Theory and Applications of NMR-Based Screening in Pharmaceutical Research

Christopher A. Lepre,[†,‡] Jonathan M. Moore,[*,†] and Jeffrey W. Peng[‡,§]

Vertex Pharmaceuticals Incorporated, 130 Waverly Street, Cambridge, Massachusetts 02139-4242; and Department of Chemistry and Biochemistry, University of Notre Dame, 251 Nieuwland Science Hall, Notre Dame, Indiana 46556-5670

Received April 1, 2004

Contents

* E-mail: Jonathan_Moore@vrtx.com.
† Vertex Pharmaceuticals Incorporated.
§ University of Notre Dame.
‡ Authors contributed equally to this work.

1. Introduction

Since the first report appeared in 1996[1] describing the use of NMR spectroscopy to screen for potential drug molecules, the field of NMR-based screening has evolved rapidly. Over the last several years, a variety of novel approaches have been introduced and have found widespread application in both pharmaceutical and academic research settings. These NMR-driven platforms have provided new pathways for inhibitor design against an increasing number of therapeutically relevant drug targets. In pharmaceutical research, NMR screening has become an important component in an integrated arsenal of biophysical, biochemical, and computational methods designed to discover and optimize drug leads.

The advent of NMR screening as a new discipline could not have come at a more opportune time. In the early and mid-1990s, NMR research in preclinical drug discovery had been relegated to a small niche, its role primarily that of answering key structural questions for that small subset of NMR-accessible (e.g. MW < 20 kDa; expressible in *E. coli*) drug targets. However, drug design programs require a substantial body of structural information beyond the simple unliganded structure of the target protein. Typically, an iterative loop of structure determination, modeling, and chemistry is carried out, followed by biochemical and cell-based assays to validate the chemical leads. In many cases, dozens of high-resolution X-ray structures of a drug target, with several distinct classes of chemical scaffolds, are necessary to achieve the potency, selectivity, pharmacokinetic, and toxicological properties required of a preclinical drug candidate. Then, and still today, NMR structural methods were simply not capable of generating structural information at a useful (<2.5 Å) resolution for drug design on a time scale compa-

10.1021/cr030409h CCC: $48.50
Published on Web 07/24/2004

Christopher Lepre received his Ph.D. in Chemistry from MIT in 1989 and completed his postdoctoral fellowship with Dr. Peter Wright in the Department of Molecular Biology at the Scripps Research Institute, where he studied protein structure determination by NMR. He joined Vertex Pharmaceuticals as a founding scientist in 1991. Chris is presently a Principal Investigator in Structural Biology, where his principal research interests include applications of NMR structural and screening methods in lead generation, library design, and kinase drug discovery.

Jeff Peng received his Ph.D. in Biophysics from the University of Michigan in 1993, developing new methods for studies of protein dynamics in the laboratory of Professor Gerhard Wagner. He then did postdoctoral research with Prof. Dr. Richard Ernst at ETH-Zürich, where he focused on small molecule dynamics and hydration. Afterward, he returned to the Boston area to join the Protein NMR Group at Vertex Pharmaceuticals, where his primary focus was NMR methods development for the study of protein–ligand interactions. After 9 years at Vertex, in August 2003 Jeff joined the Chemistry & Biochemistry faculty at the University of Notre Dame. His central research interest is investigating the relationships between molecular flexibility and function in biological problems.

rable to that of X-ray crystallography. Consequently, many structural biology groups have leveraged their NMR resources to focus on the more critical path and opportunistic approaches afforded by NMR screening. Because NMR screening can in many cases provide important clues regarding the best route to druglike lead molecules early and rapidly in a drug discovery program, these methods have all but replaced the NMR structural strategies used during the past decade in pharmaceutical research. Beyond NMR-based screening, NMR practitioners are also branching into drug development. One of the most exciting new areas of research is the field of metabonomics. These methods rely on NMR spectroscopy of biofluids such as urine, plasma, and CSF, followed by chemometric analysis of the data, to classify xenobiotic toxicities in animals and humans. These methods

Jonathan Moore received his Ph.D. in Biophysics from the University of Pennsylvania in 1985. After postdoctoral studies with Dr. Peter Wright at the Scripps Research Institute from 1986 to 1989, Jon was a Swiss National Science Foundation International Postdoctoral Fellow in the laboratory of Prof. Dr. Kurt Wuthrich at the ETH (Zurich). In 1990, Jon joined Vertex Pharmaceuticals Incorporated as a founding scientist, where he started and has led the Protein NMR Group until the present. Jon is currently a Senior Research Fellow and Head of Structural Biology at Vertex, where his research interests include applications of NMR and X-ray crystallographic techniques in drug discovery.

have been reviewed extensively[2−5] and are beyond the scope of the present work.

In this review, we will focus on both theoretical aspects and practical considerations for the most commonly used NMR-based screening approaches, combining literature review with our own experiences in the laboratory. Because there are many excellent reviews available on the subject,[6−11] we will focus on more recent literature when possible. We will begin by discussing the theoretical aspects of chemical exchange phenomena that lay at the foundation of NMR screening experiments, followed by a detailed description of commonly used ligand and receptor-based NMR approaches and their implementation in the laboratory. Finally, using a framework based on common ligand design strategies, we will review examples of how these NMR methods have been applied to generate and optimize new chemical classes of ligands for real drug targets.

2. Binding Equilibria

As a framework for interpreting and implementing NMR screening experiments, it is useful to review basic facts concerning binding equilibria. More complete treatments can be found in any introductory text in biochemistry (see e.g. ref 12) or enzyme kinetics and mechanism.[13] By presenting these well-known equations in the general forms given below, they can be interpreted in the context of multiple NMR observables. It also allows well-understood principles of biology and physical chemistry to be distinguished from new developments in NMR.

2.1. Simple One-Site Binding

The binding process underlying most NMR screening experiments can be described by

$$[\mathrm{E}] + [\mathrm{L}] \underset{k_{\mathrm{off}}}{\overset{k_{\mathrm{on}}}{\rightleftharpoons}} [\mathrm{EL}] \qquad (1)$$

Equation 1 represents a dynamic equilibrium involving three species: the free receptor E, the free ligand L, and the receptor−ligand complex EL. The unimolecular rate constant k_{off} is inversely proportional to the mean lifetime τ_B of the receptor−ligand complex. The bimolecular rate constant k_{on} measures the probability of a productive encounter between free receptor and ligand. In many textbooks and literature articles, k_{on} is often assumed to be the diffusion-limited on rate. This "constant" value cited can vary between 10^7 and 10^9 M^{-1} s^{-1}. It is important to note that the assumption of a universal rate constant for diffusion-limited encounters is an approximation that does not account for the potential complexity of intermolecular forces that may amplify or attenuate the encounter frequency.[14]

The binding affinity can be described by the temperature-dependent equilibrium dissociation constant, $K_D = [E][L]/[EL] = k_{off}/k_{on}$. Combining the definition of K_D with that of the bound receptor fraction $P_B^E = [EL]/([E] + [EL])$ yields

$$P_B^E = \frac{[L]}{[L] + K_D} \quad (2)$$

P_B^E is the fractional occupation of the receptor binding site by ligand L. Equation 2 is essentially the Langmuir isotherm rearranged.[15] Furthermore, eq 2 is a hyperbolic function of [L]; hence, increasing [L] increases P_B^E, although by progressively smaller amounts. When $[L] \ll K_D$, P_B^E is proportional to [L]. When $[L] = K_D$, the receptor is half-saturated; that is, half of the receptor molecules exist in a one-to-one complex with the ligand. When $[L] \gg K_D$, the receptor is completely saturated and $P_B^E = 1.0$. In this limit, every receptor binding site is occupied by a ligand, which on average exchanges with another distinct ligand every $\approx 1/k_{off}$ s. The significance of K_D is that ligands of weaker affinity have larger K_D and thus require the addition of more ligand to saturate the receptor binding site.

It is useful to rewrite the receptor−ligand concentration [EL] in terms of the known experimental parameters E_T, the total receptor concentration, and L_T, the total ligand concentration. Substitution of the constraints $E_T = [EL] + [E]$ and $L_T = [EL] + [L]$ into $K_D = [E][L]/[EL]$ yields

$$[EL] = {}^1/_2(E_T + L_T + K_D) - {}^1/_2\sqrt{(E_T + L_T + K_D)^2 - 4E_TL_T} \quad (3)$$

Figure 1 uses eq 3 to plot the bound receptor fraction $P_B^E = [EL]/E_T$ as a function of L_T for several values of K_D. It is clear from the plots for the right and center curves corresponding to $K_D = 50$ μM and $K_D = 500$ μM that the receptor is half-saturated when $L_T \approx K_D$. However, under conditions where $[E] \sim K_D$, for example, the curve corresponding to $K_D = 5$ μM, this is only approximate, since depletion of the ligand must also be taken into account.

The bound *ligand* fraction $P_B = [EL]/L_T$ assumes values in the range $0 \le P_B \le 1/\epsilon$, where $\epsilon = L_T/E_T$, the ligand molar excess. The upper limit for the bound ligand fraction P_B occurs in the limit of low L_T, where the receptor binding site is saturated, that is, when $[EL] = E_T$. Note that saturation conditions also occur in the opposite limit, $L_T \gg E_T$ and $L_T \gg K_D$; however, under these conditions P_B approaches its lower limit of zero. Theoretical curves illustrating P_B versus total ligand concentration L_T are shown in Figure 2 for the same three K_D values and enzyme concentration used in Figure 1. For simplicity, in this review, P_B will always refer to bound ligand fraction, while the bound receptor fraction will always be denoted by P_B^E.

By careful control of receptor and ligand concentrations, it is possible to "select" the maximum K_D for which P_B results in an observable NMR signal. Although, for $L_T \gg E_T$, reduction of the receptor concentration E_T reduces P_B more or less indepen-

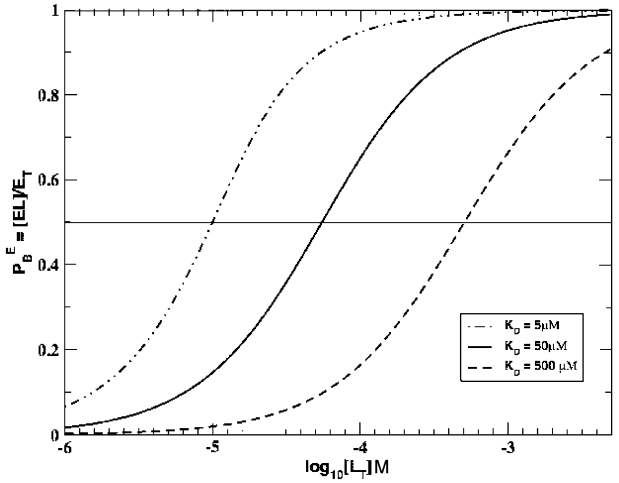


Figure 1. Simulation of the bound receptor fraction [EL]/E_T as a function of increasing ligand concentration L_T. The plot assumes the single site equilibrium [E] + [L] ⇔ [EL]. The total receptor concentration $E_T = 10$ μM. Three ligand affinities are plotted: double-dotted−dashed curve, $K_D = 5$ μM; solid curve, $K_D = 50$ μM; dashed curve, $K_D = 500$ μM. The horizontal line indicates the half-saturation condition.

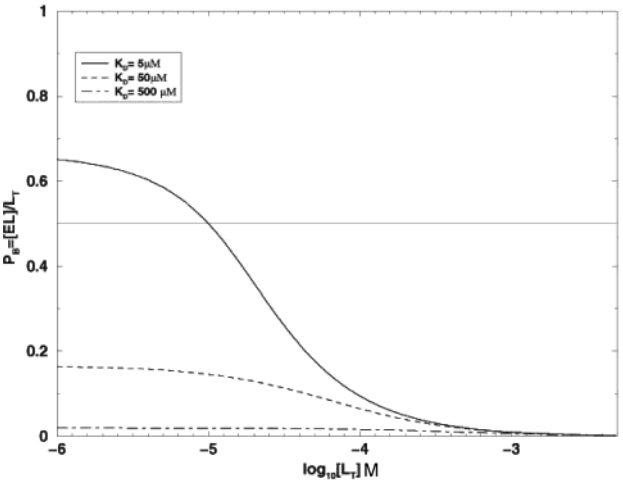


Figure 2. Simulation of the bound ligand fraction [EL]/L_T as a function of increasing ligand concentration L_T. The plot assumes the single site equilibrium [E] + [L] ⇔ [EL]. The total receptor concentration $E_T = 10$ μM. Three ligand affinities are plotted: solid curve, $K_D = 5$ μM; dashed curve, $K_D = 50$ μM; dotted−dashed curve, $K_D = 500$ μM. The horizontal line indicates the half-saturation condition.

dently of K_D, this effect may be used to reduce relative P_B values such that, for example, ligands with $K_D \sim 1$ mM can no longer be detected, while those with $K_D \sim 10-100\ \mu$M may still be observed. This "tuning" of the detection threshold is very useful in ligand-based screening experiments such as the STD experiment (section 5.5), particularly when screening large compound libraries.

2.2. Competitive Binding Equilibria

Competition binding experiments are well-established methods for determination of both ligand binding affinity and specificity to macromolecules. Recently, NMR-based protocols employing competition methods have been proposed.[16−19] These approaches expand the utility of existing experiments to characterize binding for compounds of higher affinity, for example, sub-micromolar binders, or those compounds that are sparingly soluble. For this reason, it is useful to review the binding equilibria that apply in these situations. First, we consider the simplest case, in which two ligands L and I compete for the same binding site. We take L to be a previously characterized ligand. I represents a new ligand, or "inhibitor". We distinguish their equilibrium dissociation constants using $K_D = [E][L]/[EL]$ and $K_I = [E][I]/[EI]$. The appropriate equilibrium is written

$$[EI] \underset{k^I_{off}}{\overset{k^I_{on}}{\leftrightarrow}} [E] + [I] + [L] \underset{k^L_{off}}{\overset{k^L_{on}}{\leftrightarrow}} [EL] \tag{4}$$

Addition of I reduces [EL] through competitive displacement. Using the constraints $I_T = [EI] + [I]$, $K_I = [E][I]/[EI]$, and $E_T = [E] + [EI] + [EL]$, it is possible to write the perturbed bound receptor fraction $[EL]/E_T$ as

$$P^E_{B,+I} = \frac{[L]}{[L] + K_{D,app}} \tag{5}$$

Equation 5 has the same form as the unperturbed fraction P^E_B of eq 2 if we define the apparent dissociation constant $K_{D,app}$ as

$$K_{D,app} = K_D\left(1 + \frac{[I]}{K_I}\right) = \frac{(E_T - [EL])(L_T - [EL])}{[EL]} \tag{6}$$

As expected, the factor $(1 + [I]/K_I)$ must be ≥ 1, and thus $K_{D,app} \geq K_D$ and $P^E_{B,+I} \leq P^E_B$. This higher value reflects the reduction of available receptor sites for L due to competitive interference from I.

If the inhibitor dissociation constant, $K_I = [E][I]/[EI]$ is known, then it becomes possible to estimate the K_D of L by assessing the displacement caused by I. Following Cheng and Prusoff,[20] we define I_{50} as the value of [I] that causes $P^E_{B,+I} = 0.5P^E_B$. Then, setting eq 5 equal to half of eq 2 yields

$$K_D = \frac{[L]K_I}{I_{50} - K_I} \tag{7}$$

If the new ligand L is in molar excess, then $[L] \approx L_T$. K_D can then be estimated from the values of I_{50}, K_I, and $[L] \approx L_T$.

Again, it is useful to express [EL] and [EI] in terms of the known experimental variables E_T, L_T, and I_T as well as the two dissociation constants K_I and K_D. Wang and Sigurskjold[21,22] provide the desired expressions for the fractions of bound L and I,

$$x_{EL} = [EL]/E_T = (L_T/E_T)x_E/\{(K_D/E_T) + x_E\} \tag{8}$$

$$x_{EI} = [EI]/E_T = (I_T/E_T)x_E/\{(K_I/E_T) + x_E\} \tag{9}$$

where x_E is the free receptor fraction $x_E = [E]/E_T$. Substitution of eqs 8 and 9 into the constraint $x_E + x_{EL} + x_{EI} = 1.0$ yields a cubic equation in x_E. The relevant root is

$$x_E = \frac{2\sqrt{\alpha^2 - 3\beta}\cos(\Theta/3) - \alpha}{3} \tag{10}$$

where

$$\Theta = \cos^{-1}\left\{\frac{-2\alpha^3 + 9\alpha\beta - 27\gamma}{2\sqrt{(\alpha^2 - 3\beta)^3}}\right\} \tag{11a}$$

and

$$\alpha = (1/E_T)(K_I + K_D + I_T + L_T - E_T) \tag{11b}$$

$$\beta = (1/E_T)^2\{K_I(L_T - E_T) + K_D(I_T - E_T) + K_IK_D\} \tag{11c}$$

$$\gamma = -K_IK_D/E_T^2 \tag{11d}$$

Substitution of the free fraction x_E into eqs 8 and 9 yields x_{EL} and x_{EI}, and thus [EL] and [EI].

Figure 3 plots $[EL]_{+I}/[EL]_{-I}$ as a function of I_T for four representative values of K_I (K_I = 0.1, 1.0, 10.0, and 50.0 μM) with fixed values for $E_T = 1\ \mu$M and $K_D = 50\ \mu$M. $[EL]_{+I}/[EL]_{-I}$ is simply the ratio of [EL] in the presence of inhibitor to [EL] in the absence of inhibitor. The x-axis plots I_T normalized to the fixed total ligand concentration ($L_T = 100\ \mu$M). The maximum value of $[EL]_{+I}/[EL]_{-I} = 1.0$ and occurs when no inhibitor is present. As inhibitor concentration increases, [EL] decreases, as does the ratio $[EL]_{+I}/[EL]_{-I}$. For $K_I \ll K_D$, the inhibitor I acts as a "knockoff", in the sense that it displaces L for $I_T \ll L_T$. Such an inhibitor can be a very useful reagent; if I is known to bind specifically and L is a test ligand, this effect can be used to verify the binding specificity of the test ligand. In addition, if L is a known specific binder and I is the test ligand, this effect can be used in NMR-based screens to identify previously unknown tight binding ligands.

3. Accounting for Binding-Induced Chemical Exchange

In section 2, we reviewed basic principles of simple and competitive binding equilibria. In this section, we will discuss how chemical exchange, driven by the

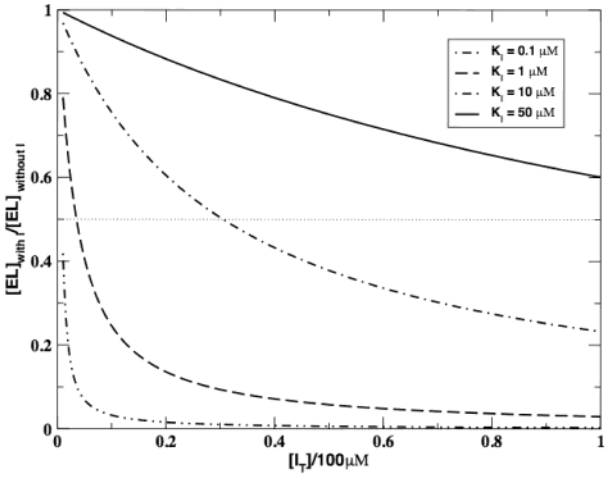


Figure 3. Simulation of competitive displacement of ligand L by inhibitor I. The plotted quantity is $[EL]_{+I}/[EL]_{-I}$, which is [EL] in the presence of I over [EL] in the absence of I. The ratio is plotted as a function of inhibitor concentration normalized to a fixed ligand concentration of $L_T = 100\ \mu M$. The total receptor concentration E_T is $1\ \mu M$, and the affinity of L is $K_D = 50\ \mu M$. The four curves correspond to four inhibitor affinities: double-dotted−dashed curve, $K_I = 0.1\ \mu M$; dashed curve, $K_I = 1\ \mu M$; dotted−dashed curve, $K_I = 10\ \mu M$; solid curve, $K_I = 50\ \mu M$. The horizontal line indicates a reduction by 50% of [EL] from the noninhibited state.

binding equilibria, modulates the NMR parameters of the free and bound states of the ligand and receptor. Observation of these modulated parameters forms the basis for all NMR screening experiments.

In the two-state equilibrium given by eq 1, ligand and receptor molecules will exist in either a free (L, E) or complexed (EL) state. In the free state, both receptor and ligand retain their intrinsic NMR parameters (e.g. chemical shifts, relaxation rates, translational diffusion coefficients). In each other's presence, the mutual binding affinity of ligand and receptor drives an exchange process that toggles both sets of molecules between the free and complexed states. At equilibrium, they adopt free and bound state populations ([E], [L], [EL]) consistent with eq 1. Under these conditions, the ligand transiently adopts NMR parameters characteristic of the typically much larger receptor. Alternatively, from the receptor's perspective, the ligand transiently perturbs the binding site microenvironment(s), which may alter distribution of conformations sampled by the ensemble of receptor molecules. In either case, the exchange modulates the NMR parameters of both molecules.

A complete discussion of how chemical exchange modifies the NMR parameters (dynamic NMR) should include a description of the modified Bloch equation formalism of Hahn, Maxwell, and McConnell.[23,24] This formalism (HMM) provides an excellent theoretical framework to describe the majority of exchange phenomena that occur in NMR screening experiments, under all exchange regimes. Such a treatment is beyond the scope of the present work but can be found in a previous review.[7] For most experiments described here, it is sufficient to consider only the case of fast exchange.

3.1. Fast Exchange Approximation

The solutions to the HMM equations[23,24] describe the behavior of system magnetization on arbitrary exchange time scales. In NMR screening practice, however, these equations are almost never solved, and fast exchange is simply assumed. This assumption is made for two reasons. First, the experimental conditions for ligand-based NMR screening are often well-suited to fast exchange. These experiments are typically carried out with $L_T/E_T > 10$, and the binding compounds, or "hits", have $K_D \geq 100\ \mu M$. If k_{on} is well-approximated by a diffusion-limited value ($10^7-10^9\ M^{-1}\ s^{-1}$), then the slowest k_{ex} values lie in the range $1000 < k_{ex} < 100\ 000\ s^{-1}$. Ligand-based NMR screening methods are primarily ^1H based; consequently, k_{ex} exceeds most differences in intrinsic ^1H relaxation rates and rotating frame precession frequencies, thus providing assurance that the fast exchange assumption is valid.

A second motivation for assuming fast exchange is the resulting algebraic simplicity. Exchange-modulated NMR parameters can be awkward.[7] However, the descriptions of these parameters under fast exchange are simple sums. Generally, the NMR parameters Q become the simple averages

$$Q_{avg} = P_B Q_B + P_F Q_F \qquad (12a)$$

$$Q_{avg} = P_B Q_B + P_F Q_F + Q_{ex} \qquad (12b)$$

Here, Q_{avg} is the exchange-averaged parameter we observe for the ligand/receptor in the presence of the receptor/ligand. Observed differences between Q_{avg} and Q_F provide a signature of receptor binding and indicate a hit in a NMR screen based on that parameter. In the case of eq 12a, Q_{avg} is a simple population-weighted average. For eq 12b, Q_{avg} has an additional offset term Q_{ex}. The form shown in eq 12a applies to those parameters Q for which chemical shift modulations are not relevant. These parameters include longitudinal autorelaxation and cross-relaxation rates, rotating-frame spin-locking autorelaxation and cross-relaxation rates, and translational diffusion coefficients. The bound state contribution in eqs 12a,b is $P_B Q_B$. The ability to detect binding with adequate sensitivity depends critically on $P_B Q_B$ being significant relative to $P_F Q_F$. However, typical screening conditions where $L_T \gg E_T$ make $P_B \ll P_F$. For this reason, it is much preferred to measure NMR parameters Q that become *amplified* in the bound state (i.e. $Q_B \gg Q_F$).

The population-weighted average of eq 12a can be related to ligand binding affinity. Since $P_F = 1 - P_B$ and $P_B = P_B^E/\epsilon$, we can write the difference $\epsilon(Q_{avg} - Q_F)$ as

$$\epsilon(Q_{avg} - Q_F) = \frac{(Q_B - Q_F)[L]}{[L] + K_D} \qquad (13)$$

The functional form of eq 13 describes the now familiar hyperbolic dose−response curve we have already seen for P_B^E in eq 2. To experimentally determine $Q_{avg} - Q_F$, the Q-related NMR signal intensities are measured in the presence of receptor

and then corrected for the free state contribution Q_F. Q_F is determined either by performing a reference experiment for the compound in the absence of receptor or by designing an experimental protocol in which $Q_F = 0$. Equation 13 demonstrates that $\epsilon(Q_{avg} - Q_F)$ increases with ligand addition and then plateaus at $(Q_B - Q_F)$ when the binding site is saturated ($L_T \gg K_D$). Provided ligand solubility is not problematic, ligand titration data can be fitted to this simple relation to estimate binding affinity. Specifically, for a large ligand excess ($\epsilon = L_T/E_T \gg 1$), we can approximate $[L] = L_T$, and a fit of ligand titration data to eq 13 provides estimates for K_D and $(Q_B - Q_F)$.

4. Ligand-Based versus Receptor-Based Screening

Screening may proceed by ligand- or receptor-based methods. Receptor-based methods observe and compare the NMR parameters of the receptor molecule resonances in the presence and absence of compound mixtures. Thus far, the receptor-based methods have focused mostly on proteins. Such methods incorporate site specific characterization afforded by assigned protein NMR spectra along with a priori knowledge of the protein's three-dimensional structure (either from X-ray or NMR) to drive lead generation. By identifying perturbations of assigned protein resonances, not only are ligands identified, but also their binding sites are localized. This site specific characterization of binding suggests strategies for fragment-based lead generation, in which lower affinity molecular fragments binding to distinct subsites can be linked or elaborated to yield higher affinity compounds.[1,25] Localization of binding sites also enables one to immediately distinguish specific from nonspecific binding. Finally, unlike ligand-based methods, receptor-based methods do not rely on fast exchange to retrieve bound state information. Observation of receptor resonances permits the characterization of both higher and lower affinity hits.

A major caveat affecting receptor-based methods is that many therapeutically important pharmaceutical targets are not amenable to NMR spectroscopy. NMR methods demand physicochemical properties of the protein target that present progressively more difficult challenges. First, milligram quantities of soluble, nonaggregated protein must be overexpressed and purified. Then, suitable expression hosts must be found that permit isotope enrichment (e.g. ^{13}C, ^{15}N, ^{2}H) critical for the resonance assignment of typically large (>30 000 Da) therapeutic targets. For example, the most popular NMR expression host, *E. coli*, is often not an option for mammalian proteins whose overexpression may prove to be toxic to the host cell. After sufficient quantities of labeled protein are available, it must be ensured that the sample is stable for the time required for sequential resonance assignment. Although new data acquisition approaches promise to accelerate resonance assignment, it can still be a relatively lengthy process (weeks) for the large *monomeric* proteins (>30 000 Da) routinely encountered in pharmaceutical research. Unfortunately, the time required for NMR assignment of such targets inevitably favors other approaches, such as X-ray crystallography, that can provide high-resolution structural information to medicinal chemistry on a faster time scale.

Alternatively, the typical implementation of ligand-based methods compares the NMR parameters of a mixture of compounds in the presence and absence of the receptor molecules. This approach renders the molecular weight of the receptor molecule irrelevant. In fact, the most powerful ligand-based approaches become more sensitive when dealing with larger receptors. Additionally, ligand observation bypasses the need to produce milligram quantities of isotope-labeled receptor. Depending on the approach, less than a milligram of *unlabeled* protein is required for these experiments (receptor concentration is often ≤ 1 μM, and no assignments are necessary). This allows the spectroscopist to evaluate new targets more rapidly. This is important not only for adapting to constantly shifting priorities in drug discovery but also for contributing on a time scale useful for chemistry and high-throughput screens.

An obvious disadvantage of ligand-based approaches is the inability to localize the binding site of the small molecule hits on the receptor. Also, ligand-based approaches rely on the exchange-mediated transfer of bound state information to the free state. This requirement biases ligand-based methods toward identification of weakly binding ligands (rapid exchange) and the use of large ligand molar excesses ($L_T/E_T \gg 1$). The consequent risk is that, under these conditions, ligand may start to occupy weaker affinity nonspecific binding sites. However, recent developments in ligand-based methods discussed below are beginning to address these difficulties.[17−19]

Both receptor- and ligand-based approaches have distinct advantages and disadvantages. Clearly, if receptor-based methods are possible (low molecular mass, efficient *E. coli* expression, available resonance assignments), then the potentially higher information content obtainable makes these the methods of choice. However, due to the scarcity of low-molecular-weight drug targets, ligand-based screening is, in general, of broader applicability and places less demands on other research disciplines and infrastructure. For these reasons, we dedicate a considerable portion of this review to discussion of ligand-based screening methods.

5. Ligand-Based NMR Screening Methods

The exquisite sensitivity of ligand NMR parameters to the differences between bound and free states forms the basis for all ligand-based screening experiments. Generally, NMR screening libraries consist of molecules with masses < 500 Da. These compounds exhibit small relaxation rates $R_1 = 1/T_1$ and $R_2 = 1/T_2$, vanishing or weakly negative 2D-NOESY cross-peaks (due to vanishing or weakly positive laboratory-frame dipole−dipole cross-relaxation rates), and large translational diffusion coefficients D_t. Bound compounds share the NMR properties of the much larger receptor (mass > 30 000 Da). Therefore, bound compounds have large R_2 (and large selective R_1), positive 2D-NOESY cross-peaks and highly

efficient spin diffusion, and smaller molecular diffusion coefficients D_t. These distinct differences imply that changes in the ligand NMR spectral parameters can be monitored as a means to assess target binding. Most ligand-based NMR experiments detect binding by one of two mechanisms: (i) exploiting the differential mobility of the ligand in the free versus bound state (hits will transiently experience the much slower rotational and translational mobility of the large receptor, manifested as altered relaxation parameters and diffusion coefficients, respectively) and (ii) exploiting a ^{1}H magnetization transfer process from the receptor (binders or hits will experience this transfer while nonbinders will not).

Most ligand-based NMR screening experiments employ ^{1}H NMR methods. Unless otherwise stated, we assume this to be the case for the following discussion. The chief relaxation mechanisms in ^{1}H NMR are ^{1}H−^{1}H dipole−dipole (DD) interactions between pairs of proton spins. For a given proton of spin order V (V could be I_z, I_x, multiple quantum coherence, etc.), a DD relaxation rate can be written as the double sum

$$R(V) = \sum_m a_m \sum_j \frac{1}{r_j^6} J_j(m\omega) \qquad (14)$$

The inner sum goes over all other distinct protons that have dipolar couplings to the proton of interest. The outer sum represents a linear combination of spectral density functions $J_j(m\omega)$ evaluated at various integral multiples m of the ^{1}H Larmor frequency ω_H. The weighting coefficients a_m depend on the spin order V (e.g. $V = I_z$, I_x) and attendant rate constant (e.g. $R_1 = 1/T_1$, NOE, $R_2 = 1/T_2$) being considered. The spectral density functions $J_j(\omega)$ are frequency distribution functions whose shapes profile the rotational motions of interproton vectors connecting the proton of interest to proton j relative to the external magnetic field $\mathbf{B}_0$. Therefore, the complexity of $J_j(\omega)$ is dictated by the nature of the molecular dynamics present. Usually, both structural and screening studies assume rigid receptor and ligand molecules that undergo isotropic tumbling, so the only "dynamics" to consider are those of overall molecular tumbling. Accordingly, each $J_j(\omega)$ becomes the familiar Lorentzian distribution function

$$J(\omega) = \frac{2}{5}\frac{\tau_c}{1 + (\omega\tau_c)^2} \qquad (15)$$

τ_c is the effective overall rotational correlation time that scales with the molecular mass. Therefore, $J(0) = {}^2/_5\tau_c$. When the ligand binds a large receptor molecule, τ_c increases and dramatically amplifies the corresponding relaxation parameter. Thus, relaxation parameters containing a $J(0)$ dependence are highly sensitive probes of binding.

As discussed, a significant practical advantage of ligand-based approaches is that isotope labels are *not* required. With this advantage, however, comes the additional challenge of selective observation of the ligand signal. Interference from receptor signals can compromise the accuracy and ease of data interpretation. Fortunately, several strategies can be employed to enhance selective observation of ligand resonances: (i) use a high L_T/E_T and exploit the ligand excess that overwhelms the receptor signals, (ii) use transverse relaxation filters (e.g. Hahn echoes, spin-locks) that preferentially eliminate the rapidly relaxing receptor resonances, and (iii) use translational diffusion filters to select against rapidly diffusing molecules. For experiments performed in H_2O, water is typically suppressed via excitation sculpting[26] or various applications of WATERGATE.[27,28] Since these methods involve spin−echo segments, they simultaneously act as relaxation filters prior to detection.

5.1. Transverse Relaxation Rates

Comparison of the ligand transverse autorelaxation rate $R_2 = 1/T_2$ in the presence and absence of receptor is the most well-established, classic NMR binding assay. R_2 is a highly attractive probe of binding due to its nearly direct dependence on the overall molecular rotational correlation time τ_c. This dependence can be seen by expression of R_2 in terms of spectral density functions. For a given ligand proton under ^{1}H−^{1}H DD relaxation, R_2 is given by

$$R_2 = \frac{\hbar^2\gamma_H^4}{8}\sum_{j=1}^{N}\frac{1}{r_j^6}\{5J_j(0) + 9J_j(\omega_H) + 6J_j(2\omega_H)\} \qquad (16)$$

Given the form of $J_j(\omega)$ in eq 15, the $J_j(0)$ dependence in eq 16 infers that R_2 has a strong dependence on τ_c. When the ligand binds to the receptor, its τ_c transiently becomes that of the large receptor. Because $\tau_{c,receptor} \gg \tau_{c,free-ligand}$, we have $R_{2B} \gg R_{2F}$.

Ligands that undergo sufficiently rapid exchange will possess exchange-averaged relaxation rates that reflect exchange-mediated transfer of R_{2B} to the free state. In the fast exchange limit, the exchange-averaged rate $R_{2,avg}$ is the simple sum, derived from the more general form of eq 12a

$$R_{2,avg} = P_B R_{2B} + P_F R_{2F} + R_{ex} \qquad (17a)$$

where

$$R_{ex} = (\Omega_F - \Omega_B)^2 P_F P_B / k_{ex} \qquad (17b)$$

The $P_B R_{2B}$ term contains the information transferred from the bound state that is characteristic of a screening hit. Note that the evidence of binding contained in $R_{2,avg}$ becomes strong when $P_B R_{2B}$ is significant relative to $P_F R_{2F}$. However, the typical screening conditions of $L_T/E_T \gg 1$ result in $P_B \ll P_F$. It is clear that because $R_{2B} \gg R_{2F}$, the ligand R_2 becomes amplified in the bound state, and $P_B R_{2B}$ can be significant even under a large excess of ligand. This is a direct consequence of the $J_j(0)$ dependence of R_2 in eq 16 and underscores the fact that relaxation parameters containing a $J_j(0)$ dependence are highly sensitive probes of binding.

$R_{2,avg}$ contains another potential signature of binding: R_{ex}. R_{ex} is the familiar line broadening arising from the nonequivalence of free versus bound chemi-

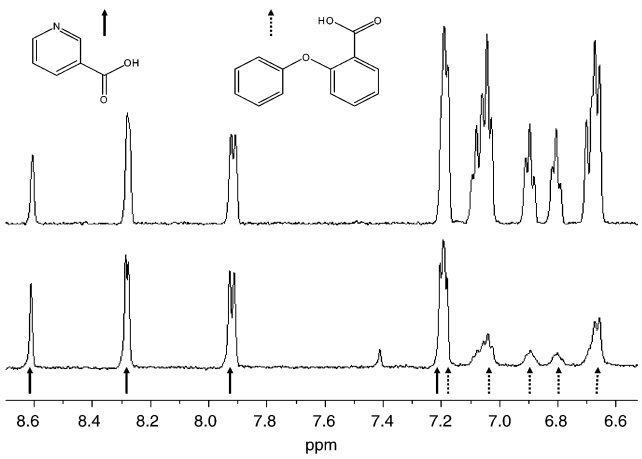


Figure 4. Use of ^{1}H line broadening (increased $R_2 = 1/T_2$) to detect binding. Shown are ^{1}H NOESY spectra of two compounds in a mixture in the absence (top trace) and presence (bottom trace) of the catalytic domain of p38 MAP kinase, a protein of 42 kDa. Resonances from nicotinic acid (top left structure) and 2-phenoxybenzoic acid (top right structure) are marked with solid and dashed arrows, respectively. The peak at 7.2 ppm consists of overlapping resonances from both compounds. Line broadening, suppression of fine structure, and attenuation of ligand resonance peak height due to the relaxation filter in the bottom spectrum indicate 2-phenoxybenzoic acid binds to p38, while nicotinic acid does not. The sample contained 1 mM ligands, 0.2 mM p38 MAP kinase, 25 mM deutero-Tris, 10% deutero-glycerol, 20 mM deutero-DTT at pD* = 8.4. Experiments were carried out at 278 K. A relaxation filter was used after the preparatory delay to attenuate broad resonances arising from the protein. Reproduced with permission from ref 11. Copyright 2001 Elsevier.

cal shifts. From eq 1, $k_{ex} = [E]k_{on} + k_{off}$. k_{ex}^{-1} can be considered as the correlation time for the complexation reaction of eq 1.[29] In this regard, the R_{ex} term also contains a $J(0)$ dependence, where the salient dynamic process is the two-state exchange instead of molecular tumbling. This becomes more explicit in $R_{1\rho} = 1/T_{1\rho}$ dispersion studies, where $R_{1\rho}$ represents the relaxation time of magnetization along the effective spin-lock field in the rotating reference frame.[30] For the purposes of NMR screening, when R_{ex} is large, binding-induced enhancements of R_2 may be observed even when $P_B R_{2B}$ is insignificant. Alternatively, R_{ex} becomes negligible if $\Omega_F \approx \Omega_B$, if k_{ex} is too large, or if the populations are not evenly weighted (i.e. $P_B P_F \ll 1$). This is often the case in ^{1}H NMR screening. Working with stoichiometric amounts of receptor and ligand minimizes the latter possibility, but this is not practical if receptor concentration is limited.

To detect small molecule binding, line shapes are compared in the presence and absence of receptor. The full-width half-maximum (fwhm) homogeneous line width of a given compound resonance is R_2/π. Its peak intensity is also proportional to $1/R_2$. Thus, binding-induced R_2 enhancements may be visible as simple broadening of proton resonance lines upon the addition of receptor. Figure 4 shows an example from our laboratory in which selective line broadening and consequent peak-height attenuation clearly reveal the binding compound in a mixture. More generally, however, observation of line broadening can be difficult, especially if the effect is small (common for $L_T/E_T \gg 1$) or there is significant spectral crowding from the receptor and/or other compounds in the mixture.

An alternative approach is to compare compound peak intensities in the presence and absence of receptor using 1D ^{1}H R_2 experiments.[31] Such experiments simultaneously reduce or eliminate receptor resonances via relaxation filtering while relaxation editing to detect compound binding. These experiments typically monitor R_2 relaxation using a spin-lock such as the Carr−Purcell−Meiboom−Gill (CPMG) pulse train or continuous wave irradiation as in a $R_{1\rho} = 1/T_{1\rho}$ experiment.[30,32,33] The pulse schemes are simply 90_x−spin-lock$_{+/-y}$−acq$_x$. Since radiofrequency irradiation is applied during the transverse relaxation period, the exchange-averaged relaxation rates differ slightly from those given in eqs 23a,b. The general expressions are quite awkward, and we write here only the fast exchange limits:[34,35]

$$R_{2,avg}^{CPMG} = P_F R_{2F} + P_B R_{2B} + \frac{P_F P_B (\Delta\Omega)^2}{k_{ex}}\left(1 - \frac{2\tanh(k_{ex}/t_{cp})}{k_{ex} t_{cp}}\right) \quad (18)$$

$$R_{1\rho,avg} = P_F R_{1\rho} F + P_B R_{1\rho} B + \frac{P_F P_B (\Delta\Omega)^2 (\sin\Theta_{rf})^2}{k_{ex}}\left(\frac{k_{ex}}{k_{ex}^2 + \Omega_{SL}^2}\right) \quad (19)$$

For $R_{1\rho,avg}$, on-resonance spin-locking corresponds to $\Theta_{rf} = \pi/2$ and $J(\omega) \approx J(\omega \pm \Omega_{SL})$, and so on-resonance $R_{1\rho} \approx R_2$. Thus, the main effect of spin-locking is to give R_{ex} a functional dependence on the spin-lock field strength ($\approx 4/t_{cp}$ for CPMG, Ω_{SL} for $R_{1\rho}$). The R_{ex} function can be quenched by sufficiently large Ω_{SL}; this is of course the basis for $R_{1\rho}$ dispersion studies aimed at measuring exchange rates. Note also that the R_{ex} function in eq 19 is the spectral density function for the two-state exchange process alluded to above. To avoid line shape distortions, it is important to suppress the effects of homonuclear scalar coupling J_{HH} during the relaxation period. In this context, the $R_{1\rho}$ experiment may be preferable, since it essentially quenches scalar coupling evolution. If the CPMG pulse train is used, then the delay t_{CP} between consecutive 180° pulses should satisfy $|4\pi J_{HH} t_{cp}| \ll 1$.[36] It is important to use sufficiently long spin-locks to eliminate receptor signals. As a benchmark, Hajduk et al.[31] noted that a 400 ms CPMG spin-lock was sufficient to completely eliminate the majority of receptor proton signals for FKBP12, a MW = 12.5 kDa protein.

Basic steps for identifying binders using R_2-edited experiments are as follows. First, the R_2 experiment is recorded for the lone compounds ("−receptor") at a set transverse relaxation delay T_{rlx}. The resulting peak intensities are proportional to $\exp[-R_{2F} T_{rlx}]$. A second R_2 experiment is then recorded on the compounds in the presence of receptor ("+receptor"). For the hits, the peak intensities are now proportional to $\exp[-R_{2,avg} T_{rlx}]$. Since $R_{2,avg} > R_{2F}$, the peak intensities of the hits will be selectively attenuated. Subtracting the +receptor spectrum from the −receptor spectrum thus reveals *only* the hits. Typically,

however, the receptor R_2 values vary (e.g. due to differential internal motion), and residual receptor signals can hinder data interpretation. To correct for this, a third control R_2 experiment on the lone receptor can be subtracted from the +receptor spectrum. In the foregoing strategy, the usual risks inherent with difference spectroscopy must be considered. Small chemical shift changes between samples or due to binding as well as instrumental instabilities due to the surrounding environment may corrupt the difference spectra. It is therefore clearly advantageous to design mixtures that minimize spectral overlap to the extent that individual compound resonances can be reliably integrated.

The amount of resonance attenuation can be used to get a coarse estimate of binding affinity. Using eqs 17a,b and 19 for $R_{2,avg}$ and $R_{1\rho,avg}$, van Dongen et al. have expressed the relative peak attenuation in terms of the bound and free relaxation rates, the net exchange rate constant k_{ex}, and the applied rf field.[37] Assuming a receptor with MW $\sim$ 15 kDa, their calculations suggest $R_{1\rho}$ spin-locks of 400 ms are sufficient for the nearly complete elimination of signals from compounds with $K_D < 500$ μM. Obviously, targets of higher molecular mass require shorter spin-locks. Tighter affinity hits will have more dramatic R_2 relaxation enhancements; thus, shorter spin-locks can "tune" the experiment for tighter binders, as long as fast exchange conditions exist.

5.2. Longitudinal Relaxation Rates

NMR screening experiments employing R_1 measurements comprise only a small subset of commonly used approaches, and for this reason they are considered only briefly here. A primary consideration in the design of any R_1-based method is the need to distinguish between and account for selective versus nonselective R_1. As described previously,[7] selective R_1, which possesses a $J(0)$ dependence, is a sensitive probe of binding, while nonselective R_1 (R_1^{NS}) is not.

To measure 1D selective R_1's, a restricted set of protons is inverted or saturated using frequency selective or other discriminatory pulse sequences. In practice, however, achieving selective spin perturbations for a large library of diverse compounds, all possessing different chemical shift patterns, is a formidable task. To achieve this, a separate inversion pulse would have to be defined for each compound. A clever approach comes from the realization that sufficient selectivity is achieved by inverting/saturating the ligand resonances relative to those of the receptor. This underlies the "reverse NOE pumping" experiments of Chen and Shapiro.[38] Another use of selective R_1 experiments is to monitor displacement of a "reporter" or "probe" molecule in competitive binding studies with higher affinity ligands.[17] Binding information comes from analysis of the same reporter molecule for a mixture of test molecules. Since the same ligand (as opposed to an entire compound library) is repeatedly observed, there is no need for the constant recalibration of selective pulse schemes.

5.3. Paramagnetic Relaxation Enhancements

A variation of the exchange-averaged relaxation parameter approach is the SLAPSTIC (spin labels attached to protein side chains as a tool to identify interacting compounds) method developed by Jahnke and co-workers.[39,40] In this approach, amplification of bound state relaxation properties is achieved through use of covalently attached spin labels on selected protein side chain types (e.g. lysine, tyrosine, cysteine, histidine, and methionine). The spin labels used include paramagnetic nitroxide radical moieties such as TEMPO. Attaching a spin label endows the receptor with the strong paramagnetic relaxation mechanism of an unpaired electron, such that compounds that bind in spatial proximity to the spin label experience enhanced proton relaxation due to the electron−proton DD interaction with the unpaired nitroxide electron. In the context of eq 12a, Q_B is enhanced *chemically* to ensure $P_B Q_B \gg P_F Q_F$. In the fast exchange limit, the paramagnetically enhanced R_2 relaxation is

$$R_{2,avg} = P_B R_{2,para} + (P_B R_{2o,B} + P_F R_{2o,F}) + R_{ex} \quad (20)$$

where $R_{2,para}$ is

$$R_{2,para} = \frac{\hbar^2 \gamma_e^2 \gamma_H^2}{8} \sum_{j=1}^{N} \frac{1}{r_j^6} \{4J_j(0) + 3J_j(\omega_H)\} \quad (21)$$

The j sum in eq 21 runs over the N spin labels in proximity to the ligand proton under consideration. $R_{2,o}$ refers to all *other* sources of relaxation such as the $^1H-^1H$ DD mechanisms already described above for "conventional" R_2. R_{ex} is the same contribution seen in eq 17. The power of SLAPSTIC lies in the comparatively large magnitude of the electron−proton DD interaction. Because $|\gamma_e/\gamma_H| \approx 658$, $R_{2,para}$ is effective over much longer interspin distances than are typical for $^1H-^1H$ relaxation. This enhancement allows a considerably lower bound ligand fraction P_B to be used than would be possible in the absence of spin labels. On the basis of their original proof-of-concept work with FKBP, Jahnke et al. estimate that the use of spin labels reduces the protein requirement by $\approx$50-fold.[39]

In practice, SLAPSTIC experiments rely on the availability of amino acid side chains near the binding site that are amenable to spin labeling. For the approach to be successful, attachment of spin labels should not compromise the structural integrity or binding properties of the target receptor. In this regard, SLAPSTIC requires considerable prior knowledge of the receptor's 3-dimensional structure. This requirement could impose significant limitations in the applicability of SLAPSTIC experiments to novel targets. SLAPSTIC is also appealing for screening second-site binders. Specifically, a spin label can be attached to a known specific ligand. This altered first ligand then serves to screen for new ligands binding in a second proximal site. Again, it must be confirmed that the addition of the spin label does not compromise either receptor binding site. A distinct advan-

tage is that the first ligand need not be present in saturating amounts, owing to the dominance of $R_{2,para}$.[40]

5.4. ^{19}F Relaxation

Although most NMR screening experiments focus mainly on ^{1}H, the ^{19}F nucleus has a number of unique properties that render it a highly effective relaxation probe for NMR screening.[41] First, and fortuitously, medicinal chemists regard ^{19}F incorporation (e.g. aromatic fluorines, trifluoromethyl groups) as an established tactic for enhancing the pharmacokinetic properties of drug leads. Second, the absence of endogenous ^{19}F in biological molecules permits clean observation of ligand spectra, thus obviating the need for relaxation filters and/or difference spectroscopy to eliminate receptor or large solvent signals. Third, ^{19}F occurs at 100% natural abundance and has a gyromagnetic ratio $\gamma_F \sim 0.94\gamma_H$; therefore, the sensitivity of ^{19}F NMR is competitive with that of ^{1}H. Finally, and most relevant for transverse relaxation methods, the chemical shift range of ^{19}F is much larger than that of ^{1}H ($\approx$900 ppm).[42] This larger chemical shift range implies high sensitivity of the ^{19}F chemical shift to changes in microenvironment. In contrast to the case of ^{1}H, a large $\Delta\Omega = |\Omega_F - \Omega_B|$ associated with ligand binding can be expected. In the context of eq 17 this suggests significant R_{ex} despite low P_B. Thus, the spectral signatures of binding can be more pronounced using ^{19}F detection.

^{19}F relaxation is also useful for secondary screening experiments aimed at estimating exchange rates and equilibrium dissociation constants.[43−45] More recently, *cross-correlation* between the ^{19}F aromatic chemical shift anisotropy (CSA) and ^{1}H−^{19}F DD relaxation mechanisms has been exploited to improve the accuracy of K_D estimates.[41]

A drawback of ^{19}F relaxation as a primary screening method is that ^{19}F nuclei are not nearly as ubiquitous in small molecules when compared to ^{1}H nuclei. This creates difficulties, for example, when attempting to choose a diverse set of compounds for a screening library. However, useful screening information can still be gained by looking at the ^{19}F relaxation of a small set of compounds. An example of this approach is the FAXS (fluorine chemical shift anisotropy and exchange for screening) strategy of Dalvit and co-workers.[46,47] In this approach, the relaxation properties of a small set of ^{19}F "spy" compounds report on the binding of a larger set of higher affinity binders via competitive displacement.

5.5. Saturation Transfer Difference (STD) Methods

Saturation transfer difference (STD) spectroscopy is a highly versatile technique for NMR-based screening. There are several advantages STD experiments possess over their counterparts that have led to widespread adoption of the method. These benefits include reduced protein (or receptor) consumption, relative ease of implementation, and applicability to large MW therapeutic targets.[48] All of these benefits will be discussed in more detail below. As the name suggests, STD takes the difference of two experiments. In a first experiment (the "on-resonance" experiment), receptor proton magnetization ($\mathbf{M}_z = 0$, and $\Delta\mathbf{M}_z = \mathbf{M}_z - \mathbf{M}_o = -\mathbf{M}_o$) is selectively saturated via a train of frequency selective rf pulses. The rf train is applied to a frequency window that contains receptor resonances but for which compound resonances are absent (e.g. 0.0 to −1.0 ppm for proteins). The saturation propagates from the selected receptor protons to other receptor protons via the vast network of intramolecular ^{1}H−^{1}H cross-relaxation pathways; this process of spin diffusion is quite efficient due to the typically large molecular weight of the receptor. As sketched in Figure 5, saturation is transferred to binding compounds via intermolecular ^{1}H−^{1}H cross-relaxation at the ligand−receptor interface. The small molecules then dissociate back into solution where the saturated state persists due to their small free state R_1 values. At the same time, more "fresh" unsaturated ligand exchanges on and off the receptor while saturation energy continues to enter the system through the sustained application of rf, thus increasing the population of saturated free ligands. A reference experiment (the "off-resonance" experiment) is then recorded that applies the identical rf train far off-resonance, such that no NMR resonances are perturbed. The "on-resonance" and "off-resonance" experiments are recorded in an interleaved fashion and subtracted. The resulting difference spectrum yields only those resonances that have experienced saturation, namely, the receptor and the binding compound resonances. Because they are present at minimal concentration, receptor resonances will usually not be visible and, if so, can be eliminated by R_2 relaxation filtering prior to detection. The result is a simple 1D ^{1}H spectrum that reveals *only* the binding compounds. Acquiring a difference spectrum vastly simplifies data interpreta-

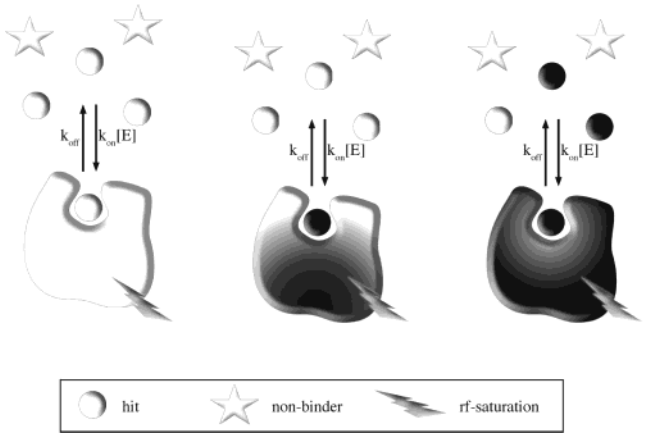


Figure 5. Detection of binding using the saturation transfer difference (STD) experiment.[48] Frequency selective irradiation (lightning bolt) causes selective ^{1}H saturation (shading) of the target receptor. Irradiation is applied for a sustained interval during which saturation spreads throughout the entire receptor via ^{1}H−^{1}H cross-relaxation (spin diffusion). Saturation is transferred to binding compounds (circles) during their residence time in the receptor binding site. The number of ligands having experienced saturation transfer increases as more ligand exchanges on and off the receptor during the sustained saturation period. Nonbinding compounds (stars) are unaffected.

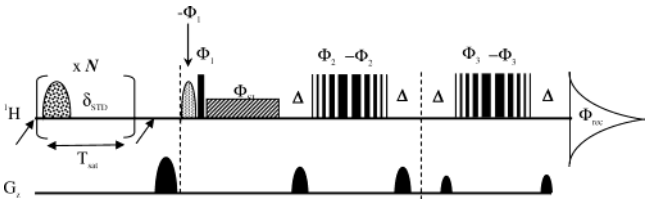


Figure 6. Example of an STD pulse sequence.[48] The upper and lower staffs show proton rf and field gradient pulses, respectively. Selective rf saturation occurs for a time $T_{SAT} = 1-3$ s via a train of N frequency selective rf pulses (e.g. 50 ms Gaussian or Seduce-1[49] 90° pulses separated by an interpulse delay of ~1 ms). Two experiments are performed, which apply saturation on and off resonance with receptor protons in an interleaved fashion. The 1D signals are stored in separate locations, and their spectral differences are taken via inverting the receiver phase. A WATERGATE-5[28] readout sequence polls the residual z-magnetization. Suppression of receptor signals is achieved by relaxation filtering during the spin−echo delay Δ and the optional $R_{1\rho} = 1/T_{1\rho}$ spin-lock. An optional water flip-back 90° pulse precedes the first hard 90° pulse (e.g. 2 ms Seduce-1 90° pulse). Phase cycling is as follows: $\Phi_1 = (16x, 16\ -x)$; $\Phi_2 = (x, y, -x, -y)$; $\Phi_3 = (4x, 4y, 4-x, 4-y)$; $\Phi_{SL} = y$; and $\Phi_{rec} = 2(x, -x, x, -x, -x, x, -x, x)$, $2(-x, x, -x, x, x, -x, x, -x)$. Φ_{rec} flips 180° between the on and off resonance spectra.

tion compared to, for example, the case of differential line broadening experiments, where line widths of the compounds with receptor present must be carefully compared with a reference spectrum of free compounds.

Figure 6 depicts an STD pulse scheme. Typical saturation trains involve N repetitions of 50 ms frequency selective pulses with Gaussian or Seduce-1 profiles.[49,50] The train lasts for a total duration of T_{sat} (typically 1−3 s), after which a pulsed field gradient ensures that only z-magnetization remains. A 90° (Φ_1) pulse (with optional antecedent water flip-back 90°, e.g. Seduce-1 pulse of 1−3 ms) then polls the z-axis. A subsequent spin-lock serves as a relaxation filter for residual receptor signals. The sequence finishes with a WATERGATE−5 block that simultaneously suppresses water while further relaxation filtering the residual receptor signal. On-resonance saturation yields a spectrum of intensity I_{SAT}, while off-resonance saturation yields the equilibrium value I_0. Appropriate phase cycling of the receiver subtracts the two intensities to yield the STD response $I_{STD} = I_0 - I_{SAT}$. Figure 7 illustrates the generation of the difference spectrum. The literature defines the fractional STD response, which is of the same form as the traditional steady-state NOE: $\eta_{STD} = (I_0 - I_{SAT})/I_0 = I_{STD}/I_0$.[50]

The STD experiment has several attractive features that warrant discussion. First, STD is ideally suited to receptors with large masses (>30 000 Da) typically encountered in drug discovery. Receptors with large molecular masses possess large rotational correlation times, τ_c, that enhance spin diffusion and, consequently, saturation transfer within the receptor and to the ligand. Second, STD experiments require relatively low concentrations of receptor (~1 μM). Assuming sufficiently fast ligand exchange, the population of saturated ligand builds up during the sustained rf pulse train (T_{SAT}) due to the small free

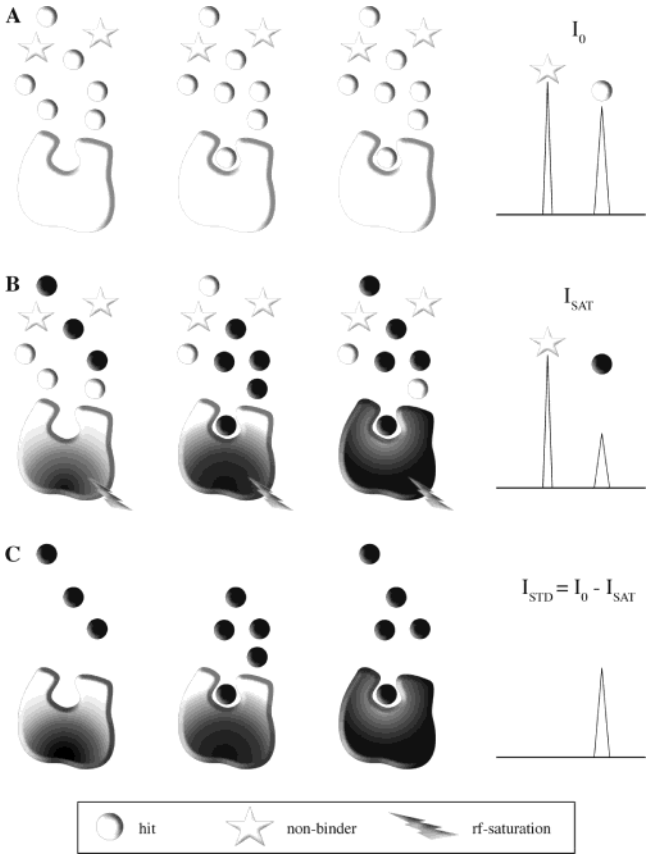


Figure 7. Schematic diagram depicting difference spectroscopy in the STD experiment. Circles and stars indicate binding and nonbinding compounds, respectively. STD involves two experiments: an off-resonance experiment and an on-resonance experiment. Top panel, A: off-resonance (reference) applies rf irradiation off-resonance from both receptor and compound protons. Detection produces spectra with intensity I_0. Middle panel, B: In the on-resonance experiment, the rf irradiation selectively saturates receptor and any binding compounds (indicated by dark shading). This manifests as the decreased signal intensity I_{SAT}. Bottom panel, C: The STD response is the spectral difference $I_{STD} = I_0 - I_{SAT}$, which yields only resonances of the receptor and binding compounds. Receptor resonances are usually invisible due to either low concentration or relaxation filtering. The STD sensitivity depends on the number of ligands receiving saturation from the receptor and can be described in terms of the average number of saturated ligands produced per receptor molecule (STD amplification factor; see main text).

state R_1 values. Through this mechanism, a small amount of saturated receptor can produce an *amplified* amount of saturated ligand. A third advantage of STD is that only the signal from the *bound* state of the ligand is observed. One does not need to correct for free state contributions Q_F (cf. eq 27a) that might otherwise complicate interpretation. This is especially advantageous when an experimental design with high ligand-to-receptor ratios $L_T/E_T \gg 1$ is employed. Figure 8 depicts an example from our own laboratory using the same receptor−ligand system as in Figure 4. The STD spectrum (lower trace) reveals only the binding compound, greatly simplifying data reduction.

Optimal setup, use, and interpretation of the STD experiments require a familiarity with the exchange processes at work and have been the subjects of

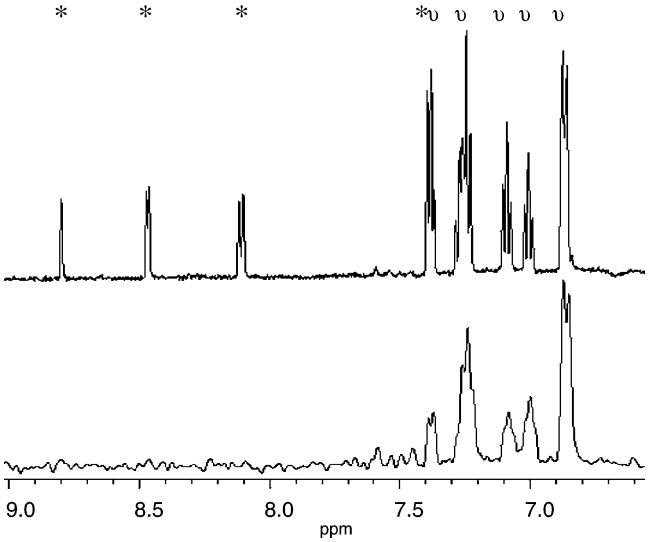


Figure 8. Example of the STD experiment using the same compound mixture and protein receptor (p38 kinase domain, 42 kDa) as in Figure 4. Top panel, A: ^{1}H NMR spectra of two compounds in the absence of receptor using the 1D version of a standard NOESY. Resonances from nicotinic acid and 2-phenoxybenzoic acid are marked with asterisks and diamonds, respectively. Bottom panel, B: Result of the STD experiment in the presence of receptor; the resonances of the binding compound (2-phenoxybenzoic acid) are present. Receptor protons are invisible due to relaxation filtering. Sample conditions were as follows: 1 mM compounds, 35 μM receptor dissolved in D_2O buffer (25 mM *d*-Tris, 10% *d*-glycerol, 20 mM *d*-DTT, pD = 8.4). Both spectra were acquired at 11.74 T. The 1D spectrum (top) was collected using a standard NOESY pulse sequence with 16K data points, 128 transients, and a relaxation delay of 3 s. The STD spectrum (bottom) was recorded at 278 K, with 2K data points and 256 transients acquired for both on- and off-resonance spectra. A 3 s train of 50 ms Gaussian pulses separated by 1 ms was used for selective receptor saturation. The proton carrier was placed at 0.74 and −20 ppm for on- and off-resonance saturation, respectively. Excitation sculpting[26] was used to eliminate residual H_2O and provide the aforementioned relaxation filtering. Reproduced with permission from ref 11. Copyright 2001 Elsevier.

several methods development studies in the literature. To better quantitate the amplification of resonances in the STD experiment, Mayer and Meyer[50] introduced the "STD amplification factor" A_{STD}. Because the STD response arises directly from the receptor−ligand complex, it is proportional to [EL]. Hence, I_{STD} can be written as $I_{STD} = C\alpha_{STD}$[EL], where C is a proportionality constant that makes the appropriate unit conversions and α_{STD} is a dimensionless scaling factor that represents the maximum STD amplification. Note that I_{STD} corresponds to Q_{avg} in eq 12a, since there is no contribution from [L] (i.e. $Q_F = 0$). The reference, or equilibrium, intensity I_0 is just proportional to L_T; thus, $I_0 = CL_T$. Then the ratio $\eta_{STD} = I_{STD}/I_0 = \alpha_{STD}$[EL]$/L_T = \alpha_{STD}P_B$. Since $P_B = P_B^E/\epsilon$, where $\epsilon = L_T/E_T$, we have

$$A_{STD} = \epsilon\eta_{STD} = \epsilon\{I_{STD}/I_o\} = \frac{\alpha_{STD}[L]}{[L] + K_D} \quad (22)$$

Equation 22 is just the hyperbolic dose−response curve given in eq 13 and is analogous to the well-known equation for the Henri−Michaelis−Menten enzymatic reaction rate v_0:[51,52]

$$v_0 = \frac{V_{max}[S]}{[S] + K_M} \quad (23)$$

Using an analogy to enzyme kinetics, A_{STD} will act in a manner similar to that of v_0: it will increase with increasing L_T until the maximum amplification α_{STD} is reached when the receptor binding site is saturated ($L_T \gg K_D$). After the point of receptor saturation, continued increase of L_T will monotonically *decrease* the fractional response $\eta_{STD} = I_{STD}/I_0$. In principle, if L_T well-approximates [L], ligand titration data can be fitted to the form of eq 22 to estimate K_D and α_{STD}.

Following the enzyme argument, A_{STD} gives the average number of saturated ligands "turned over" per receptor. As such, it provides a convenient means for gauging the inherent sensitivity of the experiment. Mayer and Meyer provide an example in which $A_{STD} = 10$, which implies that a receptor concentration of 50 μM yields an effective saturated ligand concentration of 500 μM.[50] The latter concentration is more than sufficient for sensitive detection by current high-field magnets and probes. Additional motivation for using A_{STD} is that spectra from samples having different receptor concentrations may still be compared. This would be relevant for titration and competition experiments that may involve samples with variable amounts of receptor.

Assuming a diffusion-limited on rate, the K_D range of the STD method has been estimated to be $10^{-8} < K_D < 10^{-3} > M$.[48] For weak binders having $K_D > L_T$, over half of the receptor molecules will have no ligand in the binding site. As K_D increases further, the population of the ligand−receptor complex [EL] decreases, which leads to a reduction and, ultimately, disappearance of the STD signal. For the case of strong binders, decreasing K_D increases the receptor−ligand lifetime k_{off}^{-1} and thus decreases the exchange rate constant k_{ex}. At sufficiently small K_D, the free state residence times of the ligands can exceed their free state R_1 values. Under these conditions, exchange is so slow that the free ligand magnetization "forgets" its visit to the saturated receptor by relaxing back to equilibrium at a faster rate than the receptor is able to "turn over" newly saturated ligands. The initial population of saturated ligands decreases and ultimately disappears, thereby eliminating the STD signal.

Recent applications of STD experiments demonstrate their versatility beyond simple enzyme systems. For example, Benie et al.[53] have used STD to identify ligands targeting HRV2 (human rhinovirus serotype 2), a macromolecular complex of 8.5×10^6 Da. Meinecke and Meyer have characterized the interactions of peptide ligands binding to the extracellular region of an integral membrane protein (Integrin $\alpha_{IIb}\beta_3$) reconstituted in liposomes.[54] Other examples of alternative targets include small RNA fragments[55] and macromolecules immobilized on solid-support beads.[56]

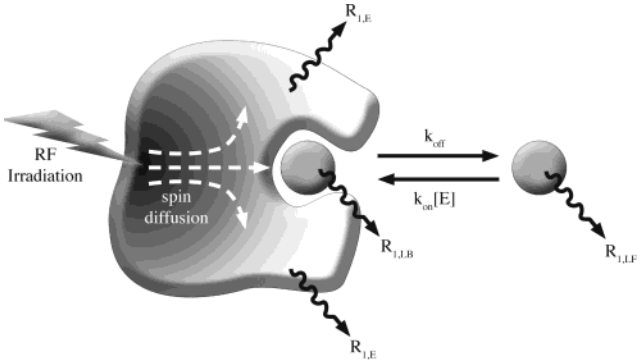


Figure 9. Schematic diagram of magnetization transfer pathways during the STD experiment. The intended pathway is initiated by irradiation (lightning bolt) that selectively saturates a subset of receptor protons. Intrareceptor spin diffusion ideally spreads the saturation condition (indicated by shading) to all receptor protons and bound ligands. Bound ligands then dissociate back into solution and add to the growing pool of free saturated ligands. Longitudinal proton relaxation of the ligand−receptor complex ($R_{1,E}$ for the receptor and $R_{1,LB}$ for the bound ligand) and free ligand ($R_{1,LF}$) limits the efficiency of the intended transfer process. Additionally, the efficiency of intrareceptor spin diffusion will vary according to the overall tumbling time of the receptor and the local proton density.

The intrinsic sensitivity of the STD experiment is limited by the efficiency of the intended spin energy transfer pathway from receptor (source) to ligand (recipient). Figure 9 summarizes some of the principle pathways that must be considered. The main factor limiting magnetization transfer at the recipient (ligand) end is the ligand R_1. Since the ligands are usually low-molecular-mass compounds (<1000 Da), the free state R_1 values are small and therefore any nonequilibrium magnetization state (like saturation) dissipates quite slowly. In contrast, when bound to the receptor, the ligand R_1 can be much larger. Hence, efficient spin energy transfer requires that the ligand dissociate from the receptor at a rate faster than the *bound state* R_1.

The source (receptor molecule) receives a constant influx of energy by the applied rf saturation train. This constant energy input is what enhances the sensitivity of STD over other methods such as transferred-NOE spectroscopy. In practice, one often assumes 100% saturation of the receptor shortly after application of the rf train due to intramolecular spin diffusion. However, the actual extent of receptor saturation depends critically on the competition between energy influxes and the various R_1 relaxation and/or "leakage" mechanisms. Jayalakshmi and Rama Krishna have recently emphasized the importance of exchange-mediated leakage.[57] Receptor protons at the ligand−receptor interface cross-relax with unsaturated ligand and solvent protons as well as other distinct saturated receptor protons. As saturation enters from the other receptor protons, it leaks away via the exchanging ligand and solvent protons. The leakage mechanism can be very important for some systems. For example, Mayer and James have demonstrated the effects of exchange-mediated leakage from solvent molecules by comparing the STD responses of RNA-binding ligands in H_2O versus D_2O.[55] The overall STD response is significantly less in H_2O due to the additional DD interactions between the RNA protons and hydration waters.

Another factor that can compromise saturation is low receptor molecular mass. Although low mass (<20 000 Da) is a rare concern for therapeutic targets, one occasionally encounters a target whose rapid tumbling leads to inefficient spin diffusion and, hence, poor saturation. In these cases, it is possible to increase the effective τ_c by applying longer saturation trains or, alternatively, by addition of viscosity-enhancing reagents (e.g. glycerol) and/or screens at lower temperatures. If such approaches are not feasible, then $R_2 = 1/T_2$ relaxation filtering methods may prove more sensitive than STD experiments.[37]

Finally, it should also be considered that weaknesses in the $^1H-^1H$ cross-relaxation network can compromise saturation efficiency. Such weaknesses can be due to local molecular motion that would effectively scale down the DD interactions or simply local paucities in proton density, such as in RNA targets.[58] Targets that have inherently low proton density are suboptimal for STD. Instead, the following approach may be much more effective.

5.6. WaterLOGSY

The waterLOGSY (water−ligand observed via gradient spectroscopy)[59,60] technique, like STD, relies on excitation of the receptor−ligand complex through a selective rf pulse scheme. However, waterLOGSY achieves this *indirectly* by selective perturbation of the bulk water magnetization as opposed to direct perturbation of receptor magnetization. The intended transfer of magnetization is therefore water → receptor → ligand. Magnetization transfer may occur via a number of mechanisms, which will be described in detail below. The original presentation of waterLOGSY proposed both selective saturation and inversion of the water resonance.[59] Selective water inversion may be achieved either by [sel-$90_{(+x/-x)}$−hard−90_{+x}], as in the WEXII method of Mori et al.,[61,62] or by [90_{+x}−sel-$180_{(+x/+y)}$−90_{+x}], as in the e-PHOGSY method of Dalvit.[63] The most recent publications favor selective inversion via e-PHOGSY; henceforth, our discussion will emphasize this case.[60]

Inverted water magnetization can be transferred to the bound ligands via three simultaneous pathways shown in Figure 10. One pathway involves direct $^1H-^1H$ cross-relaxation between the bound ligand and "bound" waters within the binding site. Such waters are bound in the sense that their receptor residence times exceed the receptor rotational correlation time. As a result, the rate constants governing DD intermolecular cross-relaxation between the bound ligand and water molecules are negative and tend to bring the ligand magnetization to the same inverted state as that of the water. A second pathway is direct cross-relaxation with exchangeable receptor NH and OH protons within the binding site. Chemical exchange of these protons with those of bulk water inverts their magnetization. These NH/OH then propagate inversion to the bound ligand protons via intermolecular DD cross-relaxation. The third pathway involves indirect cross-

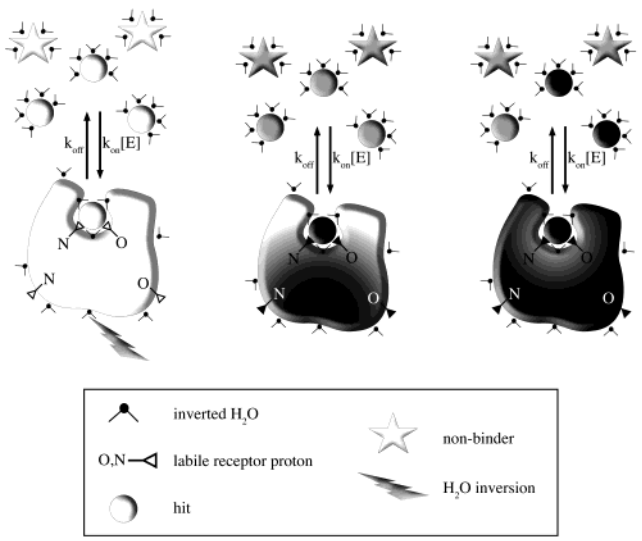


Figure 10. Magnetization transfer mechanisms underlying waterLOGSY.[59,60] Magnetization transfer from bulk water to ligand occurs via labile receptor protons within and remote from the ligand-binding site as well as from long-lived water molecules within the binding pocket. Dark gray and light gray shading indicate magnetization transfer from inverted water to ligand protons in the slow tumbling (i.e. receptor−ligand complex) and fast tumbling (i.e. free ligand) limits, respectively. Only the hits experience both types of magnetization transfer. The pool of free ligands having experienced inversion transfer from bulk water builds up as ligand continues to exchange on and off the receptor.

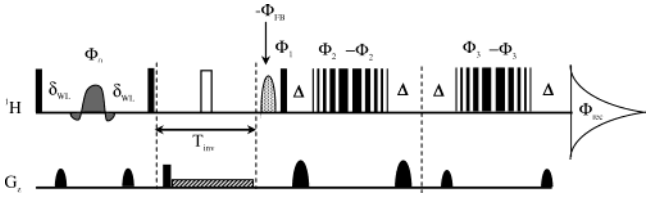


Figure 11. Example of the the preferred waterLOGSY pulse sequence of Dalvit et al.[60] The water resonance is selectively inverted on alternate scans by 90° phase shifts of the selective 180° refocusing pulse (e.g. 20 ms REBURP pulse[64] between the delays δ_{WL}). To reduce inversion of receptor protons having chemical shifts degenerate with H_2O, δ_{WL} can be adjusted to ensure decay of receptor magnetization via rapid transverse relaxation. Transfer of inverted water magnetization to both bound and free ligand occurs during T_{inv} (typically 1−3 s). A 1 ms crusher gradient (e.g. 40 G/cm) followed by a sustained weak gradient (e.g. 0.5 G/cm) is applied during T_{inv} to minimize radiation damping effects. Phase cycling is as follows: $\Phi_0 = x, y, -x, -y$; $\Phi_1 = x$; $\Phi_2 = x$; $\Phi_3 = 4x, 4y, 4-x, 4-y$; and $\Phi_{rec} = (2(x, -x), 2(-x, x))(2(x, -x), 2(-x, x))$. Pulses without explicit phase labels are along $+x$. The unshaded pulse in the middle of the T_{inv} period is an optional nonselective composite 180° pulse that serves to minimize magnetization relaxing during T_{inv}. The phase of Φ_{FB} depends on the sign of the relaxed residual H_2O magnetization at the end of T_{inv}. An optional $R_{1\rho}$ spin-lock (not shown) as in Figure 6 can be applied just after the third hard 90° pulse (with phase Φ_1) for further relaxation filtering of the receptor.

relaxation with remote exchangeable NH/OH protons via spin diffusion. The inverted magnetization is then relayed to other nonlabile spins via spin diffusion. Thus, NH/OH protons remote from the binding site act as entry points for widespread spin inversion throughout the receptor. While the above mechanism deals with transfer of magnetization to bound ligands, another mechanism exists for transfer of magnetization to free ligands, which is via chemical exchange of bulk water with exchangeable ligand protons, and consequent intramolecular ligand dipole−dipole cross-relaxation. This effect has the potential to complicate interpretation of data but can be accounted for by collecting control waterLOGSY spectra for free ligands, as discussed below.

The above magnetization transfer schemes allow binding compounds to pick up the bulk water inversion while residing in the receptor binding site. The ligands then dissociate into free solution, where, analogous to the case of the STD experiment, their perturbed magnetization state is maintained due to their small free state R_1 values. The smaller these R_1 values are, the more time is available for ligands to complex with the receptor, receive inversion transfer from the receptor−ligand complex, and then dissociate back into free solution, where they add to the growing pool of spin-inverted ligands.

Distinguishing binding from nonbinding compounds in the waterLOGSY experiment is slightly different from the case of STD and is achieved by observation of the differential cross-relaxation properties of these ligands with water. In the magnetization transfer schemes described above, the bound ligands interact directly or indirectly with inverted water spins via DD interactions with sufficiently long rotational correlation times τ_c (on account of being associated with the large receptor) to yield negative cross-relaxation rates. By contrast, nonbinders' DD interactions with water have much shorter τ_c, leading to positive cross-relaxation rates. As a consequence, binders and nonbinders display waterLOGSY peak intensities of opposite sign, thus providing an easy means to discriminate between them.

Figure 11 shows an example of a typical waterLOGSY sequence. A key feature is the use of the e-PHOGSY spin−echo at the beginning of the sequence to selectively invert the water resonance while dephasing the off-resonant spins with pulsed field gradients.[63] The selective 180° refocusing pulse is typically 5−20 ms long using, for example, REBURP amplitude modulation.[64] Concerns about inverting receptor protons having chemical shifts degenerate with bulk water (e.g. protein α protons) are addressed by setting the spin−echo delay sufficiently long to filter out receptor coherences.[65] Phase cycling of the second hard 90° pulse (Φ_4) serves to place water magnetization alternately on the $+z$ and $-z$ axes. In the $-z$ case, inverted water magnetization transfers to the compounds via the routes described above during T_{inv}. In the $+z$ case, the water magnetization is nearly at equilibrium, and nothing happens. Toggling the receiver phase in concert with that of the second hard 90° pulse yields a difference spectrum between the $+z$ water and $-z$ water conditions. The resultant signal, I_{WL} represent only those resonances having experienced magnetization transfer from water. Gradients are applied throughout T_{inv} to prevent premature return of water magnetization to the equilibrium condition via radiation damping. Large magnetizations or magnetizations with sharp lines relaxing during T_{inv} (e.g. free ligands) can lead

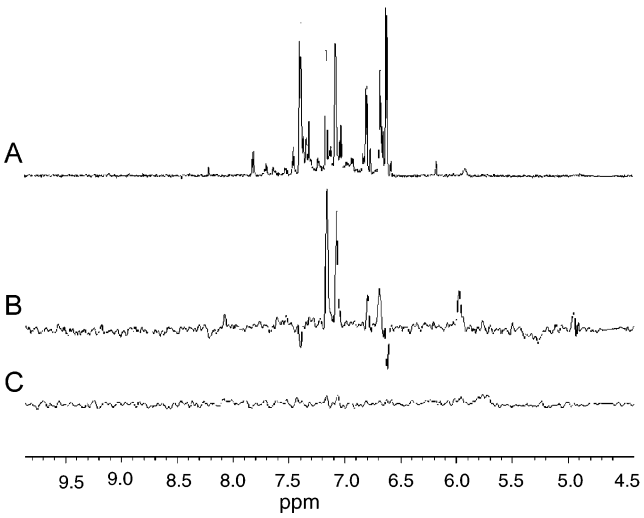


Figure 12. Comparison of waterLOGSY versus STD approaches for identifying binders of RNA. Sample and experimental conditions are 0.77 mM ligands, 38 μM RNA (P456, a 52 kDa ribozyme subdomain of 160 nucleotides), 23% H_2O at 278 K and 18.8 T. Top trace, A: reference 1D spectrum of the ligand mixture. Middle trace, B: waterLOGSY results using selective water inversion. Water inversion was achieved using a 20 ms EBURP-2 90° pulse[64] on resonance with bulk H_2O using a WEXII [sel-$90_{(+/-x)}$−hard-$90_{(+x)}$] sequence.[61,62] Resonances of negative sign are labile ligand protons. Bottom trace, C: STD with selective saturation using 50 ms 90° Gaussian pulses for 3 s on 5.5 ppm; the total experiment time is the same as that for part B. Only the most intense ligand peaks (~7.1−7.2 ppm) are barely distinguishable from noise.

to artifacts in the spectrum. The unshaded pulse in the middle of T_{inv} is an optional nonselective 180° pulse that minimizes these artifacts by keeping the large magnetizations close to their original null condition.[66] Because the nonselective pulse ideally inverts *all* magnetization, to first order, it does not disturb the prevailing differential *z*-magnetization that drives the desired cross-relaxation.

The intrinsic waterLOGSY sensitivity is limited by the efficiency of water inversion, as well as the efficiency of the intended energy transfer scheme: water → receptor → ligand. Care must taken to avoid large translational diffusion losses during the e-PHOGSY spin−echo period. For example, more closely spaced gradients of moderate strength can be used. Experimental conditions that decrease the water R_1 are desirable, since they enable more "fresh" ligand to exchange on and off the receptor. As with STD, the bound and free state R_1 values of the ligand also limit the amount of magnetization transfer. For optimal sensitivity, the exchange rates should be significantly faster than the bound state ligand R_1.

As mentioned above, receptors having inherently poor spatial proton density will experience inefficient spin diffusion and thus poor STD sensitivity.[58,67] WaterLOGSY provides an attractive alternative for such targets by using the surrounding water molecules as surrogate spins to compensate for the inherent lack of proton DD cross-relaxation pathways. For example, in our laboratory, waterLOGSY has proven to be more sensitive than STD for screening nucleic acid targets. This is illustrated in Figure 12, which compares the signals of the waterLOGSY and STD experiments run for the same amount of time on a known binder of a ribozyme domain.

Like all ligand-based screening experiments, waterLOGSY is biased toward the detection of weakly binding ligands. Ligands with tighter affinity will have correspondingly longer residence times. Similar to the case for STD, if the residence times become too long, the transferred spin inversion will vanish due to longitudinal relaxation before the ligand can dissociate back into free solution. The estimated lower limit on K_D for waterLOGSY is $K_D \approx 0.1\ \mu M$.[60]

However, unlike STD, waterLOGSY spectra reflect both the free and bound states of a small molecule ligand. In the fast exchange limit, the waterLOGSY signal can be expressed as a weighted average

$$I_{WL} = C\{[EL]\sigma_{bound} + [L]\sigma_{free}\} \qquad (24)$$

C is a proportionality constant that accounts for the appropriate unit conversions, and σ_{bound} and σ_{free} are the rate constants describing the net transfer of magnetization between water and ligand protons in the bound and free states. The linear dependence on σ_{bound} and σ_{free} reflects a first-order approximation appropriate for shorter inversion times T_{inv}. For σ_{free}, the rotational correlation time is that of a free small molecule ($\tau_{free} \ll 1$ ns). For σ_{bound}, there is an effective rotational correction time $\tau_{eff} = \tau_{res}\tau_p/(\tau_{res} + \tau_p)$, where τ_{res} is the ligand residence time and τ_p is the rotational correlation time of the receptor−ligand complex. Typically, $\tau_{eff} \gg \tau_{free}$, which leads to opposite signs for σ_{bound} and σ_{free}. If the spectrum is phased such that $[EL]\sigma_{bound}$ yields a positive peak, then $[L]\sigma_{free}$ yields a negative peak at the same chemical shift. This implies that if $L_T \gg E_T$, then the negative contribution from the free state $[L]\sigma_{free}$ can overwhelm that of bound state, resulting in a false negative. To eliminate such confusion, large ligand/receptor molar ratios should be avoided. In addition, a reference spectrum, $I_{WL,free}$, of the compounds in the absence of receptor can be recorded. The reference spectrum can then be subtracted from the original I_{WL} to better estimate the bound state contribution.

Correcting for the free state contribution also permits estimates of the ligand binding affinity. The difference $I_{WL} - I_{WL,free}$ is simply the difference $Q_{avg} - Q_F$ in eq 13. In particular, if we identify $Q_{avg} \Leftrightarrow I_{WL}$ and $Q_F \Leftrightarrow I_{WL,free}$, then

$$I_{WL} - I_{WL,free} = \frac{CE_T(\sigma_{bound} - \sigma_{free})[L]}{[L] + K_D} \qquad (25)$$

The above expression differs trivially from that of Dalvit et al.[60] in that it is based on the bound receptor fraction $P_B^E = [EL]/E_T$ whereas Dalvit and co-workers consider the free receptor fraction ($1 - P_B^E$). Fitting $I_{WL} - I_{WL,free}$ in a ligand titration to eq 25 approximating $[L] \approx L_T$ yields an estimate of K_D. Dalvit and co-workers have used this approach to estimate the binding affinity of L-Trp to human serum albumin (HSA). Their results were of the same order of magnitude as those estimated by equilibrium dialysis. On the other hand, similar results (with

larger uncertainties) were obtained by fitting "uncorrected" titration data (just I_{WL}) directly to eq 24.

5.7. Exchange-Transferred NOE

Exchange-transferred σ^{NOE} measurements were among the first parameters proposed for ligand-based NMR screening.[68−70] In NOE-based screens, intramolecular NOEs of compound mixtures are observed via 2D-NOESY spectra in the absence and presence of the receptor. Binding compounds are identified by NOE cross-peaks that have changed sign in the presence of the receptor. Nonbinders show no change upon the introduction of receptor and display either zero or negative cross-peaks with respect to the diagonal. The estimated range of binding affinities that can be probed by transferred σ^{NOE} is 100 nM $\leq K_D \leq$ 1 mM.[70]

Exchange-transferred σ^{NOE} methods have lower sensitivity when compared to STD or waterLOGSY. For example, 2D-NOESY experiments rely on a comparatively short transient perturbation of magnetization (100−500 ms) to probe for binding-induced changes of *intra*ligand magnetization transfer. By contrast, STD and waterLOGSY rely on a long period of *intermolecular* magnetization transfer between receptor and ligand that is sustained by the continuous application of rf irradiation or the long T_1 of water. This sustained period of magnetization transfer affects the growth of a large pool of ligands each "labeled" with a binding signature. This heightens the sensitivity of these experiments over that of the 2D transferred NOESY. Furthermore, STD and waterLOGSY are relatively simple 1D experiments that reduce acquisition time and data storage burdens and thus increase throughput. In our experience, the same binding information can be obtained from a 0.5 h 1D STD experiment in place of a 4 h 2D transferred NOESY. While selective 1D NOE methods are conceivable, one faces the aforementioned challenge of frequency selective inversions for a library of potentially diverse compounds. Thus, the 1D STD has largely supplanted the use of exchange-transferred σ^{NOE} as a screening tool. Nevertheless, the change of sign inherent in σ^{NOE} still makes it an effective screening tool when other screening methods give ambiguous answers. And, of course, σ^{NOE} still retains its original value as a method for determining the bioactive *conformations* of weakly binding ligands.

5.8. Competition Binding Experiments

Competitive binding studies are a well-known set of biochemical approaches for determining binding specificity and affinity. The basic expressions relevant for such studies have been given in section 2. Recently, competition binding experiments have been incorporated into several strategies for ligand-based NMR screening. Following the literature, we use I and L to denote two ligands that compete for the same receptor site. Competitive displacement by a known specific ligand I can be used to confirm the specific binding of a new ligand L identified during primary screening. Furthermore, if the inhibitor dissociation constant, $K_I = [E][I]/[EI]$, is known, then it is possible to estimate the K_D of L by using eq 7 above. Meyer and Mayer have demonstrated this approach using STD on two galactose-containing ligands (β-GalOMe and NA_2) of the 120 000 Da tetramer *Ricinus communis* agglutinin. The K_D of NA_2 was estimated to be 27 μM from a competition study with β-GalOMe (K_I = 260 μM).[50]

Competitive binding also forms the basis for a screening strategy that enables the detection of *high-affinity* ligands using ligand-based methods. As already described, a significant limitation of standard ligand-based NMR screening is its dependence on fast exchange and consequent limitation to weak binding ligands (K_D typically > 1−10 μM). Tighter binding ligands have slower exchange rates. If the exchange rates are too slow, then the bound state information relaxes away before its transfer to the free state. In that case, only free state properties for the ligand would be observed, and the tight binder could be mistaken for a nonbinder, resulting in a false negative.

Recently, several groups have proposed highly similar competitive binding strategies to extend the range of ligand-based NMR screening to include high-affinity binders.[17−19] These strategies screen for higher affinity hits I via the competitive "knockoff" effects they exert on a previously characterized *lower* affinity ligand L. Such competition strategies have already been widely used in isothermal titration calorimetry (ITC) studies, which are constrained by similar binding affinity windows.[22,71] The principle steps of the strategy are highlighted below.[17]

1. First, a known binding ligand L is identified. Various studies have described this reagent as the "reporter", "reference", "spy", or "probe" molecule. The ideal phenotype for the reporter includes high solubility, medium to weak binding affinity (K_D > 10 μM), and an NMR parameter that displays a clear signature of binding even under large ligand excess. Examples include R_2 or selective R_1 relaxation enhancements, STD, or waterLOGSY intensities. More recently, the advantages of the high sensitivity of the ^{19}F R_2 and chemical shift have been exploited by Dalvit and co-workers.[47]

2. Next, a calibration curve is determined that relates an observed magnitude of the monitored NMR parameter of the reporter molecule to the bound reporter compound fraction $P_B = [EL]/L_T$. This NMR parameter can be measured for the reporter compound in a titration of either L_T or E_T (to vary $[EL]/L_T$). It is also useful to have an independent estimate of the reporter compound $K_D = [E][L]/[EL]$ determined using an alternate method (e.g. ITC). Then, substituting K_D into eq 3 with known values of E_T and L_T allows one to convert the titrated L_T (or E_T) values into P_B. The observed NMR parameter is then plotted as a function of the reporter compound P_B. The resulting calibration curve can then correlate subsequent perturbations of the monitored NMR parameter to altered P_B.

3. For fixed values of E_T and reporter compound L_T, a mixture of test compounds is then added to

screen for displacement effects exerted on the reporter compound. Displacement manifests as a shift of the reporter compound's NMR parameters toward those intrinsic to the free state. For example, if the monitored NMR parameter is R_2 or selective R_1, then a higher affinity hit will affect a decrease. The higher affinity hits are the competitive "inhibitors" represented by [I] and I_T in eqs 4−11.

4. Using the calibration curve established in step 2, the perturbed magnitude of the NMR parameter can be used to deduce the new decreased reporter compound $P_{B,+I} = [EL]/L_T$.

5. Since $P_{B,+I}$, L_T, and E_T are known, there is sufficient information (i.e. L_T, E_T, and [EL]) to estimate the apparent reporter compound $K_{D,app}$ using eq 6 above. Once $K_{D,app}$ is fixed, eq 6 can be used to solve for K_I. Note that the free inhibitor concentration [I] can be estimated by referencing one of its peak integrals to that of the reporter compound whose concentration is known.

A competitive binding strategy is appealing because it not only identifies higher affinity binders but also estimates their affinities (K_D) via a one-point measurement. Dalvit and co-workers demonstrated this competition technique using ^{1}H R_2 and selective R_1 relaxation enhancements to probe kinase−ligand interactions. The binding affinities obtained from the "one-point" measurement correlated well with concurrent ITC studies.[17] Standard K_D determinations via exhaustive ligand titrations are not always feasible due to poor compound solubility, or deterioration of receptor integrity caused by continued additions of concentrated DMSO stock solutions. A potential drawback to this method is that it is capable only of identifying binders at the site of the reporter molecule (and not e.g. adjacent subsites that might be of interest for drug design).

Recently, a modified competitive binding approach using ^{19}F NMR has been reported. The FAXS method[47] uses a small ^{19}F-based compound library to serve as "spy" and "control" molecules. The control molecules are known nonbinders of the receptor, and changes in their ^{19}F spectra enhance the accuracy of data interpretation by accounting for changes unrelated to specific binding. As stated above, ^{19}F detection yields clean ligand selective spectra, high sensitivity to chemical shift perturbations, and much larger R_{ex} relaxation enhancements than ^{1}H detection. In particular, the inherently larger chemical shift range of ^{19}F makes the intermediate exchange regime more likely. Together, these features make FAXS highly suited for sensitive and accurate probes of binding displacement.

Another application of ^{19}F NMR goes beyond mere detection of binding. The 3-FABS method[72] measures enzymatic IC_{50} values by comparing ^{19}F signatures of CF_3-labeled substrate before and after addition of active enzyme. In settings where enzymology data are limiting, the technique could be used by chemists to directly obtain IC_{50} values of their compounds. Also, the method could be useful in academic settings, where high-throughput-screening instrumentation is not readily available.

5.9. Troubleshooting Common Problems: Aggregation and Nonspecific Binding

A basic challenge in ligand-based screening is distinguishing the spectral signatures of bona fide binding from those stemming from artifacts such as aggregation and nonspecific binding. As with ligand binding, compound aggregation increases the effective molecular mass, driving τ_c upward and enhancing the $J(0)$-dependent relaxation rates. An easy NMR check for aggregation is measurement of any τ_c-sensitive NMR relaxation parameter of the lone compound as a function of concentration. For the typically low molecular masses (<1000 Da) of test compounds, R_2 should be on the order of 1−2 s and NOE cross-relaxation rates zero or positive (2D-NOESY peaks of opposite sign relative to the diagonal) at $B_0 \sim 11.7-14.0$ T. Of course, it is also possible to use other biophysical methods, such as light scattering, to assess aggregation tendencies.

Another problem that can occur is nonspecific binding. It is important to define and distinguish "nonspecific" versus "low-affinity" binding. Here, we consider low-affinity binders to be ligands with K_D values $\geq 10\ \mu$M and in fast exchange on the chemical shift time scale. Low-affinity binders are therefore a more general class of compounds than nonspecific binders. While nonspecific binders may be low-affinity binders, the converse certainly need not be true. Low-affinity interactions may indeed be specific for well-defined sites on the receptor. In some cases, binding of a ligand to these sites will modulate receptor activity. In other cases, binding could be to a specific but distinctly separate site proximal to the active site. As will be illustrated later in this review, one of the key benefits of NMR screening is the ability to detect low-affinity, yet specific, binders that might be missed by an enzymological screen. Such ligands may then be optimized to become higher affinity novel inhibitors. The criterion for specificity for these low-affinity binders is that they bind preferentially to the targeted active site or to another site that directly modulates receptor activity (e.g. through allosteric interactions or linkage to an active site ligand). In contrast, nonspecific binders usually bind to receptor surface regions that have no direct effect on receptor activity. These interactions result from general interactions of the ligand with hydrophobic patches on the protein surface (vide infra).

Nonspecific binding effects can be a serious concern for ligand-based approaches that use large ligand-to-receptor molar ratios. High ligand concentrations encourage the occupation of lower-affinity nonspecific binding sites. As mentioned in the previous paragraph, these nonspecific binding processes correspond to adsorption to hydrophobic patches on the protein surface. High ligand concentrations can increase the likelihood of this nonspecific surface adsorption. Accordingly, Murali et al. have used dilution to identify nonspecific binding.[73] In this study, the absolute concentrations of ligand were decreased while the ligand-to-receptor ratio was maintained at a constant value. Nonspecific binding was characterized by diminishment of the NMR binding signature at lower ligand concentration. Although in the above

study ligand affinities had already been characterized and thus it was straightforward to deduce nonspecific binding, in the case of a binder of unknown affinity, this information would not be available. Undoubtedly, a preferred method to expose nonspecific binding for ligand-based approaches is to test for displacement effects upon addition of a known specific and competitive binder. In addition, HSQC-based chemical shift perturbation can be used to map the ligand binding sites, provided that resonance assignments are available. Another strategy is to specifically label the active site of the receptor in such a way that only the active site can transfer energy to a compound. Examples of this strategy include the SLAPSTIC approach,[39] in which spin labels are judiciously placed at the active site. Alternatively, selective protonation of active site residues can be introduced in otherwise deuterated receptor proteins. STD experiments on such systems provide signals only for those compounds that bind to the active site.[74] In both cases, a priori information about the active site structure is needed, and in the last example, one relies on the facile overexpression of the deuterated receptor. To reduce the risk of nonspecific binding, it is helpful to work at lower receptor and ligand concentrations. By keeping E_T low, only the higher affinity binders (those that can saturate the binding site with minimal L_T due to low K_D) will have bound ligand fractions of significance, and thus detectable NMR signals. Finally, it is worthwhile to consider alternative expressed forms of the target for screening. For example, working with full-length proteins rather than truncated domains might present fewer exposed hydrophobic patches that attract "promiscuous" compounds.

5.10. Epitope Mapping via Ligand-Based Approaches

To provide more information for lead generation beyond the simple detection of compound binding, NMR researchers have sought to determine the specific portions of the ligand and protein critical for molecular recognition. This has been referred to as "epitope mapping" in reference to the original meaning of identifying antigen regions necessary for antibody or T-cell recognition. Epitope mapping can help guide lead optimization, especially if structural information for the target is not available (e.g. transmembrane proteins). A knowledge of which parts of the ligand are involved in the binding interface can help chemists decide how binding scaffolds should be elaborated, linked, or both. Application of epitope mapping to ligands is discussed below. Characterization of receptor binding site interactions is discussed under Receptor-Based Approaches (section 6).

5.10.1. Group Epitope Mapping

Mayer and Meyer have reported the GEM (group epitope mapping) protocol for identifying binding surfaces on the ligand using STD methods.[50] This approach compares the STD response $I_{STD} = I_0 - I_{SAT}$ for different protons within a ligand. In 1D STD and 2D TOCSY-dispersed STD spectra, various I_{STD} intensities (or peak volumes for 2D spectra) are normalized to the largest I_{STD} response within the ligand. The variation of normalized STD responses is related to ligand proton proximity to the receptor binding site. Stronger STD responses are interpreted as evidence of closer contact between ligand and receptor protons, taking into account the distance dependence of the intermolecular NOE underlying the saturation transfer process. To establish proof-of-concept, Mayer and Meyer performed a GEM analysis of galactose-containing ligands (NA_2 and β-Gal-OMe) of the 120 000 Da *R. communis* lectin agglutinin I. Their results were consistent with the known binding regions of galactose-containing ligands using other methods.

There are several caveats to be aware of when interpreting GEM experimental data. First, the mean lifetime of the receptor−ligand complex k_{off}^{-1} must be sufficiently short (ligand K_D sufficiently large), or spin diffusion within the bound state will equalize the ligand proton magnetization before its dissociation into free solution. Accordingly, epitope mapping is better suited for weaker binders (large k_{off}) and/or for sufficiently short saturation transfer times T_{SAT}. Second, it has been documented by Jayalakshmi and Rama Krishna[57] that interpretation of differential STD responses purely in terms of intermolecular cross-relaxation (NOE) can be an oversimplification. Understanding how this occurs is straightforward if we recall that the STD experiment is simply a 1D truncated, driven NOE (TOE) experiment.[75] In the TOE experiment involving a simple dipolar-coupled "I−S" two-spin system, the S magnetization is saturated, and the effects on the I spin magnetization are observed as a function of time. The I magnetization reaches a steady-state that is directly proportional to the cross-relaxation rate between the two spins and inversely proportional to the I-spin $R_1 = 1/T_1$. Thus, the magnitude of the TOE steady-state response reflects a competition between both cross-relaxation and autorelaxation. In this regard, interpreting the final I spin magnetization solely in terms of cross-relaxation can lead to erroneous conclusions. Jayalakshmi and Rama Krishna[57] carried out a rigorous complete relaxation and exchange matrix approach (CORCEMA)[76] for multispin systems engaged in chemical exchange. Their analysis confirms that the STD response, as in the simpler two-spin case, reflects a competition between the two effects: (i) cross-relaxation with the saturated receptor protons powered by the applied rf; (ii) autorelaxation that would dissipate energy to the lattice.

Because GEM maps not only intermolecular cross-relaxation but also the local R_1 values, caution must be exercised when interpreting smaller STD responses within a ligand. Smaller I_{STD} values do not *necessarily* imply remote locations from the receptor; rather, they may simply imply larger local R_1 values. Confusion of these two effects could lead to incorrect or incompletely characterized binding epitopes. Individual ligand proton R_1 values depend on several factors, such as the proton moiety considered (e.g. methine, methylene), local proton density, mobility, and the ligand conformation(s). Methylene protons,

for example, are more likely to exhibit smaller STD responses due to the strong mutual dipolar R_1 relaxation between proximal geminal protons. Quantitative interpretations of differential STD effects should therefore be treated with caution. If all of the ligand protons fortuitously have similar R_1 values, then these effects could be neglected. Alternatively, if the appropriate R_1 values could be measured, then the product of the STD intensities and the R_1 values should be proportional to the effective intermolecular cross-relaxation rate constant. It should be noted that the proof-of-concept experiment by Mayer and Meyer relied heavily on aliphatic methine protons within sugar rings, where the R_1 values could conceivably be quite similar.

5.10.2. Diffusion-Based Epitope Mapping

An alternative method of epitope mapping is based on the popular BPP−STE experiment (bipolar pair pulsed field gradient stimulated echo). This well-established method is used to measure molecular translational diffusion coefficients.[77−80] Comparisons of translational diffusion coefficients of small molecules in the presence and absence of receptor can be used to screen mixtures for receptor binding. Chen and Shapiro[81] have shown that intermolecular cross-relaxation between small ligand molecules and their protein receptors during the diffusion delay can give rise to erroneous estimates of the translational diffusion coefficient of the ligands. Recently, Yan et al.[82] have exploited this phenomenon to provide a novel method for epitope mapping. This method takes advantage of what would be considered an artifact in one experiment (intermolecular cross-relaxation), to provide an elegant solution to another problem. Because the artifact stems purely from intermolecular cross-relaxation, diffusion-based epitope mapping bypasses the problem of R_1 contributions discussed in the previous section for GEM.

Yan et al. compared the results of both GEM and diffusion-based epitope mapping for dihydrofolate reductase (DHFR) with those of a known ligand, trimethoprim (TMP). While the diffusion and GEM epitope maps were in agreement for most of the TMP protons, the diffusion-based approach gave a more accurate overall map. In particular, GEM incorrectly excluded the TMP methylene protons from the binding interface while the diffusion-based approach did not. This result can be rationalized on the basis of the large local R_1 of the methylene protons, which would lead to reduced STD intensities for these protons.

A disadvantage of diffusion-based epitope mapping relative to GEM is that the experimental method is considerably more cumbersome. In particular, the BPP−STE experiment uses a stimulated gradient echo, which incurs a mandatory loss of half the available magnetization, thus resulting in lower sensitivity. Finally, measurement of the intermolecular NOE artifact requires a series of 1D BPP−STE measurements, increasing experiment time and reducing throughput. Accordingly, this method is best suited to highly soluble receptors and ligands.

6. Receptor-Based Approaches

The first NMR-based screening approach to be described in the literature, the "SAR by NMR" method, monitored changes in the $^{15}N-^{1}H$ heteronuclear correlation spectra of a target to identify hits from a mixture of compounds.[1] These methods have found widespread application, especially in the pharmaceutical industry, since they were first published. Numerous applications and "tweaks" to the method have been published[83−86] and have been covered in a number of excellent reviews.[6−10,87−89] Many of these studies are discussed in the applications sections of the present review. Despite the wealth of information generated from these methods thus far, crowded spectra and resonance assignment of large monomeric proteins (>30 000 Da) remain bottlenecks that continue to limit the utility of receptor-based screening methods. However, recent progress has begun to address some of these bottlenecks. Advances in isotope-labeling strategies and data interpretation show promise in expanding the general applicability of receptor-based approaches. Several examples of recently published methods are discussed below. In the first example, site specific isotope labeling is employed to identify a ligand binding site resonance without performing a full sequential resonance assignment.[90] The second example employs a quantitative approach to characterize chemical shift perturbations due to ligand binding and localize sites of interaction to a much higher resolution than previously demonstrated.[91] While the following approaches cannot be classified as NMR-based screening methods, they are nonetheless enabling strategies that allow the researcher to use screening hits to generate earlier and better models for ligand design.

6.1. Selective Active-Site Isotope Labeling

For receptor-based screening of larger protein targets (>30 000 Da), it would be highly advantageous if binding site resonances could be identifed without undergoing the lengthy and tedious process of making sequential resonance assignments. To address this goal, Weigelt et al. have proposed a scheme that selectively labels pairs of sequential amino acid residues.[90] Selective labeling is achieved by first identifying unique pairs of residues, for example, amino acid X and amino acid Y, which reside exclusively within the protein active site. If such a pair can be identified, then a strategy that labels amino acid X with ^{13}C and amino acid Y with ^{15}N can in principle selectively mark the active site, as these residues will be the only source of signals in a multiple resonance experiment that edits according to scalar coupling between $^{13}C{=}O$ and amide ^{15}N.

Weigelt et al. performed proof-of-concept studies on the 14 700 Da fatty acid binding protein FABP-4. First, residues Val114 and Val115 were identified as a unique pair of sequential amino acids within the FABP-4 binding pocket. Subsequent labeling with $^{13}C/^{15}N$ Val in prototrophic bacteria yielded a sample that displayed a single cross-peak in the 2D $^{15}N-^{1}H$ projection of an HNCO experiment. Upon addition

of a known binder, this cross-peak displayed the expected chemical shift perturbations. Comparison of this peak with peaks in a conventional $^{15}N-^{1}H$ HSQC spectrum of the same sample (which shows all valines) allowed the binding site Val115 to be assigned. In turn, this allowed one to pinpoint perturbations remote from the binding site.

The advantage of this method is also its drawback. Since only one cross-peak probes the binding site, it becomes impossible to account for indirect effects from location(s) distinct from the binding site. Alternatively, compensatory binding-induced changes may accidentally sum to no net perturbation, thus generating a false negative. As a precaution, the authors recommend identifying additional distinct residue pairs. Another factor to consider when choosing amino acid pairs will be the presence of amino acid types that may undergo scrambling when expressed in *E. coli*. Complications may arise when attempting to use these amino acids (e.g. Ser, Asp, Asn) even when appropriate auxotrophic strains are used.

A new approach that combines aspects of selective isotope labeling, STD NMR, and epitope mapping has been recently described by Hajduk et al.[92] This method uses STD to examine a panel of target samples that have been uniformly deuterated except for certain amino acid types. By examining the relative effects on ligand protons for differentially protonated proteins, the amino acid composition of the ligand binding site can be defined. If a high-resolution structure of the protein is known, the structure of the protein−ligand complex can be inferred from all of the above data. In this case, no protein resonance assignments are necessary. The approach was validated with two test proteins, FKBP-12 and MurA. As much as this method is elegant, it may also be impractical for most laboratories for several reasons. First, multiple protein samples must be expressed and deuterated in a bacterial host and then purified in milligram quantities. Depending on expression levels, this can be very expensive from the perspectives of both manpower and cost. Second, unless there is no structural information available for a target, the structures of weakly binding ligands do not, for most chemists, present an attractive starting point for drug design. Finally, in our experience, if a protein can be expressed in the quantities needed for these studies, then it is highly likely that the target will be amenable to crystallization and X-ray structure determination. On the other hand, these methods provide a useful alternative for those targets that evade crystallization, or for proteins where crystal packing is known to induce structural artifacts.

6.2. Ligand Localization from J-Surface Analysis of Chemical Shift Perturbations

Chemical shift perturbation experiments have been the primary source of information enabling localization of ligand interactions with the receptor. Perturbations in ^{15}N chemical shifts, while indicating which amide protons are perturbed by ligand binding, do not provide precise information regarding binding geometry. McCoy and Wyss have developed an alternative, higher resolution method for localizing ligand binding sites by using previously unexploited information contained in binding-induced protein chemical shift perturbations.[91] The method recognizes the high prevalence of aromatic rings in drug-like molecules and that chemical shift perturbations of the protein upon binding must be due in part to ring current shifts induced by the ligand. By quantifying the spatial dependence of the ligand ring current field, and the consequent effects on the local magnetic fields of neighboring spins, the ligand binding site can in principle be more accurately characterized. Since localization proceeds from an analysis of shift perturbations alone, there is no need for a lengthy structure determination of the protein−ligand complex.

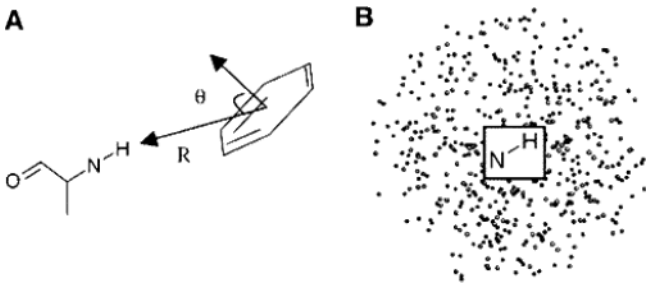


Figure 13. Schematic of ring current shifts caused by ligand aromatic groups on receptor NH protons. Left panel, A: Geometric parameters describing the influence of a ligand aromatic group on the NH chemical shift. The aromatic group is modeled as a magnetic point dipole. θ is the relative angle between a unit vector normal to the aromatic plane and passing through its center, and a vector **R** (with associated distance *R*) pointing from the ring center to the receptor NH in question. Right panel, B: The observed perturbation is quantified in terms of a maximum radius according to eq 26 in the main text. A sphere with this radius is constructed about the NH. Dots are distributed randomly throughout the sphere to represent potential ligand aromatic ring locations. Reprinted with permission from ref 91. Copyright 2002 American Chemical Society.

For receptor protons remote from a ligand aromatic ring, the local magnetic field stemming from the aromatic ring current can be approximated by that of a point or "perfect" magnetic dipole (infinitesimal current loop) located at the center of the ring.[93,94] This dipolar field can perturb the local magnetic field of a nearby protein proton *i*, causing a binding-induced chemical shift perturbation $\Delta CS(i)$, where $\Delta CS(i) = CS(i)_{+ligand} - CS(i)_{-ligand}$. If this perturbation is taken to be proportional to the perturbing ring current field, then $\Delta CS(i)$ becomes

$$\Delta CS(i) = -2(B_{dip}/R_i^3)P_2(\cos\theta_i) \qquad (26)$$

where $P_2(\cos\theta)$ is the familiar second-order Legendre polynomial $P_2(\cos\theta) = {}^1/_2(3\cos^2\theta - 1)$. R_i is the magnitude distance of a displacement vector $\mathbf{R}_i$ pointing from the ring center to the position of protein proton *i*. Both θ and **R** are illustrated in Figure 13A. B_{dip} is a proportionality constant, and θ_i is the angle between the ring plane normal (through the planar center) and $\mathbf{R}_i$. The sign and magnitude of $\Delta CS(i)$ will obviously depend on θ_i. $\Delta CS(i)$ extends from a minimum of $-2(B_{dip}/R_i^3)$ to a maximum of

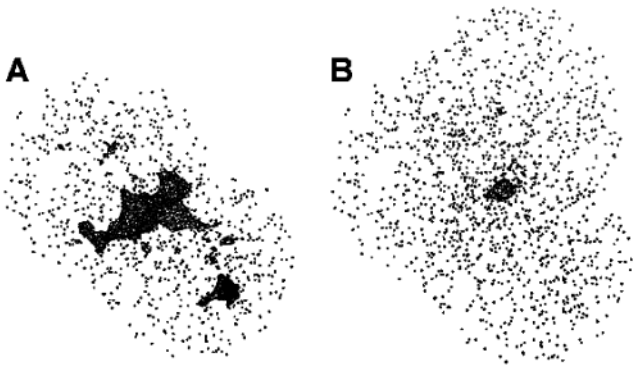


Figure 14. Consensus location of ligand aromatic ring by the method of overlapping spheres.[91] The overlap produces a localized volume for the ring location consistent with the observed chemical shift perturbations of multiple NHs. Left panel, A: Overlap of two spheres. Right panel, B: Overlap of three spheres. Reprinted with permission from ref 91. Copyright 2002 American Chemical Society.

$+1(B_{dip}/R_i^3)$. At $\theta = 54.7°$ (the magic angle), $\Delta CS(i) = 0$.

To translate the $\Delta CS(i)$ information into ligand localization, McCoy and Wyss use spherical dot-density representations; an example of which is given in Figure 13B. Specifically, spheres are constructed centered on each perturbed proton i that has a radius R_i consistent with eq 26. If the observed $\Delta CS(i) > 0$, then the maximum R_i possible must obey $\Delta CS(i) = (B_{dip}/R_{i,max}^3)$. On the other hand, if $\Delta CS(i) < 0$, then the maximum R_i must obey $\Delta CS(i) = -2(B_{dip}/R_{i,max}^3)$. The spheres are then filled randomly with dots, where each dot represents a possible location for the center of the aromatic ring. Smaller and larger $\Delta CS(i)$ will lead to larger and smaller $R_{i,max}$, respectively. To compensate for this, the same number of dots is used for each sphere. Hence, the dot-density will vary from sphere to sphere, depending on the magnitude of the $\Delta CS(i)$.

The emergence of a consensus volume created by overlap of multiple dot-density spheres from multiple perturbed protein protons serves to localize the ligand. Those $\Delta CS(i)$ consistent with a single ligand (common source of perturbation) produce a localized consensus volume with high dot-density on the surface of the protein. Figure 14 illustrates this principle. Inconsistent $\Delta CS(i)$ data (i.e. data that cannot be linked to a common perturbing source) yield diffuse dot-density data that point to no specific region of the protein. Sufficiently high dot-density is defined as being greater than 2−3 standard deviations above the mean dot-density, with the additional requirement that this be above the mean density of nonoverlapping spheres. The surfaces of these consensus volumes are called "*j*-surfaces", where *j* refers not to scalar spin−spin coupling constants but, rather, to current-density, analogous to the case of electromagnetic theory. In this case, the current-density referred to is the ligand ring current responsible for the $\Delta CS(i)$.

Ligands often contain more than one aromatic ring. Ligands with multiple rings separated by linkers of sufficient length (>5 Å) behave as linear systems; influences from separate ring systems can be added to achieve the net $\Delta CS(i)$. However, this treatment is not accurate for fused rings, and some error will be introduced into the resulting dot-density spheres. But, a key aspect of *j*-surface mapping is the reliance on *multiple* $\Delta CS(i)$'s for a given ligand. The greater the redundancy, the less vulnerable the conclusions are to approximate treatments of the current-density.

McCoy and Wyss have demonstrated the *j*-surface method using the HCV NS3 protease and helicase systems.[91] In both cases, *j*-surfaces from multiple $\Delta CS(i)$'s were able to localize ligands to surface positions consistent with those seen in X-ray crystal structures. The authors also used the HCV NS3 protease data to investigate whether *j*-surfaces could still correctly localize the ligand in the absence of sequential resonance assignments. A simulation was carried out in which the protein structure was known, and $\Delta CS(i)$ data were available for several systems with amino acid specific isotope labeling. The results were encouraging and suggested that such a procedure should be feasible. Clearly, bypassing sequential resonance assignment would greatly improve the general applicability of receptor-based screening in lead generation.

7. Applications of NMR-Based Screening

Pharmaceutical lead discovery has historically been dominated by high-throughput screening (HTS). Despite considerable advances in HTS methods, screening large numbers of compounds can still fail to deliver leads with acceptable potency, molecular diversity, novelty, or physicochemical properties. Fragment-based lead discovery has gained recent interest as an alternative approach capable of efficiently searching chemical diversity space. By screening relatively small libraries of molecular fragments, functional groups can be found that bind to each subpocket within an active site, and that information can then be used to guide synthesis. This approach circumvents the combinatorial problem that complicates screening library design. To illustrate this problem, consider a protein with two subpockets within the active site. One thousand fragments can be easily screened for binding at the two subsites, whereas a library of compounds containing 1000 different fragments at each subsite with a common connecting linker would require 1 million members (1000×1000). In fact, when possible linkage isomers are considered, the required library would be much larger. The fragment-based screen in effect models the chemistry space of the fully enumerated library without requiring the purchase or synthesis of enormous numbers of compounds. Furthermore, smaller molecules represent better starting points for medicinal chemistry because they can be built into larger, more potent compounds without exceeding the limits of certain physicochemical properties (e.g., molecular weight, polar surface area, clogP) known to correlate with oral bioavailability.[95,96] Because NMR-based methods are capable of detecting even very weak binding, they are very well-suited to screening small molecular fragments.

Independent of the experimental methods used, NMR screening applications are shaped by the strat-

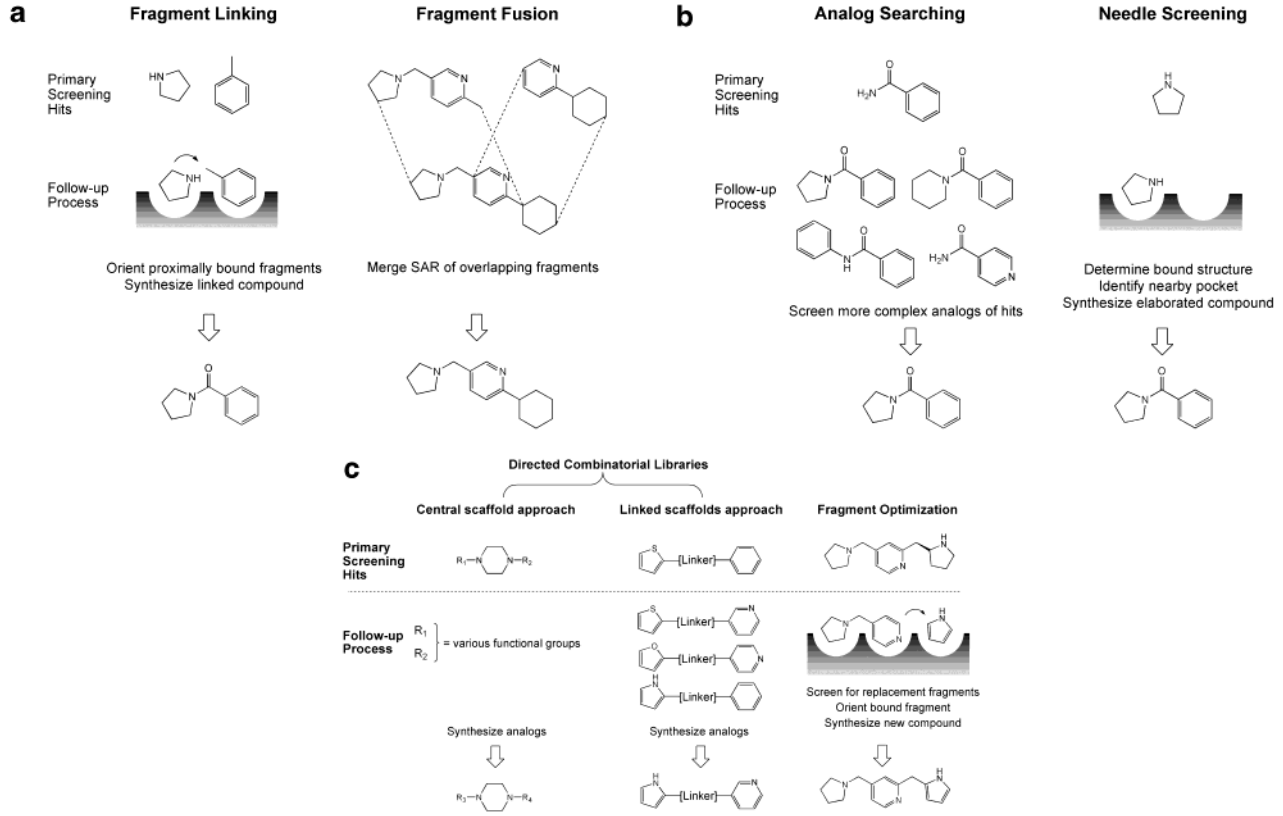


Figure 15. Basic strategies for constructing leads from NMR screening hits. (A) **Combination strategies.** In fragment linking, primary screening hits are mapped to specific binding sites and the relative orientations are determined for fragments that bind in close proximity to one another. Linked compounds are designed using structural and SAR information. In fragment fusion, the SARs for two fragments with overlapping binding sites are merged into one hybrid molecule. The relative orientations of the functional groups in the merged molecule are deduced from the topology of the fragments. In "blind" fragment fusion, there is no information regarding functional group orientations, so compounds containing random combinations of fragments are used. (B) **Elaboration strategies.** In analogue searching, elaborated analogues of the primary screening hits are found by searching databases of commercially available compounds using similarity and substructure-based computational searches. Information about active compounds is used to direct the selection of analogues for the next round of screening. Successive rounds of screening produce increasingly complex and potent inhibitors. In needle screening, small fragments capable of probing for specific interactions in the active site are screened. Structural information is obtained for bound hits, and nearby sites are identified that can be accessed by building from the starting fragment. SAR and structural information is used to guide the synthesis of more complex molecules. (C) **Variation strategies.** Using directed combinatorial libraries, in the central scaffold approach, a library of scaffolds containing multiple sites for substituents is screened. Hits are followed up by combinatorially varying the substituents. In the linked scaffolds approach, the screening library consists of compounds with two or more scaffolds connected by linkers that are amenable to combinatorial chemistry. Hits are followed up by combinatorially varying the scaffolds. Under fragment optimization, an undesirable substituent is removed from a lead compound. A screen is then carried out in the presence of the truncated lead compound to find replacement fragments that bind in the vacated site. The new fragments are oriented relative to the original scaffold, and linked compounds are synthesized.

egy used to direct information flow through the lead discovery process. The strategy dictates the selection of compounds to be screened, the choice of screening experiments and conditions, and the means by which information from the screen is used to find more desirable molecules. NMR screens are almost never run in isolation but are integrated with a wide variety of other techniques both to design screening libraries (e.g., virtual screening, physicochemical property-based selection, retrosynthetic fragmentation of known drugs) and to follow up screening hits (e.g., crystallography, enzymology, modeling, combinatorial chemistry). Many different permutations of these methods have been used; for example, HTS has been used to validate NMR screening hits, and NMR screening has been used to validate HTS hits. The integrated use of virtual screening, NMR screening, enzymology, and X-ray crystallography has proven to be particularly effective. Virtual screening is used as a filter to remove molecules with inappropriate shapes and functionality to bind to an active site, as well as to provide a relative ranking of candidates and models for bound orientations. NMR screening provides complementary experimental validation regarding which molecules actually bind and their relative affinities and binding modes.

The discussion of experimental techniques in the preceding sections was organized according to the methods used to detect ligand binding (i.e., protein vs ligand-based). For the following review of NMR screening applications, it is useful to categorize the examples according to the strategies used to develop the primary screening hits into leads. According to the classification scheme outlined previously[67,97] and illustrated in Figure 15, primary screening hits may be developed into more potent leads by (1) combining,

(2) elaborating, or (3) varying their molecular cores. Whatever the methods that are used to follow up on screening hits, the underlying strategy invariably falls into one of these categories. In practice, the distinctions between the three strategies can be subtle, and several strategies may be combined in one application. The following sections review each of the three strategies in turn and illustrate the methods used to complement the NMR screen, with emphasis on more recently published examples.

7.1. Applications Using a Combination Strategy

The combination strategy entails combining molecular fragments that have been demonstrated to bind individually to the target. This approach takes advantage of the fact that many drugs are modular, with groups that bind to distinct subpockets within an active site. When multiple, weakly binding fragments are combined into a single, more complex molecule that contacts the same set of subsites, potency can be dramatically enhanced. The combined molecule possesses a binding energy roughly equal to the sum of the binding energies of the fragments, but it pays a lower entropic cost by binding only one species. Thus, combining three fragments with millimolar affinities could create a molecule with mM × mM × mM = nM affinity.[1] In practice, the actual improvement in potency is difficult to predict due to entropic factors and changes in binding orientation upon linking.

Fragments may be combined either by linking them together or by merging chemical features from multiple fragments into one molecule. To combine fragments with a linker, one must know their relative orientations and proximities when bound. Chemical shift perturbation experiments have been used to determine the binding sites and orientations of bound ligands,[10,91,98−100] while NOE experiments can be used to identify ligand−protein and interligand contacts.[69,74,101,102] Crystallography has been widely used to determine the bound structures of weakly binding fragments, either soaked or cocrystallized with protein targets.[67,103−109] Proximity information can also be inferred from differential relaxation induced by a spin-labeled protein or ligand.[110] Alternatively, when fragments containing two or more functional groups bind at overlapping subsites, the relative position of their functional groups is apparent from the topology of the fragments, and it is possible to merge them together without having specific information about their bound orientations (Figure 15), an approach called “fragment fusion”.[69]

7.1.1. SAR by NMR

The first reported fragment linking method, SAR by NMR,[1] consists of a screen for binders at a first site, followed by a screen for binders at a second, proximal site carried out in the presence of saturating amounts of a first-site binder. Prior knowledge regarding known inhibitors is often used to bias the choice of fragments screened for the first binding site. The first application of this method used the FK506 binding protein (FKBP-12) as a model system.[1] A

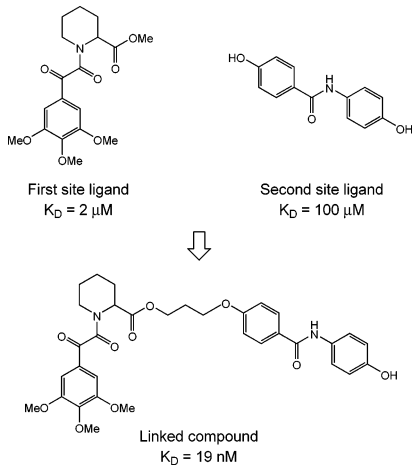


Figure 16. Construction of FKBP-12 inhibitors by fragment linking: (top) optimized first and second site ligands; (bottom) resulting linked compound.[1]

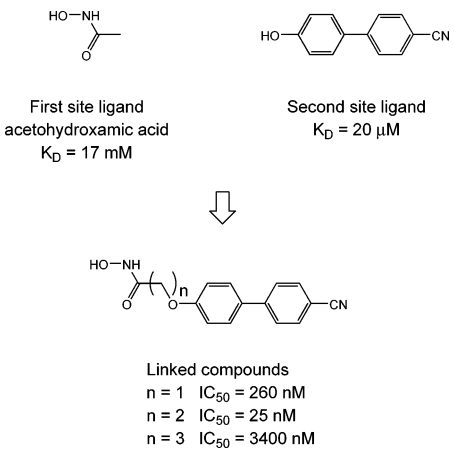


Figure 17. Construction of stromelysin inhibitors by fragment linking: (top) optimized first and second site ligands; (bottom) resulting linked compounds, with linkers containing one, two, or three methylenes.[111]

fragment with K_D = 2 μM (from a family of known inhibitors) was connected by a methylene linker to a K_D = 100 μM fragment found via ^{15}N HSQC-based NMR screening, producing an inhibitor with K_D = 19 nM (Figure 16). The linkers were designed using a model for the ternary complex calculated from ligand−protein NOEs. The same approach was used to build a stromelysin inhibitor with IC_{50} = 25 nM by linking acetohydroxamic acid (a known inhibitor with IC_{50} = 17 mM) to a 20 μM biphenyl fragment found by second site screening in the presence of saturating amounts of the hydroxamate (Figure 17).[25,111] The linked compound was more potent than expected from the sum of the binding energies of the fragments. Thermodynamic analysis[25] revealed that binding of the linked compound was augmented by cooperative contacts between the two fragments, entropy gained by reducing the number of components, and enthalpically favorable linker−protein contacts (the latter being very sensitive to the length of the linker). This example underscores the basic rules of linker design summarized by Jahnke:[110,112] (1) preserve the relative orientations of the bound fragments; (2) use a rigid linker; and (3) optimize contacts between linker and protein.

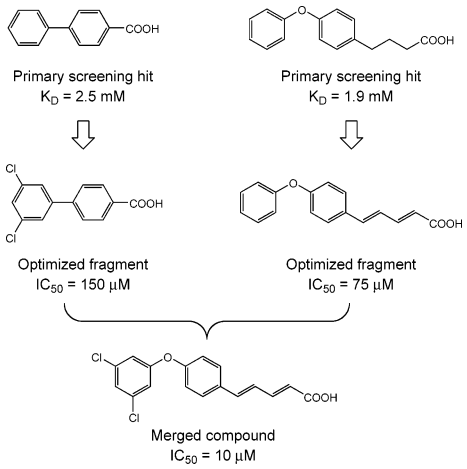


Figure 18. Construction of HPV-E2 inhibitors by fragment fusion. Primary screening hits (top) binding at overlapping sites on the protein were independently optimized (middle). The optimal features of each fragment were then merged into a hybrid molecule (bottom).[113]

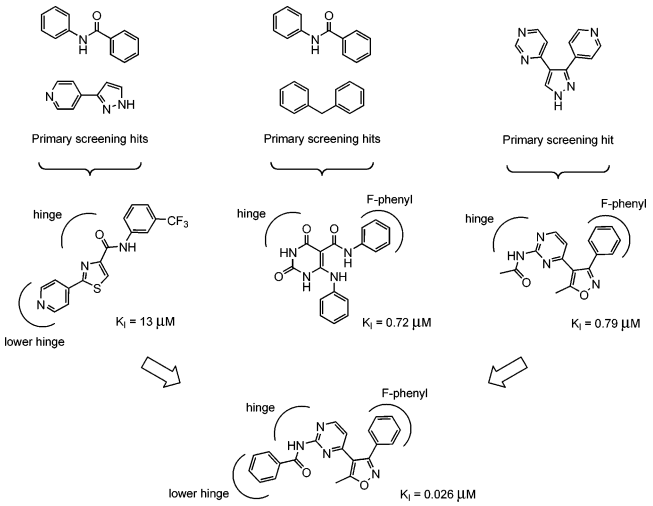


Figure 19. Construction of Jnk-3 inhibitors by fragment fusion. Primary screening hits (top row) were followed up by screening compounds containing random combinations (left and center columns) and variations (right column) of active fragments. Crystal structures of the three lead classes (middle row) revealed three subsites (hinge, fluorophenyl pocket, and lower hinge pocket) that were filled by functional groups on the leads. The thiazole amide (left column) and isoxazole (right column) classes each contact two of the three subsites; a merged molecule (bottom) was synthesized that contacts all three.[67,104]

Examples of combination by fragment fusion include the discovery of inhibitors of the DNA-binding domain of human papilloma virus E2 protein.[113] The SARs of two lead series were optimized independently and then merged into one molecule (Figure 18). Although the potency only increased by 8-fold when the fragments were linked, this case is noteworthy because it was the first published example in which the search was not biased by the use of fragments from known inhibitors. An example of "blind" fragment fusion (merging fragments without any prior information on their relative orientations) is the design of inhibitors for Jnk3 MAP kinase.[67,104] In this application of the SHAPES method, small ATP site-binding fragments with very weak (0.5−1 mM) affinities were identified by tNOE-based screening and followed up using two independent strategies (Figure 19). First, molecules containing any combination of two binding fragments were identified in a database of commercially available compounds using substructure and similarity searching. In the second strategy, the screening hits were computationally docked into the ATP site of the X-ray structure, and four fragments were found to give energetically reasonable bound orientations. Compounds containing those fragments were then docked, and the most promising were purchased. Of 300 follow-up compounds that were tested, 8 had potencies better than 20 μM, a hit rate 10-fold higher than that observed in the prior Jnk3 HTS screen. The X-ray structures of the three principal leads revealed that all placed scaffolds with hydrogen bonding groups in the site occupied by the adenosine ring of ATP, and each extended a functional group into one of two different adjacent pockets that are not contacted by ATP. This SAR was merged to create a compound contacting all three subsites that had a $K_I = 26$ nM (Figure 19). It is interesting to note that none of the three classes of lead compounds had been identified as hits in the previous HTS screen, even though all were represented in the HTS library. Presumably, the relatively low detection threshold ($IC_{50} < 30$ μM) of the enzyme assay caused those compounds to be missed. These two cases illustrate the ability of NMR screening to discover leads even after conventional HTS screening of many thousands of compounds has been unsuccessful.

Successive rounds of screening can be used to systematically construct linked inhibitors. Nonpeptidic inhibitors of the Syk C-terminal SH2 domain were recently assembled using an iterative, modular approach.[114] Separate virtual screens were used to identify fragments capable of binding to each of the three principal SH2 domain subsites: the phosphotyrosine site (pY), pY+1, and pY+3. Compounds were manually selected from highest-scoring fragments for each subsite (including pTyr-mimetic malonates for pY and carboxylates to bind to the basic lysine in pY+1) and tested for binding using a subsite-competitive surface plasmon resonance assay. The pY malonates were $IC_{50} = 6-12$ mM inhibitors (for comparison, pTyr has $IC_{50} = 14$ mM), and the best pY+1 ligand had $IC_{50} = 8$ mM. The hydrophobic Y+3 subsite did not bind fragments well, possibly due to a scarcity of residues capable of making specific interactions (such as hydrogen bonds) with ligands. Weak intermolecular NOEs between fragments bound in pY and pY+1 were used to design a linked compound from the best-binding fragments (Figure 20) that had an $IC_{50} = 350$ μM. Guided by a docking model, a hydrophobic group was then attached to interact with pY+3; this increased the potency to 38 μM. Although NMR was not used for screening in this case, it demonstrates the practicality of using a systematic, structure-based fragment linking approach.

A recent series of papers from Abbott Laboratories[115−118] describes a particularly elegant application of iterative fragment linking to the design of protein tyrosine phosphatase 1B (PTP1B) inhibitors.

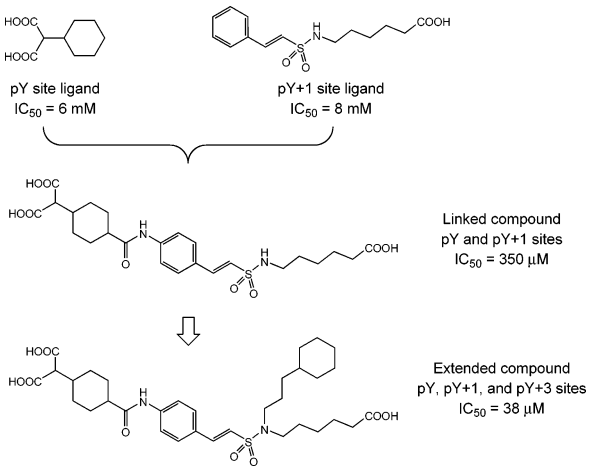


Figure 20. Iterative fragment linking approach to assembly of Syk SH2 domain inhibitors. The best hits from independent fragment screens of the pY and pY+1 subsites (top) were linked together (middle). The linked compound was then elaborated by adding a hydrophobic group to access the pY+3 pocket (bottom).[114]

The principal challenges in designing PTP1B inhibitors are to (1) find functional groups that bind to the phosphotyrosine site without introducing excessive charge on the molecule (since this reduces cell permeability) and (2) achieve selectivity over other phosphatases, particularly T-cell PTPase (the closest PTP1B homologue, with 74% identity in the catalytic domain[116]). ^{13}C and ^{15}N HSQC-based NMR screens of 10 000 small molecules (MW < 350) were carried out to search for novel phosphotyrosine mimetics that could be used as starting points for lead development.[115] A simple diaryloxamic acid was found to bind at the pTyr catalytic site with $K_I = 93\ \mu M$ (Figure 21). This lead was followed up using a variation strategy to test 2,3-dimethylaniline analogues, producing the slightly more potent 2-ethylaniline and 7-hydroxynaphthalene derivatives. The X-ray structure of the latter compound ($K_I = 17\ \mu M$) revealed that it binds with the oxamate in the phosphate site and the naphthyl in the same position as that of the phenyl ring of pTyr, causing the flexible WPD loop, which normally closes over the substrate, to remain in an "open", apo-protein-like position. This structure suggested that a linker from the naphthyl ring could be used to reach a second, noncatalytic pTyr binding pocket that has lower homology among tyrosine kinases, offering a possible means to achieving selectivity. Incorporation of a diamide linkage and polymethylene chain improved potency to 1 μM, increased bioavailability, and improved selectivity against five other phosphatases (LAR, SHP-2, CD45, CDC25, calcineurin) with the exception of TCPTP. A second NMR screen identified ligands for the second binding site, leading to the synthesis of polymethylene-linked salicylate analogues with potencies of about 20 nM (Figure 21B). Although these compounds showed some (2- to 3-fold) selectivity for PTP1B versus TCPTP, their cell permeability was poor and could only be improved by use of a methyl ester prodrug. Similarly, the 7-hydroxynaphthalene derivative was connected to a 2-naphththoic acid fragment in the second pTyr site[118] (not shown).

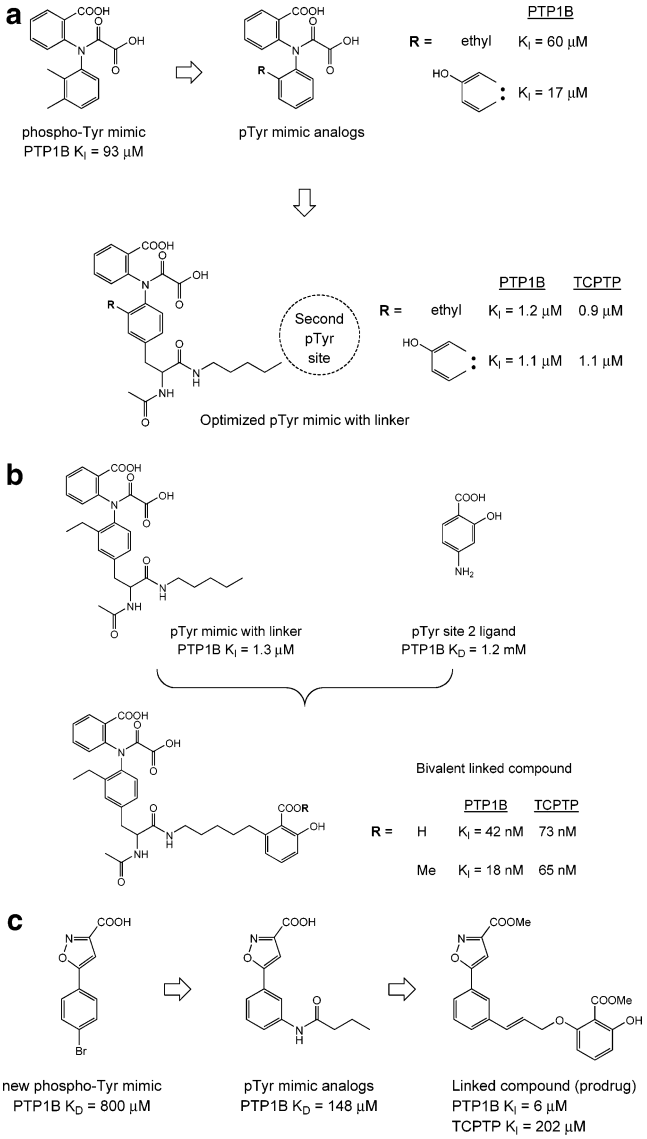


Figure 21. Iterative fragment linking approach to assembly of PTP1B inhibitors. (A) First step. The phosphotyrosine mimic identified in the primary NMR screen (upper left) was optimized by testing various analogues (upper right). A polymethylene linker was then attached (bottom) to access the second phosphotyrosine binding site.[115] (B) Second step. A second NMR screen identified fragments binding to the second pTyr site. The best of these fragments was attached to the pTyr mimic-linker compound from the first step (top row), yielding a bivalent linked compound (bottom).[116] (C) Third step. A third NMR screen identified a monoanionic pTyr mimic (left) that was optimized by testing analogues (center) and then connected to the previously identified salicylate fragment in the second pTyr site to make a bivalent linked compound (right).[117]

This linked compound had a similar potency (PTP1B $K_i = 22$ nM) to that of the salicylate-linked compound and was likewise selective against a panel of phosphatases but only 2-fold selective against TCPTP. In an effort to find novel phosphotyrosine mimetics with improved cell permeability, a fragment-optimization approach was used: a third NMR screen was carried out using only monocarboxylic- or non-carboxylate-containing fragments.[117] An isoxazole carboxylic acid (Figure 21C) was found to be a weak ($K_D = 800\ \mu M$) binder at the first pTyr site, and the X-ray structure

of a more potent (K_I = 148 μM) analogue suggested that a four atom linker would be ideal for linking to the previously identified salicylate fragment bound in the second site. The linked compounds had K_I values of 6−7 μM, and although not as potent as the previous lead series, they were approximately 30-fold selective for PTP1B versus TCPTP, were >50-fold selective against the panel of other phosphatases, and showed moderate cell permeability. The preceding example was not the first attempt by the Abbott group to find phosphotyrosine replacements with improved pharmacological properties. They previously screened 3500 compounds against the Lck SH2 domain for this purpose,[119] and even though novel phthalamate-based replacements were found, no attempts to link them using SAR by NMR were reported.

7.1.2. Bifunctional Ligands

Fragment linking has also been used to design bifunctional ligands intended to target entire gene families of proteins by incorporating common ligand mimics. This approach is a slight variation of SAR by NMR in which the first site ligand simply mimics a cofactor of the target (e.g., NADH for a reductase or ATP for a kinase). Two examples using this approach have been published.[74] "NMR-DOC" (NMR docking of compounds) uses amino acid type specific isotopic labels in an otherwise deuterated protein to simplify the spectra and facilitate screening and docking of hits, particularly for targets too large to be readily assigned. In a model system using U-^{2}H-[^{13}C,^{1}H-Met,Ile,Thr]-dihydropicolinate reductase ([MIT]-DHPR), an NADH mimic, nicotinamide mononucleotide (NMNH), was confirmed to bind to the active site using selective saturation transfer via the labeled residues. Two labeled active site residues (Thr107 and Met17) were assigned by observing which resonances shifted upon adding NADH. The bound NMNH conformation was modeled using the crystal structure of the NADH complex and NOEs from the assigned residues to the ligand. NOEs between NMNH and a bound substrate analogue, pyridine-2,6-dicarboxylate (PDC), were then used to demonstrate that relative ligand orientations could be determined. "NMR-SOLVE" (structure oriented library valency engineering) combines fragment linking with the design of targeted combinatorial libraries. For targets with two adjacent cofactor and substrate binding sites, a linker is attached to a ligand bound in the cofactor site and a combinatorial library is made of fragments that attach to the linker and bind at the substrate site. When the target belongs to a gene family of proteins with a common cofactor, the first ligand is a cofactor mimic and the second ligand confers specificity; the resulting biligand library can be used to screen for target specific inhibitors within that family. This approach was tested using selectively labeled 1-deoxy-D-xylulose 5-phosphate reductoisomerase ([MIT]-DOXPR), a homotetramer of 174 kDa for which there is no solved structure or close homologue in the Protein Data Bank. A methionine residue was identified on the basis of NOEs from bound NADH that was presumed to be located between the cofactor and substrate binding sites. Contacts between this residue and screening hits bound at the substrate site were reportedly used, in the absence of any other structural information, to design biligand DOXPR inhibitors with nanomolar affinities.

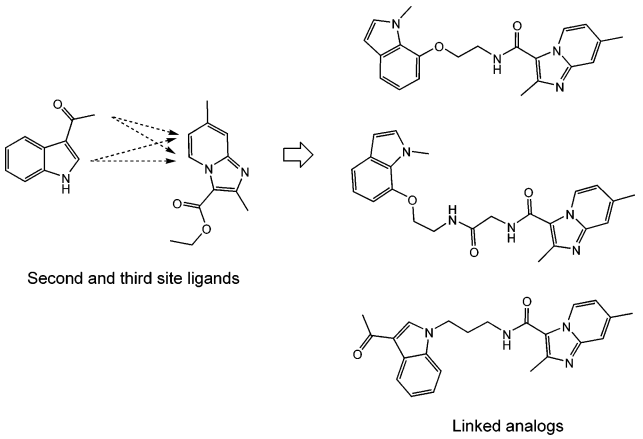


Figure 22. Orientation and linking of tubulin-binding fragments. Two fragments that bound close together at second and third sites on tubulin (left) were oriented using interligand transferred NOEs (arrows). Linked analogues were synthesized (right) and are shown ranked by binding affinity, with the most avid binder at the top.[112]

7.1.3. Spin Label-Guided Fragment Linking

Spin labels can be used in several ways to aid in the design of linkers to connect bound fragments.[110] When the SLAPSTIC method[39] is used, ligand resonances that do not undergo line broadening or chemical shift changes upon binding to a spin-labeled protein are probably not in contact with protein and thus indicate possible sites for attaching linkers. Screening for second site binders can be carried out in the presence of a spin-labeled first ligand, and the distances between the bound ligands can be calculated from the differential relaxation rate enhancements. If a site specifically spin-labeled protein sample is available (e.g., labeled at the amino terminus), then the same calculation may be used to derive distances to two unlabeled ligands.

Second-site screening using a spin-labeled first ligand has been applied to find inhibitors of tubulin.[112] A TEMPO spin label was attached to a known ligand, 3-methylpyridine, and a second site screen was carried out using $T_{1\rho}$ relaxation difference experiments to identify proximally bound fragments. No NOEs were observed between the hits and 3-methylpyridine, but two of them unexpectedly exhibited NOEs to one another, indicating that they were bound close together at second and third sites on tubulin. Several linked molecules containing the second and third site ligands were synthesized (Figure 22) and had K_D values of low micromolar or higher. A similar screening strategy was followed to find fragments binding to the anti-apoptotic protein Bcl-xL.[40,112] The starting point (Figure 23) was an IC_{50} = 180 μM bis-aryl compound that was found in an HTS screen and could not be significantly improved by medicinal chemistry. This compound was TEMPO-labeled at the amine and used for a second site screen, identifying at least one weakly binding

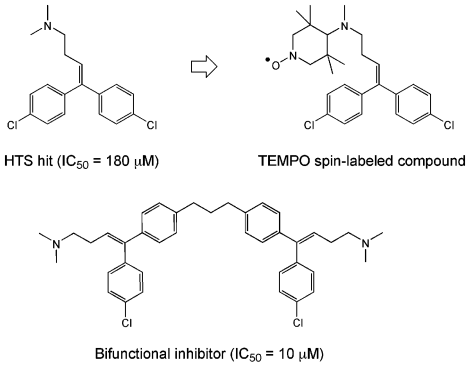


Figure 23. Design of Bcl-xL inhibitors using a spin label. The original HTS hit (upper left) was TEMPO-labeled at the amine position (upper right). Analysis of relaxation enhancement data revealed that two spin-labeled molecules bound in proximity, leading to synthesis of a bifunctional inhibitor (bottom). Adapted from refs 40 and 112.

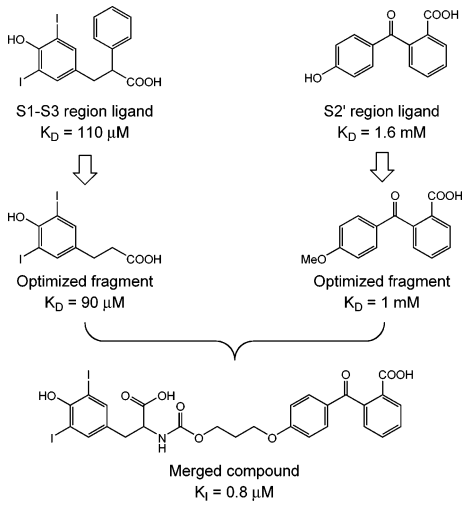


Figure 24. Construction of HCV NS3 protease inhibitors by fragment linking. Primary NMR screening hits (top row) individually targeting the S1 to S3 (left column) and S2′ (right column) subsites were independently optimized by testing analogues (middle row) and then linked together (bottom row) to make an inhibitor spanning the S3 to S2′ subsites.[120]

(1 mM) compound that bound nearby in the site targeted by Bak, the natural peptide antagonist. Surprisingly, it was also found that two of the TEMPO-labeled ligands could bind simultaneously at adjacent positions in the Bak binding site. Linked compounds containing two bis-aryl groups were synthesized on the basis of spin label and intermolecular NOE data, resulting in an $IC_{50} = 10\ \mu M$ bivalent inhibitor. It should be noted that both of the surprises in these examples involve fragments binding unexpectedly at multiple sites. This type of problem can occur for any screening method and if unrecognized will confound the fragment linking strategy.

The most recent fragment linking example at the time of this writing was the design of inhibitors targeting the hepatitis C virus protease NS3-NS4A[120] (Figure 24). Using ^{15}N HSQC experiments, 3639 compounds were screened and mapped to subsites within the active site. Of 50 primary screening hits, 16 bound in the active site: one covering the S4 to S2 subsites, six covering the S3 to S1 subsites, and

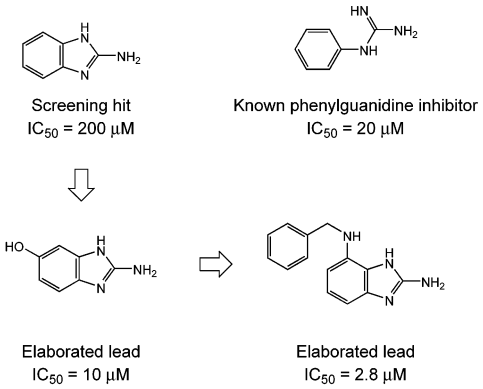


Figure 25. Design of urokinase inhibitors. A phenylguanidine-competitive NMR screening hit (upper left) was elaborated to make a more potent lead (lower left), which was further extended by attaching a phenyl ring to access the S1B pocket (lower right).[106]

nine at the S2′ subsite. The five classes of primary hits, most with millimolar affinities, were independently optimized by testing analogues, and two fragments, one covering the S3−S1 subsites and the other at the S2′ subsite, were chosen for linking. The bound structures of the fragments could not be determined by crystallographic or conventional NOE-based methods, so they were modeled by fitting the ligands to the corresponding *j*-surface derived from chemical shift perturbation data.[91] The linked compound, spanning S3 to S2′, was a substrate-competitive inhibitor with $K_I = 0.8\ \mu M$, and its pattern of chemical shift perturbations was consistent with the predicted binding mode. Although linking increased the potency by 100-fold compared to that of the most potent fragment, the increase was not as large as theoretically possible; this may be due to the use of a long, flexible linker to span the distance between the S1 and S2′ subsites.

7.2. Applications Using an Elaboration Strategy

The elaboration strategy consists of systematically building upon primary screening hits to make more complex molecules. Analogues are selected that contain functional groups or ring systems capable of making additional interactions with the target, thereby increasing potency without disrupting binding of the core scaffold. The SAR from each generation of analogues is used to bias the selection of compounds for the next, so that the process converges rapidly to more potent inhibitors. Structural information about the bound leads is often used to direct lead optimization.

7.2.1. Basic Elaboration Approaches

One of the earliest examples of an elaboration approach is the discovery of a novel urokinase inhibitor.[106] Using ^{15}N HSQC-based detection, 3000 compounds were screened in the presence of 1 mM phenylguanidine (required to prevent protein autolysis), resulting in discovery of 2-aminobenzimidazole, an $IC_{50} = 200\ \mu M$ phenylguanidine-competitive inhibitor (Figure 25). This hit was of particular interest because it is less basic ($pK_a = 7.5$) than conventional arylamidine or guanidine inhibitors ($pK_a = 9-11$)

and thus may have intrinsically higher cell permeability. Since analogous heterocyclic scaffolds (benzoxazole, benzotriazole) did not bind urokinase, the aminobenzimidazole scaffold was elaborated by addition of simple functional groups. The most potent analogue, 5-hydroxy-2-aminobenzimidazole, was equipotent with known inhibitor classes ($IC_{50} = 10\ \mu M$). The crystal structure revealed that this compound contacts the same residues as conventional inhibitors but does not access the S1B pocket, a site that is critical for the potency of other inhibitors. Attaching a phenyl group to fill this pocket improved potency by 70-fold.[121] The resulting compound is orally bioavailable and represents a novel scaffold for constructing urokinase inhibitors with improved pharmacokinetic properties.

Virtual screening has been used to select primary screening compounds and then to choose more elaborate analogues to follow up hits. An example of this is the discovery of inhibitors for the human hydroxysteroid dehydrogenase 3α-HSD.[19] In this case, a small, relatively weakly binding "reporter ligand", 2-acetylbenzofuran, was first found by NMR screening to target the active site of 3α-HSD. This fragment was used to compile a list of 713 similar, proprietary compounds that were docked into 3α-HSD. Thirty of the top-ranked compounds were then manually selected and tested for their ability to displace the reporter ligand from spin-labeled 3α-HSD. The protein spin label amplified the binding signal observed in $T_{1\rho}$-edited spectra, enabling use of protein concentrations as low as 0.2 μM. Three compounds (not shown) displaced the reporter ligand, and these were used as input for another round of virtual screening. The hits from this follow-up set reportedly had submicromolar affinities.[122]

7.2.2. Needle Screening

Needle screening is a process by which "needles", small (MW < 300) and simple fragments that can probe active site subpockets, are screened and the hits elaborated using structure-guided design. Needle libraries are constructed using modeling methods to select molecules comprising the minimum structural elements necessary to make essential interactions with the target. The first reported application was to find leads for DNA gyrase after conventional HTS had proven unsuccessful.[105] Virtual screening (docking and pharmacophore searching) was used to compile a list of 600 needles predicted to be capable of forming essential hydrogen bonds with a side chain (Asp73) and a conserved water molecule in the ATP site of the gyrase B subunit. These compounds were screened at high concentration (up to 0.5 mM) in an ATPase enzyme assay. 2400 analogues of the primary hits were screened in a follow-up ATPase assay, yielding a total of 150 weak hits that clustered into 14 structural classes. The hits were validated using a supercoiling assay and several direct binding assays: analytical ultracentrifugation, surface plasmon resonance, and ^{15}N HSQC spectroscopy. Several classes of nonspecific inhibitors were eliminated, and chemical shift perturbation experiments confirmed that seven classes bound at the ATP site. HSQC

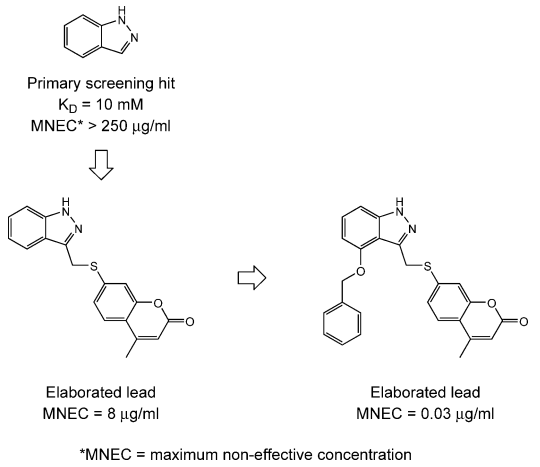


Figure 26. Discovery of DNA gyrase inhibitors using needle screening. The ATP-competitive "needle" hit (upper left) was first elaborated by addition of a group capable of hydrogen bonding to the active site Arg136 (lower left) and then further extended by adding a phenyl ring to access a nearby hydrophobic pocket (lower right).[105]

titration experiments were used to measure K_D values for the primary needle hits that ranged from 10 to 200 mM. Despite the very weak affinities, crystal structures were solved for the five most avid hits, of which indazole (Figure 26) made the most intimate protein contacts in the ATP site, apparently by displacing the conserved water molecule. The indazole scaffold was elaborated by adding a methylcoumarin group that could hydrogen bond to active site residue Arg136, increasing potency by over 30-fold. A nearby lipophilic site identified in the crystal structure was exploited by adding a second side chain; this improved the potency by an additional 270-fold. This elegant example demonstrates that virtual screening, enzymatic assays, direct binding assays, and structural information can be combined to efficiently elaborate even extremely weakly binding fragments into potent inhibitors. It is also interesting to note that HSQC-based screening for second site binders was attempted for gyrase B using saturating amounts of a phenylnitroindazole derivative shown to bind in the ATP site.[100] Of 250 small (MW < 150) needles screened, about 20% bound to gyrase B at a specific site that was mapped using chemical shift perturbations. Unfortunately, this binding site was located on the opposite face of the protein from the ATP site and was too far away to be exploited by chemistry. Thus, the elaboration strategy produced leads for a target for which SAR by NMR-style fragment linking was unsuccessful.

In one of several recent applications of needle screening, peptide deformylase (PDF) was screened in an effort to find novel, nonchelating motifs that bind in the hydrophobic S1′ pocket of the metalloprotease.[100] The diamagnetic Zn^{2+} form of the enzyme was screened using ^{15}N HSQC detection, and the hits were ranked using NMR-derived K_D values. A model for the bound orientation of one hit, a fused bicyclic aromatic compound, was then generated by restrained docking using ^{13}C HSQC chemical shift perturbation and NOESY distance data (not shown). In another example, a "zinc needle" library of putative zinc-chelating fragments was screened against

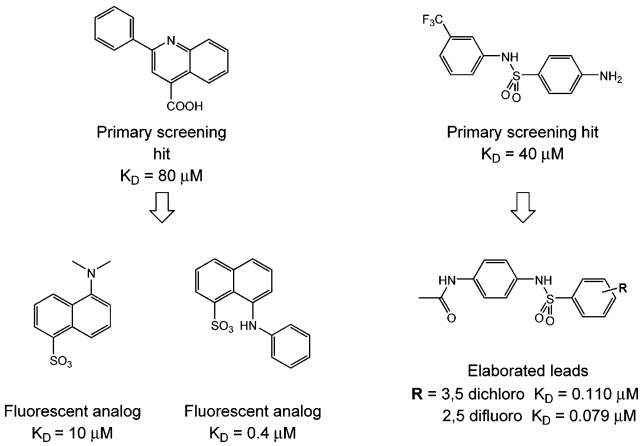
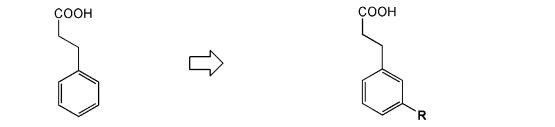


Figure 27. Discovery of FABP-4 inhibitors using the SHAPES strategy. Primary screening hits (top row) were used to guide the selection of fluorescent analogues (left column) that were used for assay development. Structural information on bound primary hits was used to select elaborated leads for secondary screening (right column) [Lepre, C.; Margolin, N.; Parker, M. Unpublished results].

the catalytic domain of human fibroblast collagenase (MMP-1) by 1D ^{1}H NMR.[100] Changes in the ^{1}H chemical shifts of the well-resolved active site histidine resonances (His218 and His222 HD1 at ~13.3 ppm) signaled binding, and ^{15}N HSQC experiments were used to confirm the binding site and measure K_D values. Although the structures of the hits were not reported, at least one novel chelator was found that had an affinity of 15 mM, which is comparable to that of the hydroxamic acid group commonly used for building metalloprotease inhibitors.

7.2.3. SHAPES Screening

The SHAPES strategy consists of ligand-based screening of a diverse library of druglike molecules combined with successive rounds of follow-up screens by HTS, NMR, or other direct-binding assay.[67,69,104] The compounds are comprised of scaffolds and side chains commonly found in known drugs, chosen to have desirable physicochemical properties and be simple enough to allow further elaboration without creating molecules too complex to serve as good leads.[67,69,123] It should be noted that a SHAPES library and needle libraries are conceptually different. While both libraries contain low-molecular-weight, druglike fragments that are capable of binding one or more subsites within a target active site, a needle library is created using virtual screening and is unique to the target under consideration. A SHAPES library, on the other hand, can be considered a "universal" library that is used for all potential targets. An example of the elaboration strategy using SHAPES screening is a study carried out on the human fatty acid binding protein (FABP-4). The primary tNOE-based screen identified 13 ligands with affinities ranging from 0.3 to 800 μM (Figure 27). The crystal structures of two primary hits soaked into apo-FABP-4 were used to direct the selection of follow-up compounds with functional groups positioned to make additional protein contacts. In addition, fluorescent analogues of the NMR hits were screened in order to find probe molecules suitable for development of a high-throughput, fluorescence-based competition assay. A second generation calorimetric screen of 134 commercially available follow-up compounds yielded nine leads with low-micromolar to nanomolar affinities. The crystal structures of seven primary and secondary screening hits were subsequently solved, mapping out the essential ligand contacts in the binding pocket and defining the pharmacophore.



R group	EC50	Selectivity vs. FABP-3
Ethyl	80 μM	10-fold
$CH(CH_3)_2$	10 μM	26-fold
benzyl	5.5 μM	10-fold

Figure 28. Elaboration of a FABP-4 screening hit into more selective leads. The primary phenyl propionate screening hit (left) was followed up by testing a series of more complex analogues (right).[37]

7.2.4. Structure-Based Screening

The group at Biovitrum independently screened FABP-4 using a library of approximately 500 small, diverse compounds and 1D $T_{1\rho}$ experiments for detection.[37] Thirty-eight of the 52 hits were categorized as having affinities of ~250 μM or better, since they were detected in less sensitive $T_{1\rho}$ experiments using a short relaxation filter. Almost all of these highest-ranked hits possessed an acidic group, including a phenyl propionate scaffold with $EC_{50} = 590$ μM that was 25-fold selective versus FABP-3 (Figure 28). Binding to the latter protein is undesirable, since it is expressed in heart and muscle, but selectivity is difficult to achieve because the FABP-4 sequence is 65% identical with FABP-3 and there are few differences in the lipid-binding pocket. The crystal structure of the phenyl propionate ligand bound to FABP-4 identified residues responsible for selectivity and guided selection of more elaborate analogues from corporate and commercial databases. On the basis of the SARs from these compounds, 12 analogues with various phenyl substituents were synthesized, leading to a 10 μM inhibitor that retained selectivity for FABP-4. These two examples illustrate that fragment-based methods are particularly effective for discovering *novel* leads: although both groups applied very similar approaches to exactly the same protein, the resulting lead series were very different.

7.3. Applications Using a Variation Strategy

In the variation strategy, systematic modifications are made to individual portions of a primary screening hit. In the simplest approach, substructure and similarity-based searches of compound databases are carried out to find analogues of primary screening hits. More sophisticated approaches require synthesis of compounds. Because the starting hits already contain multiple connected functional groups, it is

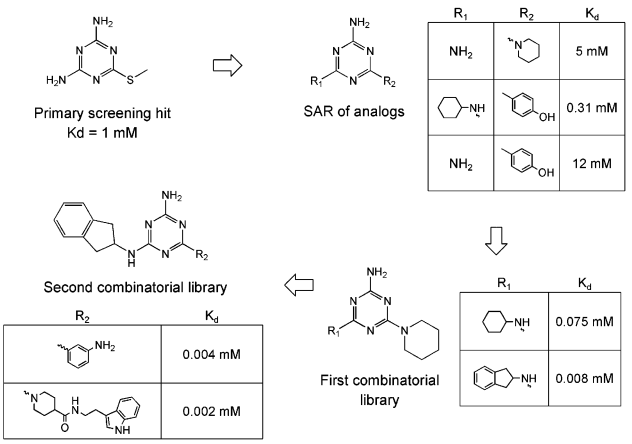


Figure 29. Design of Erm methytransferase inhibitors using a directed combinatorial strategy. The primary screening hit (upper left) contained two sites (R1, R2) for variation that were explored by testing analogues (upper right). The SAR of these analogues indicated that cycloalkyls were preferred at both R1 and R2 (table, upper right). A combinatorial library (lower right) was made with piperidine at R2 and various cycloamines at R1; this confirmed the previous SAR and identified 2-aminoindane as a preferred R1 substituent. A second combinatorial library (lower left) was then made with 2-aminoindane at R1 and various R2 substituents that produced a number of low-micromolar leads.[124]

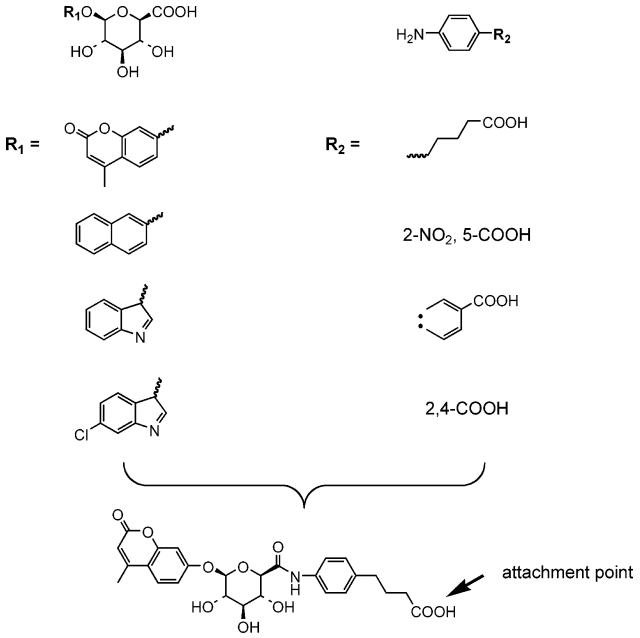


Figure 30. Combinatorial assembly of aryl glycoside PPA ligands. A combinatorial library was constructed from glucuronic acid and arylamine derivatives. The four most avid ligands contained the glucoronic acids (left column) and arylamines (right column) shown; one compound was chosen on the basis of chemical yield and availability of starting materials for attachment to a Sepharose gel (bottom).[126]

much easier to synthesize follow-up compounds than it is with the fragment linking method; the groups are already correctly oriented, and structural information is not required. The variation strategy has been implemented most effectively in two ways: (1) the design of directed combinatorial libraries and (2) fragment optimization. Combinatorial library design requires compounds that contain either a central scaffold with multiple positions for varying substituents (so that the SAR of each can be explored separately) or two scaffolds connected by a linker that is amenable to combinatorial chemistry, allowing the scaffolds to be readily replaced or modified (such as the SHAPES Linking Library[123]).

7.3.1. Directed Combinatorial Libraries

The earliest example of the variation strategy using a combinatorial approach was the discovery of triazine inhibitors of the antibiotic resistance associated protein ErmAM methyltransferase.[124] The unadorned triazine scaffold ($K_D = 1$ mM) was found in a primary ^{15}N HSQC-based screen and was competitive with a naturally occurring inhibitor, *S*-adenosyl-L-homocysteine (SAH). Testing of substituted analogues uncovered SARs that led to synthesis of a disubstituted lead with $IC_{50} = 75$ μM (Figure 29). Using parallel solution phase synthesis, 643 compounds were made to test the SARs at the two substituent positions, leading to several low-micromolar inhibitors of both ErmAM and another Erm family member, ErmC′. NMR and crystal structures of two of these compounds revealed that the triazine occupies the same site as the adenosine of SAH and identified nearby binding pockets that could be accessed to increase potency.

Peptides can be particularly amenable to a combinatorial approach if solubility does not present a problem. Virtual screening of commercially available compounds against the insulin growth factor binding protein IGFBP-5 identified several candidates to target the IGF-1 binding site.[125] One of these, *N*-FMOC-*O*-phosphotyrosine, was shown to bind at the IGF-1 site with $K_D = 1$ mM by ^{15}N HSQC mapping experiments. Since this compound is commonly used as a reagent for peptide synthesis, analogues could be readily purchased. Compounds with larger aromatic groups showed higher affinity, culminating with *N*-FMOC-*N*-BOC-tryptophan ($K_D = 43$ μM). Docking models suggested that the aromatic group mimics the Phe16 side chain from IGF-1 while the FMOC replaces two Leu side chains. This example hints at the potential usefulness of a combinatorial approach for targeting peptide binding sites in other proteins, such as proteases.

A structure-based combinatorial approach was recently used to find ligands for porcine pancreas α-amylase (PPA).[126] In this case, the goal was to create aryl glycoside ligands with suitable sites for attachment to chromatography media. Glucuronic acid (Figure 30) was chosen as a scaffold for its predicted ability to form essential hydrogen bonds in the central, deepest sugar-binding subsite of PPA. This scaffold was functionalized with aryl groups on two sides of the carbohydrate: one site for accessing adjacent sugar-binding subsites and the other for coupling to support. A 234 compound virtual library was constructed from 26 arylamines and 9 glucuronic acids, and on the basis of the docking results, 23 were manually selected for synthesis: 13 as putative ligands and 10 as negative controls. Fourteen of these compounds were soluble enough at 1 mM to screen by STD NMR, and five were hits, including three of

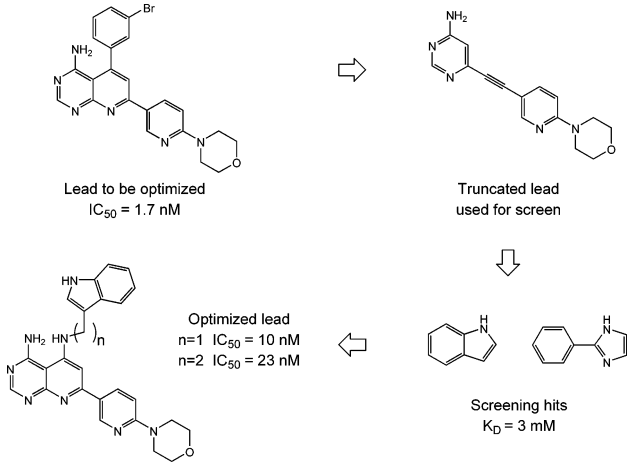


Figure 31. Fragment optimization of adenosine kinase inhibitors. To replace the bromophenyl group of the starting lead compound (upper left), a screen was carried out in the presence of a truncated scaffold (upper right). The indole screening hit (lower right) was attached to the original scaffold using methylene linkers (lower left).[121]

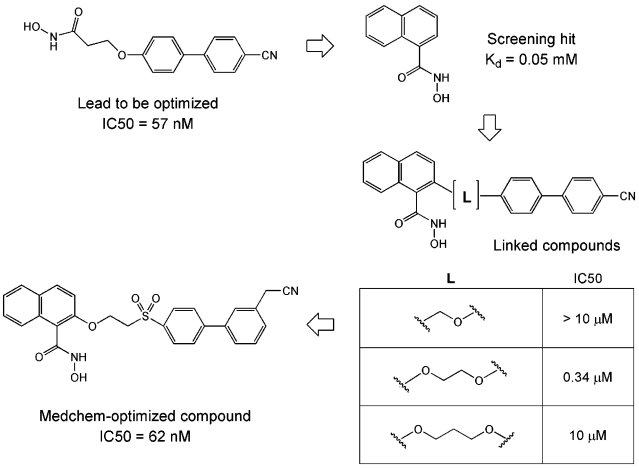


Figure 32. Fragment optimization of stromelysin inhibitors. A targeted screen for zinc ligands to replace the metabolically unstable acetohydroxamate group of the lead compound (upper left) identified naphthylhydroxamate (upper right) as a possible replacement. A series of compounds was synthesized with linkers of varying lengths (lower right), and medicinal chemistry optimization of the four atom linker and biphenyl substituent produced a potent, metabolically stable inhibitor (lower left).[131]

the most promising candidates and one negative control. One of the three best binders was successfully attached to a Sepharose gel. This functionalized matrix selectively retained PPA, and binding was reversed by acarbose, a specific competitor.

7.3.2. Fragment Optimization

Fragment optimization is used to improve existing lead compounds that suffer from unacceptable characteristics, such as low solubility, lack of novelty, poor bioavailability, toxicity, or metabolic instability. A portion of the lead molecule is chosen for selective replacement with fragments that bind to the same subsite on the protein but possess superior properties. Typically, a screen is performed in the presence of saturating amounts of an analogue of the lead compound that lacks the undesirable fragment. After fragments with desirable properties are found that bind to the subsite of interest, they are synthetically linked to the original scaffold with the aid of whatever structural information is available.

The earliest reported example of this method was the optimization of non-nucleosidic inhibitors of adenosine kinase (AK).[127] The existing lead (Figure 31) was highly potent (in vitro $IC_{50} = 1.3$ nM; cellular $IC_{50} = 43$ nM) and showed activity in animal models but suffered from low solubility and poor pharmacokinetic properties.[121] Replacements for the bromophenyl moiety were sought by screening 2000 compounds against AK in the presence of saturating amounts of a truncated lead scaffold containing a pyrimidine core. The original pyridopyrimidine core was not used as a first site ligand because a proton at the linking 5-position was judged to be capable of sterically blocking some fragments from accessing the bromophenyl subsite. Indole and 2-phenylimidazole were very weak ($K_D = 3$ mM) but attractive hits because of their inherently high solubility. ^{15}N HSQC chemical shift perturbation and NMR competition experiments using the original lead confirmed that these fragments both bound in the bromophenyl pocket. Since structural information was unavailable, a series of competition experiments with various sized scaffolds was used to estimate a pyridopyrimidine to indole distance of 3−6 Å. Two compounds were synthesized with two and three methylene linkers, one of which exhibited nanomolar potency in enzymatic and cell assays and activity in an animal model. Although it was not reported that the new compounds had improved pharmacological properties, this example proved that a fragment-based approach could test thousands of diverse bromophenyl replacements very efficiently, allowing chemistry to focus on only those fragments that actually bind to the site of interest.

The biphenyl hydroxamate-based stromelysin inhibitors described in the previous section on the combination strategy were subsequently improved using fragment optimization.[121] Although the NMR-derived lead was potent ($IC_{50} = 57$ nM), it lacked oral bioavailability due to rapid hydrolysis of the acetohydroxamate group. Fourteen putative zinc-binding fragments were screened by ^{15}N HSQC to find more stable replacements. Naphthylhydroxamate (Figure 32) was found to be a more avid zinc ligand than acetohydroxamate (50 μM vs 17 mM), and the bulky naphthyl ring was expected to hinder hydrolysis. The NOE-based structure of naphthylhydroxamate bound to stromelysin revealed that the naphthyl binds in a hydrophobic pocket near the S1 subsite, leaving the S1′ subsite open to the biaryl group. On the basis of the observed interfragment distance of about 5 Å, compounds containing two to five atom methylene/ether linkers were synthesized. The most potent of these was 6-fold weaker than the acetohydroxamate parent compound but exhibited higher oral bioavailability. The lower relative potency of the linked compound presumably originates from perturbations of protein−fragment contacts that override the inherently higher affinity of the naphthylhydroxamate fragment. Traditional medicinal chemistry optimization of the linker and biaryl substituents subse-

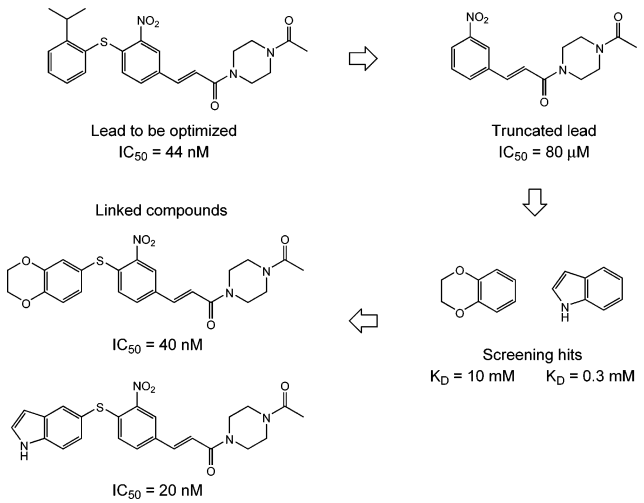


Figure 33. Fragment optimization of LFA-1/ICAM-1 inhibitors. The hydrophobic isopropylphenyl group was removed from the poorly soluble starting lead molecule (upper left) to make a truncated scaffold (upper right). Screening in the presence of the truncated scaffold identified two highly soluble candidates to replace the isopropylphenyl (lower right), and linked compounds with improved PK properties were synthesized (lower left).[121]

quently produced a molecule that was equipotent with the parent lead and had better pharmacokinetic properties. An interesting lesson from this example is that even though the linking problem is usually easier for fragment optimization than fragment linking, the previously described difficulties in predicting the potency of a linked compound on the basis of the potencies of the fragments are still common.

A series of compounds targeting the association between the cell surface receptor LFA-1 (leukocyte function-associated antigen-1) and the intracellular adhesion molecule-1 (ICAM-1) was likewise improved by fragment optimization.[121] NOE studies showed that the isopropylphenyl ring of the existing lead series bound to the highly hydrophobic I-domain allosteric site (IDAS) of LFA-1. Medicinal chemistry optimization of this series had increased the hydrophobicity, producing very potent ($IC_{50} = 44$ nM) leads with poor bioavailability, the latter problem presumably due to low aqueous solubility (Figure 33). The existence of lysine residues in the IDAS suggested that more polar replacements for the isopropylphenyl ring would be well-tolerated. Using a truncated scaffold to block the central hydrophobic pocket of the IDAS, 2500 small compounds (MW < 150 Da) were screened using ^{15}N HSQC to detect hits and measure K_D values, and a number of fragments with millimolar affinities were found. NOESY experiments for two of them, indole and benzodioxane, indicated that they bound to the IDAS in close proximity to the truncated scaffold, with their heteroatom-containing rings pointing toward solvent. Various compounds were made by connecting these fragments to the parent scaffold via a one atom linker at the 4-, 5-, or 6-position, yielding analogues equipotent with the original lead but having improved bioavailability. It is interesting to note that even though the K_D's of the two fragments differed by 30-fold, the potencies of the final linked compounds only differed by 2-fold. This discrepancy was attributed to small changes in the orientation of the bound fragments induced by introduction of the linker atom.

7.4. New Implementations of NMR Screening

7.4.1. RNA Targets

In recent years, a number of NMR screening applications have been reported that use existing experimental techniques but implement the screen in a novel manner. These examples include the screening of nucleic acid targets, multiplexing of targets, and the complementary use of NMR and HTS screening.

Since the early days of NMR screening, it has been recognized that the experimental techniques (particularly ligand-detected methods) were not limited to screening freely diffusing proteins. Using STD-based methods, ligand binding has been detected for proteins immobilized on glass beads,[56] embedded in liposomes,[54] and comprising the coat of a 8.5 MDa virus particle.[53] It is therefore not surprising that ligand-detected screens were attempted using RNA, an important antibacterial and antiviral target. In one of the first studies, the 160-nucleotide P4P6 domain of the Group 1 intron ribozyme from a thermophilic bacterium was screened against the SHAPES library, and the STD, transferred NOESY, and WaterLOGSY detection methods were compared.[58,67] WaterLOGSY was the most sensitive, presumably due to the high solvent accessibility of RNA and its 2-fold lower proton density (which reduces the spin diffusion required for high STD sensitivity). The 23 hits were counterscreened against double stranded RNA to check for nonspecific binding, and only one appeared to bind specifically. Since the SHAPES library was designed using known protein inhibitors, a higher specific hit rate might have been obtained by screening an RNA-targeted library. In a related study, STD was used to study the binding of several known antibiotics to the 50S and 30S ribosomal subunits.[7] Competition experiments demonstrated partial displacement of spectinomycin from the 30S subunit by tetracycline, but the studies were generally hampered by nonspecific binding of the antibiotics to the ribosomes. Finally, NMR chemical shift perturbation experiments have been used to confirm that compounds found in a gel shift assay to inhibit binding of the HIV-1 transactivation response (TAR) to the viral transactivation protein (Tat) were indeed binding specifically to the 5′ bulge region of TAR.[128]

RNA screening recently progressed beyond the proof-of-concept stage. The 29-mer bacterial ribosome 16S decoding region aminoacyl-tRNA site (A-site) RNA was screened by the Abbott group against the same 10 000 compound library customarily used for proteins, giving a hit rate of about 3%.[129] Perturbations of the imino proton resonances of the A-site loop residues that form the known paromomycin-binding pocket were used to detect ligand binding and determine K_D values. Several classes of primary hits were found that had K_D values ranging from 70 μM to 3 mM, and follow-up testing of synthetic analogues produced low-micromolar leads in two classes. The

most potent lead was a 2-aminopyridine derivative with 3 μM affinity; its binding to the paromomycin site in the A-site loop was confirmed using NMR competition experiments. A sparse set of seven ligand−RNA NOE restraints was subsequently used to construct a docking model for the complex. Note that unlike the first RNA screening examples cited, there were no problems with nonspecific binding reported in the latter two examples. The strategy employed was responsible for this difference: specific inhibitors were found when the screens detected functional activity or binding at a known, well-characterized site. It may be that nucleic acids are more prone to nonspecific binding than proteins; if so, it will be preferable to avoid using ligand-based methods unless independent confirmation of the binding site can be obtained.

7.4.2. Multiplexed Targets—Screening Mixtures of Receptors

Although the testing of mixtures of compounds is routine, the testing of multiple proteins in a single sample has only been reported recently. Zartler et al.[130] have recently demonstrated the possibility of screening several proteins at once via a new protocol, RAMPED-UP NMR (rapid analysis and multiplexing of experimentally discriminated uniquely labeled proteins using NMR). This approach calls for a unique isotope-labeling scheme for each protein (e.g. different amino acid selective labeling) to provide a unique spectroscopic signature. HMQC-based screening is then used to determine which, if any, of the proteins are ligated by observing perturbations of the chemical shifts of each subset of amino acids. Because only a subset of the amino acids are labeled, this is not intended as a general screening method per se; compounds that bind at an unlabeled site would be missed. An exception to this rule would be screening for compounds that disrupt protein−protein interactions, since these should perturb a large number of residues located at the complex interface and can be readily observed with sparse labeling.

Zartler et al. demonstrated RAMPED-UP NMR on a mixture of three unrelated ^{15}N selectively labeled proteins including the PTP1B phosphatase (^{15}Nϵ-Trp$_6$), k-RAS (^{15}N-Ile$_{13}$), and GFP (^{15}N-Ala$_8$). The labeling schemes yielded only 6 to 13 labeled ^{15}N−^{1}H spin systems per protein, resulting in simple 2D ^{15}N−^{1}H correlation spectra for the protein mixture. The interactions of these proteins with two known ligands were examined. One ligand was specific for PTP1B, and the other, for k-RAS. Comparisons of 2D ^{15}N−^{1}H correlation spectra of the proteins in isolation or in the mixture showed the same specific binding-induced perturbations upon the addition of ligands. The results thus show the potential for facile, unambiguous multiple protein screening. While a priori structural information is not essential for RAMPED-UP NMR, it can clearly accelerate the search for unique labeling schemes. In the example by Zartler et al., X-ray structures were available for PTP1B and GFP and for a protein highly homologous to k-RAS.

Two major challenges facing this approach are avoiding unintended protein−protein interactions and the identification of buffer conditions that are simultaneously suitable for several proteins. The first challenge can be addressed to some degree by the sequential addition of proteins followed by 2D ^{15}N−^{1}H correlation spectra. The presence of unwanted protein−protein interactions may then be reflected by spectral changes upon the addition of new protein. The second challenge of identifying buffer conditions appears to require "brute-force" searching. In particular, it is not obvious that closely related members of a given gene family will be similarly soluble under identical buffer conditions, since modest amino acid substitutions may confer quite different solubility properties.

The potential attraction of this approach is not merely to increase throughput but also to screen for compounds that are selective or antiselective for a set of similar proteins. An example would be the screening of multiple resistance mutants of a viral or bacterial protein; each mutant would be labeled at the amino acid that differs from wild type, and the screen would seek compounds that bind simultaneously to as many isoforms as possible. Another example would be counterscreening of close homologues in a family of related proteins such as kinases; each homologue would be labeled at the sites that differ from the target of interest, and the goal would be to find highly selective compounds. In some research settings, however, some or all of these strategies may be superfluous, in that comparing "selectivity" among low-affinity ($K_D \sim 10-100$ μM) binders may not be meaningful and that better suited enzyme-based counterscreening assays will be available.

7.4.3. Assay Development and Validation

NMR screening has also been integrated with high-throughput biological or affinity-based screens in order to capitalize upon the strengths of both techniques. For example, HSQC-based NMR screens have been routinely used at Abbott to validate the raw hits obtained from high-throughput enzymatic and cellular assays.[131] False positive rates as high as 97−99% were found for a variety of coupled enzymatic, fluorescence-based, and cellular assays. The mechanisms responsible for these artifactual hits include aggregation of compounds with the target, compound absorbance or emission at wavelengths that interfere with a fluorescent probe, oxidation or covalent modification of the target, and non-target-mediated inhibition in cell assays. Affinity-based mass spectroscopy techniques have gained recent popularity as flexible tools for screening large mixtures of compounds while consuming very small quantities of reagents and requiring minimal assay development. However, these methods are usually unable to determine if the hits are bound in the target active site, at a different site, or nonspecifically at multiple sites. HSQC-based hit mapping has been used to determine the specificity and location of ligand binding sites as a validation of hits from affinity-based screens.[131,132] Preliminary results indicate that, like biologically based screens,

affinity-based screens also produce very high rates of "false positives" or, more accurately, nonspecific association of ligands with the target molecule.[131]

Finally, hits from a diverse screening library can be used to develop an assay in cases where none is available. This concept has been taken a step further in the design of a library of known ligands to be used for screening proteins with unknown functions.[133] A 160 compound library was constructed from amino acids, nucleic acids, carbohydrates, steroids, cofactors, substrates, enzyme inhibitors, and other small organic molecules known to bind to proteins. This library was screened using HSQC-based methods against HI-0033, a conserved protein of unknown function from *Haemophilus influenzae* bacteria, revealing that only adenosine, deoxyadenosine monophosphate (dAMP), and *S*-adenosylhomocysteine (SAH) bind with $K_D < 10\ \mu M$ and that all bind at the same site on the protein. Fluorescein-tagged SAH was found to have a K_D value of 0.7 μM and was used to develop a fluorescence-based competition assay to screen for compounds targeting the nucleoside binding site of HI-0033.

8. Conclusions

In this review, we have made an effort to provide the most contemporary overview possible of NMR screening techniques and their application in pharmaceutical research. Since the inception of NMR screening as a distinct discipline was only 8 years ago, it is difficult to describe any single approach as dated; however, some experimental techniques and follow-up strategies have been more widely adopted than others, and we have made an effort to provide a more rigorous description of these in both the experimental and applications sections of this review. In addition to physical and technical descriptions of how these experiments may be carried out, we have provided, through recent examples, a rich context of applications in which these experiments have been applied, illustrating the potential for success when integrated into a strategy incorporating other highly enabling technologies such as virtual screening, enzymology, and X-ray crystallography. As the field has evolved, and more examples have been presented of how these methods have been used in practice, it has become clear that each target and drug design problem is unique. It is hoped that the descriptions provided and references to the literature therein will allow investigators to implement these methods in their own laboratories, as well as develop drug design strategies uniquely suited to their targets and goals.

9. Acknowledgments

The authors would like to thank Zehan Abdul-Manan for helpful discussions and John Thomson for critical review of the manuscript. We would also like to thank Evan Burrows for assistance with graphics as well as the original artwork presented in Figures 5, 7, 9, and 10, Elizabeth Palermino for assistance with the manuscript, and Daniel Wyss (Schering-Plough) for providing a manuscript prior to publication.

10. References

(1) Shuker, S. B.; Hajduk, P. J.; Meadows, R. P.; Fesik, S. W. *Science* **1996**, *274*, 1531.
(2) Nicholson, J. K.; Connelly, J.; Lindon, J. C.; Holmes, E. *Nat. Rev. Drug Discovery* **2002**, *1*, 153.
(3) Nicholson, J. K.; Lindon, J. C.; Holmes, E. *Xenobiotica* **1999**, *29*, 1181.
(4) Robosky, L. C.; Robertson, D. G.; Baker, J. D.; Rane, S.; Reily, M. D. *Comb. Chem. High Throughput Screening* **2002**, *5*, 651.
(5) Holmes, E.; Antti, H. *Analyst* **2002**, *127*, 1549.
(6) Pellecchia, M.; Sem, D. S.; Wuthrich, K. *Nat. Rev. Drug Discovery* **2002**, *1*, 211.
(7) Peng, J. W. *Prog. Nucl. Magn. Reson. Spectrosc.*, in press.
(8) Stockman, B.; Dalvit, C. *Prog. Nucl. Magn. Reson. Spectrosc.* **2002**, *41*, 187.
(9) van Dongen, M.; Weigelt, J.; Uppenberg, J.; Schultz, J.; Wikstrom, M. *Drug Discovery Today* **2002**, *7*, 471.
(10) Wyss, D. F.; McCoy, M. A.; Senior, M. M. *Curr. Opin. Drug Discovery Dev.* **2002**, *5*, 630.
(11) Peng, J. W.; Lepre, C. A.; Fejzo, J.; Abdul-Manan, N.; Moore, J. M. *Methods Enzymol.* **2001**, *338*, 202.
(12) Mathews, C. K.; Van Holde, K. E. *Biochemistry*, 2nd ed.; Benjamin/Cummings Publishing Co. Inc.: Menlo Park, CA, 1996.
(13) Fersht, A. *Enzyme structure and mechanism*, 2nd ed.; W. H. Freeman: New York, 1985.
(14) Camacho, C. J.; Vajda, S. *Curr. Opin. Struct. Biol.* **2002**, *12*, 36.
(15) Haynie, D. T. *Biological Thermodynamics*; Cambridge University Press: Cambridge, U.K., 2001.
(16) Dalvit, C.; Fasolini, M.; Flocco, M.; Knapp, S.; Pevarello, P.; Veronesi, M. *J. Med. Chem.* **2002**, *45*, 2610.
(17) Dalvit, C.; Flocco, M.; Knapp, S.; Mostardini, M.; Perego, R.; Stockman, B. J.; Veronesi, M.; Varasi, M. *J. Am. Chem. Soc.* **2002**, *124*, 7702.
(18) Siriwardena, A. H.; Tian, F.; Noble, S.; Prestegard, J. H. *Angew. Chem., Int. Ed.* **2002**, *41*, 3454.
(19) Jahnke, W.; Floersheim, P.; Ostermeier, C.; Zhang, X.; Hemmig, R.; Hurth, K.; Uzunov, D. P. *Angew. Chem., Int. Ed. Engl.* **2002**, *41*, 3420.
(20) Cheng, Y.-C.; Prusoff, W. H. *Biochem. Pharmacol.* **1973**, *22*, 3099.
(21) Wang, Z.-X. *FEBS Lett.* **1995**, *360*, 111.
(22) Sigurskjold, B. W. *Anal. Biochem.* **2000**, *277*, 260.
(23) Hahn, E. L.; Maxwell, D. E. *Phys. Rev.* **1952**, *88*, 1070.
(24) McConnell, H. M. *J. Chem. Phys.* **1958**, *28*, 430.
(25) Olejniczak, E. T.; Hajduk, P. J.; Marcotte, P. A.; Nettesheim, D. G.; Meadows, R. P.; Edalji, R.; Holzman, T. F.; Fesik, S. W. *J. Am. Chem. Soc.* **1997**, *119*, 5828.
(26) Hwang, T.-L.; Shaka, A. J. *J. Magn. Reson., Ser. A* **1995**, *112*, 275.
(27) Piotto, M.; Saudek, V.; Sklenar, V. *J. Biomol. NMR* **1992**, *2*, 661.
(28) Liu, M.; Mao, X.-A.; Ye, C.; Huang, H.; Nicholson, J. K.; Lindon, J. C. *J. Magn. Reson.* **1998**, *132*, 125.
(29) Chandler, D. *Introduction to Modern Statistical Mechanics*, 1st ed.; Oxford University Press: New York, 1987.
(30) Deverell, C.; Morgan, R. E.; Strange, J. H. *Mol. Phys.* **1970**, *18*, 553.
(31) Hajduk, P. J.; Olejniczak, E. T.; Fesik, S. W. *J. Am. Chem. Soc.* **1997**, *119*, 12257.
(32) Carr, H. Y.; Purcell, E. M. *Phys. Rev.* **1954**, *94*, 630.
(33) Meiboom, S.; Gill, D. *Rev. Sci. Instrum.* **1958**, *29*, 688.
(34) Luz, Z.; Meiboom, S. *J. Chem. Phys.* **1964**, *40*, 2686.
(35) Davis, D. G.; Perlman, M. E.; London, R. E. *J. Magn. Reson. B* **1994**, *104*, 266.
(36) Freeman, R.; Hill, H. D. W. *J. Chem. Phys.* **1971**, *55*, 1985.
(37) van Dongen, M.; Uppenberg, J.; Svensson, B.; Lundbäck, T.; Åkerud, T.; Wikström, M.; Schultz, J. *J. Am. Chem. Soc.* **2002**, *124*, 11874.
(38) Chen, A.; Shapiro, M. J. *J. Am. Chem. Soc.* **2000**, *122*, 414.
(39) Jahnke, W.; Rudisser, S.; Zurini, M. *J. Am. Chem. Soc.* **2001**, *123*, 3149.
(40) Jahnke, W.; Perez, L. B.; Paris, G.; Strauss, A.; Fendrich, G.; Nalin, C. M. *J. Am. Chem. Soc.* **2000**, *122*, 7394.
(41) Peng, J. W. *J. Magn. Reson.* **2001**, *153*, 32.
(42) Gerig, J. T. *Prog. Nucl. Magn. Reson. Spectrosc.* **1994**, *26*, 293.
(43) London, R. E.; Gabel, S. A. *J. Am. Chem. Soc.* **1994**, *116*, 2562.
(44) London, R. E.; Gabel, S. A. *J. Am. Chem. Soc.* **1994**, *116*, 2570.
(45) Luck, L. A.; Vance, J. E.; O'Connell, T. M.; London, R. E. *J. Biomol. NMR* **1996**, *7*, 261.
(46) Dalvit, C.; Flocco, M.; Veronesi, M.; Stockman, B. J. *Comb. Chem. High Throughput Screening* **2002**, *5*, 605.
(47) Dalvit, C.; Fagerness, P. E.; Hadden, D. T.; Sarver, R. W.; Stockman, B. J. *J. Am. Chem. Soc.* **2003**, *125*, 7696.
(48) Mayer, M.; Meyer, B. *Angew. Chem., Int. Ed. Engl.* **1999**, *38*, 1784.
(49) McCoy, M. A.; Mueller, L. *J. Am. Chem. Soc.* **1992**, *114*, 2108.
(50) Mayer, M.; Meyer, B. *J. Am. Chem. Soc.* **2001**, *123*, 6108.

(51) Henri, V. *Lois generales de l'action des diastases*; Hermann: Paris, 1903.
(52) Michaelis, M.; Menten, M. *Biochem. Z.* **1913**, *19*, 333.
(53) Benie, A. J.; Moser, R.; Bäumi, E.; Blaas, D.; Peters, T. *J. Am. Chem. Soc.* **2003**, *125*, 14.
(54) Meinecke, R.; Meyer, B. *J. Am. Chem. Soc.* **2001**, *44*, 3059.
(55) Mayer, M.; James, T. L. *J. Am. Chem. Soc.* **2002**, *124*, 13376.
(56) Klein, J.; Meinecke, R.; Mayer, M.; Meyer, B. *J. Am. Chem. Soc.* **1999**, *121*, 5336.
(57) Jayalakshmi, V.; Rama Krishna, N. *J. Magn. Reson.* **2002**, *155*, 106.
(58) Johnson, E. C.; Feher, V. A.; Peng, J. W.; Moore, J. M.; Williamson, J. R. *J. Am. Chem. Soc.* **2003**, *125*, 15724.
(59) Dalvit, C.; Pevarello, P.; Tato, M.; Veronesi, M.; Vulpetti, A.; Sundstrom, M. *J. Biomol. NMR* **2000**, *18*, 65.
(60) Dalvit, C.; Fogliatto, G.; Stewart, A.; Veronesi, M.; Stockman, B. *J. Biomol. NMR* **2001**, *21*, 349.
(61) Mori, S.; Abeygunawardana, C.; van Zijl, P. C.; Berg, J. M. *J. Magn. Reson. B* **1996**, *110*, 96.
(62) Mori, S.; Berg, J. M.; van Zijl, P. C. *J. Biomol. NMR* **1996**, *7*, 77.
(63) Dalvit, C. *J. Magn. Reson. B* **1996**, *112*, 282.
(64) Geen, H.; Freeman, R. *J. Magn. Reson.* **1991**, *93*, 93.
(65) Grzesiek, S.; Bax, A. *J. Biomol. NMR* **1993**, *3*, 627.
(66) Stott, K.; Keeler, J.; Van, Q. N.; Shaka, A. J. *J. Magn. Reson.* **1997**, *125*, 302.
(67) Lepre, C. A.; Peng, J.; Fejzo, J.; Abdul-Manan, N.; Pocas, J.; Jacobs, M.; Xie, X.; Moore, J. M. *Comb. Chem. High Throughput Screening* **2002**, *5*, 583.
(68) Henrichsen, D.; Ernst, B.; Magnani, J. L.; Wang, W.-T.; Meyer, B.; Peters, T. *Angew. Chem., Int. Ed. Engl.* **1999**, *38*, 98.
(69) Fejzo, J.; Lepre, C. A.; Peng, J. W.; Bemis, G. W.; Ajay; Murcko, M. A.; Moore, J. M. *Chem. Biol.* **1999**, *6*, 755.
(70) Mayer, M.; Meyer, B. *J. Med. Chem.* **2000**, *43*, 2093.
(71) Sigurskjold, B. W.; Berland, C. R.; Svensson, B. *Biochemistry* **1994**, *33*, 10191.
(72) Dalvit, C.; Ardini, E.; Flocco, M.; Fogliatto, G. P.; Mongelli, N.; Veronesi, M. *J. Am. Chem. Soc.* **2003**, *125*, 14620.
(73) Murali, N.; Jarori, G. K.; Landy, S. B.; Rao, B. D. *Biochemistry* **1993**, *32*, 12941.
(74) Pellecchia, M.; Meininger, D.; Dong, Q.; Chang, E.; Jack, R.; Sem, D. S. *J. Biomol. NMR* **2002**, *22*, 165.
(75) Wagner, G.; Wüthrich, K. *J. Magn. Reson.* **1979**, *33*, 675.
(76) Moseley, H. N. B.; Curto, E. V.; Rama Krishna, N. *J. Magn. Reson., Ser. B* **1995**, *108*, 243.
(77) Tanner, J. E. *J. Chem. Phys.* **1970**, *52*, 2523.
(78) Gibbs, S. J.; Johnson, C. S., Jr. *J. Magn. Reson.* **1991**, *93*, 395.
(79) Dötsch, V.; Wider, G.; Wüthrich, K. *J. Magn. Reson. A* **1994**, *109*, 263.
(80) Chen, A.; Johnson, C. S., Jr.; Lin, M.; Shapiro, M. J. *J. Am. Chem. Soc.* **1998**, *120*, 9094.
(81) Chen, A.; Shapiro, M. J. *J. Am. Chem. Soc.* **1999**, *121*, 5338.
(82) Yan, J.; Kline, A. D.; Mo, H.; Zartler, E. R.; Shapiro, M. J. *J. Am. Chem. Soc.* **2002**, *124*, 9984.
(83) Ross, A.; Senn, H. *Drug Discovery Today* **2001**, *6*, 583.
(84) Ross, A.; Schlotterbeck, G.; Klaus, W.; Senn, H. *J. Biomol. NMR* **2000**, *16*, 139.
(85) Hajduk, P. J.; Augeri, D. J.; Mack, J.; Mendoza, R.; Yang, J.; Betz, S. F.; Fesik, S. W. *J. Am. Chem. Soc.* **2000**, *122*, 7898.
(86) Hajduk, P. J.; Gerfin, T.; Boehlen, J.-M.; Haberli, M.; Marek, D.; Fesik, S. W. *J. Med. Chem.* **1999**, *42*, 2525.
(87) Moore, J. M. *Curr. Opin. Biotechnol.* **1999**, *10*, 54.
(88) Hajduk, P. J.; Meadows, R. P.; Fesik, S. J. *Q. Rev. Biophys.* **1999**, *32*, 211.
(89) Roberts, G. C. *Drug Discovery Today* **2000**, *5*, 230.
(90) Weigelt, J.; van Dongen, M.; Uppenberg, J.; Schultz, J.; Wikström, M. *J. Am. Chem. Soc.* **2002**, *124*, 2446.
(91) McCoy, M. A.; Wyss, D. F. *J. Am. Chem. Soc.* **2002**, *124*, 11758.
(92) Hajduk, P. J.; Mack, J. C.; Olejniczak, E. T.; Park, C.; Dandliker, P. J.; Beutel, B. A. *J. Am. Chem. Soc.* **2004**, *126*, 2390.
(93) Moyna, G.; Zauhar, R. J.; Williams, H. J.; Nachman, R. J.; Scott, A. I. *J. Chem. Inf. Comput. Sci.* **1998**, *38*, 702.
(94) Pople, J. A. *J. Chem. Phys.* **1956**, *24*, 1111.
(95) Teague, S. J.; Davis, A. M.; Leeson, P. D.; Oprea, T. *Angew. Chem., Int. Ed. Engl.* **1999**, *38*, 3743.
(96) Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeny, P. J. *Adv. Drug Delivery Rev.* **1997**, *23*, 3.
(97) Lepre, C. A. In *BioNMR Techniques in Drug Research*; Zerbe, O., Ed.; Wiley-VCH: Weinheim, Germany, 2002.
(98) Medek, A.; Hajduk, P. J.; Mack, J.; Fesik, S. W. *J. Am. Chem. Soc.* **2000**, *122*, 1241.
(99) McCoy, M. A.; Wyss, D. F. *J. Biomol. NMR* **2000**, *18*, 189.
(100) Klaus, W.; Senn, H. In *BioNMR in Drug Research*; Zerbe, O., Ed.; Wiley-VCH: Weinheim, Germany, 2003.
(101) Li, D.; DeRose, E. F.; London, R. E. *J. Biomol. NMR* **1999**, *15*, 71.
(102) Lugovskoy, A. A.; Degterev, A. I.; Fahmy, A. F.; Zhou, P.; Gross, J. D.; Yuan, J.; Wagner, G. *J. Am. Chem. Soc.* **2002**, *124*, 1234.
(103) Liepinsh, E.; Otting, G. *Nat. Biotechnol.* **1997**, *15*, 264.
(104) Fejzo, J.; Lepre, C.; Xie, X. *Curr. Top. Med. Chem.* **2002**, *2*, 1349.
(105) Boehm, H. J.; Boehringer, M.; Bur, D.; Gmuender, H.; Huber, W.; Klaus, W.; Kostrewa, D.; Kuehne, H.; Luebbers, T.; Meunier-Keller, N.; Mueller, F. *J. Med. Chem.* **2000**, *43*, 2664.
(106) Hajduk, P. J.; Boyd, S.; Nettesheim, D.; Nienaber, V.; Severin, J.; Smith, R.; Davidson, D.; Rockway, T.; Fesik, S. W. *J. Med. Chem.* **2000**, *43*, 3862.
(107) Nienaber, V. L.; Richardson, P. L.; Klighofer, V.; Bouska, J. J.; Giranda, V. L.; Greer, J. *Nat. Biotechnol.* **2000**, *18*, 1105.
(108) Carr, R.; Jhoti, H. *Drug Discovery Today* **2002**, *7*, 522.
(109) Lesuisse, D.; Lange, G.; Deprez, P.; Benard, D.; Schoot, B.; Delettre, G.; Marquette, J. P.; Broto, P.; Jean-Baptiste, V.; Bichet, P.; Sarubbi, E.; Mandine, E. *J. Med. Chem.* **2002**, *45*, 2379.
(110) Jahnke, W. *ChemBioChem* **2002**, *3*, 167.
(111) Hajduk, P. J.; Sheppard, G.; Nettesheim, D. G.; Olejniczak, E. T.; Shuker, S. B.; Meadows, R. P.; Steinman, D. H.; Carrera, G. M., Jr.; Marcotte, P. A.; Severin, J.; Walter, K.; Smith, H.; Gubbins, E.; Simmer, R.; Holzman, T. F.; Morgan, D. W.; Davidsen, S. K.; Summers, J. B.; Fesik, S. W. *J. Am. Chem. Soc.* **1997**, *119*, 5818.
(112) Jahnke, W.; Florsheimer, A.; Blommers, M. J.; Paris, C. G.; Heim, J.; Nalin, C. M.; Perez, L. B. *Curr. Top. Med. Chem.* **2003**, *3*, 69.
(113) Hajduk, P. J.; Dinges, J.; Miknis, G. F.; Merlock, M.; Middleton, T.; Kempf, D. J.; Egan, D. A.; Walter, K. A.; Robins, T. S.; Shuker, S. B.; Holzman, T. F.; Fesik, S. W. *J. Med. Chem.* **1997**, *40*, 3144.
(114) Niimi, T.; Orita, M.; Okazawa-Igarashi, M.; Sakashita, H.; Kikuchi, K.; Ball, E.; Ichikawa, A.; Yamagiwa, Y.; Sakamoto, S.; Tanaka, A.; Tsukamoto, S.; Fujita, S.; Tatsuta, K.; Maeda, Y.; Chikauchi, K. *J. Med. Chem.* **2001**, *44*, 4737.
(115) Liu, G.; Szczepankiewicz, B. G.; Pei, Z.; Janowick, D. A.; Xin, Z.; Hajduk, P. J.; Abad-Zapatero, C.; Liang, H.; Hutchins, C. W.; Fesik, S. W.; Ballaron, S. J.; Stashko, M. A.; Lubben, T.; Mika, A. K.; Zinker, B. A.; Trevillyan, J. M.; Jirousek, M. R. *J. Med. Chem.* **2003**, *46*, 2093.
(116) Liu, G.; Xin, Z.; Liang, H.; Abad-Zapatero, C.; Hajduk, P. J.; Janowick, D. A.; Szczepankiewicz, B. G.; Pei, Z.; Hutchins, C. W.; Ballaron, S. J.; Stashko, M. A.; Lubben, T. H.; Berg, C. E.; Rondinone, C. M.; Trevillyan, J. M.; Jirousek, M. R. *J. Med. Chem.* **2003**, *46*, 3437.
(117) Liu, G.; Xin, Z.; Pei, Z.; Hajduk, P. J.; Abad-Zapatero, C.; Hutchins, C. W.; Zhao, H.; Lubben, T. H.; Ballaron, S. J.; Haasch, D. L.; Kaszubska, W.; Rondinone, C. M.; Trevillyan, J. M.; Jirousek, M. R. *J. Med. Chem.* **2003**, *46*, 4232.
(118) Szczepankiewicz, B. G.; Liu, G.; Hajduk, P. J.; Abad-Zapatero, C.; Pei, Z.; Xin, Z.; Lubben, T. H.; Trevillyan, J. M.; Stashko, M. A.; Ballaron, S. J.; Liang, H.; Huang, F.; Hutchins, C. W.; Fesik, S. W.; Jirousek, M. R. *J. Am. Chem. Soc.* **2003**, *125*, 4087.
(119) Hajduk, P. J.; Zhou, M.-M.; Fesik, S. W. *Biorg. Med. Chem. Lett.* **1999**, *9*, 2403.
(120) Wyss, D. F.; Arasappan, A.; Senior, M. M.; Wang, Y.-S.; Beyer, B. M.; Njoroge, F. G.; McCoy, M. A. *J. Med. Chem.*, in press.
(121) Huth, J. R.; Sun, C. *Comb. Chem. High Throughput Screening* **2002**, *5*, 631.
(122) Rudisser, S.; Jahnke, W. *Comb. Chem. High Throughput Screening* **2002**, *5*, 591.
(123) Lepre, C. A. *Drug. Discovery Today* **2001**, *6*, 133.
(124) Hajduk, P. J.; Dinges, J.; Schkeryantz, J. M.; Janowick, D.; Kaminski, M.; Tufano, M.; Augeri, D. J.; Petros, A.; Nienaber, V.; Zhong, P.; Hammond, R.; Coen, M.; Beutel, B.; Katz, L.; Fesik, S. W. *J. Med. Chem.* **1999**, *42*, 3852.
(125) Kamionka, M.; Rehm, T.; Beisel, H. G.; Lang, K.; Engh, R. A.; Holak, T. A. *J. Med. Chem.* **2002**, *45*, 5655.
(126) Baumann, H.; Ohrman, S.; Shinohara, Y.; Ersoy, O.; Choudhury, D.; Axen, A.; Tedebark, U.; Carredano, E. *Protein Sci.* **2003**, *12*, 784.
(127) Hajduk, P. J.; Gomtsyan, A.; Didomenico, S.; Cowart, M.; Bayburt, E. K.; Solomon, L.; Severin, J.; Smith, R.; Walter, K.; Holzman, T. F.; Stewart, A.; McGaraughty, S.; Jarvis, M. F.; Kowaluk, E. A.; Fesik, S. J. *J. Med. Chem.* **2000**, *43*, 4781.
(128) Lind, K. E.; Du, Z.; Fujinaga, K.; Peterlin, B. M.; James, T. L. *Chem. Biol.* **2002**, *9*, 185.
(129) Yu, L.; Oost, T. K.; Schkeryantz, J. M.; Yang, J.; Janowick, D.; Fesik, S. W. *J. Am. Chem. Soc.* **2003**, *125*, 4444.
(130) Zartler, E. R.; Hanson, J.; Jones, B. E.; Kline, A. D.; Martin, G.; Mo, H.; Shapiro, M. J.; Wang, R.; Wu, H.; Yan, J. *J. Am. Chem. Soc.* **2003**, *125*, 10941.
(131) Hajduk, P. J.; Burns, D. J. *Comb. Chem. High Throughput Screening* **2002**, *5*, 613.
(132) Moy, F. J.; Haraki, K.; Mobilio, D.; Walker, G.; Powers, R.; Tabei, K.; Tong, H.; Siegel, M. M. *Anal. Chem.* **2001**, *73*, 571.
(133) Hajduk, P. J.; Betz, S. F.; Mack, J.; Ruan, X.; Towne, D. L.; Lerner, C. G.; Beutel, B. A.; Fesik, S. W. *J. Biomol. Screening* **2002**, *7*, 429.

CR030409H