# NMR experiments for lead generation in drug discovery

Jeffrey W. Peng*, Jonathan Moore, Norzehan Abdul-Manan

*Corresponding author. Address: Department of Chemistry and Biochemistry, 251 Nieuwland Science Hall, University of Notre Dame, Notre Dame, IN 46556, USA. Tel.: +1-574-631-2983; fax: +1-574-631-6652.

E-mail address: jpeng@nd.edu (J.W. Peng).

**Contents**

1. Introduction ............................................................................. 2 25
2. Binding equilibria ........................................................................ 2 27
   2.1. Simple one-site binding .......................................................... 2 27
   2.2. Relationship to enzyme kinetic parameters .............................. 2 28
   2.3. Competitive binding equilibria ............................................... 2 29
3. Accounting for binding-induced chemical exchange ..................... 2 30
   3.1. Modified Bloch equations ...................................................... 2 30
   3.2. Fast-exchange approximation ............................................... 2 33
4. Ligand-based versus receptor-based screening ............................ 2 33
5. Ligand-based NMR screening methods ....................................... 2 34
   5.1. Transverse relaxation rates ................................................... 2 35
   5.2. Longitudinal relaxation rate .................................................. 2 37
   5.3. Paramagnetic relaxation enhancements ................................. 2 38
   5.4. 19F relaxation ................................................................. 2 38
   5.5. Saturation transfer difference methods .................................. 2 39
   5.6. WaterLOGSY ................................................................. 2 43
   5.7. Exchange-transferred NOE ................................................... 2 45
6. Competition binding studies .................................................... 2 46
7. Beyond binding: epitope mapping via ligand-based approaches ....... 2 48
   8.1. Group epitope mapping ....................................................... 2 48
   8.2. Diffusion-based epitope mapping ......................................... 2 50
9. Receptor-based approaches ....................................................... 2 51
   9.1. Selective active-site isotope labeling ..................................... 2 51
   9.2. Screening mixtures of receptors .......................................... 2 51
   9.3. Ligand localization from J-surface analysis of chemical shift perturbations ........................................ 2 52
10. Specific example of NMR screening: targeting the ribosome ........... 2 53
11. Conclusions and challenges remaining ...................................... 2 54
Acknowledgements ..................................................................... 2 55
References ............................................................................... 2 55

**Keywords:** Structure–activity; Atomic resolution; Drug discovery; Protein-ligand interactions; NMR screening

---

## 1. Introduction

Three critical stages of drug discovery are lead generation, lead optimization, and preclinical development. The first stage identifies starting compounds that demonstrate specific activity against a therapeutic target as well
as dynamic structure–activity relationships (SAR). The second stage optimizes in vitro potency through iterative cycles of design-synthesize-assay, often while being guided by three-dimensional (3D) molecular structures at atomic resolution. It is now well established that the 3D structures of receptor–ligand complexes can accelerate the process of driving a compound to nM (\(10^{-9}\) M) specific potency. The third stage optimizes the physicochemical properties dictating cellular and in vivo potency and efficacy, as well as oral bioavailability. These properties include solubility and absorption, distribution, metabolic stability, excretion, and toxicity (ADMET).

Initially (ca. first half of the 1990s) many NMR research groups assumed that biomolecular NMR would have a decisive impact on lead optimization (second stage) through its ability to determine protein solution structures. It is now clear that while protein structure remains an important aspect of drug discovery, the initial predictions concerning the impact of NMR were overly optimistic. The optimism stemmed largely from the understandable enthusiasm among the protein NMR community over the methodological advances that were extending the molecular weight limit for sequential resonance assignment and solution structure determination beyond \(\sim 10,000\) Da. These events diverted attention away from the actual bottlenecks and compressed time-lines of drug discovery, as well as the simultaneous advances made by other biophysical methods, such as X-ray diffraction and mass spectrometry.

The chief goal of pharmaceutical research is the rapid and profitable design of drugs specific for therapeutic targets in vivo, not the design of in vitro inhibitors specific for NMR-friendly proteins. Drug design is a vastly more complex problem than inhibitor design, and pharmaceutical laboratories must be especially rigorous when choosing the appropriate biophysical methods with regards to speed, sensitivity, and material requirements. As such, pharmaceutical research acts as a ‘proving ground’ for competing biophysical methods. While the speed of NMR solution structure determination continues to increase, it still remains a relatively slow, detail-oriented technique with high demands on material and expertise. Additionally, a consensus on how the quality of NMR structures should be quantified is still lacking, thus making them difficult to use by other disciplines. Finally, since NMR has still not been able to leverage solution structures in any significant way to produce better drugs, these features have, at the time of this writing, nearly eliminated structural bio-NMR from the pharmaceutical arena. In light of this, an instructive review discussing the consequences of high promises from high-maintenance technologies in drug discovery has been written recently by Gershell and Atkins [1]. We note, however, that this marginalization of structural NMR in the pharmaceutical industry may be linked to the choice of targets traditionally pursued therein—large (typically \(>30,000\) Da) soluble enzymes for which structure determination by NMR is usually too time-consuming. Recent studies of integral membrane proteins suggest that the impact of structural NMR in pharmaceutical research could revive, should the expression and purification of these important and under-exploited proteins become more facile [2–4].

To contribute effectively to pharmaceutical research, biological NMR has re-invented itself away from the narrow structural focus of the early 1990s. This has been achieved by focusing on the first and third stages of drug discovery: NMR screening to accelerate lead generation, and metabonomics to address critical bottlenecks in drug development. This review focuses on the application of NMR for lead generation. Excellent discussions of the principles and techniques of metabonomics can be found in a number of authoritative reviews [5,6].

In 1996 Abbott Laboratories unveiled an NMR-driven method for lead discovery based on perturbations of protein heteronuclear NMR spectra upon the addition of compound mixtures (‘SAR-by-NMR’) [7]. This transformed the outlook of many protein NMR spectroscopists, who realized that NMR’s ability to act as a binding assay in lead generation could rejuvenate its impact in pharmaceutical research. This induced a proliferation of NMR publications proposing new screening methods. The past six years have seen the emergence of a short list of preferred methods, which are discussed in this review.

NMR may assist in either primary or secondary screening applications. Primary screening refers to the initial screen of a therapeutic target using a large compound library. Such binders are typically ‘weak’ by enzymological standards, with equilibrium dissociation constants \(K_D\) in the range \(100\) \(\mu\)M < \(K_D\) < 1 mM. Usually, the objective of primary NMR screening is to identify a set of weakly binding molecular scaffolds or fragments (‘hits’), whose subsequent elaboration or linking can yield promising lead candidates. The NMR hits can also assist in silico screening as well as guide the design of chemical databases and/or combi-chem libraries [8]. NMR primary screening strengthens its impact when performed prior to high-throughput enzymological screening (HTS) and structure determination. If enzyme assays are simply not feasible, then, clearly, NMR screening becomes of higher importance.

Secondary screening usually refers to more detail-oriented ‘follow-up’ experiments on a smaller number of compounds with intermediate binding affinities. Such compounds may have been flagged as hits by other biophysical methods. Secondary screening by NMR can validate these hits to reduce false positives and provide ‘rescue’ operations for compounds that demonstrate good in vitro potency, yet, poor pharmacokinetic properties. It can also supply supplementary information not available from other techniques, such as bound ligand mobility [9].

The value of the NMR screen depends critically on the quality and scope of the underlying experimental methods and the library design. This review will focus on the NMR methods. Library design will not be discussed here.
Instead, we refer readers to lucid reviews in the literature (see e.g. Lepre, 2001 [10,11]).

The organization of this review is as follows. We first review some elementary principles of binding equilibria underlying the logic of the NMR screening experiments. We then discuss how the chemical exchange associated with the binding equilibria can modulate NMR parameters, thereby making these parameters probes of ligand binding that can drive screening. We pay particular attention to the typical approximations and assumptions made by NMR screeners about the kinetics of the binding exchange. We then discuss in detail some of the more popular screening experiments that have emerged in the last ~6 years. We finish with some discussions of ambitious new applications of NMR screening, and the future outlook of its applicability in pharmaceutical research.

2. Binding equilibria

Before discussing the various NMR screening experiments, it is useful to review some basic facts concerning binding equilibria. More in-depth discussions can be found in any introductory biochemistry text (see e.g. Mathews and van Holde, 1996 [12]) By presenting the simple formulae below, we not only organize results for experimental control variables, but also distinguish well known facts of biology and chemistry from new developments in NMR technique.

2.1. Simple one-site binding

The binding process underlying the design of most NMR screening experiments can be described by the elementary step:

\[ [E] + [L] \rightarrow_{k_{on}}^{k_{off}} [EL] \]  

Eq. (1) represents a dynamic equilibrium involving three species: the free receptor, [E], the free compound [L], and the receptor–ligand complex [EL]. The unimolecular rate constant \( k_{off} \) is inversely proportional to the mean lifetime, \( \tau_B \), of the receptor–ligand complex. The bimolecular rate constant \( k_{on} \) is a measure of the probability of a productive encounter between free receptor and ligand. The NMR literature often fixes \( k_{on} \) at a constant diffusion-limited value. Curiously, the ‘constant’ value invoked seems to vary between \( 1 \times 10^3 \) to \( 1 \times 10^9 \) M\(^{-1}\) s\(^{-1}\). Use of a ‘universal’ constant for diffusion-limited encounters is of course a gross approximation that does not account for the potential complexity of intermolecular forces that may amplify or attenuate the encounter frequency [13].

The binding affinity can be quantified 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

\[ P_B^E = \frac{[L]}{[L] + K_D}. \]  

\( P_B^E \) is the fractional occupation of the receptor binding site by ligand ‘L’, and the complementary free receptor fraction is \( P_B^{EF} \), where \( P_B^E + P_B^{EF} = 1 \). Note that \( P_B^E \) 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; i.e. 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 \( P_B^E = 1.0 \). Ligands of weaker affinity have larger \( K_D \) and thus require the addition of more ligand to saturate the receptor binding site. Note that Eq. (2) is essentially the Langmuir binding isotherm which describes the general behavior for simple saturation binding processes [14]. This will become evident in subsequent sections as we will repeatedly revisit Eq. (2) in different guises.

It is desirable to express the receptor–ligand concentration \([EL]\) in terms of the experimental control variables, which include the total receptor concentration \( E_T \) and the total ligand concentration \( L_T \). Straight substitution of the constraints \( E_T = [E] + [L_T] = [EL] + [L] \) into \( K_D = [E][L]/[EL] \) yields

\[ [EL] = \frac{1}{2}(E_T + L_T + K_D) - \frac{1}{2}\sqrt{(E_T + L_T + K_D)^2 - 4E_T L_T}. \]

Fig. 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 \( K_D \).
Recognizing that

Thus, the initial rate

These 'dose–response' curves verify that the receptor is

The bound ligand fraction \( P_B = [EL]/LT \) lies in the range

The upper limit occurs when the receptor binding site becomes saturated; i.e. when \([EL] = ET\). Fig. 1 re-emphasizes that ligands having weaker binding affinity (larger \( K_D \)) require larger \( L_T \) to saturate the receptor binding site. By judicious control of \( e = L_T/ET \), one can establish a ceiling value on \( K_D \) for which \( P_B \) remains significant. Reducing \( ET \) reduces the \( P_B \) of the weaker binders first. This fact allows one to 'tune' the detection threshold of screening experiments to select for tighter binders.

2.2. Relationship to enzyme kinetic parameters

Numerous pharmaceutical targets are soluble enzymes. It is therefore useful to review the parameters of elementary enzyme kinetics and relate them to the binding parameters encountered in NMR screening.

For a simple one-substrate, one-product catalytic mechanism, we have the series of elementary steps

\[
[E] + [S] \xrightleftharpoons[k_{\text{on}}]{k_{\text{off}}} [ES] \rightarrow [E] + [P]
\]  

Eq. (4) represents a mechanism involving four species:

the free enzyme receptor \([E]\), the free substrate \([S]\), and the enzyme-substrate complex \([ES]\), and the product \([P]\). The enzymatic reaction rate (sometimes called ‘velocity’) is \( v = d[P]/dt \). One usually focuses on the initial rate wherein \([P] \approx 0 \). In this limit, we have

\[
v_0 = \left. \frac{d[P]}{dt} \right|_{[P]=0} = k_{\text{cat}}[ES]. \tag{5a}
\]

Recognizing that \([ES] = ETP_{ES} \), where \( ET \) is the total enzyme and \( P_{ES} \) is the fraction of bound enzyme, we can recast Eq. (5a) as

\[
v_0 = k_{\text{cat}}ETP_{ES} = V_{\text{max}}P_{ES}. \tag{5b}
\]

Thus, the initial rate \( v_0 \) is proportional to the bound enzyme fraction \( P_{ES} \); it achieves its maximum value \( V_{\text{max}} = k_{\text{cat}}ET \) when \( P_{ES} = 1.0 \).

Typically, one expresses \( v_0 \) in terms of \([S]\) and \( ET \) using the approach of steady-state kinetics. In this approach, one assumes that the total substrate \( S_T \) is much greater than the total enzyme \( ET \) (a reasonable assumption in enzyme assays). As a consequence, \([ES]\) will be approximately constant for the majority of the reaction time course. Thus, we make the steady-state approximation that \( d[ES]/dt = k_{\text{on}}[E][S] - (k_{\text{off}} + k_{\text{cat}})[ES] \approx 0 \). Since \( ET = [E] + [ES] \), the bound enzyme fraction \( P_{ES} = [ES]/ET \) becomes

\[
P_{ES} = \frac{[S]}{[S] + KM} \tag{6}
\]

where \( KM \) is the Michaelis constant \( KM = (k_{\text{on}}^S + k_{\text{cat}})/k_{\text{off}}^S \).

Note that \( P_{ES} \) has the same hyperbolic form as the Langmuir \( P_B \) given in Eq. (2). Plugging the above expression for \( P_{ES} \) into Eq. (5b) produces the well known Henri-Michaelis-Menten hyperbolic form \([15,16]\)

\[
v_0 = \frac{V_{\text{max}}[S]}{[S] + KM}. \tag{7}
\]

Although \( KM = [E][S]/[ES] \), it is generally not equivalent to the equilibrium dissociation constant for the ‘E’–‘S’ interaction, \( K_S \). Instead, \( KM = (k_{\text{on}}^S + k_{\text{cat}})/k_{\text{off}}^S \) while \( K_S = k_{\text{off}}^S/k_{\text{cat}}^S \). Only in the special case that \( k_{\text{on}}^S \gg k_{\text{cat}}^S \), does \( KM \) reduce to \( K_S \). Equivalently, we must recall that \([E] \) and \([ES]\) represent steady-state concentrations and not necessarily equilibrium concentrations. Analogous to \([L] \) and \( KD \) in Eq. (2), increasing \([S]\) increases \( v_0 \), although with diminishing returns. When \([S] \ll KM \), \( v_0 \) is directly proportional to \([S]\). When \([S] = KM \), then \( v_0 = V_{\text{max}}/2 \). Finally, when \([S] \gg KM \), \( v_0 \) saturates at its upper limit of \( V_{\text{max}} \). Thus, enzyme-substrate mechanisms characterized by larger \( KM \) require a greater amount of substrate to reach \( V_{\text{max}} \).

If we now introduce a competitive inhibitor ‘C’, then we obtain the competitive inhibitory mechanism

\[
[EC] \xrightleftharpoons[k_{\text{cat}}^l]{k_{\text{on}}^l} [E] + [C] \xrightleftharpoons[k_{\text{off}}^l]{k_{\text{off}}^l} [ES] \xrightarrow{k_{\text{cat}}^l} [E] + [P]. \tag{8}
\]

where \([C]\) and \([EC]\) represent the free inhibitor and the enzyme-inhibitor complex, respectively. In the context of competitive inhibition, NMR screening tries to identify the chemical building blocks of ‘C’ through equilibrium binding measurements of \([E]\) in the presence of compound mixtures. We avoid the symbol ‘I’ to denote the inhibitor since the NMR screening literature has unfortunately already chosen ‘I’ to indicate another non-substrate ligand that competes with ‘L’, the latter of which is generally another NMR screening hit and not an enzyme substrate (vide infra). If we assume that \([C], [E], \) and \([EC]\) are at equilibrium, then we can define \( KC = [E][C]/[EC] \). The total amount of enzyme is \( ET = [E] + [EC] + [ES] \), where \([EC] = [E][C]/KC \) and, as before, \([E] = KM[ES]/[S] \). This leads to a perturbed bound enzyme fraction

\[
P_{ES,+C} = \frac{[S]}{[S] + \left(1 + \frac{[C]}{KC}\right)KM}. \tag{9}
\]

Plugging \( P_{ES,+C} \) into Eq. (5b) above leads to a perturbed reaction rate

\[
v_{+C} = \frac{V_{\text{max}}[S]}{[S] + \left(1 + \frac{[C]}{KC}\right)KM}. \tag{10}
\]
Note that $v_{+C}$ has the same form as the unperturbed rate $v_0$ (cf. Eq. (7)) if we define an apparent Michaelis constant

$$K_{M,app} = \left(1 + \frac{[C]}{K_C}\right)K_M. \quad (11)$$

Since $(1+[C]/K_C)$ must be $\approx 1$, we must have $K_{M,app} \approx K_M$, $P_{ES+C} \leq P_{ES}$, and $v_{+C} \leq v_0$. The increase of $K_{M,app}$ over $K_M$ means that more substrate must be added to achieve $V_{\text{max}}/2$ as in the presence of ‘C’ than in its absence. This is reasonable since a competitive inhibitor ‘ties up’ some of the available enzyme active sites. A competitive inhibitor leaves $V_{\text{max}} = k_{cat}E_T$ unaffected, and if sufficient amounts of substrate can be added, the original $V_{\text{max}}$ will be observed.

A common metric of inhibitor potency is the IC$_{50}$. The IC$_{50}$ is the (free) inhibitor concentration $[C]$ at which the targeted enzymatic reaction rate (or target receptor binding) is reduced by 50%. Cheng and Prusoff have given the relationship between IC$_{50}$ and $K_C$ by setting $v_0 = 2v_{+C}$ and solving for ‘[C]’ [17]. The result is

$$\text{IC}_{50} = K_C \left(1 + \frac{[S]}{K_M}\right). \quad (12)$$

Eq. (12) shows explicitly that the IC$_{50}$ depends on not only $K_C$, but also on the substrate concentration $[S]$ and $K_M$. Therefore, IC$_{50}$ values from different assays cannot be compared unless one accounts for potential differences in $[S]$ and $K_M$. A deeper complication is that different compounds can adopt different inhibitor strategies (e.g. ‘non-competitive’ in which the inhibitor binds to a site remote from the active site and thus forms complexes with both $[E]$ and $[ES]$, and ‘un-competitive’, in which the inhibitor binds just to $[ES]$). In turn, this requires the use of different functional relationships between IC$_{50}$ and $K_C$. Cheng and Prusoff have also given the relationships for non-competitive and un-competitive cases [17].

The more fundamental metric of competitive inhibitor potency is simply $K_C = \frac{[E][C]}{[EC]}$. $K_C$ is merely the equilibrium dissociation constant $K_D$ (cf. mechanism 1 and Eq. (2)) applied to the inhibitor–enzyme interaction. Smaller $K_C$ values imply more potent competitive inhibitors. While $K_C$ is an equilibrium dissociation constant, it is more typically determined from enzyme kinetic experiments in which enzyme, competitor, and substrate are present and steady-state conditions are assumed. NMR and other biophysical methods (e.g. equilibrium dialysis, fluorescence quenching, isothermal titration calorimetry (ITC), surface plasmon resonance) provide an alternative way to assess $K_C$ by performing equilibrium measurements involving only the enzyme and inhibitor.

### 2.3. Competitive binding equilibria

Competitive ligand binding experiments are well-known tools for ascertaining binding specificity and affinity. The competitive inhibition formalism of Section 2.2 provides the algebraic formulae necessary for interpreting competitive ligand binding experiments. We consider two ligands, ‘L’ and ‘I’ that compete for occupation of the same binding site of a given receptor. We take ‘L’ to be a previously characterized ligand, and ‘I’ a new competitive ligand. We distinguish their equilibrium dissociation constants using $K_D = [E][L]/[EL]$ and $K_I = [E][I]/[EI]$. The appropriate binding equilibrium is

$$\frac{[EI]}{[I]} = \frac{[E] + [I] + \frac{K_I}{K_D}}{[EL]} \quad (13)$$

In Section 2.3, we were interested in the initial enzymatic reaction rate $v_0$ in the presence versus absence of inhibitor ‘C’. Here, we are interested in the bound fraction receptor factor $[EL]/E_T$ in the presence versus absence of the competitive ligand ‘I’. In fact, the two problems are essentially the same as they are both concerned with the concentration of a receptor–ligand complex: $[ES]$ in the enzyme reaction rate case and $[EL]$ here. We can therefore expect the same formulae to be applicable. In particular, it is readily apparent that Eq. (13) is just the competitive mechanism of Eq. (8) revisited with different labels for the chemical species and the final catalytic step omitted. Thus, we can immediately use Eq. (9) of Section 2.2 to write the perturbed bound receptor fraction $[EL]/E_T$ by using the substitutions $[C] \Rightarrow [I]$, $[S] \Leftrightarrow [L]$, and $[ES] \Leftrightarrow [EL]$ to get

$$P_{I}^{B} = \frac{[L]}{[L] + K_{D,app}}, \quad (14)$$

where

$$K_{D,app} = K_D(1 + [I]/K_I) = \frac{(E_T - [EL])(L_T - [EL])}{[EL]} \quad (15)$$

As before, the factor $(1 + [I]/K_I)$ must be $\approx 1$, and thus $K_{D,app} \approx K_{D}$ and $P_{I}^{B} \approx P_{I}^{B}$. This reflects the reduction of available receptor sites for ‘L’ due to competitive interference from ‘I’.

Following the development of Eq. (12), we can define here an IC$_{50}$ value for the competitive binding experiment as the value of [I] that is consistent with $2P_{I}^{B} + 1 = P_{I}^{B}$. Swapping $K_D$ for $K_M$ and $K_I$ for $K_C$ in Eq. (12) and then rearranging yields

$$K_D = \frac{[L]K_I}{\text{IC}_{50} - K_I} \quad (16)$$

If ‘L’ is in molar excess, then $[L] \approx L_T$. Therefore, $K_D$ can be estimated from a knowledge of IC$_{50}$, $K_I$, and $[L] \approx L_T$.

It is desirable to write $[EL]$ and $[EI]$ in terms of the experimental control variables $E_T$, $L_T$, and $T_I$ as well as the two dissociation constants $K_I$ and $K_D$. The analyses of Wang and Sigurskajold provide the desired expressions for the fractions of bound ‘L’ and ‘I’[18,19]. In particular,

$$x_{EL} = [EL]/E_T = (L_T/E_T)x_L/([K_D/E_T] + x_E) \quad (17a)$$
\[ x_{EI} = \frac{[EI]}{[ET]} = \frac{(I_T/E_T) x_E ((K_I/E_I) + x_E)} \]

where \( x_E \) is the free receptor fraction \( x_E = [E]/[E_T] \). Substitution of Eqs. (17a) and (17b) into the constraint \( x_E + x_{EI} + x_{EL} = 1 \) yields a cubic equation in \( x_E \). The relevant root is

\[ x_E = \frac{2\sqrt{\alpha^2 - 3\beta \cos(\theta/3)} - \alpha}{3} \]  

where

\[ \Theta = \cos^{-1} \left( \frac{-2\alpha^2 + 9\alpha \beta - 27\gamma}{2\sqrt{(\alpha^2 - 3\beta)^3}} \right) \]  

and

\[ \alpha = \frac{1}{[ET]}(K_I + K_D + I_T + L_T - E_T) \]

\[ \beta = \frac{1}{[ET]} \left( (K_I(L_T - E_T) + K_D(I_T - E_T) + K_D K_I) \right) \]

\[ \gamma = -K_K D/\gamma_D \]

Plugging the free fraction \( x_E \) into Eqs. (17a) and (17b), yields \( x_{EI} \) and \( x_{EL} \), and thus [EL] and [EI].

Fig. 2 plots \( P_{B,LT}^E/P_{B}^E = [EL]_{LT}/[EL]_{LT} \) as a function of \( I_T \) for various values of \( K_D \) and \( K_I \). Note that \( [EL]_{LT}/[EL]_{LT} \) is simply the ratio of [EL] in the presence of competitor to [EL] in the absence of competitor. The x-axis plots \( I_T \) normalized to the fixed total ligand concentration \( L_T = 100 \mu M \). In addition, the figure assumes \( E_T = 1 \mu M \) and \( K_D = 50 \mu M \). The figure plots four competitor affinities including \( K_I = 0.1, 1.0, 10.0, \) and \( 50.0 \mu M \). The maximum value of \( [EL]_{LT}/[EL]_{LT} \) is unity and corresponds to the complete absence of competitor. Addition of competitor decreases [EL] and the ratio \( [EL]_{LT}/[EL]_{LT} \) drops. For \( K_I < K_D \), the competitor ‘I’ acts as a ‘knock-off’ in the sense that it displaces L for \( I_T < L_T \). If ‘I’ is the known specific binder and ‘L’ is the test ligand, this effect can be used to verify specific binding. Conversely, if ‘L’ is the known specific binder and ‘I’ is the test ligand, this effect can screen for hitherto unknown tight binders (vide infra).

3. Accounting for binding-induced chemical exchange

We now review the effects the chemical exchange processes underlying the binding equilibria given by Eq. (1). When the receptor and ligand molecules are free, they retain their intrinsic NMR parameters (e.g. chemical shifts, relaxation rates, translational diffusion coefficients). However, in each other’s presence, their mutual binding affinity drives a two-state exchange process that can toggle both sets of molecules between the free and complexed states. At equilibrium, they adopt the free and bound state populations ([EI], [LI], [EL]) consistent with Eq. (1). From the ligand’s perspective, it transiently adopts the NMR parameters characteristic of the typically much larger receptor–ligand complex. 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. Under conditions of fast and intermediate exchange, the exchange modulates the NMR parameters of both molecules. It is the ability to experimentally distinguish these exchange-modulated parameters from those of the free state that enables NMR screening.

3.1. Modified Bloch equations

To appreciate how exchange modifies the NMR parameters, it is useful to highlight some basic aspects of ‘dynamic’ NMR. The modified Bloch equation formalism of Hahn, Maxwell, and McConnell (HMM) forms an appropriate theoretical framework to explain and predict the majority of exchange phenomena relevant to NMR screening [20,21]. In this formalism, we consider \( M \) nuclear spins exchanging amongst \( p \) states such that the total number of exchanging magnetizations to be considered is \( N = pM \). We therefore, specify the state of equilibrium, longitudinal, and transverse magnetization using the \( N \)-dimensional vectors \( \mathbf{M}_0, \mathbf{M}_0(t), \) and \( \mathbf{M}_0(t) \). The instantaneous longitudinal magnetization state is defined by \( \mathbf{M}_0(t) = [M_{01}(t), M_{02}(t), ... M_{0p}(t), ... M_{0N}(t)] \) where \( M_{0j}(t) = \gamma h \cdot \mathbf{J}' \text{Tr}[I_j \rho(t)] \) and \( I_j \) is the Zeeman spin operator for spin ‘ \( j \)’. \( \rho(t) \) is the spin density operator of the system. \( \mathbf{J}' \) is the number of ‘ \( j' \) spins per unit volume, and ‘ \( \text{Tr} \)’ is the trace operation. The equilibrium longitudinal magnetization \( \mathbf{M}_0(t) \) follows by setting \( \sigma(t) = \sigma_0 = \exp(-H_0/k_B T)/Z \), where \( H_0 \) is the equilibrium spin Hamiltonian and \( Z \) is the corresponding partition function. The deviation of the prevailing
magnetization from equilibrium is thus $\Delta M_i = M_i(t) - M_0$. The instantaneous magnetization state is described by $M_i(t) = [M_{a1}(t), M_{a2}(t), \ldots M_{aj}(t), \ldots M_{an}(t)]$, where $M_{aj}(t) = y \Re \{I_{aj,\alpha}(t)\}$. $M_j(t)$ denotes the complex transverse magnetization $M_q(t) + iM_p(t)$, and $I_{aj} = I_{aj} - iI_{aj}$. We assume that the individual exchange-coupled states have fixed intrinsic relaxation rates and rotating frame precession frequencies. We can then describe the time rate-of-change of $M_i(t)$ and $M_j(t)$ under the joint effects of relaxation, precession, and exchange using the systems of ordinary linear differential equations written as

$$\frac{d}{dt}(\Delta M_i) = - (R + K)(\Delta M_i)$$

$$\frac{d}{dt}(\Delta M_j) = - (R + K) - i\Omega (\Delta M_i).$$

In Eqs. (20) and (21), $R$ is an $N \times N$ matrix that describes magnetization transfer due to auto- and cross-relaxation. $\Omega$ is a diagonal matrix of angular precession frequencies for the $N$ transverse magnetizations in the rotating frame. Most importantly, the $N \times N$ matrix $K$ describes magnetization transfers due to chemical exchange. Its elements are rate constants for ‘jump’ transitions between discrete chemical states. In each individual ‘reaction’ $i \Rightarrow j$ to state $j$. Taking $M_{aj}$ to be proportional to the equilibrium population of the ‘$j$’th’ state, $P_{aj}$, the condition of detailed balance (viz. each individual ‘reaction’ $i \Rightarrow j$ is at equilibrium) implies $M_{aj}k_{ij} = M_{ij}k_{ji}$. The chemical equilibrium condition renders $K_{ai} = 0$. Note that the $K$ matrix assumes that an exchange event itself does not ‘flip’ the spin state. As a consequence, the relaxation and exchange kinetics of different coherence orders can be treated independently from one another using the same mathematical structure as Eq. (21) (e.g. see the multiple-quantum coherence treatment by Kloiber and Konrat [22]).

The solutions of the HMM Eqs. (20) and (21) yield the general temporal behavior of the free and bound receptor and ligand magnetizations. The solution involves first symmetrizing and then diagonalizing the $R$, $\Omega$, and $K$ matrices using standard techniques for solving eigenvalue problems (see e.g. Mathews and Walker [23]). The resulting eigenvalues and eigenvectors yield the exchange-modulated precession frequencies, relaxation rates, and amplitudes of the NMR signals. This process is perhaps best illustrated by considering Eq. (21) for the simplest case of a single uncoupled spin belonging to a ligand that exchanges between the free and receptor-bound states. The HMM Eq. (21) then simplifies to

$$\frac{d}{dt}\begin{bmatrix} M_{+1} \\ M_{+B} \end{bmatrix} = - \begin{bmatrix} R_{2F} & 0 \\ 0 & R_{2B} \end{bmatrix} + \begin{bmatrix} P_{B}k_{ex} & -P_{F}k_{ex} \\ -P_{B}k_{ex} & P_{F}k_{ex} \end{bmatrix} \begin{bmatrix} \Omega_{F} \\ 0 \end{bmatrix} \begin{bmatrix} M_{+1} \\ M_{+B} \end{bmatrix}.$$  \hspace{1cm} (22a)

The subscripts $F$ and $B$ indicate the free and bound states, respectively. The relaxation matrix $R$ has only diagonal elements including the transverse relaxation rate constants $R_{2F} = 1/T_{2F}$ and $R_{2B} = 1/T_{2B}$. The exchange matrix $K$ includes $k_{on}[E] = k_{ex}P_{B}$, where $P_{B}$ is the bound ligand fraction $[EL]/[LT]$, and $k_{off} = k_{ex}P_{F}$. The only asymmetric matrix is $K$. To bring all matrices to symmetric form, one can effect the similarity transformation on Eq. (22a)

$$S^{-1}(R + K - i\Omega)S,$$

where $S = \text{diag}([1/\sqrt{[L]}], [1/\sqrt{[EL]}]$ and $S^{-1} = \text{diag}([1/\sqrt{[L]}], [1/\sqrt{[EL]}]$. Then Eq. (22a) becomes

$$\frac{d}{dt}\begin{bmatrix} M_{+F}/\sqrt{[L]} \\ M_{+B}/\sqrt{[EL]} \end{bmatrix} = - \begin{bmatrix} R_{2F} - i\Omega_{F} & 0 \\ 0 & R_{2B} - i\Omega_{B} \end{bmatrix} \begin{bmatrix} P_{B}k_{ex} & -P_{F}k_{ex} \\ -P_{F}k_{ex} & P_{B}k_{ex} \end{bmatrix} \begin{bmatrix} M_{+F}/\sqrt{[L]} \\ M_{+B}/\sqrt{[EL]} \end{bmatrix}.$$  \hspace{1cm} (22b)

Since all matrices of Eq. (22b) are now symmetric, they can be brought simultaneously to diagonal form using standard methods of orthogonal matrix transformations obtained from an orthorhombic basis of eigenvectors with attendant eigenvalues. If we consider the usual NMR screening scenario in which $z = \text{L}_{F}/\text{L}_{T} \gg 1$, then $P_{F} = P_{B}$. The general HMM solutions for the transverse relaxation rate of the majority fraction ($P_{F}$) are:

$$R_{2F}^{\text{HMM}} = 1/2 \left( R_{2F} + R_{2B} + k_{ex} - 1/\sqrt{2} \sqrt{[\psi + \sqrt{\psi^{2} + z^{2}}}] \right)$$  \hspace{1cm} (23a)

$$\zeta = 2(\Omega_{F} - \Omega_{B})(R_{2F} - R_{2B} + k_{ex}(P_{B} - P_{F}))$$

$$\psi = (R_{2F} - R_{2B} + k_{ex}(P_{B} - P_{F}))^{2} - (\Omega_{F} - \Omega_{B})^{2}$$

$$+ 4P_{B}P_{F}k_{ex}^{2}.$$  \hspace{1cm} (23c)

While generally valid, the above expressions are rather awkward. Fortunately, the condition $P_{F} \gg P_{B}$ permits use of the simpler Swift–Connick formula:

$$R_{2F}^{\text{SC}} = P_{F}R_{F} + P_{P}P_{B}k_{ex} \left[ R_{2F}(R_{2B} + P_{B}k_{ex} + (\Omega_{B})^{2}) \right]/(R_{2B} + P_{B}k_{ex}^{2} + (\Omega_{B})^{2}).$$  \hspace{1cm} (24)

As with the HMM Eqs. (23a)–(23c), the Swift–Connick expression $R_{2F}^{\text{SC}}$ of Eq. (24) is valid over all exchange time scales.

The rate expressions simplify further when we consider the extreme limits of ‘slow’ and ‘fast’ exchange. One gauges the ligand exchange rapidity using the net rate constant $k_{ex} = k_{on}[E] + k_{off}$ (cf. Eq. (1)), and receptor exchange rapidity using $k_{ex}^{2} = k_{on}[L] + k_{off}$. For a given NMR parameter $Q$ where $Q_{B}$ and $Q_{F}$ are, respectively, the intrinsic bound and free state parameters, the exchange is referred to as slow, intermediate, or fast if $k_{ex}$ is much
less than, approximately equal to, or much greater than \( |Q_B - Q_F| \). As an aside, the physical significance of \( k_{ex} \) and \( k_{bi} \) in Eq. (1) goes beyond being a mere speedometer for exchange: they represent both the relaxation rate constant for equilibration of non-equilibrium concentrations, as well as the inverse correlation times for the individual exchange process at the molecular level [24].

In the limit of slow exchange, the matrix \( \mathbf{K} \) acts as a small perturbation to \( \mathbf{R} + i\mathbf{Q} \) in Eq. (22b). Equivalently, \( \sqrt{P_P P_B k_{ex}} \ll (R_{2F} - R_{2B}) \) and \( |\Omega_{ex} - \Omega_B| \). Following standard linear algebra perturbation theory, only the diagonal elements of \( \mathbf{K} \) are significant. The free and bound ligand signals then retain their intrinsic precession frequencies \( \Omega_F \) and \( \Omega_B \) and have modified relaxation rates

\[
\begin{align*}
R_{2F,\text{slow}} &= R_{2F} + P_B k_{ex} \quad (25a) \\
R_{2B,\text{slow}} &= R_{2B} + P_F k_{ex} . \quad (25b)
\end{align*}
\]

As stated, typical screening conditions are \( P_B \ll P_F \) and \( R_{2B} \gg R_{2F} \). This means \( R_{2F,\text{slow}} \) will scarcely deviate from \( R_{2F} \), and the bound state signal will be essentially undetectable. It also means slow exchange induced by high-affinity binding may be hard to distinguish from a complete lack of binding. If we are restricted to observing the free ligands, then slow-exchange is an inefficient means for conveying information from the bound state.

In the limit of fast exchange on the chemical shift and relaxation time scales, \( \mathbf{R} + i\mathbf{Q} \) now becomes a small perturbation to \( \mathbf{K} \). One observes only a single signal with angular precession frequency \( \Omega_{avg} = P_F \Omega_F + P_B \Omega_B \) with an averaged relaxation rate, \( R_{2,avg} \), given by [25]

\[
R_{2,avg} = P_B R_{2B} + P_F R_{2F} + R_{ex} \quad (26a)
\]

where

\[
R_{ex} = (\Omega_F - \Omega_B)^2 P_F P_B / k_{ex} . \quad (26b)
\]

Eqs. (26a) and (26b) show that the bound state information is encoded within the averaged relaxation rate of the single resonance. In the limit of very fast exchange, the line broadening term \( R_{ex} = 0 \).

Fig. 3 pools the above Eqs. (23)–(26). The figure assumes a ligand ‘L’ exchanging on and off a single receptor site with \( R_{2F} = 5 \, \text{s}^{-1} \), \( R_{2B} = 60 \, \text{s}^{-1} \), \( E_T = 50 \, \mu\text{M} \), \( L_T = 1 \, \text{mM} \), \( \epsilon = L_T/E_T = 20 \), \( k_{on} = 1 \times 10^8 \, \text{M}^{-1} \, \text{s}^{-1} \), and \( (\Omega_F - \Omega_B)/2\pi = 100 \, \text{Hz} \). The horizontal axis is \( \log_{10}(k_{eff}) \); thus, faster exchange (weaker binding) proceeds to the right. For reference, the lowest flat trace represents the free state \( R_{2F} = 5 \, \text{s}^{-1} \). The solid peaked curve represents \( R_{2F}^{\text{HMM}} \) from Eqs. (23a)–(23c); following this curve illustrates how the transfer of bound state information to the free state varies as a function of the exchange rate. At slow exchange (extreme left), the residence time \( k_{ex}^{-1} \) is long. \( R_{2F}^{\text{HMM}} \) scarcely deviates from the free state (i.e. unbound state) value and we receive essentially no information concerning the bound state. As we move to the right, the exchange rate increases and so does the efficiency of information transfer from the bound to free state. The increase in \( R_{2F}^{\text{HMM}} \) reflects the exchange-relayed transfer of the intrinsic bound state \( R_{2B} \) to the free state and enhanced relaxation stemming from the toggling of chemical shifts \( \Omega_F \leftarrow \Omega_B \). To appreciate

![Fig. 3. Simulation of transverse relaxation rate constant \( R_2 \) for a single uncoupled ligand spin as a function of the off-rate, \( k_{eff} \) using the Eqs. (23)–(26) in the main text. The figure assumes \( L_T = 1 \, \text{mM} \), \( E_T = 50 \, \mu\text{M} \), \( k_{on} = 1 \times 10^8 \, \text{M}^{-1} \, \text{s}^{-1} \), \( \Delta \Omega/2\pi = 100 \, \text{Hz} \), \( EF = 5 \, \text{s}^{-1} \), and \( R_{2B} = 60 \, \text{s}^{-1} \). Slow and fast exchange limits prevail on the left and right respectively. Curve traces are as follows: solid flat trace = \( R_{2F} \); solid peaked trace = \( R_{2F}^{\text{HMM}} \), sigmoidal diamond trace = \( R_{2F}^{\text{HMM}} \) with \( \Delta \Omega = 0 \); dashed-dotted curve = fast exchange result \( R_{2,fast} \); long dashes = fast exchange \( R_{2,fast} \) with \( \Delta \Omega = 0 \); open circle trace = Swift–Connick \( R_{2,SC} \).](image-url)
the contribution of the former effect, we can consider the case when \( \Delta \Omega = \Omega_q - \Omega_R = 0 \). Then \( \Omega_{Q,HMM} \) becomes the sigmoidal trace of diamonds that shows a point of inflection near \( k_{ex} = |R_{2q} - R_{2b}| \). This point can be thought of as the intermediate exchange condition on the relaxation time scale. On the other hand, \( \Omega_{Q,HMM} \) for \( \Delta \Omega \neq 0 \) reaches a maximum near \( k_{ex} = \Delta \Omega \) (shift coalescence). This reminds us that the assessment of exchange rapidity depends on the NMR parameter being considered (e.g. relaxation versus chemical shift). With even faster exchange, the transfer of bound state information to the free state is consistent with the simpler fast-exchange expressions of Eqs. (26a) and (26b). This is evidenced by the increasingly good congruence of Eqs. (26a) and (26b) (slope of dashed-dots) to \( \Omega_{Q,HMM} \) for \( k_{off} > \Delta \Omega \). As \( k_{off} \) increases further, the \( \Omega_{ex} \), contribution fades and \( \Omega_{Q,HMM} \) becomes exclusively the population weighted average of free state and bound state rates \( P_R R_{2b} + P_b R_{2q} \) indicated by the nearly flat trace of long dashes. The drop at the highest \( k_{off} \) simply reflects the inevitable decrease of \( P_R \) due to the fact that \( \Omega_{Q,HMM} \) now approaches \( \Omega_{LT} \) (receptor binding site is no longer saturated). At this stage, we start to lose bound state information again. The Swift—Connick expression of Eq. (24) is given by the curve of open circles. Comparing this to the solid \( \Omega_{Q,HMM} \) curve illustrates the accuracy of the Swift—Connick approximation for \( \Omega_{LT} \) (i.e. \( P_R \gg P_b \)).

### 3.2. Fast-exchange approximation

As stated, the solutions to the HMM Eqs. (20) and (21) yield the behavior of system magnetization on arbitrary exchange time scales. In practice, however, one almost never solves Eqs. (20) and (21) in NMR screening and simply assumes fast exchange. The reasons for this are two-fold. First, the experimental conditions of ligand-based NMR screening are often conducive to fast exchange. One typically works with \( \epsilon = \frac{L_T}{T_1} \gg 1 \), and the binders or ‘hits’ have \( K_D \geq 100 \mu M \). If \( k_{ex} \) is well-approximated by a diffusion-limited value in the range of \( 10^7 - 10^9 \text{ M}^{-1} \text{ s}^{-1} \), then the slowest \( k_{ex} \) lie in the range \( 1000 < k_{ex} < 100,000 \text{ s}^{-1} \). NMR screening relies mainly on \( ^1 \text{H} \) NMR, and this \( k_{ex} \) exceeds most differences in intrinsically \( ^1 \text{H} \) relaxation rates and rotating frame precession frequencies, and thus gives some reassurance that the fast exchange assumption is tenable.

The second motivation for assuming fast exchange is the resulting algebraic simplicity. It is painfully clear from the above that exchange-modulated NMR parameters can be awkward. However, the fast exchange results of Eqs. (26a) and (26b) are simple sums. More generally, under fast exchange, NMR parameters \( Q \) become the simple averages

\[
Q_{avg} = P_R Q_R + P_b Q_b.
\]

\[
Q_{avg} = P_R Q_R + P_b Q_b + Q_{ex}.
\]

\( Q_{avg} \) is the exchange-averaged parameter we observe for ligand/receptor in the presence of receptor/ligand. Its distinctiveness from \( Q_F \) is the signature of receptor binding that flags a hit. In the case of Eq. (27a), we see \( Q_{avg} \) is a simple population-weighted average. In the case of Eq. (27b), that average has an additional offset term \( Q_{ex} \); an example of this case has already been given by the transverse relaxation expression, Eq. (26a) and (26b). The form of Eq. (27a) is most commonly seen for those parameters \( Q \) for which chemical shift modulations are irrelevant (e.g. longitudinal auto- and cross-relaxation rates, rotating-frame spin-locking auto- and cross-relaxation rates, translational diffusion coefficients). The common bound state contribution in Eqs. (27a) and (27b) is \( Q_R Q_B \). Our ability to detect binding sensitively depends crucially on \( \theta_R Q_B \) being significant relative to \( \theta_R Q_B \). But the typical screening conditions of \( L_T \gg E_T \) make \( \theta_R \ll \theta_F \). We therefore prefer to measure NMR parameters \( Q \) that become amplified in the bound state (i.e. \( \theta_R \gg \theta_F \)).

The simple population-weighted average of Eq. (27a) can be related to the binding affinity. Since \( P_R = 1 - P_b \) and \( P_b = P_b^0/e \), we can write the difference \( \epsilon(Q_{avg} - Q_F) \) as

\[
\epsilon(Q_{avg} - Q_F) = \frac{(Q_R - Q_F)(L)}{[L] + K_D}.
\]

Eq. (28) is a key formula for NMR screening because the most popular ligand-based NMR screening experiments usually assume fast exchange and therefore, invoke the above formula for titration analysis. The form of Eq. (28) describes the now familiar hyperbolic Langmuir dose—response curve we have already seen for \( P_R^0 \) and \( P_b^0/e \) in Eqs. (2) and (14), respectively, and for \( v_q \) and \( v_{q,c} \) in Eqs. (7) and (10), respectively. To experimentally assess \( Q_{avg} - Q_F \), one measures the \( Q \)-related NMR signal intensities in the presence of receptor, and then corrects for the free-state contributions \( Q_F \). One can determine \( Q_F \) either by performing a ‘reference’ experiment of the lone compound, or by designing experiments for which \( Q_F = 0 \). Eq. (28) shows that \( \epsilon(Q_{avg} - Q_F) \) increases with ligand addition, reaches half of its maximal value at \( [L] = K_D \), and then plateaus at its maximal value \( (Q_R - Q_F) \) when the binding site is saturated (\( L_T \gg K_D \)). Provided ligand solubility is not limiting, one can fit ligand titration data to this simple relation to estimate binding affinity. More specifically, for a large ligand excess (\( \epsilon = L_T/E_T \gg 1 \)), we can safely approximate \( [L] = L_T \), and a fit of ligand titration data to Eq. (28) yields estimates for \( K_D \) and \( (Q_R - Q_F) \).

### 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 exclusively on proteins. Such methods exploit the site-specific information afforded by assigned protein NMR spectra along with a priori knowledge of its 3D structure (either from X-ray diffraction or NMR) to drive lead generation. By identifying binding-induced perturbations of assigned protein resonances, one not only identifies ligands, but also localizes their binding sites. This suggests strategies for fragment-based lead generation, in which lower affinity molecular fragments binding to distinct sub-sites can be linked or elaborated to yield higher affinity compounds [7,26]. The ability to track perturbations to assigned resonances of a protein with known structure also distinguishes between specific versus non-specific binding. Finally, unlike ligand-based methods, receptor-based methods do not rely on fast exchange to retrieve bound state information. Thus, receptor observation permits the characterization of both higher and lower affinity hits.

A major caveat affecting receptor-based methods is that pharmaceutical research chooses protein targets based on their therapeutic potential, and not their compliance with NMR spectroscopy. NMR methods demand physico-chemical properties of the protein target that can be difficult to realize in a timely cost-effective manner for therapeutically relevant systems. For example, one must confront the basic challenge of over-expressing and purifying mg quantities of soluble, non-aggregating protein. Then, there is the additional challenge of finding suitable expression hosts that permit isotope enrichment (e.g. $^{13}$C, $^{15}$N, $^2$H) critical for the resonance assignment of the typically large (>30,000 Da) therapeutic targets. One should recall that the most popular NMR expression system, _E. coli_, is often not an option for mammalian targets whose over-expression may prove to be toxic to the host cell. Even after sufficient quantities of labeled material are available, one must ensure that the sample is stable for the time required for sequential resonance assignment using a battery of multi-resonance experiments. 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. Presently, the required NMR time expenditures for such targets often dictate use of other approaches that can contribute to medicinal chemistry faster.

Ligand-based methods compare the NMR parameters of a mixture of compounds in the presence and absence of the receptor molecules. This renders the size of the receptor molecule irrelevant. In fact, the most powerful ligand-based approaches become more sensitive when dealing with larger molecular weights. Additionally, ligand observation bypasses the challenges of producing mg quantities of isotope-labeled receptor suitable for multi-dimensional NMR. In turn, this provides the spectroscopist with the flexibility to evaluate new targets more rapidly. This is important not only for adapting to an ever-shifting priority landscape, but also for contributing on a time scale useful for chemistry and high-throughput screens. Finally, as discussed below, the most popular method for NMR screening works well with minimal receptor concentration-often ≤1 μM, thus reducing the burden on protein expression and purification.

A main disadvantage of ligand-based approaches is the inability to a priori localize the binders on the receptor. Additionally, ligand-based approaches rely on the exchange-mediated transfer of bound-state information to the free state. This biases ligand-based methods towards weakly binding ligands (rapid exchange) and large ligand molar excesses ($e = L_I/E_T \gg 1$). The consequent risk is that the abundance of ligand may start to occupy weaker affinity non-specific binding sites. However, recent developments in ligand-based methods discussed below are starting to gnaw at these difficulties [27–29].

Clearly, both receptor and ligand-based approaches have distinct advantages and disadvantages. At the time of this writing, however, ligand-based screening is of broader applicability and places less demands on other disciplines and infrastructure. These qualities make it better adapted to the pharmaceutical research setting. We therefore, spend the majority of this review discussing ligand-based screening methods.

5. Ligand-based NMR screening methods

The exquisite sensitivity of ligand NMR parameters to the bound versus free states underlies the rationale for ligand-based screening. Generally, the compounds of NMR screening libraries are small molecules with masses <1000 Da. Therefore, the free compounds are characterized by small relaxation rates $R_1 = 1/T_1$, $R_2 = 1/T_2$, vanishing or weakly negative 2D-NOESY cross-peaks due to vanishing or weakly positive dipole–dipole laboratory-frame cross-relaxation rates, and large translational diffusion coefficients, $D_t$. Bound compounds share the NMR relaxation properties of the receptor which is often a protein molecule with mass >30,000 Da. Thus, bound compounds have large $R_2 = 1/T_2$, positive 2D-NOESY cross-peaks and highly efficient spin-diffusion, and smaller molecular diffusion coefficients, $D_t$. These clear differences suggest that one can watch for changes in the ligand NMR relaxation parameters to assay for binding. Generally, the majority of ligand-based NMR experiments assay for binding by: (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, which manifest as altered relaxation parameters and diffusion coefficients, respectively; (ii) exploiting a $^1$H magnetization transfer process whose origin is the receptor; binders or ‘hits’ will experience this transfer while non-binders will not.
The majority of ligand-based screening NMR experiments use 1H NMR. Unless otherwise stated, we will assume this to be the case. The chief relaxation mechanisms for 1H NMR are 1H–1H dipole–dipole (DD) interactions between pairs of proton spins. For a given proton in a state of spin order 'V' (e.g. 'V' could be I_x, I_y, multiple-quantum coherence, etc.), a DD relaxation rate can be written as the double sum

\[ R(V) = \sum_m a_m \sum_j \frac{1}{r_j^6} J_m(mw). \]  

(29)

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 \( J_m(mw) \) evaluated at various integral multiples 'm' of the 1H Larmor frequency \( \omega_{1H} \). The weighting coefficients, \( a_m \), depend on the spin order \( V \) (e.g. \( V = I_x, I_y \)) and attendant rate constant (e.g. \( R_1 = 1/T_1 \), NOE, \( R_2 = 1/T_2 \)) being considered. The spectral density functions \( J_\omega(\omega) \) 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 \( B_0 \). Therefore, the complexity of \( J_\omega(\omega) \) is dictated by the nature of molecular dynamics present. Usually, both structural and screening studies assume rigid receptor and ligand molecules that undergo isotropic tumbling. Thus, the only 'dynamics' are those of overall tumbling. Accordingly, each \( J_\omega(\omega) \) becomes the familiar Lorentzian distribution function

\[ J_\omega(\omega) = \frac{2 \tau_c}{\pi} \frac{\omega_{1H}}{1 + (\omega_{1H} \tau_c)^2}, \]  

(30)

\( \tau_c \) is the effective overall rotational correlation time which scales with the molecular mass. Therefore, \( J(0) = 2/5 \tau_c \). When the ligand binds to a large receptor molecule, \( \tau_c \) increases and thereby amplifies dramatically the corresponding relaxation parameter. Thus, relaxation parameters containing a \( J(0) \) dependence are highly sensitive probes of binding (vide infra).

As stated, a major selling point of ligand-based approaches to non-NMR research groups is that isotope labels are not needed. This presents a challenge for selective observation of the ligand signal; interference from receptor signals can obviously compromise the accuracy and ease of data interpretation. For selective ligand observation, one can: (i) use a high \( \epsilon = L_1/E_2 \) and exploit the ligand excess that overwhemls the receptor signals; (ii) use transverse relaxation-filters (e.g. Hahn echoes, spin-locks) that preferentially eliminate the rapidly relaxing receptor resonances; (iii) use translational-diffusion filters to select against rapidly diffusing molecules. For experiments performed in H2O, water suppression is typically achieved via excitation sculpting [30] or various incarnations of WATERGATE [31,32]. Because these methods involve spin-echo segments, they simultaneously act as relaxation filters prior to detection. We now turn to descriptions of specific ligand-based NMR strategies.

5.1. Transverse relaxation rates

Comparisons of ligand transverse auto-relaxation rates \( R_2 = 1/T_2 \) in the presence and absence of receptor are among the most venerable of NMR binding assays. \( R_2 \) is a highly attractive probe of binding due to its approximately direct dependence on the overall molecular rotational correlation time \( \tau_r \). This can be seen through the dependence of \( R_2 \) on spectral density functions. For a given ligand proton under 1H–1H DD relaxation, \( R_2 \) is

\[ R_2 = \frac{\hbar^2 g_H^2}{8} \sum_{j=1}^N \frac{1}{r_j^6} \left( 5 J_0(0) + 9 J_2(\omega_{1H}) + 6 J_3(2\omega_{1H}) \right). \]  

(31)

Given the form of \( J_\omega(\omega) \) in Eq. (30), the \( J(0) \) dependence in Eq. (31) means that \( R_2 \) has a strong dependence on \( \tau \). When the ligand binds to the receptor, its \( \tau_c \) temporarily becomes that of large receptor. Because \( \tau_c \text{ receptor} \gg \tau_c \text{ free–ligand} \), we have \( R_2 \text{ bound} \gg R_2 \text{ free} \).

Ligands undergoing sufficiently rapid binding exchange will yield exchange-averaged relaxation rates that reflect the exchange-mediated transfer of \( R_2 \text{ bound} \) to the free state. In the fast exchange limit, the exchange-averaged rate, \( R_{2\text{ avg}} \), is the simple sum already shown in Eq. (26a). The \( P_B R_{2\text{ bound}} \) term is the information transferred from the bound state that indicates a hit. Note that the evidence of binding becomes strong when \( P_B R_{2\text{ bound}} \) is significant relative to \( P_F R_{2\text{ free}} \). But the typical screening conditions of \( e = L_1/E_2 \gg 1 \) make \( P_B \ll P_F \). The advantage of \( R_{2\text{ bound}} \gg R_{2\text{ free}} \) now becomes clear; because the ligand \( R_2 \) becomes amplified in the bound state, \( P_B R_{2\text{ bound}} \) can be significant even under a large excess of ligand. This is a direct consequence of the \( J(0) \) dependence of \( R_2 \) in Eq. (31), and underscores the general fact that relaxation parameters having a \( J(0) \) dependence are sensitive probes of binding.

\( R_{2\text{ avg}} \) contains another signature of binding: \( R_{ex} \) (cf. Eq. (26b)). \( R_{ex} \) is the familiar relaxation enhancement arising from the non-equivalence of free versus bound chemical shifts. As stated before, \( k_{ex} = [E]k_{on} + k_{off} \). We should recall that \( k_{ex}^{-1} \) is essentially the correlation time for the complexation reaction of Eq. (1) [24]. Thus, \( R_{ex} \) is really just another \( J(0) \) dependence in disguise where the salient dynamic process is the two-state exchange instead of molecular tumbling. This becomes more explicit in \( R_{1\text{ fast}} = 1/T_{1\text{ fast}} \) dispersion studies, where \( R_{1\text{ fast}} \) represents the relaxation time of magnetization along the effective spin-lock field in the rotating reference frame [33]. Under large \( R_{ex} \), one may observe binding induced enhancements of \( R_2 \) even when \( P_B R_{2\text{ bound}} \) is insignificant. On the other hand, \( R_{ex} \) becomes negligible if \( \Omega_B = \Omega_{1\text{ fast}} \) is too large, or if the populations are lopsided (i.e. \( P_B P_F \ll 1 \)). Such often seems to be the case in 1H NMR screening. Working with stoichiometric amounts of receptor and ligand minimizes the latter
possibility. However, this is not always practical if receptor concentration is limited (e.g. to 1 μM).

To detect binders, one can compare compound lineshapes in the presence and absence of the receptor. The full-width at half-maximum height (FWHM) homogeneous linewidth of a given resonance is \( R_2/\pi \). Its peak height is also proportional to \( 1/R_2 \). Thus, binding-induced \( R_2 \) enhancements may be visible by simple line-broadening of proton resonance lines upon the addition of receptor. Fig. 4 shows an example from our own laboratory in which selective line broadening and consequent peak-height attenuation reveal clearly the binding compound in a mixture. More generally, however, line-broadening can be difficult to observe, especially if the effect is small, or if there is significant spectral crowding from the receptor and the other compounds in the mixture.

Alternatively, one can compare peak heights in the presence and absence of receptor using 1D \(^1H\) \( R_2 \) experiments [34]. Such experiments simultaneously relaxation-editing away the receptor resonances while relaxation-editing for binding compounds in a mixture. Popular experiments actually monitor \( R_2 \) relaxation using a spin-echo pulse train or continuous wave irradiation as in a \( R_1p \) experiment [33,35,36]. The pulse schemes are simply 90°-x-\( \gamma \)-Acq. Since radio-frequency (rf) irradiation is applied during the transverse relaxation period, the exchange-averaged relaxation rates differ slightly from those given in Eqs. (23a)–(23c). Specifically, one solves Eq. (22b) using boundary conditions imposed by the CPMG refocusing 180° pulses [37] or adding terms related to the applied \( R_{1p} \) rf-field magnitude and tilt-angle [38].

The general expressions are quite awkward and we write here only the fast exchange limits [39,40]:

\[
R_{2,\text{avg}}^{\text{CPMG}} = P_x R_{2F} + P_{\text{RL}} R_B \\
+ \frac{P_x P_{\text{RL}} (\Delta \Omega)^2}{k_{\text{ex}}} \left( 1 - \frac{2\tan(h(k_{\text{ex}}t_{cp}/2))}{k_{\text{ex}}t_{cp}} \right)
\]

\[
R_{1p,\text{avg}} = P_x R_{1p} + P_{\text{RL}} R_{1pB} \\
+ \frac{P_x P_{\text{RL}} (\Delta \Omega)^2 \sin(\Theta_{1p})^2}{k_{\text{ex}}} \left( \frac{k_{\text{ex}}}{k_{\text{ex}}^2 + \Delta \Omega_{\text{SL}}} \right)
\]

For \( R_{1p,\text{avg}} \), on-resonance spin-locking corresponds to \( \Theta_{1p} = \pi/2 \) and \( J(\omega) \approx J(\omega \pm \omega_{\text{SL}}) \), and so on-resonance \( R_{1p} \approx R_2 \). Thus the main effect of spin-locking is to give \( R_{ex} \) a functional dependence on the spin-lock field strength (\( \approx 4l_{\text{cp}} \) for CPMG, \( \Omega_{\text{SL}} \) for \( R_{1p} \)). The \( R_{ex} \) function can be quenched by sufficiently large \( \Omega_{\text{SL}} \); this is of course the basis for \( R_{1p} \) dispersion studies aimed at measuring exchange rates. Note also that the \( R_{ex} \) function in Eq. (33) is the spectral density function for the two-state exchange process alluded to above (vide supra). To avoid lineshape distortions, it is important to suppress the effects of homonuclear scalar coupling \( J_{\text{HH}} \) between protons during the relaxation period. In this context, the \( R_{1p} \) experiment may be preferable since it essentially quenches scalar coupling evolution. If instead one uses the CPMG pulse train, then the delay \( t_{cp} \) between consecutive 180° pulses should satisfy \( |4\pi\omega_{\text{HI}}l_{\text{cp}}| \ll 1 \) [41]. It is important to use sufficiently long spin-locks to eliminate receptor signals. As a benchmark, Hajduk et al. noted that a 400 ms CPMG spin-lock was sufficient for the complete elimination of the majority of receptor proton signals on the FKBP-12 protein (FKBP is an immunophilin with MW = 12.5 kDa that binds the immunosuppressant drug FK-506) [34].

Basic steps for identifying binders using \( R_2 \)-edited experiments are as follows. One first records the \( R_2 \) experiment for the lone compounds (‘− receptor’) for a set transverse relaxation delay, \( T_{\text{rel}} \). The resulting peak intensities are proportional to \( \exp(-R_{2F} T_{\text{rel}}) \). One then records a second \( R_2 \) experiment on the compounds in the presence of receptor (‘+ receptor’). For the hits, the peak intensities are now proportional to \( \exp(-R_{2B} T_{\text{rel}}) \). Since \( R_{2,\text{avg}} \approx R_{2F} \), 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, one must obviously be wary of the usual risks inherent with difference spectroscopy. Small chemical shift changes between samples or because of 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 one can reliably integrate individual compound resonances.

The amount of resonance attenuation can be used to get a coarse estimate of binding affinity. Using Eqs. (26a,b) and (33) for $R_{2,avg}$ and $R_{1,avg}$, 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 [42]. Their calculations suggest $R_1$ spin-locks of 400 ms are sufficient for the nearly complete elimination of signals from compounds with $K_D < 500 \mu M$. These calculations assumed a receptor molecular weight of 14,700 Da, equal to that of their test system, the human adipocyte fatty acid binding protein FABP4. 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.

5.2. Longitudinal relaxation rate

When considering $R_1 = 1/T_1$ measurements as a binding assay, inter-proton cross-relaxation (NOE) makes it important to distinguish between selective and non-selective $R_1$ measurements. Selective $R_1$ values are sensitive probes of binding while non-selective $R_1$ values are not. One sees this by considering longitudinal relaxation in a system of two ‘unlike’ $^1$H spins, I and S. The Solomon equation describing the longitudinal relaxation of $\langle I_1 \rangle$ is [43]

$$\frac{d(I_1 - I_0)}{dt} = -R_{1,IS}(I_1 - I_0) - \sigma_{IS}(S_x - S_0).$$

(34)

$R_{1,IS}$ is the selective longitudinal relaxation rate constant for spin ‘I’ due to the IS DD interaction, and is given by

$$R_{1,IS} = \frac{h^2}{4\gamma^2 IS} [J_{IS}(0) + 3J_{IS}(\omega_{H}) + 6J_{IS}(2\omega_{H})],$$

(35)

and $\sigma_{IS}$ is the familiar (NOE) cross-relaxation rate constant

$$\sigma_{IS} = \frac{h^2}{4\gamma^2 IS} \{6J_{IS}(2\omega_{H}) - J_{IS}(0)\}.$$ 

(36)

Note both $R_{1,IS}$ and $\sigma_{IS}$ have a $J(0)$ dependence and are therefore, sensitive reporters of binding. By adding and subtracting $\sigma_{IS}(I_1 - I_0)$, and assuming $I_0 = S_x$, Eq. (34) can be rewritten in the manner of Kalk and Berendsen [44]

$$\frac{d(I_1 - I_0)}{dt} = -(R_{1,IS} + \sigma_{IS})(I_1 - I_0) + \sigma_{IS}(I_x - S_2).$$ 

(37)

Consider an experiment in which we force $I_x = \alpha I_0$ and $S_2 = \beta I_0$ at time $t = 0$. The constants $\alpha$ and $\beta$ dictate the initial conditions; for example, they are $-1$ for inversion, 0 for saturation, and +1 for doing nothing. Then the initial rate of relaxation for $\langle I_1 \rangle$ is:

$$\frac{d(I_1 - I_0)}{dt} \bigg|_{t=0} = -(R_{1,IS} + \sigma_{IS})(I_1 - I_0) + \sigma_{IS}(I_x - S_2).$$ 

(38)

Equ. (38) shows that different initial conditions will produce different apparent relaxation rates for $I_1$. If we perturb $I_x$ and $S_2$ identically, then we do a non-selective experiment in which $\alpha = \beta$. Eq. (38) shows that the initial rate of recovery is then $R_{1,IS} = (R_{1,IS} + \sigma_{IS})$.

$$R_{1,IS} = \frac{h^2}{4\gamma^2 IS} [3J_{IS}(\omega_{H}) + 12J_{IS}(2\omega_{H})].$$ 

(39)

$R_{1,IS}$ has no $J(0)$ dependence due its cancellation in the $R_{1,IS} + \sigma_{IS}$ sum; it is therefore an insensitive reporter of binding. On the other hand, if we invert only $I_0$, then we do a selective ‘I’ experiment in which $\alpha = -1$ and $\beta = 1$. The initial rate of relaxation is just (the selective) $R_{1,IS}$ given by Eq. (35) above, which retains the $J(0)$ dependence. Thus, to achieve selective $R_1$ values (for detecting binding) we must differentially perturb the salient magnetizations, either by selective inversions (as in 1D NOE experiments) or at different times (as in 2D -NOE experiments).

Henceforth, we restrict our attention to selective $R_1$ (Eq. (35)). Under fast two-state exchange, we observe an averaged selective $R_1$ (cf Eq. (27a))

$$R_{1,avg} = P_F R_{1F} + P_B R_{1B}. $$

(40)

For a collection of ligand protons, the bound and free (selective) $R_1$ of a given proton is just the sum of $R_{1,IS}$ rate constants as given in Eq. (35)

$$R_1 = \frac{h^2}{4} \sum_{j=1}^{N} \frac{1}{r_j^2} [J_{j}(0) + 3J_{j}(\omega_{H}) + 6J_{j}(2\omega_{H})].$$ 

(41)

Again, the ‘j’ sum counts the N proximal protons that are separated by a distance $r_j$ from the proton under scrutiny. In going from the free to the bound state, both $J_j(0) = 2\pi/5$ and $N$ will increase. Both increases are responsible for $R_{1B} \gg R_{1F}$, and thus make $R_{1,avg}$ a sensitive probe of binding even when $P_B \ll P_F$.

To measure 1D selective $R_1$ values, one inverts or saturates a restricted set of protons using frequency-selective or other discriminatory pulse sequences. A practical obstacle is achieving selective spin perturbations for a large library of diverse compounds. Clearly, one does not want to define a separate inversion pulse for each compound. A clever alternative 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 [45]. In this experiment, one applies a pulse scheme consisting of two $90^\circ$ pulses bracketing a Hahn spin-echo or CPMG pulse train applied along ± $y$. The echo or train duration
is of sufficient length to selectively eliminate the receptor transverse magnetization via $R_2$ relaxation. This simultaneously saturates the receptor protons while selectively inverting those of the ligand. After a relaxation delay, a final 90° pulse reads out the resulting magnetization. Toggling the phase of the second 90° pulse places the ligand magnetization alternately on the $\pm z$-axis, and like toggling of the receiver phase yields the ligand selective $R_1$. Compounds that bind the receptor are expected to have enhanced ligand $R_1$ values and thus greater residual signal after the subtraction.

Another clever use of selective $R_1$ experiments is to monitor displacement of a ‘reporter’ or ‘probe’ molecule in competitive binding studies with higher affinity ligands [28]. Binding information comes from analysis of the same reporter molecule for a mixture of test molecules. Since one repeatedly observes the same ligand (as opposed to an entire compound library), there is no need for the incessant re-calibration of selective pulse schemes.

### 5.3. Paramagnetic relaxation enhancements

A variation of the exchange-averaged relaxation parameter approach is the Spin Labels Attached to Protein Side chains as a Tool to identify Interacting Compounds (SLAPSTIC) method developed by Jahnke and co-workers [46,47]. In this approach, one forces amplification of the bound state relaxation properties via covalently attached spin labels to protein side chains such as lysine, tyrosine, cysteine, histidine, and methionine. Such spin labels include paramagnetic organic nitroxide radicals such as TEMPO. In the context of Eq. (27a), one chemically enhances the second-order approach is the Spin Labels Attached to Protein

A distinct advantage is that the first ligand need not be expected to have enhanced ligand $R_1$ values and thus greater residual signal after the subtraction.

Another clever use of selective $R_1$ experiments is to monitor displacement of a ‘reporter’ or ‘probe’ molecule in competitive binding studies with higher affinity ligands [28]. Binding information comes from analysis of the same reporter molecule for a mixture of test molecules. Since one repeatedly observes the same ligand (as opposed to an entire compound library), there is no need for the incessant re-calibration of selective pulse schemes.

### 5.3. Paramagnetic relaxation enhancements

A variation of the exchange-averaged relaxation parameter approach is the Spin Labels Attached to Protein Side chains as a Tool to identify Interacting Compounds (SLAPSTIC) method developed by Jahnke and co-workers [46,47]. In this approach, one forces amplification of the bound state relaxation properties via covalently attached spin labels to protein side chains such as lysine, tyrosine, cysteine, histidine, and methionine. Such spin labels include paramagnetic organic nitroxide radicals such as TEMPO. In the context of Eq. (27a), one chemically enhances the second-order approach is the Spin Labels Attached to Protein Side chains as a Tool to identify Interacting Compounds (SLAPSTIC) method developed by Jahnke and co-workers [46,47]. In this approach, one forces amplification of the bound state relaxation properties via covalently attached spin labels to protein side chains such as lysine, tyrosine, cysteine, histidine, and methionine. Such spin labels include paramagnetic organic nitroxide radicals such as TEMPO. In the context of Eq. (27a), one chemically enhances $P_B$ to ensure $P_BQ_B \gg P_RQ_R$. The spin label endows the receptor with the strong paramagnetic relaxation mechanism of an unpaired electron. Compounds that bind in spatial proximity of the spin label now experience enhanced proton relaxation due the electron–proton DD interaction with the unpaired nitroxide electron. In the fast exchange limit, the paramagnetically enhanced $R_2$ relaxation is

$$R_{2,\text{avg}} = P_BR_{2,\text{para}} + (P_BR_{2,\text{B}} + P_RR_{2,\text{F}}) + R_{\text{ex}}$$

where $R_{2,\text{para}}$ is

$$R_{2,\text{para}} = \frac{h^2y_B^{-2}\mu_B^2}{8} \sum_{j=1}^{N} \frac{1}{J_j} \{4J_j(0) + 3J_j(\omega_B)\}.$$  

The ‘j’ sum in Eq. (43) runs over the $N$ spin labels in proximity of the ligand proton under consideration. $R_{2,\text{para}}$ refers to all other sources of relaxation such as the $^1$H–$^1$H DD mechanisms already described above for ‘conventional’ $R_2$. $R_{\text{ex}}$ is the same contribution seen in Eq. (26b). The power of SLAPSTIC lies in the comparatively colossal magnitude of the electron-proton DD interaction. Because $|\gamma_e/\gamma_B|$ is $\approx 658$, $R_{2,\text{para}}$ is effective over much longer interspin distances than typical for $^1$H–$^1$H relaxation. This gives the freedom to use considerably lower bound ligand fractions $P_B$ than would be possible in the absence of spin labels. Based on 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 [48].

Obviously, one relies on the availability of amino acid side chains amenable to spin labeling near the binding site. Moreover, one must ensure that the attachment of spin labels does not compromise the structural integrity or binding properties of the receptor. Thus SLAPSTIC requires considerable advance knowledge of the receptor’s 3D structure. This may limit the swiftness by which SLAPSTIC can attack novel targets. SLAPSTIC is also appealing for screening second-site binders. Specifically, one can attach a spin-label to known specific ligand. This altered first ligand then serves to screen for new ligands binding in a second proximal site. Again, one must confirm that the addition of the spin label does not compromise the receptor binding site. A distinct advantage is that the first ligand need not be present in saturating amounts owing to the dominance of $R_{2,\text{para}}$ [46].

### 5.4. $^{19}$F relaxation

As seen above, NMR screening experiments mainly focus on $^1$H. However, the $^{19}$F nucleus has a number of unique properties that render it a highly effective relaxation probe for NMR screening (see e.g. Peng [49]). First, quite indifferent to the needs of NMR spectroscopy, medicinal chemistry already regards $^{19}$F incorporation as an established tactic for enhancing the pharmacokinetic properties of leads (e.g. aromatic fluorines, tri-fluoromethyl groups). One therefore, does not meet the resistance typically encountered when advocating $^{13}$C, $^{15}$N, or $^2$H labeling. Second, the lack of endogenous $^{19}$F in biological molecules implies clean ligand selective 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 $\gamma_F \sim 0.94 \gamma_H$; the sensitivity of $^{19}$F NMR is therefore quite competitive with that of $^1$H. Finally, and most relevant for transverse relaxation methods, the chemical shift range of $^{19}$F ($\approx 900$ ppm) is much larger than that of $^1$H [50]. This implies high sensitivity of the chemical shift to changes in microenvironment. In contrast to $^1$H, one can expect large $\Delta \Omega = |\Omega_F - \Omega_H|$ associated with binding. In the context of Eqs. (26a) and (26b) this suggests significant $R_{\text{ex}}$ in spite of low $P_B$. Thus, the symptoms of binding can be more intense using $^{19}$F detection.

$^{19}$F relaxation is also useful for secondary screening experiments aimed at estimating exchange rates and equilibrium dissociation constants [51–53]. 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 [49].

The obvious drawback of $^{19}$F relaxation studies is the lack of ubiquity when compared to $^1$H. However, one can
still gain useful screening information 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 [54]. 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 methods

Saturation Transfer Difference (STD) spectroscopy has become the most popular method for screening due to small amounts of receptor required, the relative ease of implementation, and its compatibility with large therapeutic targets [55]. As its name suggests, STD takes the difference of two experiments. In a first experiment (‘on-resonance’ experiment), one selectively saturates the various receptor proton magnetizations ($M_z = 0$, and $\Delta M_z = M_z - M_0 = -M_0$) via a train of frequency-selective rf pulses. The rf train is applied to a frequency window that contains receptor resonances, but is devoid of resonances from the compounds (e.g. 0.0 to –1.0 ppm for proteins). The saturation propagates from the point of application 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 Fig. 5, binding compounds pick up this saturation via inter-molecular $^1$H–$^1$H cross-relaxation at the ligand–receptor interface. They then dissociate back into free solution where the saturated state persists due to the small free state $R_1$ values. At the same time, more ‘fresh’ ligand exchanges on and off the receptor while saturation energy continues to enter the system through the sustained application of rf. Thus, saturated free ligands accrue during the saturation time. One then records a complementary reference experiment (‘off-resonance’ experiment) that applies the identical rf train far off-resonance; thus, 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. This includes precisely the receptor and the binding compound resonances. The receptor resonances will not be visible on account of their minimal concentration or because of $R_2$ relaxation-filtering just prior to detection. The result is a simple 1D $^1$H spectrum that reveals only the binding compounds; this vastly simplifies data interpretation.

Fig. 6 depicts a typical STD pulse scheme. Typical saturation trains involve N repetitions of 50 ms frequency-selective pulses with Gaussian or Seduce-1 profiles [56,57]. 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° ($\Phi_1$) pulse (with optional

---

**Fig. 5. Detection of binders using the Saturation Transfer Difference (STD) experiment [55].** Frequency selective irradiation (lightning bolt) cause selective $^1$H saturation (shading) of the target receptor (e.g. protein, nucleic acid). The 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 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. Non-binding compounds are unaffected (stars).
antecedent water flip-back 90°, e.g. Seduce-1 pulse of 1–3 ms) then reads the z-axis magnetization. A subsequent spin-lock relaxation filters out residual receptor signals. The sequence finishes with a WATERGATE-5 block that simultaneously suppresses the water resonance while further relaxation-filtering away residual receptor signal. On-resonance saturation yields a spectrum of intensity $I_{\text{sat}}$, 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 $I_{\text{STD}} = I_0 - I_{\text{sat}}$.

Fig. 7 shows the generation of the difference spectrum. The fractional STD response, which is of the same form as the traditional steady-state NOE, is given by $\eta_{\text{STD}} = (I_0 - I_{\text{sat}})/I_0 = I_{\text{STD}}/I_0$ [57].

STD has several attractive features that deserve emphasis. First, STD is ideally suited to the large receptor masses (>30,000 Da) typically encountered in drug discovery. Large masses imply large rotational correlation times $\tau_c$ that enhance spin-diffusion and, hence, saturation transfer within the receptor and to the ligand. Second, STD requires relatively low amounts of receptor (~1 μM) when sufficiently rapid ligand exchange, the population of saturated ligand builds up during $T_{\text{sat}}$ due to the small free-state (i.e. unbound) $R_l$ values, and the sustained application of the rf train. Thus, a small amount of saturated receptor can produce an amplified amount of saturated ligand. A third advantage of STD is that one observes only signal from the bound state of the hits. 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 one considers the very high ligand-to-receptor ratios $\varepsilon = L/T_E \gg 1$. Fig. 8 depicts an example from our own laboratory using the same receptor–ligand system as in Fig. 4. As stated, the STD spectrum (lower trace) reveals only the binding compound; this obviously simplifies data reduction.

To better quantify the STD amplification effect, Mayer and Meyer have introduced an ‘STD amplification factor’ $A_{\text{STD}}$. The STD response results from the receptor–ligand complex and is therefore, proportional to [EL]. Hence, $I_{\text{STD}}$ can be written as $I_{\text{STD}} = C_0 I_{\text{STD}}[\text{EL}]$, where ‘C’ is a proportionality constant that makes the appropriate unit conversions, and $\sigma_{\text{STD}}$ is a dimensionless scaling factor that represents the maximum STD amplification. Note that $I_{\text{STD}}$ corresponds to $Q_{\text{FEG}}$ in Eq. (27a) since there is no contribution from [L] (i.e. $Q_T = 0$). The reference, or equilibrium intensity, $I_0$ is just proportional to $L_T$; thus, $I_0 = C_L T_{\text{STD}}$. Then the ratio $I_{\text{STD}}/I_0 = \sigma_{\text{STD}}[\text{EL}]/L_T = \sigma_{\text{STD}} P_B$. Since $P_B = P_{\text{H}}/\varepsilon$, where $\varepsilon = L/T_E$, we have

$$A_{\text{STD}} = \varepsilon \eta_{\text{STD}} = \varepsilon (I_{\text{STD}}/I_0) = \frac{\sigma_{\text{STD}}[\text{L}]}{[\text{L}] + K_D}.$$  

Eq. (44) is just the hyperbolic dose–response curve given in Eq. (28), and seen previously for the Henri-Michaelis-Menten enzymatic reaction rate $v_0$ (cf.Eq. (7)). Therefore, $A_{\text{STD}}$ will act in an analogous manner to $v_0$: it will increase with increasing $L_T$ until one reaches maximum amplification $\sigma_{\text{STD}}$ when the receptor binding site is saturated ($L_T \gg K_D$). Of course, continued increase of $L_T$ will monotonically decrease the fractional response $\eta_{\text{STD}} = I_{\text{STD}}/I_0$. In principle, if $L_T$ well approximates [L], one can fit ligand titration data to the form of Eq. (44) to estimate $K_D$ and $\sigma_{\text{STD}}$.

Following the enzyme analogy, $A_{\text{STD}}$ 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 $A_{\text{STD}} = 10$. This implies that a receptor concentration of 50 μM yields an effective saturated ligand concentration of 500 μM. Of course, the latter concentration is more than sufficient for sensitive detection by current high-field magnets and probes. Another motivator for using $A_{\text{STD}}$ is that spectra from samples having different receptor concentrations may still be compared. This would be salient for titration and competition.
experiments that may involve samples with variable amounts of receptor (vide infra).

The $K_D$ range of the STD method has been estimated to be $10^{-3} > K_D > 10^{-8}$ M$^{-1}$ [55]. For weak binders having $K_D > L_T$, over half of the receptor molecules have no ligand in the binding site. As $K_D$ increases further, the population of ligand–receptor complex [EL] decreases; this leads to the diminishment and ultimate disappearance of the STD signal. In the case of strong binders, decreasing $K_D$ increases the receptor–ligand lifetime, $k_{off}$, 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. The exchange is then so slow that the free ligand magnetization ‘forgets’ its visit to the saturated receptor, and relaxes back to equilibrium at a faster rate than the receptor is able to ‘turn over’ a newly saturated ligand. The population of saturated ligands diminishes and ultimately disappears, thus killing the STD signal.

Recent STD applications illustrate its versatility and scope. For example, Benie et al. have used STD to identify ligands targeting HRV2 (human rhinovirus serotype 2), a macromolecular complex of $8.5 \times 10^6$ Da [58]. 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 [59]. Other examples of exotic targets include small RNA fragments [60] and macromolecules immobilized on solid-support beads [61].

The intrinsic sensitivity of the STD experiment is limited by the efficiency of the intended spin energy transfer.

---

Fig. 7. Schematic depicting difference spectroscopy in the STD experiment. Circles and stars indicate binding and non-binding compounds, respectively. STD involves two experiments: an off-resonance and 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 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).
pathway from receptor (source) to ligand (recipient). Fig. 9 summarizes some of the principle pathways to consider. The main factor limiting this transfer at the recipient (ligand) end is the ligand $R_1$. Since the ligands are usually low mass compounds ($<1000$ Da), the free state $R_1$ values are small and therefore, any non-equilibrium magnetization state (like saturation) dissipates quite slowly. In contrast, when bound to the receptor, the ligand $R_1$ can be much larger as explained above (cf. Eq. (41)). Hence, efficient spin energy transfer requires the ligand to dissociate from the receptor at a rate faster than the bound state $R_1$.

The source (i.e. receptor molecule) receives a constant influx of energy by the applied rf saturation train; in fact, the constant energy input is what enhances the sensitivity of STD over other methods like transferred-NOE. One often assumes 100% receptor saturation shortly after application of the rf train due to intramolecular spin diffusion. In reality, the actual extent of receptor saturation depends on the competition between this influx and the various $R_1$ relaxation and/or 'leakage' mechanisms. Jayalakshmi and Rama Krishna have recently emphasized the importance of exchange-mediated leakage [62]. Receptor protons at the ligand–receptor interface cross-relax with unsaturated ligand and solvent protons as well as other distinct saturated receptor protons. While saturation enters from the other receptor protons, it leaks away via the exchanging ligand and solvent protons. In particular, 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$_2$O versus D$_2$O [60]. The overall STD response is significantly less in H$_2$O due to the additional DD interactions between the RNA protons and hydration waters.

Another factor that can compromise saturation is low molecular mass of the receptor. 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. One can compensate by applying longer saturation trains. Alternatively, one can add viscosity-enhancing reagents (e.g., glycerol) and/or screen at lower temperatures, which can increase the effective $\tau_c$. If such remedies are not feasible, then $R_2 = 1/T_2$ relaxation filtering methods as described above (Section 5.1) may prove more sensitive than STD [42].

Finally, weaknesses in the $^1\text{H} - ^1\text{H}$ cross-relaxation network can compromise the 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. Clearly, targets that have inherently low proton density are suboptimal for STD. Instead, the following method may be more effective.

### 5.6. WaterLOGSY

A method closely related to STD is Water–Ligand Observed via Gradient SpectroscopY (waterLOGSY) [63, 64]. Like STD, the receptor–ligand complex is selectively ‘tagged’ through a selective rf pulse scheme. However, waterLOGSY tags indirectly by selective perturbation of the bulk water magnetization as opposed to directly perturbing receptor magnetization, as in STD. The intended transfer of energy is therefore, water $\rightarrow$ receptor $\rightarrow$ ligand. The original presentation of waterLOGSY proposed both selective saturation and inversion of the water resonance [63]. Selective water inversion may be achieved by either [sel-90(+$\pi$/2)-hard 90+$\pi$], as in the WEXII method of Mori et al. [65, 66], or by [90+$\pi$-sel-180(+$\pi$/+$\pi$/2)-90+$\pi$], as in the e-PHOGSY method of Dalvit [67]. The most recent publications favor selective inversion via e-PHOGSY; henceforth, our discussion will emphasize this case [64].

Inverted water magnetization transfers to the bound ligands via three simultaneous strategies schematized in Fig. 10. One strategy involves direct $^1\text{H} - ^1\text{H}$ cross-relaxation between the bound ligand and ‘bound’ water molecules within the binding site. Such water molecules are bound in the sense that their receptor residence times exceed the receptor rotational correlation time. As a result, the rate constants governing DD inter-molecular cross-relaxation between the bound ligand and water molecules are negative and tend to bring the ligand magnetization to the same inverted state as the water. A second strategy is direct cross-relaxation with exchangeable receptor NH/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 strategy involves indirect cross-relaxation with remote exchangeable NH/OH protons via spin diffusion. The inverted magnetization is then relayed to other non-labile spins via spin-diffusion. Thus, NH/OH protons remote from the binding site act as entry points for widespread spin inversion throughout the receptor.

The above magnetization transfer schemes allow binding compounds to pick up the bulk water inversion while residing in the receptor binding site. They then dissociate into free solution where they maintain their perturbed magnetization state due to their small free-state R1 values. The smaller these 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 can add to the growing pool of spin-inverted ligands.

Binding compounds are distinguished from non-binding compounds by their differential cross-relaxation properties with water. In the magnetization transfer scheme above, the binders interact directly or indirectly with inverted water spins via DD interactions with sufficiently long rotational correlation times τc (on account of being associated with the large receptor) to yield negative cross-relaxation rates. By contrast, non-binders’ DD interactions with water have much shorter τc leading to positive cross-relaxation rates. As a consequence, binders and non-binders display water-LOGSY peak intensities of opposite sign, thus providing an easy means for distinction.

Fig. 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 [67]. The selective 180° refocusing pulse is typically 5–20 ms long using e.g. REBUP amplitude modulation [68]. Concerns about inverting receptor protons having chemical shifts degenerate with that of bulk water (e.g. protein α protons) are alleviated by setting the spin-echo delay sufficiently long to filter out receptor coherences [69]. Phase cycling of the selective 180° pulse (∆FB) serves to place water magnetization alternately on the + z and − z axis. In the − z case, inverted water magnetization transfers to the compounds via the routes described above during Tinv. In the + z case, the water magnetization is nearly at equilibrium and so nothing happens. Toggling the receiver phase in concert with that of the selective 180° pulse yields a difference spectrum between the + z and − z water conditions. The resultant signal, IWL, represent only those resonances having experienced magnetization transfer from water. Gradients are applied throughout Tinv to prevent premature return of water magnetization to the equilibrium condition via radiation damping. Large magnetizations or magnetizations with sharp lines relaxing during Tinv (e.g. free ligands) can lead to artifacts in the spectrum. The unshaded pulse in the middle of Tinv (in Fig. 11) is an optional non-selective 180° pulse that minimizes these artifacts by keeping the large magnetizations close to their original null condition [70]. Because the non-selective 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 both 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 (e.g. one can use more closely spaced gradients of moderate strength). Experimental conditions that decrease the water R1 (increase T1) are desirable.

![Fig. 11. Example of the preferred waterLOGSY pulse sequence of Dalvit et al. [64]. The water resonance is selectively inverted on alternate scans by 90° phase shifts of the selective 180° refocusing pulse (e.g. 20 ms REBUP pulse [68]) between the delays δWL. To reduce inversion of receptor protons having chemical shifts degenerate with H2O, δWL 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 Trec (typically 1–3 s). A 1 ms crusher gradient (e.g. 40 Gauss/cm) followed by a sustained weak gradient (e.g. 0.3 Gauss/cm) is applied during Tinv to minimize radiation damping effects. Phase cycling is as follows: ∆FB = x, y, −x, −y; ∆FRec = x, x, y, x; ∆F1 = 4x, 4y, 4−x, 4−y; ∆F2 = (2x, −x, y, −y)/(2x, −x, y, −y). Pulses without explicit phase labels are along + x. The unshaded pulse in the middle of the Trec period is an optional non-selective composite 180° pulse that serves to minimize magnetization relaxing during Trec. The phase of ∆FRec depends on the sign of the relaxed residual H2O magnetization at the end of Trec. An optional R1 spin-lock (not shown) as in Fig. 6 can be applied just after the third hard 90° pulse (with phase ∆F1) for further relaxation filtering of the receptor.](image-url)
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 stated above, receptors having inherently poor proton density will experience inefficient spin diffusion and thus poor STD sensitivity. 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 clearly by Fig. 12, which compares the signals of the waterLOGSY and ‘traditional’ STD experiments run for the same amount of time on a known binder of the ribozyme target.

Like all ligand-based screening experiments, waterLOGSY is biased towards the detection of weakly-binding ligands. Ligands with tighter affinity will have correspondingly longer residence times. 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 $K_D \approx 0.1 \mu M$ [64].

Unlike STD, waterLOGSY spectra reflect both free and complexed states of a hit. In the fast exchange limit, the waterLOGSY signal can be expressed as a weighted average

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

‘C’ is a proportionality constant that accounts for the appropriate unit conversions, and $\sigma_{bound}$ and $\sigma_{free}$ 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 $\sigma_{bound}$ and $\sigma_{free}$ reflects a first-order approximation appropriate for shorter inversion times $T_{inv}$. Note that $\sigma_{bound}$ and $\sigma_{free}$ have the NOE cross-relaxation form of Eq. (48a) given below in Section 5.7. For $\sigma_{free}$, the rotational correlation time is that of a small free compound ($\tau_{free} \ll 1$ ns). For $\sigma_{bound}$, there is an effective rotational correction time $\tau_{eff} = \tau_{res} \tau_p/(\tau_{res} + \tau_p)$, where $\tau_{res}$ is the ligand residence time and $\tau_p$ is the rotational correlation time of the receptor–ligand complex. Typically, $\tau_{eff} \gg \tau_{free}$ leading to opposite signs for $\sigma_{bound}$ and $\sigma_{free}$. 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 means that if $I_T \gg I_F$, then the negative contribution from the free state $[L] \sigma_{free}$ will overwhelm that of the bound state, resulting in a false negative. To avoid such confusion, one can avoid large ligand-to-receptor molar ratios. One can also record a reference spectrum, $I_{WL,free}$ of the compounds in the absence of receptor. 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. (28). In particular, if we identify $Q_{avg} \approx I_{WL}$ and $Q_F \approx I_{WL,free}$, then

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

The above expression differs trivially from that of Dalvit et al. [64] in that it is based on the bound receptor fraction $P_B^R = [EL]/E_T$ whereas Dalvit and co-workers consider the free receptor fraction, $(1 - P_B^R)$. Fitting $I_{WL} - I_{WL,free}$ in a ligand titration to Eq. (46) 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 Serine Albumin (HSA). Their results are of the same order of magnitude as those estimated by equilibrium dialysis. Similar results, albeit, with larger uncertainties, were obtained by fitting ‘uncorrected’ titration data (just $I_{WL}$) directly to Eq. (45).

5.7. Exchange-transferred NOE

The DD cross relaxation rate constants $\sigma^{NOE}_{average}$ (labouratory frame) and $\sigma^{ROE}_{average}$ (rotating frame) between two protons ‘i’ and ‘j’ under fast exchange become the population-weighted averages

$$\sigma^{NOE}_{avg} = P_F \sigma^{NOE}_F + P_B \sigma^{NOE}_B.$$  

\[\text{(47a)}\]
\[ \sigma_{\text{avg}}^{\text{ROE}} = P_F \sigma_{F}^{\text{ROE}} + P_B \sigma_{B}^{\text{ROE}}. \] (47b)

Of course, these averages are the basis for the well-known exchange-transferred NOE and ROE experiments. \( \sigma_{F}^{\text{NOE/ROE}} \) and \( \sigma_{B}^{\text{NOE/ROE}} \) are the intrinsic free and bound cross-relaxation rates. In terms of spectral density functions, they have the form

\[ \sigma_{ij}^{\text{ROE}} = \frac{kT}{4} \frac{1}{\rho_{ij}} \{6J_{ij}(2\omega_{t}) - J_{ij}(0)\} \] (48a)

\[ \sigma_{ij}^{\text{NOE}} = \frac{kT}{4} \frac{1}{\rho_{ij}} \{2J_{ij}(0) + 3J_{ij}(\omega_{t})\}. \] (48b)

Both cross-relaxation rates also have strong dependencies on \( \tau_c \) that can be exploited as screening probes. The \( \sigma_{\text{NOE}}^{\text{ROE}} \) cross-relaxation rate constant is particularly useful since it flips sign when the ligand hops from the free state to the bound state. This can be seen in Eq. (48a). In the free state, we have small \( \tau_c \) satisfying \( \omega_{t}\tau_c \ll 1 \), \( J_{ij}(2\omega_{t}) \approx J_{ij}(0) \) and \( \sigma_{ij}^{\text{NOE}} \approx +5J_{ij}(0) \). In the bound state, we have large \( \tau_c \) satisfying \( \omega_{t}\tau_c \gg 1 \), \( J_{ij}(2\omega_{t}) \approx 0 \), and \( \sigma_{ij}^{\text{NOE}} \approx -J_{ij}(0) \). By contrast, \( \sigma_{ij}^{\text{ROE}} \) does not flip sign, and thus, is less definitive.

Exchange-transferred \( \sigma_{\text{NOE}}^{\text{ROE}} \) measurements were among the first parameters proposed for ligand-based NMR screening [8,71,72]. In NOE-based screens, one monitors intra-ligand NOEs of compound mixtures via 2D-NOESY spectra in the absence and presence of the receptor. Binders are identified by NOE cross peaks that have flipped sign upon introduction of the receptor. Non-binders 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 \( \sigma_{\text{NOE}}^{\text{ROE}} \) is 100 nM \( \leq K_D \leq 1 \text{ mM} \) [72].

Exchange-transferred \( \sigma_{\text{NOE}}^{\text{ROE}} \) 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 intraligand magnetization transfer. By contrast, STD and waterLOGSY rely on a long period intermolecular magnetization transfer between receptor and ligand that is sustained by the continuous application of rf irradiation or the long \( T_1 \) of water protons. This sustained period of magnetization transfer prompts the growth of a large pool of ligands each ‘labeled’ with a binding signature. This heightens the sensitivity of these experiments over the 2D transferred NOESY. Furthermore, STD and waterLOGSY are relatively simple 1D experiments that reduce acquisition time and data storage burdens and increase throughput. In our experience, the same binding information can be gleaned 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 \( \sigma_{\text{NOE}}^{\text{ROE}} \) as a screening tool. Nevertheless, the change of sign inherent in \( \sigma_{\text{NOE}}^{\text{ROE}} \) still makes it an effective screening tool when other screening methods give ambiguous answers. And, of course, \( \sigma_{\text{NOE}}^{\text{ROE}} \) retains its original value as a method for determining the bioactive conformations of weakly binding ligands.

### 6. Competition binding studies

Competitive binding studies are part of general biochemical tactics for determining binding specificity and affinity. The basic expressions important for such studies have been given above in Section 2.2 and Section 2.3. There has been a recent resurgence of competition binding studies in ligand-based NMR screening efforts. Following the emerging literature, we use ‘I’ and ‘L’ to denote two ligands that compete for binding to the same receptor site. Competitive displacement by a known specific ligand ‘I’ can be used to confirm the specific binding of new ligand ‘L’ uncovered during primary screening. Furthermore, if one knows the inhibitor dissociation constant, \( K_I = [E][I]/[EI] \), then one can estimate the \( K_D \) of ‘L’ by using Eq. (16) above. Meyer and Mayer have demonstrated this approach using STD on two galactose-containing ligands (\( \beta \)-GalOMe and NA2) of the 1,20,000 Da tetramer, \( Ricinus communis \) agglutinin. The \( K_D \) of NA2 was estimated to be 27 \( \mu \text{M} \) from a competition study with \( \beta \)-GalOMe (\( K_I = 260 \text{ mM} \)).

Competitive binding studies can also form the basis for a screening strategy that enables the detection of high-affinity ligands using ligand-based methods. As forementioned, a glaring limitation of ‘normal’ ligand-based NMR screening is its reliance on fast exchange and its consequent limitation to weak binders (\( K_D \) typically \( > 100 \mu \text{M} \)). Tighter binders 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, one would observe only free state properties for the ligand and mistake the tight binder for a non-binder. The two-state exchange behavior of transverse relaxation \( R_2 \) provides a good example of this scenario. In the slow exchange limit given by Eqs. (25a) and (25b), the ratio of bound to free peak heights is proportional to \( (R_{2F} + P_Bk_{ex})(R_{2B} + P_Bk_{ex}) \). Given that \( P_B \ll P_F \), and \( R_{2B} \gg R_{2F} \), the bound state signal will be practically unobservable compared to the free state signal. Additionally, since \( P_Bk_{ex} \ll 1 \), the free state \( R_{2F\text{avg}} \) will scarcely deviate from \( R_{2F} \). This increases greatly the probability that \( R_2 \)-based screening experiments will observe only the ‘free’ peak and draw a false negative.

Recently, several groups have proposed essentially identical competitive binding strategies to extend the range of ligand-based NMR screening to include high-affinity binders, and thereby avoid false negative scenarios like the one just described [27–29]. This strategy screens for higher affinity hits ‘I’ via the competitive ‘knock-off’
effects they exert on a previously characterized lower affinity ligand ‘L’. Such competition strategies have already been in wide use in isothermal titration calorimetry (ITC) studies, which are constrained by similar binding affinity windows [19,73]. Below, we highlight the principle steps of the strategy [28].

1. First, one identifies a known ligand deemed the ‘reporter’, ‘reference’, ‘spy’, or ‘probe’ compound ‘L’. The ideal phenotype includes high solubility, medium to weak binding affinity \( (K_D > 10 \mu M) \), 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 water-LOGSY 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 [54].

2. Next, one establishes a calibration curve that relates an observed magnitude of the monitored NMR parameter of the reporter molecule to the bound reporter compound fraction, \( P_B = [EL]/L_T \). One can measure this NMR parameter for the reporter compound in a titration of either \( L_T \) or \( E_T \) (to vary \([EL]/L_T\)). One should also have an independent estimate of the reporter compound \( K_D = [E][L]/[EL] \) (e.g. using ITC). Then, plugging \( K_D \) 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 \( P_B \). One then plots the observed NMR parameter as a function of the reporter compound \( P_B \). The result is a calibration curve that can relate subsequent perturbations of the monitored NMR parameter to altered \( P_B \).

3. For fixed values of \( E_T \), and reporter compound \( L_T \), one can now add a mixture of test compounds to screen for displacement effects exerted on the reporter compound. Displacement is manifested as a shift of the reporter compound’s NMR parameters towards those intrinsic of 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. (13)–(20).

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 \( P_{B,+1} = [EL]/L_{T} \).

5. Since \( P_{B,+1} \), \( L_T \), and \( E_T \) are known, one has sufficient information (i.e. \( L_T, E_T \), and \([EL]\)) to estimate the value of \( K_{D,\text{app}} \) for the reporter compound using Eq. (15). Once \( K_{D,\text{app}} \) is fixed, Eq. (15) 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.

The competitive binding strategy is appealing because it not only identifies novel higher affinity binders, but also estimates their affinities \( (K_I) \) via a one-point measurement. Exhaustive ligand titrations are not always feasible due to poor compound solubility or receptor damage caused by continued additions of concentrated DMSO stock solutions. 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 for the tight binders obtained from the ‘one-point’ measurement extend to the nM range, and correlate well with concurrent ITC studies [28]. More recently, they have demonstrated the power of \(^{19}\)F NMR (the FAXS method) for these studies using with a small \(^{19}\)F-based library to serve as ‘reporter’ and ‘control’ molecules [54]. The control molecules are known non-binders 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 \( R_2 \) relaxation enhancements than \(^{1}H \). In particular, the inherently larger chemical shift span 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.

7. Addressing basic problems: aggregation and non-specific binding

A basic challenge in ligand-based screening includes distinguishing the spectral signatures of bona fide binding from similar ones stemming from artifacts like aggregation and non-specific binding. As with ligand binding, compound aggregation increases the effective molecular mass, thus driving \( \tau_c \) upwards and enhancing the \( J(0) \)-dependent relaxation rates. An easy NMR check for aggregation is by measurement of any \( \tau_c \)-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 \( B_0 \sim 11.7–14.0 \) Tesla. Of course, one could use any other biophysical methods (e.g. light scattering) to assess aggregation tendencies.

Another difficult problem is non-specific binding. We should distinguish between the often-seen binding classifications ‘non-specific’ versus ‘low-affinity’. Here, we consider low-affinity binders to be ligands with \( K_D \) values \( > 10 \mu M \) and in fast exchange on the chemical shift time scale. ‘Low-affinity’ binders are therefore a more general class of compounds than ‘non-specific’ binders. In particular, non-specific binders may be low-affinity binders, the converse need not be true. 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). In contrast, non-specific binders bind to receptor surface regions that have no direct effect on receptor activity.

Non-specific 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 non-specific binding sites. Presumably, these non-specific binding processes correspond to adsorption to hydrophobic patches on the protein surface. High ligand concentrations can increase the likelihood of this non-specific surface adsorption. Accordingly, Murali et al. have used dilution studies to expose non-specific binding [74]. Specifically, they revealed non-specific binding by diminishment of the NMR binding signature as the absolute concentrations were decreased while the ligand-to-receptor ratio were maintained at a constant value. Undoubtedly, the best way to expose non-specific binding is to test for displacement effects as described above upon the addition of a known specific and competitive binder. 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 are judiciously placed spin labels in the SLAPSTIC approach [47]. Alternatively, one can introduce selective protonation of active site residues in otherwise deuterated receptor proteins. STD experiments on such systems provide signal only for those compounds that bind to the active site [75]. In both cases, one needs a priori information about the active site structure. And in the last example, one relies on the facile over-expression of deuterated receptor. To reduce the risk of non-specific binding, one can work at lower receptor concentrations. By keeping E2 low, only the higher affinity binders (those that can saturate the binding site with minimal L/2 due to low K0) will have bound ligand fractions of significance.

8. Beyond binding: epitope mapping via ligand-based approaches

To further accelerate lead generation, NMR spectroscopists 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 lacking (e.g. transmembrane proteins). A knowledge of which parts of the ligand are involved in the binding interface can obviously help chemists make informed decisions as to how binding scaffolds should be elaborated, linked, or both.

8.1. Group epitope mapping

Mayer and Meyer have developed the Group Epitope Mapping (GEM) protocol for discerning binding surfaces on the ligand using STD methods [57]. This approach compares the STD response I STD = I0 - I sat for different proteins within a ligand. In 1D-STD spectra, this amounts to normalizing the various I STD to the largest I STD response within the ligand. For 2D TOCSY-dispersed STD spectra, one carries out analogous normalization for peak volume integrals. In either case, one interprets the variation of normalized STD responses in terms of ligand proton proximity to the receptor. Stronger STD responses are interpreted as evidence of closer contact between ligand and receptor protons. This interpretation acknowledges the distance dependence of the inter-molecular NOE underlying the saturation transfer process. To establish proof-of-concept, Mayer and Meyer performed a GEM analysis of galactose-containing ligands (NA2 and β-GalOMe) of the 1,20,000 Da Ricinus communis lectin agglutinin I. The results were consistent with the known binding regions of galactose-containing ligands using other methods.

An important GEM caveat is that the mean lifetime of the receptor–ligand complex, k-1 off, must be sufficiently short (ligand K0 sufficiently large). Otherwise, spin diffusion within the bound state will tend to 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.

Another very important caveat has recently been underscored by Jayalakshmi and Rama Krishna [62]. Specifically, the authors emphasize that an interpretation of differential STD responses purely in terms of inter-molecular cross-relaxation (NOE) is an oversimplification. Indeed, so beguiling is the description of the STD as a screening experiment that one often forgets that it is simply a 1D truncated, driven NOE (TOE) experiment[76]. This being the case, it is useful to review the TOE behavior of a simple system of two unlike dipolar-coupled spins, ‘I’ and ‘S’. Suppose we saturate spin ‘S’ at time T sat = 0, and observe the effects on spin ‘I’. Then the time-evolution of (I2) is

\[
(I_2)(T_{sat}) = I_0 + \{1 - \exp(-R_{II}T_{sat})\}(S_0g^{NOE}/R_{II}).
\]  

At long times, T sat → ∞, the steady-state difference (I0 - I(∞)) is

\[
(I_0 - I(\infty)) = -S_0g^{NOE}/R_{II}.
\]

Thus, the magnitude of the steady-state response of Eq. (50) reflects a competition between both cross-relaxation (S0gNOE in the numerator) and longitudinal auto-relaxation (R II in the denominator). If R II is large, \( I_2 \) will approach the steady state more rapidly and with a steady-state intensity of lesser magnitude.
Obviously, interpreting $(I_0 - I_s(\infty))$ purely in terms of cross-relaxation (the numerator) would be inaccurate.

The STD experiment is the TOE experiment applied to a much larger system of proton spins that experience two-state exchange in addition to relaxation. To see how auto-relaxation enters the STD intensity, we essentially seek the multi-spin version of Eq. (50). Jayalakshmi and Rama Krishna have recently done this using a complete relaxation and exchange matrix approach (CORCEMA) [77], and we sketch their results for the fast-exchange limit. Specifically, we start from the modified Bloch Eqs. (20), and separate out those receptor spins $S$ from $\Delta M_z(t)$ that are kept saturated and put them in a constant forcing term $Q_{\text{sat}}$. Eq. (20) then becomes

$$\frac{d}{dt}(\Delta M_z) = -(R + K)(\Delta M_z) + Q_{\text{sat}}.$$  

(51)

For STD, we are primarily interested in the steady-state solution, which is

$$I_{\text{STD}} = -\langle \Delta M_z(\infty) \rangle = \langle M_0 - M_z(\infty) \rangle = -(R + K)^{-1}Q_{\text{sat}}.$$  

(52)

When the ligand binding exchange is fast with respect to both $R$ and $\Omega$, the exchange matrix $K$ dominates and $R$ can be treated as a small perturbation [77,78]. This reduces the dimensionality of Eq. (51). Working in the eigenbase of $K$, then $R + K$ can be replaced by the exchange-averaged relaxation matrix $(R)_e$, where the angled brackets indicate the result of fast exchange-averaging. The steady-state solution simplifies to

$$I_{\text{STD-fast}} = -\langle \Delta M_z(\infty) \rangle = \langle M_0 - M_z(\infty) \rangle = -(R)^{-1}(Q_{\text{sat}}),$$  

(53)

where

$$\langle M_z \rangle = \begin{bmatrix} M_{\text{el}} + M_{\text{el}'} \\ M_{\text{el}'} + M_{\text{el}}' \end{bmatrix},$$  

$$\langle R \rangle = \begin{bmatrix} p_F R_{\text{el}} + p_B R_{\text{el}}' \\ p_B R_{\text{el}}' + p_F R_{\text{el}}' \end{bmatrix},$$  

$$\langle Q_{\text{sat}} \rangle = \begin{bmatrix} R_{\text{el}} S_0' \\ R_{\text{el}'} S_0 + R_{\text{el}'} S_0' \end{bmatrix}. $$  

(54)

In the notation of Jayalakshmi et al.[62] and Moseley et al.[77], the ‘$L$’, ‘$E$’, and ‘$LE$’ subscripts refer to the free ligand, free receptor, and receptor–ligand complex, respectively, while superscript primes indicate the bound state. The ‘elements’ of $\langle \Delta M_z \rangle, \langle Q_{\text{sat}} \rangle, \langle R \rangle$, are themselves sub-vectors and sub-matrices. For example in $\langle R \rangle$, the exchange-averaged relaxation rate matrix, $R_{\text{el}}$ and $R_{\text{el}'}$ are $m \times m$ matrices describing ligand auto- and cross-relaxation in the free and bound states. $R_{\text{el}}$ and $R_{\text{el}'}$ are the $n \times n$ analogs describing receptor proton relaxation. $R_{\text{el}'}(n \times m)$ and its transpose $R_{\text{el}'}^T(m \times n)$ describe the ‘interesting’ STD cross-relaxation-intermolecular ligand–receptor cross-relaxation. In accordance with fast exchange on the chemical shift time scale, the vectors $\langle M_z \rangle$ and $\langle Q_{\text{sat}} \rangle$ sum the free and bound magnetizations. The ‘forcing term’ $\langle Q_{\text{sat}} \rangle$ contains the matrices $R_{\text{el}S}, R_{\text{el}S'}$, and $R_{\text{el}S'}$. These reflect STD cross-relaxation between the saturated receptor protons and the other (unsaturated) receptor protons in the free state $(R_{\text{el}S})$, cross-relaxation between the saturated receptor protons and the unsaturated ligand protons in the bound state $(R_{\text{el}S'})$, and cross-relaxation between the saturated receptor protons and the other (unsaturated) receptor protons in the bound state $(R_{\text{el}S'})$. The vector $S_0'$ represents the equilibrium magnetization of the bound-state saturated receptor protons.

The STD solutions (Eqs. (52) and (53)) above are the multipih analogies of Eq. (50). Note that $Q_{\text{sat}}$ plays the role of $S_0\rho^{\text{NOE}}$ given in Eq. (50). In particular, it contains the intermolecular cross-relaxation rates motivating the use of STD for epitope mapping. Its elements consist of the cross-relaxation rate constants between the saturated receptor protons and the ligand. These cross-relaxation rate constants increase in magnitude with closer receptor–ligand contact; thus lending credence to the concept that ligand protons involved in the receptor–ligand interface yield greater $I_{\text{STD}}$.

However, the ‘denominator’ $R$ or $(R)$ contributes as well. While $R$ contains the ‘interesting’ inter-molecular cross-relaxation rates, $R_{\text{el}'}$, it also contains the lossy auto-relaxation rates $R_{\text{el}}$ and $R_{\text{el}'}$. Large auto-relaxation elements in $R$ due to large elements in the bound state $R_{\text{el}}$ will tend to reduce $I_{\text{STD}}$. Therefore, as with the single spin case, Eqs. (52) and (53) show that the STD response reflects a competition between the two effects: (i) cross-relaxation with the saturated receptor protons powered by the applied rf, signified by $Q$ or $\langle Q_{\text{sat}} \rangle$ in the numerator and; (ii) auto-relaxation that would dissipate energy to the lattice, signified by $R$ or $(R)$ in the denominator.

Thus, GEM maps not only intermolecular cross-relaxation, but the local $R_{\text{el}}$ values as well. As a consequence, one must exercise caution when interpreting smaller STD responses within a ligand. Smaller $I_{\text{STD}}$ values do not necessarily imply remote locations from the receptor; rather, they may simply imply larger $R_{\text{el}}$ values. Clearly, confusion of these two situations could lead to erroneous binding epitopes. The individual ligand proton $R_{\text{el}}$ values depend on the proton moiety considered (e.g. methine, methylene, etc.), local proton density, mobility, and the ligand conformation(s). Methylene protons are apt to show smaller STD responses due to the strong mutual dipolar $R_{\text{el}}$ relaxation between the proximal geminal protons. Quantitative interpretations of differential STD effects should therefore be treated with caution. If all of the ligand protons fortuitously have similar $R_{\text{el}}$ values, then these concerns are not important. Alternatively, if one could measure the appropriate $R_{\text{el}}$ values, then the product of the STD intensities and the $R_{\text{el}}$ values should be proportional to the effective inter-molecular cross-relaxation rate constant. We note 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.

### 8.2. Diffusion-based epitope mapping

Yan et al. [79] have proposed an alternative scheme for epitope mapping that uses the popular Bipolar Pair Pulsed field gradient STimulated Echo experiment (BPP-STE), which is often used to measure translational diffusion coefficients [80–83]. This experiment monitors translational molecular diffusion by placing the relevant magnetization along the rotating frame $z$-axis. The final signal intensity $I(G)$ is

$$I(G) = I_0 \exp \left[-D_t F^2 G^2\right]. \quad (55)$$

$D_t$ is the molecular (translational) diffusion coefficient and $G$ is the gradient strength. $F^2$ varies with the type of diffusion experiment chosen; for the BPP-STE experiment

$$F^2 = \frac{\gamma^2}{\delta^2} \left( T_{\text{diss}} - \delta / 3 - \tau / 2 \right). \quad (56)$$

The proton gyromagnetic ratio is $\gamma_H$, the gradient pulse length is $\delta$, the gradient recovery time is $\tau$, and the diffusion time is $T_{\text{diss}}$. To measure $D_t$, one records a series of 1D experiments that vary the gradient strength $G$. The resulting peak intensities $I(G)$ can then be fit as a function of $G^2$ to extract $D_t$. In particular, one can exploit the fact that

$$\ln(I(G))/F^2 = \ln I_0/F^2 - D_t G^2. \quad (57)$$

A plot of $\ln(I(G))/F^2$ versus $G^2$ should yield a straight line with negative slope $D_t$. Note that the diffusion time is fixed; only $G$ is varied. Thus, if the decay of peak intensity reflects only translational diffusion, then all ligand resonances should yield the same $D_t$.

Recent studies by Chen and Shapiro have observed the contrary [84]. Some resonances show deviations from the expected behavior of Eq. (57) for long $T_{\text{diss}}$. This gives $D_t$, an apparent dependence on the value chosen for $T_{\text{diss}}$. This impossibility points to a process unaccounted for during $T_{\text{diss}}$. Chen and Shapiro have identified this process as intermolecular NOE cross-relaxation. Specifically, during $T_{\text{diss}}$ of the BPP-STE experiment, the receptor and ligand magnetization lies along the $z$-axis. The magnitude and sign of magnetization varies along the $z$-axis according to the gradient-encoded spatial label $\exp(\gamma G Z / 2)$. During this time, $R_1$ auto-relaxation and NOE cross-relaxation occur. The $R_1$ relaxation destroys the spatial labels. However, intermolecular NOE can exchange the spatial labels. As a result, ligands can adopt the spatial labels of the large receptor and thus appear to diffuse more slowly. The result is an apparent dependence of the diffusion coefficient on $T_{\text{diss}}$, with smaller apparent diffusion coefficients observed for longer $T_{\text{diss}}$.

Clearly, intermolecular NOE during the $T_{\text{diss}}$ can lead to erroneous estimates of $D_t$. However, Yan et al. have recognized that the guilty intermolecular NOE transfers arise precisely from the ligand-receptor contacts sought by epitope mapping [79]. Ligand resonances displaying the largest inter-molecular NOE ‘artifacts’ represent those protons in closest contact with the receptor. Thus, one can exploit the NOE ‘artifacts’ of the diffusion experiment to provide a novel experiment for epitope mapping. To define the map, one measures the deviation of intensity from Eq. (57) for each resonance. The deviation can be quantified by fitting the $\ln(I(G))/F^2$ versus $G^2$ data with the three-parameter form

$$\ln(I(G))/F^2 = \alpha + \beta G^2 + \kappa G^4. \quad (58)$$

The fitted parameter $\kappa$ measures the deviation from Eq. (57). Diffusion epitope mapping amounts to fitting $\kappa$ for all ligand resonances. The map emerges after normalizing the $\kappa$ values by the largest one and correlating it with the ligand structure.

A key advantage of diffusion epitope mapping is that it bypasses the aforementioned problems of $R_1$ mentioned above for GEM. Yan et al. demonstrate this by comparing the epitope map of dihydrofolate reductase (DHFR) with trimethoprin (TMP) using diffusion and GEM. The diffusion and GEM epitope maps are in agreement for most of the TMP protons. However, they disagree strongly for the TMP methylene protons. In particular, GEM excludes these protons from the receptor binding site due to their anomalously low STD responses. In stark contrast, the diffusion map suggests intimate contact between the methylene protons and DHFR. The large body of literature on the DHFR-TMP system supports the diffusion epitope map. The discrepancy can be explained by considering the effects of $R_1$. Yan et al. observed exceptionally large $R_1$ values for the methylene protons. As stated, weak STD responses can be expected if the local $R_1$ is exceptionally large. Yan et al. also showed a strong inverse correlation between the strength of the STD signals and the $R_1$ values, suggesting $R_1$ weighting of the GEM epitope map. In contrast, no such correlation was observed for the diffusion-based methods. Thus, the diffusion-based approach successfully avoids $R_1$ weighting of the epitope map and thus has a significant advantage over GEM.

Like GEM, diffusion epitope maps require sufficiently fast exchange to ensure that spin-diffusion within the bound ligand state does not equalize all magnetizations. Additionally, the intrinsic sensitivity of the BPP-STE experiment is much lower than the STD experiment; the mere fact that it uses a stimulated gradient echo incurs a mandatory loss of half the available magnetization. Finally, measurement of the intermolecular NOE ‘artifact’ requires a series of 1D BPP-STE measurements, thus reducing throughput. Accordingly, this method is best suited to highly soluble receptors and ligands.
9. Receptor-based approaches

Crowded spectra and resonance assignment of large monomeric proteins (>30,000 Da) have been bottlenecks limiting the utility of receptor-based screening methods. However, recent advances have begun to relax some of these bottlenecks. Advances in isotope-labeling strategies and data interpretation show promise in expanding the general applicability of receptor-based approaches in pharmaceutical research by removing the pre-requisite of sequential resonance assignment to locate ligand binding sites.

9.1. Selective active-site isotope labeling

For larger protein targets (>30,000 Da), one would like the option of honing in on binding site resonances immediately without first spending large amounts of time making sequential resonance assignments. Such an option would greatly accelerate the process of ligand localization. To achieve this goal, Wiegelt et al. have demonstrated a selective labeling scheme that labels pairs of sequential amino acid residues. The premise is that one can identify unique pairs of residues (e.g. amino acid X and amino acid Y) that reside exclusively within the protein active site. If this is the case, then a strategy that labels amino acid X with $^{13}$C and amino acid Y with $^{15}$N will selectively mark the active site. More specifically, these residues will be the only source of signals in a multi-resonance experiment that edits according to scalar coupling between $^{13}$C=O and amide $^{15}$N.

Wiegelt et al. have performed proof-of-concept studies of this strategy on the 14,700 Da fatty acid binding protein FABP4. In particular, they identified Val114 and Val115 as a unique pair of sequential amino acids within the FABP4 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 spectrum. Furthermore, this cross peak displayed the expected chemical shift perturbations upon addition of a known binder. Comparison of this peak with those seen in a conventional $^{13}$N–$^1$H HSQC spectrum (which shows all Valines) thus 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 difficult 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. To guard against the latter situation, Wiegelt et al. recommend identifying additional distinct pairs of residues.

9.2. Screening mixtures of receptors

Zartler et al. have recently demonstrated the possibility of screening several proteins at once via a new protocol, Rapid Analysis and Multiplexing of Experimentally Discriminated Uniquely Labeled Proteins using NMR (RAMPED-UP NMR) [86]. The protocol calls for a unique isotope-labeling scheme for each protein (e.g. different amino acid selective labeling, or more exotic approaches such as the described in Section 9.1) to provide a unique spectroscopic identifier and facilitate rapid binding analysis. The attraction of this approach is not only to increase throughput, but also the ability to investigate binding selectivity for a set of proteins and investigate protein-protein interactions. For example, screening a set of proteins can help identify highly selective binders that can discriminate between closely related proteins from a given gene family. Alternatively, one could screen for binders that target some motif common to a set of proteins to yield a broad-spectrum inhibitor.

Zartler et al. have demonstrated RAMPED-UP NMR on a mixture of three unrelated $^{15}$N selectively labeled proteins including the PTP1b phosphatase ($^{15}$Ne-Trp6), k-RAS ($^{15}$N-Ile13), and GFP($^{15}$N-Ala8). The labeling schemes yield only 6 to 13 labeled $^{15}$N–$^1$H spin systems per protein, thus yielding simple 2D $^{15}$N–$^1$H correlation spectra for the protein mixture. For proof-of-concept, the interactions of these proteins with two known ligands were investigated. 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 the avoidance of 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 in identical buffer conditions since modest amino-acid substitutions may confer quite different solubility properties.
9.3. Ligand localization from J-surface analysis of chemical shift perturbations

McCoy and Wyss have developed an alternative method for localizing ligand binding sites by exploiting available yet, hitherto, untapped information contained in binding-induced protein chemical shift perturbations [87]. The method recognizes the high prevalence of aromatic rings in drug-like molecules. It follows that the 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, one can localize the ligand binding site. Note that the localization proceeds from an analysis of shift perturbations alone; a lengthy structure determination of the protein–ligand complex is not necessary.

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[88,89]. This dipolar field can perturb the magnetic dipole (infinitesimal current loop) located at the current can be approximated by that of a point or ‘perfect’ magnetic dipole (infinitesimal current loop) located at the center of the ring[88,89]. This dipolar field can perturb the local magnetic field stemming from the aromatic ring (with associated distance 2Ri (with associated distance Ri) pointing from the ring center to the position of protein proton ‘i’. Both θ and R are illustrated in Fig. 13a. B_dip is a proportionality constant and θ is the angle between the ring plane normal (through the planar center) and Ri. The sign and magnitude of ΔCS(i) will obviously depend on θ, ΔCS(i) extends from a minimum of −2(B_dip/Ri^2) to maximum of +1(B_dip/Ri^2). At θ = 54.7° (magic angle), ΔCS(i) = 0.

To translate the ΔCS(i) information into ligand localization, McCoy and Wyss use spherical dot-density representations; an example of which is given in Fig. 13b. Specifically, one constructs spheres centered on each perturbed proton ‘i’ that has a radius Ri consistent with Eq. (59). If the observed ΔCS(i) > 0, then the maximum Ri possible must obey ΔCS(i) = (B_dip/Ri^2). On the other hand, if ΔCS(i) < 0, then the maximum Ri must obey ΔCS(i) = −2(B_dip/Ri^2). The spheres are then filled randomly with dots, where each dot represents a possible location for the center of the aromatic ring. Clearly, smaller and larger ΔCS(i), will lead to larger and smaller Ri, respectively. To compensate for this, one uses the same number of dots for each sphere. Hence, the dot density will vary from sphere to sphere, depending on the magnitude of the ΔCS(i).

Fig. 13. Schematic of ring current shifts caused by ligand aromatic groups on receptor NH protons[87,88]. (A) (left panel): 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 Ri) pointing from the ring center to the receptor NH in question. (B) (right panel): The observed perturbation is quantified in terms of a maximum radius according to Eq. (59) 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 by McCoy and Wyss [87]. Copyright 2002 American Chemical Society.

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 ΔCS(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. Fig. 14 illustrates this principle. Inconsistent ΔCS(i) data (i.e. cannot be linked to a common perturbing source) yield diffuse dot-density 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 non-overlapping spheres. The surfaces of these consensus volumes are ‘j-surfaces’, where ‘j’ refers not to scalar spin-spin coupling constants, but, rather, current-density, as seen often in electromagnetic theory. The current-density referred to, of course, is the ligand ring current responsible for the ΔCS(i).

Fig. 14. Consensus location of ligand aromatic ring by the method of overlapping spheres [87]. The overlap produces a localized volume for the ring location consistent with the observed chemical shift perturbations of multiple NHs. (A) (left panel): Overlap of two spheres. (B) (right panel): Overlap of three spheres. Reprinted with permission by McCoy and Wyss [87]. Copyright 2002 American Chemical Society.
Ligands will often contain more than one aromatic ring. Ligands with multiple rings separated by linkers of sufficient length (>5 Å) behave as linear systems; one can add their separate influences to achieve the net \( \Delta \text{CS}(i) \). However, this treatment is not accurate for fused rings, and some error will be introduced into the resulting dot-density spheres. However, a key aspect of j-surface mapping is the reliance on multiple \( \Delta \text{CS}(i) \) for a given ligand. The greater such 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 [87]. In both cases, j-surfaces from multiple \( \Delta \text{CS}(i) \) were able to localize ligands to surface positions consistent with those seen in X-ray crystal structures. They 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. They simulated the situation in which the protein structure is known, and one has \( \Delta \text{CS}(i) \) data on several systems with amino-acid specific isotope labeling. The results are quite encouraging and suggest 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.

10. Specific example of NMR screening: targeting the ribosome

The ribosome, which is a ribozyme composed of both protein and RNA molecules, is perhaps the epitome of an elaborate catalytic molecular machine. To provide an example of cutting-edge NMR screening applications and illuminate some of the key challenges therein, we sketch here some of our recent efforts at Vertex Pharmaceuticals, Inc. to screen the bacterial ribosome.

The 70S bacterial ribosome macromolecular complex is a protein-translation machine that consists of 50S and 30S subunits with molecular weights of \( 1.6 \times 10^6 \) and \( 0.9 \times 10^6 \) Da, respectively. The bacterial ribosome is a target for many antibiotics such as macrolides, lincosamides and aminoglycosides, which exert their antibacterial action by binding to different sites on the 50S and 30S subunits [90–94]. Recent X-ray structures of complexes of the 50S and 30S subunits with several antibiotics have unambiguously revealed the binding modes of these antibiotics [95–101]. The structural data indicate that the antibiotics make a large majority of binding interactions to RNA rather than the ribosomal proteins.

There is burgeoning interest in targeting bacterial ribosomes as an avenue towards novel antibiotics; the growing resistance towards established antibiotics among bacteria is a high-profile therapeutic concern. Previous NMR studies of ligand binding to bacterial ribosomes include transferred-NOE studies by Girault and co-workers [102,103] as well as Stockman and co-workers [104]. Here, we describe the use of STD (Section 5.5) and WaterLOGSY (Section 5.6) to study the binding of small ligands to both the 50S and 30S bacterial ribosomal subunits isolated from *Deinococcus radiodurans*. Given the many accessible RNA and protein binding pockets on the large surface area of the ribosome, it is of great interest to determine whether these relatively simple NMR approaches can identify specific binding sites of small ligand molecules on the ribosome.

As proof-of-concept, we initiated binding studies on the isolated 50S and 30S ribosomal subunits with antibiotics whose modes of interactions with the ribosome have been well established both structurally and biochemically. The compounds chosen for binding studies on the 50S subunit included chloramphenicol, the macrolides tylosin and erythromycin, and the lincosamides clindamycin and lincomycin. Binding studies with the 30S subunit made use of the aminoglycoside spectinomycin and the four-ringed tetracycline (cf. Fig. 15). Exhaustive studies compared STD versus waterLOGSY (with both water saturation and inversion) methods. In addition, systematic searches for the optimum irradiation frequency, temperature, ligand:protein ratio, \( \text{D}_2\text{O}:\text{H}_2\text{O} \) ratio and buffer conditions were carried out to ensure that specific binding activities would not be missed by poor choices of experimental parameters and sample conditions. Perturbation of bulk water via WATERLOGSY gave the highest sensitivity binding spectra. Although the superiority of ‘water inversion’ waterLOGSY has been touted for highly hydrated system such as RNA-ligand interactions, we found, surprisingly, no significant improvement in S/N in the use of inversion as compared to simple selective saturation of the water resonance.

For the reason mentioned above, the large majority of the binding studies on the ribosomal subunits saturated the water resonance using the STD pulse scheme. Results from the 50S subunit binding studies showed the expected

![Chemical structures of antibiotics](image)

Fig. 15. Chemical structures of antibiotics used in competition binding studies with 30S ribosomal subunits isolated from *Deinococcus radiodurans*. Tetracycline (1) spectinomycin (2).
binding of the antibiotics tested. Only a single set of resonances were observed for all binders, suggestive of fast exchange between the free and bound states. However, we were not able to prove that any of the observed binding activities were specific for a single subunit binding site. To address binding specificity of erythromycin ($K_D \sim 10^{-6}$ M to 50S [105]) and lincomycin ($K_D \sim 10^{-6}$ M to 50S [102]), separate competition experiments were carried out using tylosin ($K_D \sim 10^{-6}$ M to 50S [106]) and clindamycin ($K_D \sim 10^{-6}$ M to 50S [102]) to displace bound erythromycin and lincomycin, respectively. Even at a 10-fold excess of competitor, there were no decreases in the STD responses of either bound erythromycin or lincomycin indicative of specific binding. Instead, the STD spectra from both the macrolide and lincosamide studies indicated that both ligand and competitor were interacting with the 50S subunit during the NMR acquisition. One explanation for this observation is that a majority of the ligand and competitor molecules bind to readily accessible hydrophobic patches on the large ribosomal subunit in a non-specific manner. In addition, the fact that ligands and competitors in both cases (erythromycin/tylosin and clindamycin/lincomycin pairs) have very similar $k_{on}/k_{off}$ compounded the problems in observing competitive binding activities in these studies (see e.g. Fig. 2).

In contrast to the 50S subunit studies, our competition binding studies with the 30S subunit implied specific binding activity for bound spectinomycin with tetracycline used as competitor. Upon addition of equimolar (100 μM) tetracycline to bound spectinomycin a 30% decrease in the STD signal for spectinomycin was observed (Fig. 16). Doubling the amount of tetracycline competitor reduced the total spectinomycin STD signal by at least 50%. This suggested partial displacement of bound spectinomycin to the 30S subunit by tetracycline is consistent with recent X-ray structures [96,97,100] These data show that the major binding site responsible for tetracycline antibiotic action is located within the A-site [96,100] and the tetracycline binding site overlaps partially with spectinomycin binding site on the 30S subunit [97].

The above studies demonstrate that the STD technique is very sensitive in detecting binding of small ligands to large macromolecular complexes. However, the limitation of this technique is that it also detects non-specific binding to the many low affinity binding pockets present on the ribosomal subunit. While it is possible to observe specific binding on a large macromolecule as demonstrated in the 30S subunit studies, binding results obtained with large macromolecular assemblies should always be treated with caution.

11. Conclusions and challenges remaining

This review has focused mainly on the methodologies underlying current NMR screening efforts to accelerate lead generation in pharmaceutical research. The focus has been primarily on ligand-based (as opposed to receptor-based) methods, since these more readily accommodate the realities of pharmaceutical research. In either case, both approaches seek to provide useful binding data in a timely manner.
fashion that can be readily used by other disciplines contributing to the lead generation process. Given the prevalence of NMR screening applications presented in other recent reviews and conferences, NMR screening shows tantalizing promise for being not only a visible but essential tool for lead generation in industry. Furthermore, its appeal has spilled over to the academic forum where the research foci of structural biology often parallel those of the pharmaceutical industry.

The above discussion deals more with NMR spectroscopists perception of their own contribution to lead generation. Equally important is the perception by non-NMR research groups within the lead generation process. The ultimate sustainability of NMR screening depends on its ability to integrate with ‘downstream’ disciplines. In turn, this will depend on so-called ‘thoughtful’ design of both NMR screening experiments and libraries. Here ‘thoughtful’ is not meant imply the existence of ‘thoughtless’ screening methodologies. Rather, it implies an acute awareness of the bottlenecks of drug design, not inhibitor design, and of the importance of making the NMR results timely, easily understood, and amenable to follow-up by other disciplines. The initial rapid growth and current languishment of structural NMR in pharmaceutical research has illustrated quite clearly the risk of presenting lots of complicated data too late to be useful in time-sensitive projects.

A challenging and exciting frontier for NMR screening lies with those receptor targets that resist rapid structure determination by classical methods. Such receptors include integral membrane proteins and molecular machines, such as the bacterial ribosome discussed above. These systems are challenging not only on account of their size and difficulty of sample preparation, but also their dynamic behavior. Here, probing the preferences of receptor–ligand interaction preferences by NMR screening methods can be immensely useful while waiting for more detailed information concerning the receptor structure. Our initial work on the bacterial ribosome above foreshadows some of the difficulties that may be anticipated in such systems. At the same time, however, NMR hardware continues to improve in sensitivity and resolution. Obvious examples are the increasing availability of cryogenically cooled probes and higher-field magnets. It is conceivable that such improvements will enable NMR to become a powerful and essential tool for designing novel therapeutics targeting these most challenging systems.

Acknowledgements

The authors wish to acknowledge and thank Jasna Fejzo, Scott Harbeson, Christopher Lepre, David Pearlman, Jennifer Pocas, John Thomson, and Xiaoling Xie for their contributions, expertise, and generosity that have made this review possible. We especially thank John Thomson for critical reading and helpful suggestions that improved significantly the content and analysis of the review. J.W.P further thanks those who have provided insight and encouragement during the writing of this review, including: Michael Badia, Steve Bellon, Guy Bemis, Kevin Cottrell, Rosario Cestari-Murphy, Maureen Dwyer, John Fulghum, Brian Hare, Marc Jacobs, Sandy Lauterbach, Stephan Lienn, Mark Namchuk, Cheryl Schairer, Celia Schiffer, and John van Drie.

References
