorly soluble ligands can only be detected if they have high binding affinity, since only then do they complex the protein to a high degree.

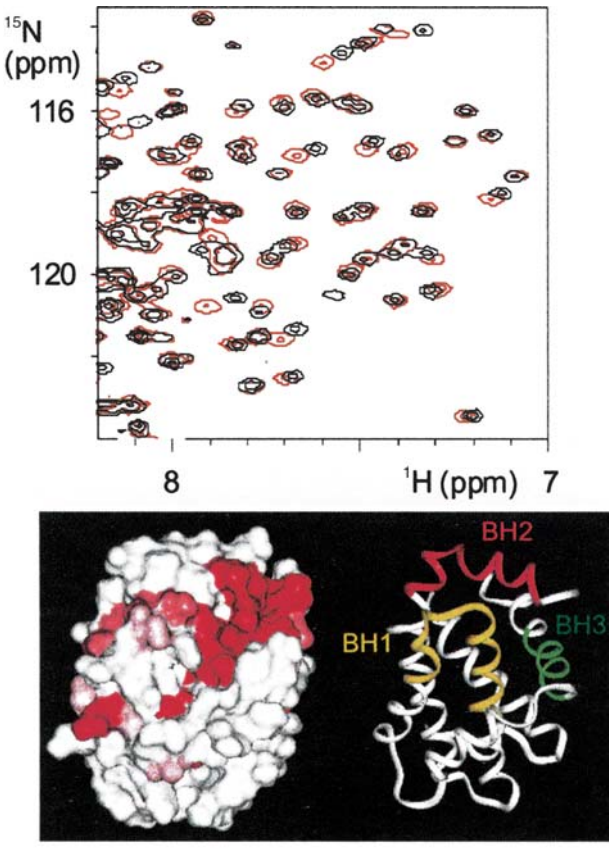

Figure 2. Protein-detected NMR screening. Chemical shift changes of the protein (here Bcl-x_L) in the absence (black) and presence (red) of ligands indicate molecular interactions. Mapping these changes on to the protein surface identifies the site of interaction.

Ligand observation

The requirement for large amounts of isotopically labeled protein is often a no-go criterion for target proteins that do not express well in *Escherichia coli*. Therefore, ligand-observation techniques are employed that require only small amounts of unlabeled protein. Ligand-observation techniques observe only resonances of the unbound ligand, but are able to detect the ligand history if the ligand dissociates within the time scale of the experiment. Was the ligand bound to the protein for some time during the 100–1000 ms before observation? If so, then effects from this past binding event have been transferred to the unbound ligand and can be detected there.

Which effects are transferred from the bound to the free ligand? During the time when a ligand is bound to a protein, the ligand temporarily does not behave like a small molecule, but like the protein, a large molecule: it tumbles more slowly. Therefore, its NMR properties are temporarily the ones of the protein: it shows faster relaxation, slower diffusion, and large and negative NOEs (fig. 3). If ligand spectra measuring a certain NMR property are recorded in the absence and presence of protein, observ-

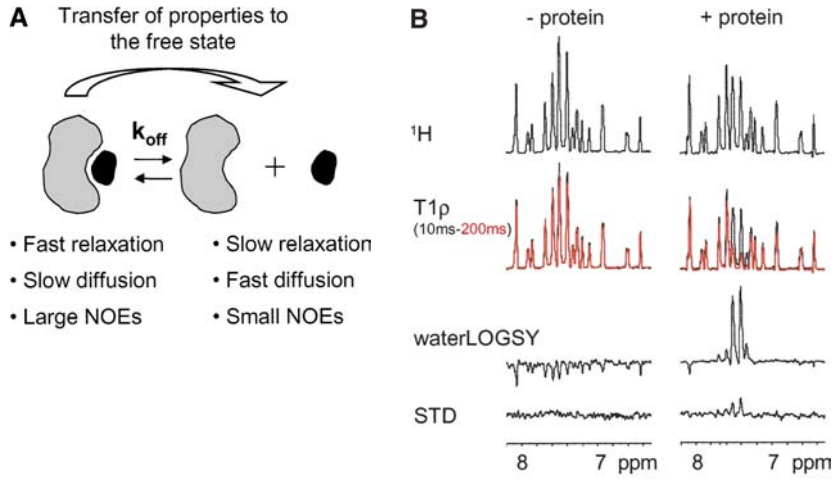

Figure 3. Ligand-detected NMR screening. (*A*) When a small compound is bound to target protein, it temporarily adopts the properties of a large molecule, such as fast relaxation, slow diffusion and large NOEs. Nonbinding compounds are and behave like small molecules. Only unbound ligand is observed, but the history of ligand binding has been transferred to the unbound ligand, so that binding compounds can be discriminated from nonbinding compounds. (*B*) Example spectra of ligand-observed NMR screening: T1ϱ relaxation, waterLOGSY and STD. The upper spectra show simple ^{1}H pulse-acquire spectra for comparison.

able differences indicate binding of the ligand to target protein.

For example, an increase of the ligand relaxation rate in the presence of protein, witnessed by the faster attenuation of the ligand signal at long relaxation periods, indicates binding [37, 38, 40, 41, 44, 46, 52]. This effect can be drastically increased by attaching spin labels to the protein, as in the SLAPSTIC experiment [53]. Similarly, a decrease in translational diffusion rates leads to the same conclusion [52, 54]. Other methods are based on the measurement of NOEs. An inversion of the sign of cross-peaks in the 2D NOE spectroscopy (NOESY) spectrum in the presence of protein indicates binding [55, 56]. A transfer of NOE magnetization from protein to ligand upon protein saturation [37, 41, 57] in a saturation transfer difference (STD) experiment [37, 40, 41, 46, 57] is a sensitive detection method, as is the magnetization transfer from bulk water to bound ligand in the water-ligand observation with gradient spectroscopy (water LOGSY) experiment [58]. The reverse NOE pumping experiment [59] can identify binding compounds as well. There are also several other, less widely used ligand-observation methods which will not be discussed here. The methods differ mainly by the quantity of protein needed, and by the required ligand solubility. Other reviews describe these experiments in more detail [7, 10, 15]. Not only can soluble proteins be screened by using ligand-observation techniques, but so can immobilized proteins [60], integral membrane proteins [61] or even intact viruses [62], as well as RNA [32, 51, 63].

The main advantage of ligand-observation techniques is the significantly lower consumption of protein which does not need to be isotopically labeled (down to 10 µg per sample). This also allows screening of poorly expressing proteins or of proteins expressed in organisms other than *E. coli*. Disadvantages are the lower information content of the spectra: the protein binding site cannot be deduced, and in fact nonspecific binding is not directly discriminated. Hits from ligand-observed NMR techniques should therefore be checked by HSQC, if labeled protein is available, or by competition experiments as shown below. Another major drawback of ligand-observation techniques is the need for significant ligand dissociation during the experiment. This essentially lets tightly binding ligands or even medium binding affinity ligands with slow binding kinetics appear nonbinding (false negative). In order to alleviate this severe shortcoming, competition-based methods have been developed.

Reporter screening

A possibility to observe tightly binding ligands is the usage of stoichiometric concentrations of protein and ligand, which leads to observable line broadening of the ligand even without ligand dissociation [37]. However, the large amounts of protein needed (about a milligram per sample) for this experiment limits its use to validation of single hits, rather than screening hundreds or thousands of compounds.

As an alternative method, competition assays (‘NMR reporter screening’) have been developed [64–66] to detect tightly binding ligands with small protein amounts (down to 10 µg per sample). For this method, a known, weakly binding ligand (the ‘spy molecule’, ‘probe’, or ‘reporter ligand’) is added to the mixture of protein and test compounds. It is not the binding of the test compounds that is detected, but the binding of the reporter ligand. If the reporter ligand is bound even in the presence of test compounds, none of the test compounds has

comparable or higher binding affinity. In contrast, if the reporter ligand is not bound anymore, one or more test compounds have displaced it and therefore must have comparable or higher binding affinity (fig. 4). Even poorly soluble test compounds can be detected in this way, if their affinity is high enough. Advantages of reporter screening are not only the ability to detect tightly binding ligands, but also the elimination of non-active-site binders: if the reporter ligand binds to the active site, every compound displacing it should also bind to the active site. As an additional bonus, reporter screening experiments are particularly easy to analyze, and the degree of displacement allows the binding affinity of test compounds to be calculated based on the known affinity of the reporter ligand [64, 67].

Signal overlap from the reporter ligand and test compounds presents a problem since the reporter ligand must have at least one nonoverlapped signal for interpretation. ^{13}C-labeling of the reporter ligand at one position [65] or more sophisticated NMR experiments [64] are a possible remedy. A promising approach is the use of ^{19}F-spectroscopy, which is a nucleus very suitable for NMR screening [7, 44, 46, 68]. Combination with competition experiments is a very elegant method to circumvent signal overlap with test compounds, most of which are non-fluorinated [18, 69, 70]. In this strategy, a fluorine-containing reporter ligand is identified by screening of a library of ^{19}F-containing compounds, or by weakening of a known ^{19}F-containing high-affinity ligand. This ^{19}F-containing reporter ligand can then be used for screening any test compounds even if they do not contain ^{19}F. Selective observation of a ^{13}C-labeled or ^{19}F-containing reporter ligand in the presence of test compounds parallels the selective observation of ^{15}N- or ^{13}C-labeled protein in the presence of test compounds, as used in protein observation techniques.

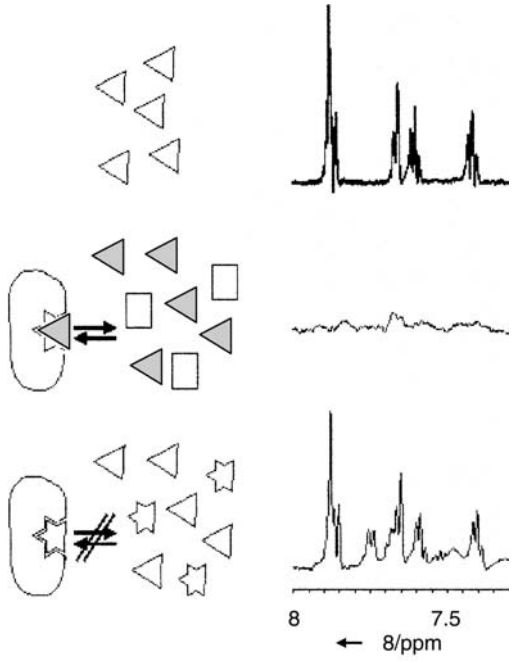

Figure 4. The principle of reporter screening. The reporter ligand (triangles) binds to the protein active site with moderate affinity. This leads to severely broadened lines in the presence of protein (middle panel), whereas the reporter ligand in the absence of protein has sharp lines (upper panel). In the presence of nonbinding test compounds, the reporter ligand is still bound and shows broad lines (middle panel). If at least one of the test compounds has higher binding affinity, it can displace the reporter ligand, which is visible by its sharpened lines (lower panel). Reprinted with permission from Jahnke et al. [65].

Hit validation by NMR

NMR techniques for the detection and characterization of protein-ligand interactions can be benefically exploited for the validation and characterization of HTS hits.

The outcome of HTS is hits, not leads. In order to convert hits into leads, upon which a chemistry program is based, the compounds must be investigated more closely. In the beginning, binding of the HTS hits to the target protein is usually validated using independent biochemical and biophysical methods. NMR in its most stripped-down version is an independent binding assay, and can profitably be used to validate the binding of HTS hits to the target protein. Since HTS hits sometimes bind so strongly that exchange is slow on the NMR time scale [often for IC50 (concentration at which 50% inhibition is observed) < 1 µM], competitive techniques such as NMR reporter screening are important. In general, the NMR assay is a very robust binding assay with very low potential for false positives if unspecific binding can be ruled out by control experiments.

Experience shows that surprisingly many HTS hits actually do not bind to the target protein [11]. In some cases, this is due to a faulty or not fully validated assay. In other cases, it is due to poor quality of the protein employed for HTS. In still other cases, it is due to the complex nature of some HTS assays. Cell-based assays or coupled assays like ELISA (enzyme-linked immunosorbent assay) are necessarily complex and contain several other proteins in addition to the target protein. The assay response is positive if any of the proteins conveying the signal is inhibited, not necessarily the target protein. In still other cases, the tested compounds interfere with the method of detection; for example, fluorescent compounds can cause artifacts in fluorescence-based assays.

In most of these cases, the compound has to be discarded. NMR for hit validation is a technically simple 'service function' which helps to separate genuine and promising hits from false positives, as illustrated in the following examples involving different target proteins:

1) The HTS yielded compounds with IC50s in the low micromolar range. Simple proton NMR spectra showed

significant chemical shift changes and broader lines of these compounds in the presence of target protein. Binding in the low micromolar range was supported by NMR: green light for the compounds.

2) Some HTS hits were found to precipitate the target protein, rather than forming a soluble complex. Similarly, some compounds chemically react with the protein (e. g. at active-site cysteine residues).
3) HTS using a new format screened 300,000 compounds and resulted in 40 hits. ^{15}N-labeled target protein was available to check binding. No chemical shift changes were observed. Hits were flagged as false positive, probably due to a faulty assay.
4) An HTS hit was active in all follow-up functional assays, and a small-scale medicinal chemistry program had already been initiated and appeared to yield a structure-activity-relationship. Surprisingly, NMR later found that the HTS hit did not bind to the target protein. Instead, an impurity was identified in small amounts which covalently bound to and inactivated the target protein to cause the positive response. The apparent SAR of the compounds was caused by varying amounts of the impurity.
5) A fluorescence-based high throughput screen for a protease target measured the increase of fluorescence as a function of inhibitor concentration, which is caused by cleavage of a fluorescent-tagged substrate by the protease. An HTS hit was also active in follow-up assays, but NMR showed that its binding to the target protease was weak and nonspecific. Instead, it was shown by NMR that the HTS hit interacts with the tagged substrate and leads to precipitation of the substrate. The lack of fluorescence at higher compound concentrations, which was ascribed to inhibition of the protease, was then actually due to precipitation of the tagged substrate. Control experiments using another assay format and untagged substrate confirmed that the compound was inactive.

Case A is a clear 'go' for the medicinal chemistry program aimed at optimizing the hit. Case B is ambiguous. In principle this mode of action may be acceptable, but it is likely that selectivity problems will occur since this hit class may also precipitate other proteins. All other cases, C–E, clearly indicate compounds not worth touching. In such cases, it is essential to know as early as possible about the false positive responses.

In a recent analysis [11], the ratio of NMR-validated hits to the total number of HTS hits was not higher than 3% in five cases, and 30–40% in only two cases. For several targets, the mechanism of false-positive detection was identified, and was due to compound fluorescence in a fluorescence polarization assay, covalent protein modification or the complexity of a cell-based assay. Clearly, the validated binders have been immediately prioritized for follow-up.

In addition to the pure yes/no answer of binding, NMR can characterize the binding of hits to the target protein. For example, competition with a known inhibitor can be investigated, and relative K_Ds can be determined [61, 65, 67, 71, 72]. The ligand-binding epitope can be determined [71, 73], and the binding site on the protein can be mapped by HSQC experiments if isotopically labeled protein is available [48], or by competition experiments with known ligand fragments [67]. These further characterizations are described under lead optimization.

Hit generation by NMR

The ability of NMR to detect and characterize molecular interactions between proteins and small molecules has been emphasized several times. It obviously follows that these techniques cannot only be used to validate compounds from HTS, as described above, but also to identify compounds as ligands for a given target which have not previously been identified by some other type of screen. In 1989, Petros et al. [40] composed a small, focused library of compounds as potential ligands for the human plasminogen kringle 4 domain and screened them using chemical shift perturbations in 2D correlation spectroscopy (COSY) NMR. They detected and characterized binding by measuring the binding affinity of ligands, protein-ligand NOEs and structural changes within the kringle domain. Similar screening studies with different focused libraries of low molecular weight compounds were presented by Thewes et al. [41] and Rejante et al. [42]. Rossi et al. [74] describe NMR screening of compounds for their ability to bind to DNA, using 1D ligand-observation techniques. In this paper, NMR screening is proposed as 'a standard method for large-scale screening of molecules' [74]. Several years later, the Abbott group elaborated on these approaches and developed a method called 'SAR-by-NMR' in which low-affinity ligands for a particular target are identified by the chemical shift perturbations they induce in the protein spectrum, and two of those weak ligands are linked to form a high-affinity ligand [49].

The following section starts with an illustration of the concept of fragment-based screening. Thereafter, follow-up strategies are discussed that are needed after low-affinity ligands have been identified, and published examples of successful NMR screening projects are presented.

Fragment-based screening

Fragment-based screening, an alternative approach for hit generation, has gained significant popularity in recent years. Most HTS activities follow classical hit-finding strategies and screen a large library of typically 10^5–10^6 compounds, all of which are (or should be) fully assembled potential lead candidates. In this approach, the IC50 detection limit is typically in the order of ≤ 10 µM, and the

likelihood of finding good-quality hits can be quite low. In contrast, fragment-based approaches attempt to assemble a lead compound piece-by-piece in a modular way. The library is typically small (10^2–10^4 compounds) and contains only fragments of potential leads with molecular weight usually < 250 Da. This library must be carefully designed and should contain diverse, druglike molecules that are readily amenable for chemical modification or linkage [75–77]. Since the fragments to be screened are small, the binding affinities of initial hits are generally small, with K_D in the order of 10 µM to 1 mM.

Why should a fragment-based screening approach be pursued, if fragment libraries are so small and only weakly binding hits will be identified anyway? By the nature of fragment-based screening, one actually deals with a very large, albeit virtually very large, library. Often enzymes and other proteins have rather large ligand binding sites which are divided into subpockets (fig. 5). A fragment can a priori bind to either of the subpockets. If the library contains 1000 compounds and weak binding to a target protein having three subpockets can be detected, the virtual library consists of 1000 · 1000 · 1000 = 1 billion assembled compounds. This calculation assumes that fragments for all three subpockets can be identified and are subsequently linked to form a single 'assembled compound'. Rather than screening a billion compounds, only a thousand compounds are screened, and out of the virtual library, only those compounds are actually synthesized that contain fragments with binding affinity. Fragment-based screening effectively accesses all combinations of fragments while avoiding the combinatorial explosion.

Another advantage of screening small fragments rather than complex molecules is that the chance of finding hits is actually higher for small fragments, provided a method is used that is sensitive to detect weak binding. This is because with larger molecules, the chance of obtaining a mismatch (which sometimes completely abrogates binding) increases more rapidly than the chance of obtaining a beneficial interaction [78]. At compound complexity levels that produce affinities measurable by HTS, the fraction of compounds that bind to the target without severe mismatches may already be very low, so that few or no useful hits result from HTS (fig. 6, right bold curve). In contrast, less complex compounds have higher chances to fit without severe mismatches, so that with a sensitive detection method the hit rate of these less complex compounds is significantly higher (fig. 6, left bold curve). Therefore, weakly binding ligands can be readily identified in cases where HTS cannot identify a single low micromolar hit out of several 10^5 compounds. Low-affinity hits are then converted into leads by one of the follow-up strategies described below.

Follow-up strategies

The NMR techniques described above provide hits, not leads, often with weak affinities in the range 10 µM to 1 mM. In order to turn an NMR hit into a lead, based upon which significant medicinal chemistry efforts are started, its affinity must be increased to the low micromolar, or better submicromolar range. The first step to achieve this should always include exploration of the structure-activity relationship around the primary hit by testing, as much as possible, close analogs of the primary hit and measuring its affinity. This may result in an optimized hit, sometimes with 10–100 times higher potency. Next, as much structural information as possible on the interaction be-

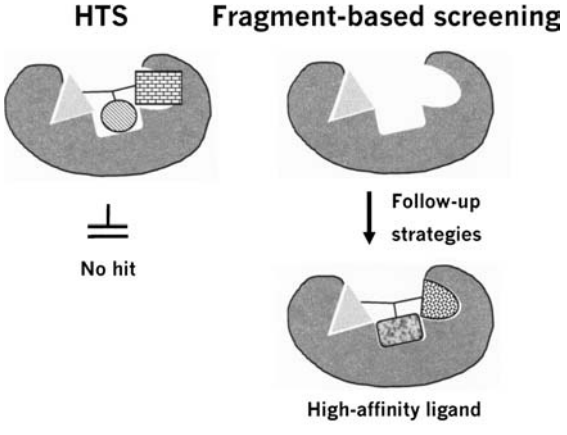

Figure 5. The concept of fragment-based screening versus HTS. A compound in the archive may contain parts that match the receptor. However, if other parts of the compound contain severe mismatches, this compound will not be a hit in HTS (left panel). Fragment-based NMR screening with sensitivity for weak interactions will identify the matching fragment (right panel, top), and by following one of several follow-up strategies, an optimized high-affinity compound can be built (right panel, bottom).

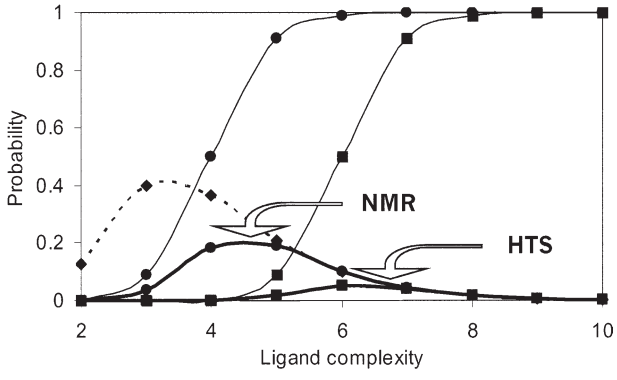

Figure 6. Fragment-based screening has a higher hit rate than conventional HTS. The dashed curve represents the probability of a compound with increasing molecular complexity to specifically bind to a target protein. This curve is based on a crude molecular recognition model [78]. The thin solid curves represent the probabilities that this binding can be observed by HTS (squares) or by NMR (circles). The probability to detect a specifically binding compound is the product of both curves and is shown by the thick solid lines. NMR-detected fragment-based screening thus has a higher hit rate than HTS. Adapted and extended from Hann et al. [78].

tween optimized hit and target protein should be collected. Ideally, the structure of the target protein complexed to the NMR hit should be determined by crystallography. This is sometimes possible if a moderately potent ligand is well water-soluble, so that its solubility is large compared with its K_D. If crystallization is not possible, or in addition to crystallization, NMR can be used to gather structural information. If a ligand-observation technique was used for NMR screening but isotopically labeled protein is available, HSQC experiments are useful to identify the binding site of the NMR hit on the protein. NOESY experiments may identify protein-ligand NOEs upon which a crude docking model of the complex can be generated. On the ligand side, STD experiments can, for ligands with a fast dissociation rate, identify the ligand epitope that is in contact with protein, which is useful information for the construction of a docking model and for follow-up activities [71, 73].

Upon gathering this structural information, the NMR hit can be further followed up by one of the following strategies [5, 17, 76] (fig. 7). In a linked-fragment approach, a second ligand is identified by second-site screening [49, 79], which binds simultaneously and in the vicinity to the NMR hit (the 'first ligand'). Linking both ligands then results in a significant increase in binding affinity if the linker has been well designed. In a merged-fragment approach, information about binding epitopes from several ligands is used to construct a merged ligand which exploits binding interactions from two or more distinct ligands. Those ligands cannot be linked since they bind at overlapping binding sites. Both the linked-fragment approach and the merged-fragment approach rely heavily on structural characterization of the binding of ligands to the target protein. In a third strategy, hits from NMR screening are used as input to design directed libraries around these hits, perhaps using combinatorial chemistry. Structural information on the complex between the hit and the target protein is desirable, but not mandatory, since NMR screening is essentially used as a filter to select the most promising scaffold for combinatorial chemistry. In the elaboration strategy, hits from NMR screening are used to select analogous compounds, e. g. identified by a similarity search, for further follow-up studies such as computational docking and scoring [80], or testing by NMR reporter screening or a biochemical assay. The follow-up strategy of choice depends on the nature of the NMR hits, the nature of the protein binding site, available chemistry capacity and the amount of available experimental information.

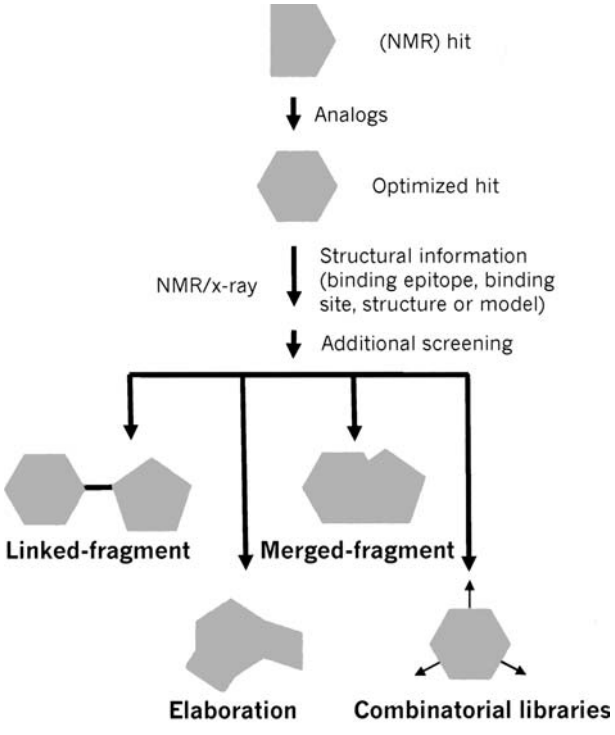

Figure 7. Follow-up strategies of fragment-based screening by NMR. After optimization of an NMR hit, structural information is gathered by NMR or X-ray, and additional screening is carried out to design high-affinity hits in a linked-fragment or merged-fragment approach, by elaboration or by the design of combinatorial libraries. For details, see text.

Applications of NMR screening

Several publications describing tangible contributions of NMR screening in lead discovery have appeared in the literature. In some cases, NMR screening was able to identify hits that could be transformed into leads even for targets for which HTS failed. In other words, these particular research programs would have been terminated without NMR. In the following section, some examples of successful NMR screening applications are described for each of the four follow-up strategies mentioned above.

The linked-fragment approach for fragment-based lead design was initially proposed as a computational method (see e. g. [81]), and was later experimentally realized by linking two identified fragments on the basis of the experimental structure of the ternary complex. Classical examples are ligand discovery for FK506 binding protein (FKBP) [49] and stromelysin [82, 83]. More recently, the linked-fragment approach was used for the discovery of a potent and selective inhibitor of protein tyrosine phosphatase 1B (PTP-1B), an enzyme that downregulates the insulin receptor and is therefore widely viewed as promising target for type II diabetes [84]. A lead compound for PTP-1B must be non-phosphate based, competitive, specific and reversible, but the identification of a such a compound by HTS faced problems. The NMR approach for ligand discovery started with site-directed NMR screening of 10,000 compounds to look for compounds that target the catalytic site. Site-specific screening was achieved by directly monitoring the resonance of Ile219, a residue that is located in the active site. The result of this and further

steps is shown in figure 8. A diaryloxamic acid was identified as active-site ligand with a K_D of around 100 µM (step 1). Chemical optimization yielded naphthyloxamic acid with fourfold improved affinity (step 2). At this stage, the structure of PTP-1B complexed with the naphthyloxamic acid was solved. The structure confirmed that the ligand is bound to the active site. It further showed that PTP-1B is in an 'open' conformation which allowed ligand access to a larger binding site. Most important, it suggested a direction to extend the ligand in order to obtain higher affinity and higher specificity. After initial failures, an attached diamido chain led to a compound with $K_D = 1.1$ µM (step 3). Analysis of the X-ray structure of this complex showed surprisingly that the naphthyl ring system had flipped by 180°, to let the diamido chain reach out to a second, 'noncatalytic' binding site. This noncatalytic binding site is much less conserved among tyrosine phosphatases and was expected to give better selectivity. Site-directed second-site screening of 10,000 compounds, this time following the resonance of Met258 in the noncatalytic binding site, identified several fragments that bound weakly ($K_D > 1$ mM) to the noncatalytic binding site. Chemically linking one of them to the existing ligand yielded the final compound with a K_i of 22 nM (step 4).

The linked-fragment approach in this case served to combine a fragment that gave affinity (steps 1–3) with a fragment that conveyed specificity in addition to affinity (step 4). By specifically monitoring residues in the catalytic site and the noncatalytic site, respectively, site-directed NMR screening was possible. It is further noteworthy that in this PTP-1B example, the linker was identified before the second-site ligand. In most other cases reported so far [49, 79, 82], the first-site ligand and the second-site ligand were identified first, and only then was a connecting linker designed.

The merged-fragment approach is exemplified by NMR-based discovery of lead compounds to inhibit DNA binding of the human papillomaviros E2 protein [85], a potential target against human papillomavirus infections, which can cause anogenital warts and cervical carcinomas. Conventional HTS of 100,000 compounds did not identify hits with activities better than 10 µM. NMR-based lead discovery started by producing large amounts of ^{15}N-labeled E2 protein (DNA binding domain) and screening of 2000 compounds by the HSQC method. Three classes of compounds were initially found to bind weakly to the E2 protein ($K_D \sim 1$ mM), and on two distinct binding sites. Compounds from the first two classes bound near the DNA recognition helix and inhibited DNA binding at high concentrations, whereas compounds from the third class bound to the β-barrel region of E2 protein and did not inhibit DNA binding at high concentrations. The first two classes of compounds were investigated further, and a series of analogs was synthesized. This led to compounds with IC50 values of 150 µM and 75 µM, respectively (fig. 9). Both optimized compounds caused similar chemical shift changes in the HSQC spectrum of E2, indicating that both classes bound to overlapping locations on E2. Therefore, linking both compounds was not attempted. Instead, it was decided to merge the structure-activity relationship of both compounds, reasoning that if both compounds bind at very similar locations, their SAR could be transferred. In practice, the dichloro substituents of one compound were transferred to the other, and with a concomitant change of the position of the butadiene spacer, a compound with an IC50 of 10 µM was discovered (fig. 9). Another example of the merged-fragment strategy is provided by work on Jnk3 inhibitors (see below) [17].

The design of directed combinatorial libraries based on NMR screening data and structural information shows

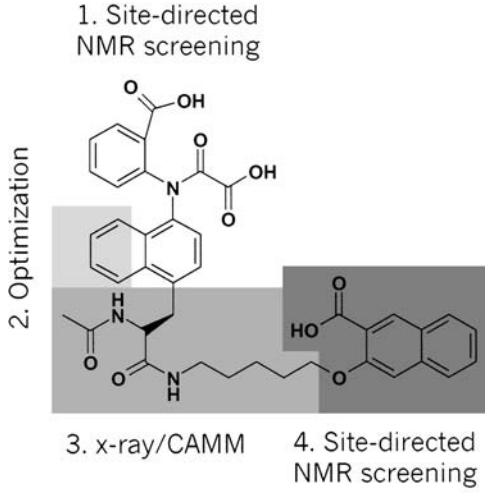

Figure 8. The design and discovery of ligands for PTP-1B by a linked-fragment approach [84].

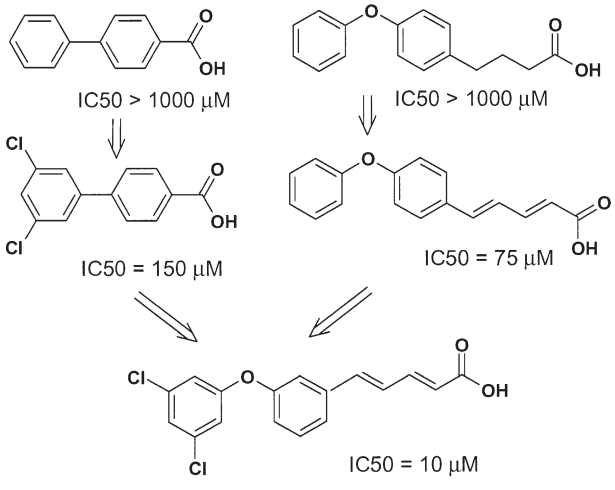

Figure 9. The design and discovery of ligands for human papillomavirus E2 protein [85].

the power of NMR also in combination with other methods. Combinatorial chemistry is able to generate compound libraries with 10^4 or more compounds. This appears high, but considering the large number of 'reasonable' drug candidates, in the order of 10^{62} [86], it is clear that even combinatorial libraries do not cover a significant fraction of possible drug candidates. In other words, combinatorial library design can vastly benefit from structural input, or from starting points identified by NMR screening. Conversely, combinatorial chemistry can greatly enhance the power of NMR screening. The synergies between both methods are described in the discovery of inhibitors of Erm methyltransferase, a potential target for the treatment of infections caused by pathogenic bacteria that are resistant to macrolide-lincosamide-streptogramin type B antibiotics [87]. Lead discovery was initiated by screening of a compound library against ^{15}N-ErmAM using the HSQC method. A number of compounds were found to bind to the active site of ErmAM with affinities around 1 mM, among them a member of the triazine family (fig. 10). Since compounds containing the triazine scaffold are readily amenable to optimization by parallel synthesis, a triazine library was generated for compound optimization. Triazines have three possible attachment points for side chains (fig. 10), and each of them was optimized separately. Starting from limited knowledge of an SAR, 232 compounds were synthesized with different substituents at R1, and 2-aminoindan was picked as the most active compound (IC50 = 8 μM). Keeping this R1 substituent, 411 compounds with different substituents at R2 were synthesized. However, none of these compounds increased activity more than twofold. Similarly, no improvements were obtained by varying the substituent at R3. Still, the best compounds resulting from this combination of parallel synthesis on an NMR-identified scaffold had K_I values in the low micromolar range. The structures of ErmAM and ErmC′ complexed with two ligands have been solved by NMR and X-ray crystallography, and may provide suitable impetus for further improvement.

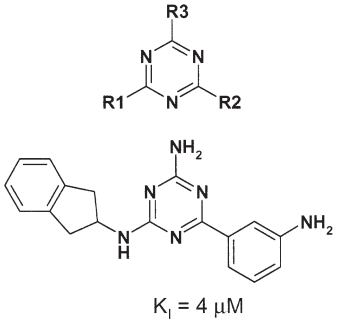

Figure 10. Design of combinatorial libraries based on NMR-detected scaffolds: Erm methyltransferase [87].

Synergies between NMR and combinatorial chemistry can also be exploited using bi-ligand targets, i.e. enzymes that require a common ligand in addition to the substrate. Important common ligands are NAD^+/$NADP^+$ for oxidoreductases, and ATP for kinases. Inhibitors for such bi-ligand targets are sought by combination of NMR screening, NMR-based structural analysis and combinatorial chemistry [19, 88]. Inhibitors of bi-ligand targets are considered to be composed of a common ligand mimetic and a substrate mimetic. The first step aims for the discovery of common ligand mimetics by site-specific NMR screening. Since the resonance assignments are generally not known and the target protein are usually too large for complete resonance assignment, only specific amino acid types, typically Met, Ile and Thr, are ^{13}C-methyl labeled and protonated, whereas the rest of the protein is deuterated. This leads to a much simplified spectrum with better NMR properties. HSQC and NOESY spectra upon addition of the common ligand then identify the residues in vicinity to the common ligand, and a screen for potential common ligand mimetics is performed by monitoring the very same resonances. Hits are structurally characterized in order to determine their binding orientation and to identify possible attachment points for the substrate mimetic. Subsequently, combinatorial libraries are created with the identified common ligand mimetics and a variety of potential substrate mimetics. Since the common ligand binding site is similar in all members of the enzyme family, such libraries are useful for a variety of enzyme family members [19, 88].

The most widely used follow-up strategy is the elaboration strategy. Here, hits from NMR screening are modified in a variety of ways, often guided by structural information and molecular modeling. This strategy has been successfully applied to urokinase [89], *E. coli* A-site RNA [51], 3α-HSD [65], adipocyte fatty acid binding protein FABP4 [32, 90] and other targets. Vertex's SHAPES strategy [91] belongs in this class, and one example will be described in the following.

Kinases are key enzymes in signal transduction pathways and are increasingly being recognized as important drug targets [92–94]. Most known kinase inhibitors compete with ATP for the ATP binding site in the catalytic domain. Chemically, most ATP competitive inhibitors are based on a phenylamino-pyrimidine scaffold, and novel classes of kinase inhibitors are highly desirable [95]. Recently, submicromolar ATP-competitive inhibitors of Jnk3 kinase were discovered by NMR-based approaches after HTS had failed to generate viable hits [17]. The basis of their strategy is the SHAPES library, which consists of a few hundred small compounds that comprise the most common drug scaffolds and are synthetically accessible [7, 75, 91]. This library was screened against Jnk3 kinase using ligand-observation techniques. 17 weakly binding

ligands were identified, and there was evidence for 13 of them to bind specifically to the ATP binding site. Since no crystal structures with these weakly binding ligands could be solved, broad substructure and similarity searches were carried out. The analogues were prioritized by virtual screening, and about 100 high-scoring compounds were selected around one of the various scaffolds of primary NMR hits (elaboration strategy). In addition, about 200 compounds were selected that comprised two or more of these scaffolds. This corresponds to the merged-fragment strategy mentioned above. From the weakly binding primary NMR hit (a pyrazole), the elaboration strategy resulted in similar compounds (isoxazoles) with more than 1000-fold higher potency (K_I = 790 nM; fig. 11). Chemical optimization could take advantage of the X-ray structures of Jnk3 complexes with this lead compound and of others found by the merged-fragment strategy. In particular, a subsite was detected that was occupied by some lead compounds and by the noncleavable ATP analog AMP-PNP, but not by compounds of the isoxazole class. Exploiting this subsite in a medicinal chemistry program led to a drastic increase in potency, so that the compounds could eventually be optimized to a K_I of less than 20 nM [17].

Discussion: NMR in lead finding, and alternatives to NMR

Why is NMR such a robust detection method for protein-ligand interactions? There are several reasons. On the one hand, the NMR assay is simple, since generally only protein and compounds are present. This eliminates the danger of undesired interactions of assay components other than the target protein. On the other hand, NMR makes it possible to monitor the integrity, concentration and purity of target protein and compounds: unfolded protein, poorly soluble or aggregating compounds and impurities in compounds will be detected. The dispersion of NMR resonances allows the selective detection of single compounds in the presence of others. Likewise, isotopic labeling of the protein allows one to selectively observe it even in the presence of high ligand concentrations. Nonspecific binding is the biggest source for false positives, but can be detected by competition experiments or with isotopically labeled protein. All of this renders the assay conditions maximally controlled and adds to a robust outcome.

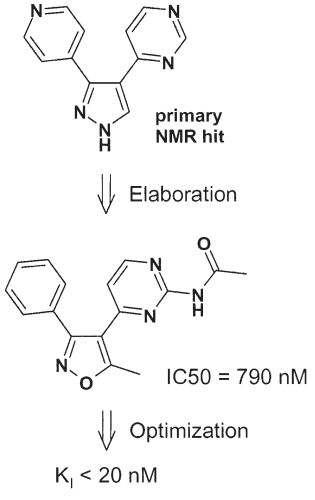

Figure 11. Elaboration and optimization of NMR SHAPES hits that resulted in high-affinity ligands for JNK3 kinase [17].

NMR is a robust assay sensitive for weak interactions and is thus well suited for fragment-based screening. It is not the only method, however, other technologies have been proposed for detection of weak interactions. Among them are X-ray crystallographic screening [96], screening by mass spectroscopy [97], possibly coupled with affinity selection [98] and preferably followed by NMR validation [11, 98], virtual screening [99] (preferably followed by NMR validation [80]), disulfide tethering [100], surface plasmon resonance and needle screening [101]. For low-throughput investigation of only a few compounds, isothermal titration calorimetry is very useful. Chemistry-driven fragment-based screening approaches have also been proposed [102–105]. Clearly, combinations of NMR with one or several of these methods has the highest chance for success.

Lead optimization

Once a lead compound has been identified, a medicinal chemistry program is initiated with the aim to improve the properties of the lead compound. Besides mere binding affinity, other properties to be considered are selectivity and the pharmacological and pharmacokinetic profile [12]. NMR can provide information that is useful for lead optimization in three different areas, which are discussed in the following section: structural and dynamic information of protein-ligand complexes, reverse screening, and detection and prevention of plasma protein binding.

Structural characterization of protein-ligand complexes

The use of structural information for lead optimization (structure-based drug design) was put forward in the early 1990s and has since experienced stages between hype and hope. Although pure de novo design of inhibitors based on structure still appears unrealistic in most cases, there are a number of drugs on the market or in clinical development for which structural information played an essential role [106, 107]. At most pharmaceutical companies, structural information is an important contributor to a drug design program, and this information is being sought and utilized whenever a target appears structurally feasible.

NMR can be used like X-ray crystallography to determine structures of biomolecules and their complexes with ligands [108]. Both methods are complementary, and each has its own strengths and weaknesses. For structure determination of large proteins or the complex of a large protein with a tightly bound inhibitor, structure determination by X-ray crystallography is generally faster. In particular, when a crystallizing protein construct has been found and the first complex structure has been solved, crystallography is often very fast in solving structures of other complexes. On the other hand, smaller proteins sometimes do not readily crystallize and can be more rapidly solved by NMR. Another strength of NMR is structure determination of weakly bound ligands, even if the receptor is large. This results in the bound conformation of the ligand, which can be used to generate a docking model of the complex, if information about the binding site is available or if protein-ligand NOEs have been detected. The NMR methods used for this type of structure determination are transferred NOE spectroscopy [109] and the measurement of transferred cross-correlated relaxation rates [110–113]. The term 'transferred' refers to the fact that the resonances of the free ligand are observed, but with the properties of the bound state transferred to the free state. This requires a fast dissociation rate of the ligand. This is why this type of analysis is only possible for weakly binding ligands, and is intrinsically connected to the fact that only weak or modest binding of ligands can be detected by ligand-observation techniques (see above).

Whereas full 3D structure determination by NMR of larger proteins or proteins tightly complexed with ligands is still a time-consuming undertaking, partial structural information can be rapidly obtained by NMR. For example, the ligand-binding epitope can be determined by STD methods [71]. This gives clues for the medicinal chemist about where to extend on a given lead, and can also be helpful for building a docked structure of the complex. On the protein side, analysis of chemical shift perturbations upon complexation can be used to rapidly generate a crude model of the complex [114]. Detection of a few protein-ligand NOEs also helps tremendously in building such a docking model, but requires ^{13}C-labeled protein and partial assignments. Competition experiments using known ligands of varying size which bind to different subpockets can be used to map the binding site on the protein of interesting leads [67]. Information about hydrogen bonding to backbone or side chains can also be rapidly gained by NMR [115, 116]. These examples show that although a full 3D structure determination by NMR may be time consuming, specific pieces of information are rapidly obtained and provide important information for lead optimization.

An alternative approach for lead optimization is the deliberate restriction of conformational flexibility in the free ligand. If the correct, i.e. bioactive conformation is frozen in the free ligand, enhanced binding affinity will result due to the reduced entropic cost for binding. If the wrong conformation is frozen, binding will be severely weakened or fully abolished. In the case of peptidic ligands, conformational flexibility can be reduced by cyclization of the peptide [117]. Introduction of a D-amino acid defines the conformation of a cyclic peptide well, and moving the D-amino acid through the sequence of a cyclic penta- or hexapeptide presents the binding epitope in modified ways. This approach has been successfully used for the design of ligands for the $\alpha_V\beta_3$ integrin. The binding epitope RGD (Arg-Gly-Asp) was presented in different but well-defined conformations by introduction and movement of a D-amino acid. The cyclic peptide c(RGDfV) where f refers to D-phenylalanine, is the most potent compound discovered by this 'spatial screening' method [118, 119]. Furthermore, this constrained pentapeptide could be further optimized by N-methylation [120]. The resulting N-methylated pentapeptide is one of the most potent and selective inhibitors of vitronectin binding to the $\alpha_V\beta_3$ integrin. It prevents tumor-induced angiogenesis and is currently in phase II clinical trials for the treatment of tumor metastasis of patients with renal or colon cancer. In principle, the conformation of the active constrained peptide can be used as structural input for depeptidization in hopes of discovering a nonpeptidic compound that retains activity [121].

Dynamic characterization of protein-ligand complexes

One of the key strengths of biomolecular NMR is its ability to characterize protein dynamics on a time scale from pico/nanoseconds to milliseconds. This can be achieved for protein backbone or side chain atoms and with a spatial resolution of individual amino acid residues [122–127]. Outcomes of this analysis are the frequencies (correlation time) of fast internal motions and their amplitudes. The latter are often referred to as the square of the order parameter. This order parameter ranges from 0 to 1 and indicates the correlation of a particular bond vector with the overall motion of the protein. Assumption of a specific motional model such as the diffusion-in-a-cone model then enables a physically realistic picture of the motion. For example, an order parameter of 0.8 for an N-H moiety, as is commonly measured for 'rigid' residues without particular flexibility, translates into a 20° variation of the backbone N-H vector orientation.

The physically most interesting parameter is conformational entropy, which is related to the order parameter but is not deduced without ambiguity [127]. Conformational entropies can be measured with a resolution of individual amino acid residues, in contrast to conventional thermodynamic measurements such as titration calorimetry. It has often been observed that several conformations of a protein receptor exist in solution, out of which one is

frozen upon complex formation with a ligand [127]. Flexibility is then reduced by complex formation, and the corresponding entropic cost is compensated by favorable enthalpic interactions. This entropy-enthalpy compensation has indeed been observed in many cases and is often restricted to the ligand binding site. There are also other cases, however, in which flexibility of certain regions increases upon ligand binding [128]. This clearly stabilizes complex formation, but the origin and consequenes of such behavior are subject to speculation. It appears as if specific interactions such as hydrogen bonds or salt bridges require rigid conformations, whereas hydrophobic interactions can in some cases lead to increased flexibility [128].

Not only do protein experience changes in conformational entropy upon ligand binding; the ligand does also. It is generally believed that the ligand has high conformational entropy in the unbound state, and loses most of it upon binding. This would correspond to a large loss in entropy that impairs binding. This was realized early, and it is now generally accepted that it is advantageous to conformationally restrict a ligand in its unbound form in order to pre-form in solution the bioactive conformation [117]. However, a ligand can also retain some flexibility when it is bound. For example, it was shown that a tightly bound ligand for matrix metalloproteinase 1 (MMP-1) adopts multiple conformations in the bound form. Two of these different bound conformations exchange slowly on the chemical shift time scale so that two sets of signals of the bound ligand are observed [129].

While characterization of protein and ligand dynamics by NMR is well developed and represents a clear asset of NMR, this information cannot yet be fully exploited in drug discovery. So far, dynamic characterization of protein-ligand complexes by NMR has had limited impact on drug design, and it seems that the ability of molecular modeling and chemistry to integrate and exploit this information remains to be developed.

Reverse screening

An alternative method of lead optimization is fragmentation of the lead compound into component molecules, and screening for alternative fragments which are then incorporated back into the original lead [130]. This strategy is useful if a single fragment of a lead is suspected to add little or negative contributions to binding affinity, or to be responsible for poor physicochemical, pharmacokinetic or toxicological properties. The principle of reverse screening is shown in figure 12. The screening for alternative fragments follows the same concept as NMR screening for lead generation. Here it is especially important to apply site-specific screening in order to be sure that the replacement binds competitively at the same binding site as the original fragment. This approach has been successfully applied to lead optimization of adenosine kinase inhibitors [130] and of inhibitors of the matrix metalloproteinase stromelysin (MMP-3) [131]. In the latter case, 1-naphthyl hydroxamate was identified as replacement for alkylhydroxamate as the zinc-chelating group. This replacement led to better pharmacokinetic properties of the improved lead compound. Similar results were obtained for LFA-1 [12].

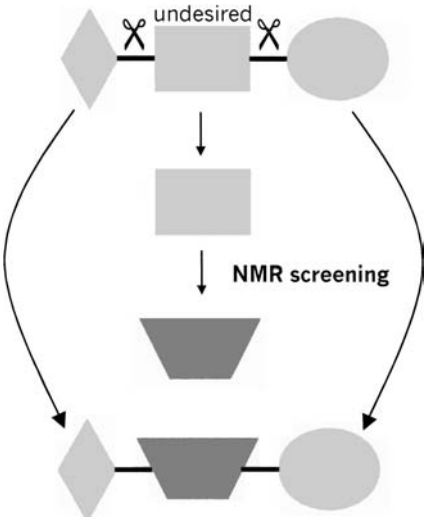

Figure 12. Principle of reverse screening for lead optimization. The lead compound is fragmented and NMR screening identifies a replacement for the undesired fragment, which is then chemically incorporated into the original lead.

Detection and prevention of plasma protein binding

One of the reasons a lead compound that shows good in vitro activity performs poorly in in vivo experiments is binding to plasma proteins such as human serum albumin (HSA). While some binding to plasma proteins may be useful to retain a constant pool of compound and protect the drug from enzymatic attack, excessive binding is undesired since it represents a compound sink that severely reduces efficacy of the drug [132]. When serum albumin binding is a severe problem, lead optimization programs are primarily aimed at reducing binding affinity to albumin, rather than at increasing binding affinity to the target. Structures of the complex between albumin and the lead compound can help to reach this goal. In this case, the purpose of structural work is to design out binding affinity, which can in some cases be achieved by mere addition of a methyl group, if the correct position is known. HSA contains three homologous domains I–III where each domain is divided into two subdomains, A and B. However, many drugs bind to one of two binding sites of HSA, located in subdomains IIA and IIIA (termed Sudlow sites I and II). Sudlow site II is the main binding site for the hydrophobic aromatic moieties common in drugs and may therefore be in many cases the most important target site to reduce HSA binding. Subdomain III con-

taining this site has been cloned and isotopically labeled, and its structure in complex with diflunisal, a cyclooxygenase inhibitor with antiinflammatory activity, has been determined by NMR [133]. On the basis of this structure, diflunisal analogs were designed and synthesized. Some of them show a more than 100-fold reduced affinity to subdomain III of HSA while retaining at least some activity against cyclooxygenase-2. Interestingly, however, affinity to full-length HSA was only reduced by a factor of 10, indicating that other HSA binding sites also contribute to compound binding.

This example shows that NMR-derived structural information cannot only be used to design in binding affinity for a desired target, but also to design out binding affinity to undesired proteins such as HSA. Two papers by Dalvit et al. show that binding affinities of compounds to HSA can elegantly and rapidly be measured by competition experiments using tryptophane as marker for Sudlow site II, or warfarin as marker for Sudlow site I [134, 135].

Conclusions

NMR is an extremely versatile biophysical technique, and its applications can be tailored to answer a variety of different questions. In most pharmaceutical companies, NMR detection and characterization of molecular interactions, for example between a target protein and low molecular weight compounds, is considered more attractive than the structure determination aspect. The abilities to deal with weak interactions and to structurally characterize binding events are unique features of NMR which have important applications in all phases of drug discovery:

- In early stages, the drugability of a potential target can be estimated by screening a small diverse library for binders. A high hit rate suggests a high likelihood for the presence of a binding site for low molecular weight compounds. If a functional genomics library of naturally occurring ligands is screened, the nature of the hits may suggest the function of the target protein, which is often unknown if it comes from a functional genomics project.
- For hit validation, NMR is an extremely valuable method because it is robust and remarkably little susceptible to false-positive or false-negative detection, if data are properly interpreted. Separating promising compounds from false positives is the main task of this NMR application, and several real-life examples have been described where NMR provided evidence whether a compound should or should not be promoted to lead compound. In some cases, no other method has the ability to generate sufficient data for this decision. Hit validation by NMR is an area of high impact in pharmaceutical research.
- For hit generation, NMR is currently one of the methods of choice for fragment-based approaches. NMR screening (or SAR-by-NMR) has a growing track record of discovering submicromolar lead candidates even for difficult targets where conventional HTS failed. In addition, NMR can identify novel scaffolds in cases where high-affinity ligands are known but have undesired properties. The principles of NMR screening, several strategies for the follow up of compounds, and some illustrative examples have been described.
- For lead optimization, a full 3D structure determination is often more rapidly achieved by X-ray crystallography, but partial information such as the conformation of a weakly bound ligand can be rapidly generated by NMR and is quite valuable for molecular modeling and medicinal chemistry. In addition, reverse screening is an attractive concept for lead optimization.

Hits are fine, leads are better, but what about drugs? Are there any drugs on the market or in clinical development that were identified or optimized or otherwise strongly influenced by NMR? The answer is yes: two compounds from NMR screening have reached or are about to reach clinical development ([12]). Among those are compounds that target MMP-3 [136] and LFA-1 [137]. In addition, Merck (Germany) is developing the N-methylated c(RGDfV) cyclic peptide that was identified by spatial screening.

Those compounds identified by NMR that are now in clinical trials are the most spectacular proof of the value of biomolecular NMR in drug discovery. But even aside from those compounds, biomolecular NMR has a strong and positive impact on drug discovery programs.

Frequently asked questions

What is the limit in protein size for NMR?

For structure determination, the limit is around 30–40 kDa, although larger proteins can be solved in favorable cases. Transverse relaxation optimized spectroscopy (TROSY) techniques are indispensable for work with larger (deuterated) proteins [138, 139]. For resonance assignment, the limit can be up to 80 kDa [140], but typically proteins >40–50 kDa become difficult due to signal overlap. This limit also holds for the detection and characterization of molecular interactions, if protein is observed. For ligand observation (NMR screening, ligand epitope mapping, competition experiments, and so on) the protein size does not matter.

Do I have to produce isotopically (^{13}C, ^{15}N) labeled protein for an NMR study?

For structure determination or resonance assignment of all but the smallest proteins, yes. Recommended labeling is ^{15}N for proteins <10 kDa, ^{13}C and ^{15}N for proteins <20 kDa and ^{13}C, ^{15}N and ^{2}H for larger proteins. Note that TROSY techniques work particularly well with deuterated proteins. For ligand observation (NMR screening, ligand epitope mapping, competition experiments, and so on) and for structure determination of weakly bound ligands, no isotope labeling of the protein is necessary.

Is isotopically labeled protein harmful?

No, no, no! ^{13}C and ^{15}N are stable isotopes and are not radioactive.

What are the requirements for an NMR study?

Structure determination by NMR needs about 200 µl of well-behaved protein (>90% purity) at a concentration of 0.4–1 mM (about 10–40 mg/ml). Determination of optimal sample conditions usually requires additional amounts. For an NMR screening project, several milligrams of unlabeled protein at a concentration of a few hundred nM to 10 µM and purity >80–90% are required. Availability of isotopically labeled protein for follow-up studies, and the 3D structure of the protein are desirable. Known ligands are an advantage for competition experiments.

What is so special about the NMR assay?

The NMR assay is a universal binding assay that needs no adaptation to a particular protein. It is very robust since only protein and potential ligands are present, since quality, purity and concentration of protein and ligands are tightly monitored, and since NMR allows the selective observation of an interesting species (e.g. protein, reporter ligand) even in the presence of high concentrations of other compounds. The NMR assay is sensitive for weak interactions with K_D up to 1–10 mM. This is because even a small percentage of bound ligand can be reliably detected. For the same reason, and unlike for other techniques, weak affinities can be detected even if the concentrations of protein and ligands are far below K_D.

What are the risks of false-positive/false-negative detection of interactions?

A potential pitfall for false-positive detection of protein interactions with a low molecular weight ligand is nonspecific compound binding to the protein. This risk can be minimized by the use of proper control experiments such as competitive binding assays (which requires a known tightly binding active-site ligand) or HSQC experiments (which require ^{15}N-labeled protein). Self-aggregating compounds can also cause a false-positive response, but this is easily detected in control experiments without protein.

Tightly binding compounds with low dissociation rates or medium affinity ligands with slow binding kinetics are subject to false-negative detection, unless competition experiments (requiring a weakly binding reporter ligand) or experiments involving (sub-)stoichiometric compound concentrations are carried out. Another pitfall for false-negative detection is low compound solubility, if protein detection as in the HSQC method is employed. However, compound solubility can be easily measured by ^{1}H spectra.

Should I involve X-ray or NMR in my project?

If possible, both! NMR and X-ray crystallography are complementary. Particularly for lead generation, our experience shows that ideal synergies between both techniques involve identification of ligands by NMR and structure determination of the complex by crystallography.

Acknowledgements. We thank our colleagues from the Novartis NMR group, Drs M. Blommers, C. Fernandez, S. Rüdisser, U. Hommel, B. Cutting, X. Zhang, and Y.-C. Li for continuous interactions and discussions. Dr P. Flörsheim and A. Widmer are thanked for excellent project collaborations and computational support. Discussions with and support from Dr H. Kubinyi are gratefully acknowledged.

1 Rudin M. and Weissleder R. (2003) Molecular imaging in drug discovery and development. Nat. Rev. Drug Discov. **2:** 123–131
2 Nicholson J. K., Connelly J., Lindon J. C. and Holmes E. (2002) Metabonomics: a platform for studying drug toxicity and gene function. Nat. Rev. Drug Discov. **1:** 153–161
3 Shockcor J. P. and Holmes E. (2002) Metabonomic applications in toxicity screening and disease diagnosis. Curr. Top. Med. Chem. **2:** 35–51
4 Stockman B. J. (1998) NMR spectroscopy as a tool for structure-based drug design. Prog. NMR Spectrosc. **33:** 109–151
5 Hajduk P. J., Meadows R. P. and Fesik S. W. (1999) NMR-based screening in drug discovery. Q. Rev. Biophys. **32:** 211–240
6 Diercks T., Coles M. and Kessler H. (2001) Applications of NMR in drug discovery. Curr. Opin. Chem. Biol. **5:** 285–291
7 Peng J. W., Lepre C. A., Fejzo J., Abdul-Manan N. and Moore J. M. (2001) Nuclear magnetic resonance-based approaches for lead generation in drug discovery. Methods Enzymol. **338:** 202–230
8 Ross A. and Senn H. (2001) Automation of measurements and data evaluation in biomolecular NMR screening. Drug Discovery Today **6:** 583–593
9 Pellecchia M., Sem D. S. and Wuthrich K. (2002) NMR in drug discovery. Nat. Rev. Drug Discov. **1:** 211–219
10 Stockman B. J. and Dalvit C. (2002) NMR screening techniques in drug discovery and drug design. Prog. NMR Spectrosc. **41:** 187–231
11 Hajduk P. J. and Burns D. J. (2002) Integration of NMR and high-throughput screening. Comb. Chem. High Throughput Screen. **5:** 613–621

12 Huth J. R. and Sun C. (2002) Utility of NMR in lead optimization: fragment-based approaches. Comb. Chem. High Throughput Screen. **5:** 631–643
13 Wyss D. F., McCoy M. A. and Senior M. M. (2002) NMR-based approaches for lead discovery. Curr. Opin. Drug Discovery Dev. **5:** 630–647
14 Zerbe O. (Ed.) (2003) BioNMR in drug research, Wiley, Weinheim
15 Meyer B. and Peters T. (2003) NMR spectroscopy techniques for screening and identifying ligand binding to protein receptors. Angew. Chem. Int. Ed. Engl. **42:** 864–890
16 Zartler E. R., Yan J., Mo H., Kline A. D. and Shapiro M. J. (2003) 1D NMR Methods in ligand-receptor interactions. Curr. Top. Med. Chem. **3:** 25–37
17 Fejzo J., Lepre C. and Xie X. (2003) Application of NMR screening in drug discovery. Curr. Top. Med. Chem. **3:** 81–97
18 Coles M., Heller M. and Kessler H. (2003) NMR-based screening technologies. Drug Discovery Today **8:** 803–810
19 Sem D. S. and Pellecchia M. (2001) NMR in the acceleration of drug discovery. Curr. Opin. Drug Discovery Dev. **4:** 479–492
20 Shapiro M. J. and Wareing J. R. (1999) High resolution NMR for screening ligand/protein binding. Curr. Opin. Drug Discovery Dev. **2:** 396–400
21 Hopkins A. L. and Groom C. R. (2002) The druggable genome. Nat. Rev. Drug Discov. **1:** 727–730
22 Kubinyi H. (2003) Drug research: myths, hype and reality. Nat. Rev. Drug Discov. **2:** 665–668
23 Clore G. M. and Gronenborn A. M. (1998) NMR structure determination of proteins and protein complexes larger than 20 kDa. Curr. Opin. Chem. Biol. **2:** 564–570
24 Staunton D., Owen J. and Campbell I. D. (2003) NMR and structural genomics. Acc. Chem. Res. **36:** 207–214
25 Montelione G. T., Zheng D., Huang Y. J., Gunsalus K. C. and Szyperski T. (2000) Protein NMR spectroscopy in structural genomics. Nat. Struct. Biol. **7:** 982–985
26 Minn A. J., Velez P. A., Schendel S. L., Liang H., Muchmore S. W., Fesik S. W. et al. (1997) Bcl-xL forms an ion channel in synthetic lipid membranes. Nature **385:** 353–357
27 Lomasney J. W., Cheng H.-F., Wang L.-P., Kuan Y. S., Liu S. M., Fesik S. W. et al. (1996) Phosphatidylinositol 4,5-bisphosphate binding to the pleckstrin homology domain of phospholipase C-.delta.1 enhances enzyme activity. J. Biol. Chem. **271:** 25316–25326
28 Wright P. E. and Dyson H. J. (1999) Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm. J. Mol. Biol. **293:** 321–331
29 Fletcher C. M., McGuire A. M., Gingras A. C., Li H., Matsuo H., Sonenberg N. et al. (1998) 4E binding proteins inhibit the translation factor eIF4E without folded structure. Biochemistry **37:** 9–15
30 Dames S. A., Martinez-Yamout M., De Guzman R. N., Dyson H. J. and Wright P. E. (2002) Structural basis for Hif-1.alpha./CBP recognition in the cellular hypoxic response. Proc. Natl. Acad. Sci. USA **99:** 5271–5276
31 Steinmetz M. O., Kammerer R. A., Jahnke W., Goldie K. N., Lustig A. and Van Oostrum J. (2000) Op18/stathmin caps a kinked protofilament-like tubulin tetramer. EMBO J. **19:** 572–580
32 Lepre C. A., Peng J., Fejzo J., Abdul-Manan N., Pocas J., Jacobs M. et al. (2002) Applications of SHAPES screening in drug discovery. Comb. Chem. High Throughput Screen. **5:** 583–590
33 Hajduk P. J., Betz S. F., Mack J., Ruan X., Towne D. L., Lerner C. G. et al. (2002) A strategy for high-throughput assay development using leads derived from nuclear magnetic resonance-based screening. J. Biomol. Screen. **7:** 429–432
34 Stockman B. J., Farley K. A. and Angwin D. T. (2001) Screening of compound libraries for protein binding using flow-injection nuclear magnetic resonance spectroscopy. Methods Enzymol. **338:** 230–246
35 Fischer J. J. and Jardetzky O. (1965) Nuclear magnetic relaxation study of intermolecular complexes. The mechanism of penicillin binding to serum albumin. J. Am. Chem. Soc. **87:** 3237–3244
36 Meadows D. H., Roberts G. C. and Jardetzky O. (1969) Nuclear magnetic resonance studies of the structure and binding sites of enzymes. 8. Inhibitor binding to ribonuclease. J. Mol. Biol. **45:** 491–511
37 Balaram P., Bothner-By A. A. and Breslow E. (1973) Nuclear magnetic resonance studies of the interaction of peptides and hormones with bovine neurophysin. Biochemistry **12:** 4695–4704
38 De Marco A., Laursen R. A. and Llinas M. (1986) 1H-NMR spectroscopic manifestations of ligand binding to the kringle 4 domain of human plasminogen. Arch. Biochem. Biophys. **244:** 727–741
39 Hammond S. J., Birdsall B., Feeney J., Searle M. S., Roberts G. C. and Cheung H. T. (1987) Structural comparisons of complexes of methotrexate analogues with *Lactobacillus casei* dihydrofolate reductase by two-dimensional 1H NMR at 500 MHz. Biochemistry **26:** 8585–8590
40 Petros A. M., Ramesh V. and Llinas M. (1989) 1H NMR studies of aliphatic ligand binding to human plasminogen kringle 4. Biochemistry **28:** 1368–1376
41 Thewes T., Constantine K., Byeon I. J. and Llinas M. (1990) Ligand interactions with the kringle 5 domain of plasminogen. A study by 1H NMR spectroscopy. J. Biol. Chem. **265:** 3906–3915
42 Rejante M. R., Byeon I. J. and Llinas M. (1991) Ligand specificity of human plasminogen kringle 4. Biochemistry **30:** 11081–11092
43 Ramesh V., Frederick R. O., Syed S. E., Gibson C. F., Yang J. C. and Roberts G. C. (1994) The interactions of *Escherichia coli* trp repressor with tryptophan and with an operator oligonucleotide. NMR studies using selectively 15N-labelled protein. Eur. J. Biochem. **225:** 601–608
44 Dwek R. A. (1973) Nuclear Magnetic Resonance (N.M.R) in Biochemistry. Applications to enzyme systems, Oxford Univ. Press, New York
45 Jardetzky O. and Roberts G. C. K. (1981) NMR in Molecular Biology, Academic Press, New York
46 Craik D. J. and Higgins K. A. (1990) NMR studies of ligand-macromolecule interactions. Annual Reports on NMR Spectroscopy **22:** 61–138
47 Feeney J. (1994) NMR studies of protein ligand interactions. NATO ASI Series, Series H: Cell Biology **87:** 115–154
48 Otting G. (1993) Experimental NMR techniques for studies of protein-ligand interactions. Curr. Opin. Struct. Biol. **3:** 760–768
49 Shuker S. B., Hajduk P. J., Meadows R. P. and Fesik S. W. (1996) Discovering high-affinity ligands for proteins: SAR by NMR. Science **274:** 1531–1534
50 Hajduk P. J., Augeri D. J., Mack J., Mendoza R., Yang J., Betz S. F. et al. (2000) NMR-Based screening of proteins containing 13C-labeled methyl groups. J. Am. Chem. Soc. **122:** 7898–7904
51 Yu L., Oost T. K., Schkeryantz J. M., Yang J., Janowick D. and Fesik S. W. (2003) Discovery of aminoglycoside mimetics by NMR-based screening of *Escherichia coli* A-site RNA. J. Am. Chem. Soc. **125:** 4444–4450
52 Hajduk P. J., Olejniczak E. T. and Fesik S. W. (1997) One-dimensional relaxation- and diffusion-edited NMR methods for screening compounds that bind to macromolecules. J. Am. Chem. Soc. **119:** 12257–12261
53 Jahnke W., Ruedisser S. and Zurini M. (2001) Spin label enhanced NMR screening. J. Am. Chem. Soc. **123:** 3149–3150

54 Chen A. and Shapiro M. J. (1999) Affinity NMR. Anal. Chem. **71:** 669A–675A
55 Lian L. Y., Barsukov I. L., Sutcliffe M. J., Sze K. H. and Roberts G. C. (1994) Protein-ligand interactions: exchange processes and determination of ligand conformation and protein-ligand contacts. Methods Enzymol. **239:** 657–700
56 Meyer B., Weimar T. and Peters T. (1997) Screening mixtures for biological activity by NMR. Eur. J. Biochem. **246:** 705–709
57 Mayer M. and Meyer B. (1999) Characterization of ligand binding by saturation transfer difference NMR spectroscopy. Angew. Chem. Int. Ed. Engl. **38:** 1784–1788
58 Dalvit C., Pevarello P., Tato M., Veronesi M., Vulpetti A. and Sundstrom M. (2000) Identification of compounds with binding affinity to proteins via magnetization transfer from bulk water. J. Biomol. NMR. **18:** 65–68
59 Chen A. and Shapiro M. J. (2000) NOE pumping as high-throughput method to determine compounds with binding affinity to macromolecules by NMR. J. Am. Chem. Soc. **122:** 414–415
60 Klein J., Meinecke R., Mayer M. and Meyer B. (1999) Detecting binding affinity to immobilized receptor proteins in compound libraries by HR-MAS STD NMR. J. Am. Chem. Soc. **121:** 5336–5337
61 Meinecke R. and Meyer B. (2001) Determination of the binding specificity of an integral membrane protein by saturation transfer difference NMR: RGD peptide ligands binding to integrin alphaIIbbeta3. J. Med. Chem. **44:** 3059–3065
62 Benie A. J., Moser R., Bauml E., Blaas D. and Peters T. (2003) Virus-ligand interactions: identification and characterization of ligand binding by NMR spectroscopy. J. Am. Chem. Soc. **125:** 14–15
63 Mayer M. and James T. L. (2002) Detecting ligand binding to a small RNA target via saturation transfer difference NMR experiments in D(2)O and H(2)O. J. Am. Chem. Soc. **124:** 13376–13377
64 Dalvit C., Flocco M., Knapp S., Mostardini M., Perego R., Stockman B. J. et al. (2002) High-throughput NMR-based screening with competition binding experiments. J. Am. Chem. Soc. **124:** 7702–7709
65 Jahnke W., Floersheim P., Ostermeier C., Zhang X., Hemmig R., Hurth K. et al. (2002) NMR reporter screening for the detection of high-affinity ligands. Angew. Chem. Int. Ed. Engl. **41:** 3420–3423
66 Siriwardena A. H., Tian F., Noble S. and Prestegard J. H. (2002) A straightforward NMR-spectroscopy-based method for rapid library screening. Angew. Chem. Int. Ed. Engl. **41:** 3454–3457
67 Mayer M. and Meyer B. (2000) Mapping the active site of angiotensin-converting enzyme by transferred NOE spectroscopy. J. Med. Chem. **43:** 2093–2099
68 Peng J. W. (2001) Cross-correlated 19F relaxation measurements for the study of fluorinated ligand-receptor interactions. J. Magn. Reson. **153:** 32–47
69 Dalvit C., Flocco M., Veronesi M. and Stockman B. J. (2002) Fluorine-NMR competition binding experiments for high-throughput screening of large compound mixtures. Comb. Chem. High Throughput Screen. **5:** 605–611
70 Dalvit C., Fagerness P. E., Hadden D. T., Sarver R. W. and Stockman B. J. (2003) Fluorine-NMR experiments for high-throughput screening: theoretical aspects, practical considerations and range of applicability. J. Am. Chem. Soc. **125:** 7696–7703
71 Mayer M. and Meyer B. (2001) Group epitope mapping by saturation transfer difference NMR to identify segments of a ligand in direct contact with a protein receptor. J. Am. Chem. Soc. **123:** 6108–6117
72 Dalvit C., Fasolini M., Flocco M., Knapp S., Pevarello P. and Veronesi M. (2002) NMR-based screening with competition water-ligand observed via gradient spectroscopy experiments: detection of high-affinity ligands. J. Med. Chem. **45:** 2610–2614
73 Yan J., Kline A. D., Mo H., Shapiro M. J. and Zartler E. R. (2003) The effect of relaxation on the epitope mapping by saturation transfer difference NMR. J. Magn. Reson. **163:** 270–276
74 Rossi C., Donati A. and Sansoni M. R. (1992) Nuclear magnetic resonance as a tool for the identification of specific DNA-ligand interaction. Chem. Phys. Lett. **189:** 278–280
75 Lepre C. A. (2001) Library design for NMR-based screening. Drug Discovery Today **6:** 133–140
76 Lepre C. (2003) Strategies for NMR screening and library design. In: BioNMR in Drug Research, pp. 391–415, Zerbe O. (ed.), Wiley, Weinheim
77 Jacoby E., Davies J. and Blommers M. J. (2003) Design of small molecule libraries for NMR screening and other applications in drug discovery. Curr. Top. Med. Chem. **3:** 11–23
78 Hann M. M., Leach A. R. and Harper G. (2001) Molecular complexity and its impact on the probability of finding leads for drug discovery. J. Chem. Inf. Comput. Sci. **41:** 856–864
79 Jahnke W., Florsheimer A., Blommers M. J., Paris C. G., Heim J., Nalin C. M. et al. (2003) Second-site NMR screening and linker design. Curr. Top. Med. Chem. **3:** 69–80
80 Rudisser S. and Jahnke W. (2002) NMR and in silico screening. Comb. Chem. High Throughput Screen. **5:** 591–603
81 Verlinde C. L. M. J., Rudenko G. and Hol W. G. J. (1992) In search of new lead compounds for trypanosomiasis drug design: a protein structure-based linked-fragment approach. J. Comput.-Aided Mol. Des. **6:** 131–147
82 Hajduk P. J., Sheppard G., Nettesheim D. G., Olejniczak E. T., Shuker S. B., Meadows R. P. et al. (1997) Discovery of potent nonpeptide inhibitors of stromelysin using SAR by NMR. J. Am. Chem. Soc. **119:** 5818–5827
83 Olejniczak E. T., Hajduk P. J., Marcotte P. A., Nettesheim D. G., Meadows R. P., Edalji R. et al. (1997) Stromelysin inhibitors designed from weakly bound fragments: effects of linking and cooperativity. J. Am. Chem. Soc. **119:** 5828–5832
84 Szczepankiewicz B. G., Liu G., Hajduk P. J., Abad-Zapatero C., Pei Z., Xin Z. et al. (2003) Discovery of a potent, selective protein tyrosine phosphatase 1B inhibitor using a linked-fragment strategy. J. Am. Chem. Soc. **125:** 4087–4096
85 Hajduk P. J., Dinges J., Miknis G. F., Merlock M., Middleton T., Kempf D. J. et al. (1997) NMR-based discovery of lead inhibitors that block DNA binding of the human papillomavirus E2 protein. J. Med. Chem. **40:** 3144–3150
86 Bohacek R. S., McMartin C. and Guida W. C. (1996) The art and practice of structure-based drug design: a molecular modeling perspective. Med. Res. Rev. **16:** 3–50
87 Hajduk P. J., Dinges J., Schkeryantz J. M., Janowick D., Kaminski M., Tufano M. et al. (1999) Novel inhibitors of Erm methyltransferases from NMR and parallel synthesis. J. Med. Chem. **42:** 3852–3859
88 Pellecchia M., Meininger D., Dong Q., Chang E., Jack R. and Sem D. S. (2002) NMR-based structural characterization of large protein-ligand interactions. J. Biomol. NMR **22:** 165–173
89 Hajduk P. J., Boyd S., Nettesheim D., Nienaber V., Severin J., Smith R. et al. (2000) Identification of novel inhibitors of urokinase via NMR-based screening. J. Med. Chem. **43:** 3862–3866
90 Van Dongen M. J. P., Uppenberg J., Svensson S., Lundbaeck T., Kerud T., Wikstroem M. et al. (2002) Structure-based screening as applied to human FABP4: a highly efficient alternative to HTS for hit generation. J. Am. Chem. Soc. **124:** 11874–11880
91 Fejzo J., Lepre C. A., Peng J. W., Bemis G. W., Ajay, Murcko M. A. et al. (1999) The SHAPES strategy: an NMR-based ap-

proach for lead generation in drug discovery. Chem. Biol. **6:** 755–769
92 Fabbro D., Ruetz S., Buchdunger E., Cowan-Jacob S. W., Fendrich G., Liebetanz J. et al. (2002) Protein kinases as targets for anticancer agents: from inhibitors to useful drugs. Pharmacol. Ther. **93:** 79–98
93 Traxler P. (2003) Tyrosine kinases as targets in cancer therapy – successes and failures. Expert Opin. Ther.Targets. **7:** 215–234
94 Capdeville R., Buchdunger E., Zimmermann J. and Matter A. (2002) Glivec (STI571, imatinib), a rationally developed, targeted anticancer drug. Nat. Rev. Drug Discov. **1:** 493–502
95 Traxler P. and Furet P. (1999) Strategies toward the design of novel and selective protein tyrosine kinase inhibitors. Pharmacol. Ther. **82:** 195–206
96 Blundell T. L., Jhoti H. and Abell C. (2002) High-throughput crystallography for lead discovery in drug design. Nat. Rev. Drug Discov. **1:** 45–54
97 Swayze E. E., Jefferson E. A., Sannes-Lowery K. A., Blyn L. B., Risen L. M., Arakawa S. et al. (2002) SAR by MS: a ligand based technique for drug lead discovery against structured RNA targets. J. Med. Chem. **45:** 3816–3819
98 Moy F. J., Haraki K., Mobilio D., Walker G., Powers R., Tabei K. et al. (2001) MS/NMR: a structure-based approach for discovering protein ligands and for drug design by coupling size exclusion chromatography, mass spectrometry and nuclear magnetic resonance spectroscopy. Anal. Chem. **73:** 571–581
99 Abagyan R. and Totrov M. (2001) High-throughput docking for lead generation. Curr. Opin. Chem. Biol. **5:** 375–382
100 Braisted A. C., Oslob J. D., Delano W. L., Hyde J., McDowell R. S., Waal N. et al. (2003) Discovery of a potent small molecule IL-2 inhibitor through fragment assembly. J. Am. Chem. Soc. **125:** 3714–3715
101 Boehm H.-J., Boehringer M., Bur D., Gmuender H., Huber W., Klaus W. et al. (2000) Novel inhibitors of DNA gyrase: 3D structure based biased needle screening, hit validation by biophysical methods and 3D guided optimization. A promising alternative to random screening. J. Med. Chem. **43:** 2664–2674
102 Lewis W. G., Green L. G., Grynszpan F., Radic Z., Carlier P. R., Taylor P. et al. (2002) Click chemistry in situ: acetylcholinesterase as a reaction vessel for the selective assembly of a femtomolar inhibitor from an array of building blocks. Angew. Chem. Int. Ed. Engl. **41:** 1053–1057
103 Otto S., Furlan R. L. and Sanders J. K. (2002) Recent developments in dynamic combinatorial chemistry. Curr. Opin. Chem. Biol. **6:** 321–327
104 Maly D. J., Choong I. C. and Ellman J. A. (2000) Combinatorial target-guided ligand assembly: identification of potent subtype-selective c-Src inhibitors. Proc. Natl. Acad. Sci. USA **97:** 2419–2424
105 Ramstrom O. and Lehn J. M. (2002) Drug discovery by dynamic combinatorial libraries. Nat. Rev. Drug Discov. **1:** 26–36
106 Kubinyi H. (1998) Structure-based drug design. Chimica Oggi **16:** 17–22
107 Kubinyi H. (1998) Combinatorial and computational approaches in structure-based drug design. Curr. Opin. Drug Discovery Dev. **1:** 16–27
108 Wüthrich K. (1995) NMR – this other method for protein and nucleic acid structure determination. Acta Crystallogr. D. Biol. Crystallogr. **D51:** 249–270
109 Ni F. (1994) Recent developments in transferred NOE methods. Prog. NMR Spectrosc. **26:** 517–606
110 Carlomagno T., Felli I. C., Czech M., Fischer R., Sprinzl M. and Griesinger C. (1999) Transferred cross-correlated relaxation: application to the determination of sugar pucker in an aminoacylated tRNA-mimetic weakly bound to EF-Tu. J. Am. Chem. Soc. **121:** 1945–1948
111 Blommers M. J. J., Stark W., Jones C. E., Head D., Owen C. E. and Jahnke W. (1999) Transferred cross-correlated relaxation complements transferred NOE: Structure of an IL-4R-derived peptide bound to STAT-6. J. Am. Chem. Soc. **121:** 1949–1953
112 Carlomagno T., Blommers M. J., Meiler J., Jahnke W., Schupp T., Petersen F. et al. (2003) The high-resolution solution structure of epothilone A bound to tubulin: an understanding of the structure-activity relationships for a powerful class of antitumor Agents. Angew. Chem. Int. Ed. Engl. **42:** 2511–2515
113 Peng J. W. (2003) New probes of ligand flexibility in drug design: transferred (13)C CSA-dipolar cross-correlated relaxation at natural abundance. J. Am. Chem. Soc. **125:** 11116–11130
114 McCoy M. A. and Wyss D. F. (2002) Spatial localization of ligand binding sites from electron current density surfaces calculated from NMR chemical shift perturbations. J. Am. Chem. Soc. **124:** 11758–11763
115 Gargaro A. R., Frenkiel T. A., Nieto P. M., Birdsall B., Polshakov V. I., Morgan W. D. et al. (1996) NMR detection of arginine-ligand interactions in complexes of *Lactobacillus casei* dihydrofolate reductase. Eur. J. Biochem. **238:** 435–439
116 Grzesiek S., Cordier F. and Dingley A. J. (2001) Scalar couplings across hydrogen bonds. Methods Enzymol. **338:** 111–133
117 Kessler H. (1982) Peptide conformations. Part 19. Conformation and biological effects of cyclic peptides. Angew. Chem. **94:** 509–520
118 Haubner R., Finsinger D. and Kessler H. (1997) Stereoisomeric peptide libraries and peptidomimetics for designing selective inhibitors of the $\alpha_v\beta_3$ integrin for a new cancer therapy. Angew. Chem. Int. Ed. Engl. **36:** 1374–1389
119 Gottschalk K.-E. and Kessler H. (2002) The structures of integrins and integrin-ligand complexes: implications for drug design and signal transduction. Angew. Chem. Int. Ed. Engl. **41:** 3767–3774
120 Dechantsreiter M. A., Planker E., Mathae B., Lohof E., Hoelzemann G., Jonczyk A. et al. (1999) N-Methylated cyclic RGD peptides as highly active and selective $\alpha_v\beta_3$ integrin antagonists. J. Med. Chem. **42:** 3033–3040
121 Gottschling D., Boer J., Schuster A., Holzmann B. and Kessler H. (2002) Combinatorial and rational strategies to develop non-peptidic $\alpha_4\beta_7$-integrin antagonists from cyclic peptides. Angew. Chem. Int. Ed. Engl. **41:** 3007–3011
122 Peng J. W. and Wagner G. (1994) Investigation of protein motions via relaxation measurements. Methods Enzymol. **239:** 563–596
123 Palmer A. G. III, Kroenke C. D. and Loria J. P. (2001) Nuclear magnetic resonance methods for quantifying microsecond-to-millisecond motions in biological macromolecules. Methods Enzymol. **339:** 204–238
124 Kay L. E. (1998) Protein dynamics from NMR. Nat. Struct. Biol. **5:** 513–517
125 Fushman D. and Cowburn D. (2001) Nuclear magnetic resonance relaxation in determination of residue-specific 15N chemical shift tensors in proteins in solution: protein dynamics, structure and applications of transverse relaxation optimized spectroscopy. Methods Enzymol. **339:** 109–126
126 Skrynnikov N. R., Millet O. and Kay L. E. (2002) Deuterium spin probes of side-chain dynamics in proteins. 2. Spectral density mapping and identification of nanosecond time-scale side-chain motions. J. Am. Chem. Soc. **124:** 6449–6460
127 Stone M. J. (2001) NMR relaxation studies of the role of conformational entropy in protein stability and ligand binding. Acc. Chem. Res. **34:** 379–388
128 Zidek L., Novotny M. V. and Stone M. J. (1999) Increased protein backbone conformational entropy upon hydrophobic ligand binding. Nat. Struct. Biol. **6:** 1118–1121

129 Moy F. J., Chanda P. K., Chen J., Cosmi S., Edris W., Levin J. I. et al. (2002) Impact of mobility on structure-based drug design for the MMPs. J. Am. Chem. Soc. **124:** 12658–12659
130 Hajduk P. J., Gomtsyan A., Didomenico S., Cowart M., Bayburt E. K., Solomon L. et al. (2000) Design of adenosine kinase inhibitors from the NMR-based screening of fragments. J. Med. Chem. **43:** 4781–4786
131 Hajduk P. J., Shuker S. B., Nettesheim D. G., Craig R., Augeri D. J., Betebenner D. et al. (2002) NMR-based modification of matrix metalloproteinase inhibitors with improved bioavailability. J. Med. Chem. **45:** 5628–5639
132 Peters T. Jr (1996) All about Albumin: Biochemistry, Genetics and Medical Applications, Academic Press, San Diego
133 Mao H., Hajduk P. J., Craig R., Bell R., Borre T. and Fesik S. W. (2001) Rational design of diflunisal analogues with reduced affinity for human serum albumin. J. Am. Chem. Soc. **123:** 10429–10435
134 Dalvit C., Flocco M., Stockman B. J. and Veronesi M. (2002) Competition binding experiments for rapidly ranking lead molecules for their binding affinity to human serum albumin. Comb. Chem. High Throughput Screen. **5:** 645–650
135 Dalvit C., Hadden D. T., Sarver R. W., Ho A. M. and Stockman B. J. (2003) Multi-selective one dimensional proton NMR experiments for rapid screening and binding affinity measurements. Comb Chem High Throughput Screen. **6:** 445–453
136 Wada C. K., Holms J. H., Curtin M. L., Dai Y., Florjancic A. S., Garland R. B. et al. (2002) Phenoxyphenyl sulfone N-formylhydroxylamines (retrohydroxamates) as potent, selective, orally bioavailable matrix metalloproteinase inhibitors. J. Med. Chem. **45:** 219–232
137 Liu G., Huth J. R., Olejniczak E. T., Mendoza R., DeVries P., Leitza S. et al. (2001) Novel p-arylthio cinnamides as antagonists of leukocyte function-associated antigen-1/intracellular adhesion molecule-1 interaction. 2. Mechanism of inhibition and structure-based improvement of pharmaceutical properties. J. Med. Chem. **44:** 1202–1210
138 Pervushin K. V. (2003) Transverse relaxation optimized spectroscopy. Biological Magnetic Resonance **20:** 3–34
139 Wüthrich K. and Wider G. (2002) Transverse relaxation-optimized NMR spectroscopy with biomacromolecular structures in solution. Encyclopedia of Nuclear Magnetic Resonance **9:** 468–477
140 Tugarinov V., Muhandiram R., Ayed A. and Kay L. E. (2002) Four-dimensional NMR spectroscopy of a 723-residue protein: chemical shift assignments and secondary structure of malate synthase G. J. Am. Chem. Soc. **124:** 10025–10035