NMR Screening Methods for Drug Discovery

Matthew D. Shortridge and Robert Powers[1]
Department of Chemistry, University of Nebraska-Lincoln, Lincoln, NE 68588-0304

Abstract. Drug discovery is a challenging endeavor with a high failure rate. This is further complicated by the fact that each disease has its own unique set of obstacles to developing a safe and effective therapy. Flexible and robust analytical methods are critical for efficiently solving these issues, where nuclear magnetic resonance (NMR) plays an important role in nearly every stage of the drug discovery process. This review will highlight recent developments in the application of NMR as a screening tool for drug discovery that includes: library design, various ligand-affinity screening techniques, rapid determination of protein-ligand co-structures, and the functional annotation of proteins to the discovery of new therapeutic targets.

Keywords. Drug discovery, NMR ligand-affinity screens, library design, rapid co-structures, functional annotation

Introduction

Drug discovery is a uniquely complex problem in science and medicine [1, 2]. This is further complicated by the fact that each disease is distinct and requires its own efficient strategy to successfully develop novel drugs [3]. Additionally, drugs primarily act through a binding interaction to a therapeutic target that either modulates or alters its biological activity [4], changes the dynamics of a protein and disrupts its normal function, [5] or interferes with critical protein-protein interactions in important signaling pathways [6]. Identifying compounds that exhibit any of these desirable activities while simultaneously demonstrating *in vivo* efficacy in the absence of toxic side-effects is an extremely challenging endeavor. Thus, an important component of the drug discovery process is the verification that a small molecule actually binds the protein target in a selective and biologically relevant fashion. This is an especially critical issue given that the vast majority of chemical leads are typically identified by a high-throughput (HTS) [7-11] or *in silico* structure-based [12-14] screens.

Compounds that lack a confirmed correlation between functional activity and a direct binding interaction with the protein target routinely emerge from HTS [15]. These "false-positives" generate an HTS response through a number of undesirable mechanisms, such as: protein aggregation, protein denaturation, protein precipitation, micelle formation, chemical modification of the protein, non-specific binding, promiscuous binding or interference with other reagents of the assay [15-19]. In effect,

[1] Corresponding Author. Robert Powers, Department of Chemistry, University of Nebraska-Lincoln, 722 Hamilton Hall, Lincoln, NE 68588-0304, USA. E-mail: rpowers3@unl.edu

an HTS assay does not generally provide any information on the mechanism of inhibition. Similarly, *in silico* screens only increase the likelihood that the resulting list of predicted high-affinity binders actually contains compounds that bind the protein target. In general, an *in silico* screen enriches the number of active compounds by approximately 10 to 1000 fold [20-22] relative to the 0.1 to 0.5% typical hit-rate observed in an HTS assay using a random chemical library [23]. Obviously, HTS and structure-based *in silico* screens require a follow-up experimental assay to confirm the observed or predicted biological activity of a chemical lead. It is also essential that the observed inhibition occurs through a productive interaction in the protein's active-site. Clearly, obtaining a rapid experimental structure for the predicted protein-ligand complex is a valuable means to confirm a relevant biological activity for the ligand [24]. While not guaranteeing success, this requirement increases the likelihood that a chemical lead can be evolved into an effective drug.

From target selection to pre-clinical trials, nuclear magnetic resonance (NMR) has established itself as an invaluable tool for the chemist working in the drug discovery industry [25, 26]. The flexibility provided by NMR comes from various molecular probes that include chemical shift changes, relaxation parameters (T_1, T_2), through space interactions (nuclear Overhauser effects, NOE), and diffusion rates. Each parameter is sensitive to the local chemical and physical environment and provides structural information at atomic resolution for both small (<1000 Da) and large (> 1000 Da) biological molecules. The utility of NMR for drug discovery is continually expanding as evident by the recent development of various novel NMR methods that include the identification of new therapeutic targets [27, 28], the efficient measurement of binding affinities [29], fragment-based approaches to ligand affinity screens [30], and the rapid determination of protein [31] and protein-ligand co-structures [24, 32]. NMR is also an invaluable technique for monitoring changes in cellular metabolism caused by a disease state or due to drug treatments [33].

The ongoing implementation of these NMR technologies is significantly enhancing the likelihood of drug discovery successes [30]. Since NMR is an inherently flexible technique with a diverse array of experimental methods, NMR is not restricted to a particular target or system. Thus, NMR is routinely used for the identification of drug leads to treat various ailments that range from cardiovascular disorders to infectious diseases. A number of recent drug discovery efforts utilizing NMR techniques have reported the identification of high affinity inhibitors ($\leq$ μM) against a variety of potential therapeutic targets. For example, Ekonomiuk *et al.* identified [4-(carbamimidoylsulfanylmethyl)-2,5-dimethylphenyl]-methylsulfanylmethanimidamide as an inhibitor ($K_D \sim 40$ μM) of the non-structural 3 protease (NS3pro) from the West Nile virus using *in silico* screening coupled with NMR validation [34]. Similar results have been reported using other NMR screening techniques that include: saturation transfer difference (STD)-NMR spectroscopy to identify a bivalent ligand MLM that inhibits ($K_D \sim 3.3$ μM) cholera toxin B pentamer CTB_5 [35], a diffusion-edited NMR screen to identify (glucosamine-aminoethoxy)triphenyltin that inhibits ($K_I \sim 2.5$ μM) immune response target human vaccinia H1-related phosphatase (VHR) [36], 2D ^{1}H-^{15}N HSQC screen to identify isoquinolinone inhibitors ($K_D \sim 5$ μM) of the MDM2–p53 interaction [37], a fragment based NMR screen to identify an indole-analog inhibitor ($K_D \sim 500$ μM) of the ZipA/FtsZ complex [38] and an *in silico* screen with NMR validation to identify a *N*-(2-acetylphenyl)benzamide analog inhibitor ($K_D \sim 11$ μM) of the *Wnt* signaling pathway [39].

Despite these NMR screening successes, HTS is still the primary component of the early-phase drug discovery process for all major pharmaceutical companies [1]. The obvious appeal of HTS is the routine and rapid ability to screen hundreds of thousands to millions of compounds against a therapeutic protein target. The throughput of HTS-NMR clearly pales in comparison to HTS. However, fragment-based chemical libraries, a small collection (hundreds to thousands) of low molecular-weight (≤ 200-300 Da) compounds with drug-like characteristics [40], are being used to address this discrepancy. Fragment-based chemical libraries maximize ligand efficiency [41], have a 10-1,000 times higher hit rate [42], more efficiently cover structural space [43], and generate better quality leads with a higher success rate [44]. Simply increasing the size of the library screened by HTS doesn't necessarily correlate with an increase in validated chemical leads that make it through the clinic to become drugs [45-48]. Thus, an important element of NMR ligand-affinity screens is the incorporation of fragment-based chemical libraries.

This review will highlight the various NMR ligand affinity screening methods that have been successfully used in drug discovery programs to identify drug-like chemical leads. Additionally, high throughput methods for generating accurate ligand bound co-structures from screening data and the benefits of applying NMR screening techniques to functionally annotate novel proteins will be discussed.

1. NMR Methods to Detect Ligand Binding

NMR ligand-affinity screening methods complement structural biology efforts by validating chemical leads prior to initiating a structure-based drug design program [49-54]. NMR screening techniques, such as SAR by NMR [55], RAMPED-UP NMR [56], STD-NMR [57], and NMR-SOLVE [58] (Table 1), were developed to identify ligands that bind a therapeutic target in a biologically relevant manner by observing chemical shift changes in two-dimensional (2D) ^{1}H-^{15}N HSQC spectra. Other techniques have been developed that utilize saturation transfer differences (STD) [57, 59], line-broadening changes [60-63], diffusion rate changes [64], ^{19}F NMR [62, 65], spin labels [63], and transfer NOEs [66]. More recently, multi-step approaches to NMR screening join complementary techniques to increase throughput and minimize resource usage [67]. Methods such as Multi-Step NMR [67] , MS/NMR [68] and FAST-NMR [27, 28] combine one-dimensional (1D) ^{1}H NMR line-broadening experiments, mass spectroscopy or 2D ^{1}H-^{15}N HSQC chemical shift perturbations to identify and qualitatively rank binding interactions from small molecule libraries.

Screening chemical libraries by NMR can be divided into two general categories, ligand based and target based. Ligand based NMR experiments benefit from lower protein and compound concentrations, higher throughput and simplicity of data analysis [54]. Target based methods primarily benefit from binding site information and accessibility to a wider range of binding affinities [54]. The two methods complement each other for rapid drug discovery, lead optimization and functional annotation. In this section, the various chemical libraries and NMR screening methods currently used to detect small molecule binding to identify chemical leads will be discussed.

Table 1. Comparison of various NMR screening methods [26].

Screening Technique	Method of Detecting Ligand Binding	Labeled Protein?	Protein-Ligand Co-structure?	Limited by Protein MW?	Ref.
3-FABS	Chemical shift changes, requires fluorinated ligands	No	No	No	[65]
Affinity NMR	Change in translational diffusion	No	No	No	[64]
AIDA-NMR	Line-broadening change (T_2) due to protein-protein complex formation, labeled protein or Trp reporter in ligand binding site	Yes/No	Yes/No	Yes	[60, 61]
FAST-NMR	Line-broadening change (T_2) & chemical shift changes	Yes	Yes	Yes	[27, 28]
FAXS	Line-broadening change (T_2) due to ligand competition, requires fluorinated ligands	No	No	No	[62]
INPHARMA	Transfer nuclear Overhauser effect (NOE)	No	Yes	No	[69]
MS/NMR	Retention on size-exclusion column & chemical shift changes	Yes	Yes	Yes	[68]
Multi-Step NMR	Line-broadening change (T_2) & chemical shift changes	Yes	Yes	Yes	[67]
NOE pumping	Transfer nuclear Overhauser effect (NOE)	No	No	No	[66]
RAMPED-UP NMR	Chemical shift changes, screening multiple proteins	Yes	No	Yes	[56]
SALMON	Saturation transfer difference from solvent	No	No	No	[70]
SAR by NMR	Chemical shift changes	Yes	Yes	Yes	[55]
SLAPSTIC	Line-broadening change (T_2) due to protein spin label	Yes	No	No	[63]
SMILI-NMR	In-cell chemical shift changes	Yes	Yes	Yes	[71]
STD NMR	Saturation transfer difference from protein	No	No	No	[57]
STINT-NMR	In-cell chemical shift changes	Yes	No	Yes	[72, 73]
TINS	Line-broadening change (T_2) due to binding to an immobilized protein target	No	No	Yes	[74]
WaterLOGSY	Saturation transfer difference from solvent	No	No	No	[59]

1.1 Library Design for High-Throughput NMR Screening

Chemical space of biomedical compounds is vast, and while estimates on the exact size varies greatly, the likely number of drug-like compounds is > 10^8 and exceeds the total number of known compounds [75-77]. Obviously, screening the entirety of chemical space of drug-like molecules by experimental methods is not feasible. Additionally, *in silico* screening methods are prone to high false hit rates (~49%) [78] and only enrich the hit-rates in focused libraries [12-14]. An *in silico* screen is still an invaluable resource since it can increase the hit-rate by 10 to 1000 fold [20-22] over a random search that typically yields only a 0.1 to 0.5% hit-rate [23]. But, false-positive and false negative rates are still very significant problems. Therefore, an efficient library design is critical to generating active drug-like leads through an effective sampling of

chemical space. Often this requires a chemical library that maximizes structural diversity [79-84], filtering compounds with drug-like properties [85-87], and screening with appropriate mixtures sizes [88, 89]. The use of mixtures creates additional issues that include: solubility and stability problems, chemical reactivity, and ligand competition. A number of reviews have discussed in detail the design of chemical libraries [42, 90-93] and these approaches will only be briefly summarized here.

HTS-NMR screening methods typically rely on a chemical library of low molecular weight (< 1000 Da) compounds or chemical fragments to screen against a therapeutic protein target. An essential consideration in the design of an NMR screening library is efficiently sampling the diversity of chemical space to ensure the identification of chemical leads that can be evolved into a drug. Methods to define chemical diversity incorporate ~ 1,600 molecular descriptors [94] that includes, among many examples, molecular weight (MW), number of rings, rotatable bonds, heteroatoms, electronegativity, number of hydrogen bond donors and acceptors, pKa values, ClogP (lipophilicity), ClogS (aqueous solubility), polar surface area, functional group counts, geometrical, connectivity, and topology descriptors [95], information indices [96], eigenvalue-based indices [97], and molecular fingerprints [98]. The molecular fingerprinting methods represent chemical structures as a bit string that attempts to encode all structural and chemical elements of the compound [99]. It has been shown that a compound with greater than 85% fingerprint similarity to an active compound will exhibit similar biological activity [100]. Molecular fingerprinting methods define a neighborhood of compounds with similar biological activity (Figure 1). Therefore it is often redundant to have multiple compounds from the same fingerprint defined neighborhood within a single library. In essence, designing a screening library that maximizes chemical diversity simply requires adding compounds that have unique molecular descriptors relative to other members of the library, such as distinct molecular fingerprints.

A major source of failure in clinical trials is the observation of toxic side-effects [101, 102]. Thus, compounds that comprise a screening library should exhibit drug-like characteristics [85-87]. Essentially, compounds that share similar physiochemical features with known drugs are predicted to have a reduced likelihood of exhibiting toxic side effects. Fundamentally, there is a greater chance of converting an HTS chemical lead with known drug-like properties into a successful drug than a completely novel chemical class [88]. Three main classifications corresponding to physiochemical properties, common functional groups and common chemical structures are often used to define drug-likeness. Lipinski's "rule of five" [86, 87] is the most widely used approach to identify drug likeness by describing the common characteristics of known drugs. These features include a molecular weight less than 500 Da, less than 5 hydrogen bond donors, less or equal to 10 hydrogen bond acceptors, and a logP of less than or equal to 5. Lipinski's "rule of five" is a general predictor for compound solubility and permeability that is related to bioactivity and bioavailability [103]. Recently, machine learning programs have been developed to predict the Absorption, Distribution, Metabolism, and Excretion (ADME) properties of chemical leads to determine drug-likeness [104-106]. Similar approaches have been applied to predict likely reactive compounds that lead to false positives in HTS assays [107]. Thus, a chemically diverse library is further filtered to select compounds with drug-like characteristics.

A fragment-based approach to library design incorporates chemical diversity

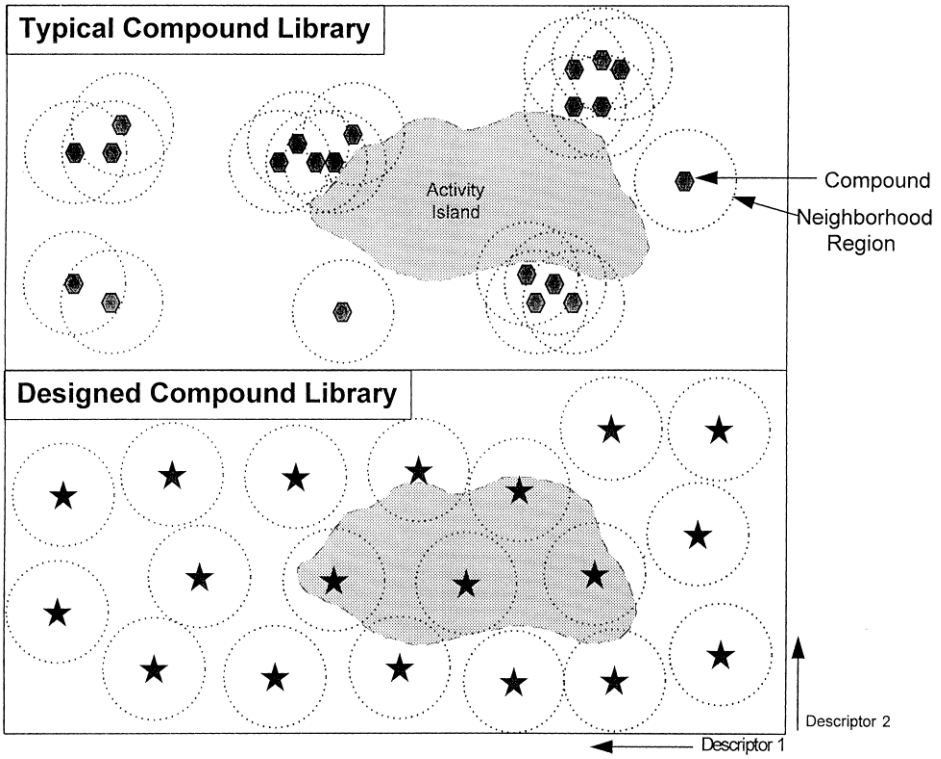

Figure 1. Neighborhood behavior of compound similarity for two arbitrary molecular descriptors (x,y) can be an efficient way of developing a chemical library to increase diversity and drug-like properties. A test compound (hexagon, star) can have a cluster of similar compounds (open circles) with similar biological activity. For typical random libraries (top) it is inefficient to screen compounds with overlapping neighborhood similarities because one compound has already probed that search space. However, exploiting ligand similarity can be an efficient way to sample and explore the activity island of a particular target (bottom) (Reprinted with permission from reference [100], Copyright 1996 by the American Chemical Society).

and drug-like characteristics to create an efficient chemical library optimized for HTS-NMR. The fragment based approach has been extensively reviewed [108-111], but essentially uses the perspective that a chemical lead can be broken down into the sum of its parts or fragments. Basically, a fragment-based library contains small molecular-weight compounds ($\leq$ 200-300 Da) that correspond to fragments of known drugs, exhibit drug-like characteristics and have high aqueous solubility. A comparable approach is using a fragment-based library composed of small biologically active compounds [112]. Drug-like fragments identified as hits from an HTS-NMR assay are chemically linked and further optimized to generate leads. Both the linking of the fragments and their relatively small size significantly improves the coverage of chemical space compared to large, random chemical libraries. The number of low MW drug-like compounds is significantly reduced from the estimate of $> 10^8$ total biomedical compounds [43]. Correspondingly, a higher-percentage of the possible low MW drug-like compounds is represented by a practical-sized library. Also, the linking of fragments indicates that a combinatorial combination of the screened compounds are

actually examined [44]. The application of fragment-based libraries has resulted in an increased success rate and higher quality drug-like chemical leads [30].

To increase throughput, chemical libraries are commonly screened as mixtures of compounds. Simply, as the mixture size increases, the number of NMR experiments required to screen the entire library decreases proportionally. The end result is a dramatic increase in the efficiency (> 3-5 fold) of the HTS-NMR screen. The binding ligand(s) are expected to be identified through a second deconvolution experiment. Unfortunately, the number of compounds per screening mixture increases the number of necessary deconvolution experiments and significantly reduces the screening efficiency. This is an important issue because the deconvolution step might overwhelm and eventually eliminate any efficiency benefits of screening with mixtures. Quantitative measures of the optimal mixture size (OMS) for an NMR screen have been described [88, 89].

A major factor in determining the optimal mixture size is the overall hit rate of the HTS-NMR assay. For large random chemical libraries, an average hit rate is on the order of 0.1 to 0.5% [23], while focused libraries have hit rates ranging from 0.7 to 20% [113]. For focused libraries with large hit rates, multiple active compounds or "hits" within a mixture becomes a significant problem [112, 114, 115]. The probability of getting x hits within a mixture of n compounds can be approximated with a binomial distribution with p representing the probability or hit-rate [88].

$$f(x) = \left(\frac{n!}{x!(n-x)!} \right) p^{x} (1-p)^{n-x} \tag{1}$$

Based on eqn. 1, a 10% hit rate would result in 26% of mixtures containing ten compounds having more than one hit per mixture. Obviously, this is a significant problem because the deconvolution of each hit requires collecting ten additional NMR spectra. Thus, the total number of NMR experiments needed to confirm binding would significantly increase, negating the initial decrease in total experiments needed to screen the entire library. For a library of 1,000 compounds divided into 100 mixtures of 10 compounds and a 10% hit rate would initially require 100 NMR experiments to identify 26 mixtures with active compounds. An additional 260 NMR spectra would then be required to deconvolute the 26 mixtures. A total of 360 NMR spectra are required to screen the entire library and identify the 100 active compounds. Clearly this is an improvement over screening the entire library as singletons (1,000 NMR spectra), but is it the optimal approach?

A hypergeometric distribution was proposed as an alternative method to identify the optimal mixture size by minimizing the total number of deconvolution steps [89]. The process involves creating a set of mixtures (N) from a compound library by selecting n compounds from the library until all compounds have been used with a total of N/n mixtures:

$$P(X = x) = h(x : n, M, N) = \left(\frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}} \right) \tag{2}$$

where M is the total number of hits present in a mixture, P is the probability of a mixture containing at least one hit and x is the number of hits present within the mixture. The problem is analogous to the classic urn problem, where an urn contains a fixed number of two differently colored balls and the above equation describes the probability of pulling out one color for a given sample size. In an NMR screen absent of deconvolution, the total number of experiments needed would simply scale by (N/n). However, taking deconvolution into account, the total number of NMR experiments (T) needed to screen the entire library and confirm binding scales by:

$$T = \left(\frac{N}{n}\right) + (N)(P) \tag{3}$$

The results of these statistical analyses draw three important conclusions. It is always more efficient to screen a library with mixtures that avoid deconvolution. If the screening library is a traditional random library with a low hit-rate, the optimal mixture size with deconvolution is approximately 20 compounds. Conversely, for focused libraries with high hit rates ($\geq$ 5%), the optimal mixture size is small enough ($\leq$ 5) that it is preferable to screen with mixtures that avoid deconvolution.

1.2 Ligand Based HTS-NMR Screening Methods

Ligand focused high-throughput NMR screening (HTS-NMR) methods monitor changes in the ligand's NMR spectrum as a result of complex formation. Generally, 1D ^{1}H NMR techniques that compare the ligand in the free-state relative to the bound state are employed. The chemical libraries are typically screened once in the free-state with the data stored for later comparisons. These reference NMR spectra are also used to verify the purity, solubility and stability for each compound in the library. Ligand detection methods require significantly less of the protein target ($\leq$ 5 μM), which is typically the limiting factor for HTS-NMR; do not require the protein to be isotopically enriched; and can be completed relatively fast (a few minutes per NMR spectrum). These advantages make ligand detected screening methods highly desirable for HTS-NMR drug discovery. The only major drawback is the lack of binding site information. Does the ligand bind specifically to the protein's known active-site or functional epitope? Is the interaction biologically relevant?

There are various ligand detected methods that exploit a variety of NMR parameters to probe ligand binding activity. These include chemical shift changes (^{19}F NMR), T_1 and T_2 relaxation times, diffusion rates, saturation transfer differences (STD), transfer NOEs, as well as spin-labels and competition experiments. The variety of ligand detected methods allow for the selection of a technique that is optimal for screening a specific protein target or chemical library. A compilation of ligand detected NMR screening methods is highlighted in Table 1. In this section, two of the most commonly used screening methods, line broadening experiments and STD experiments will be discussed in detail.

One of the simplest ligand based screening methods is the line broadening experiments originally described by Jaradzey *et al.* [116, 117], where a decrease in ligand peak intensity was observed for penicillin in the presence of serum albumin. Specific binding of penicillin to serum albumin was determined to be the only mechanism causing the observed decrease in ligand signal after an increase in

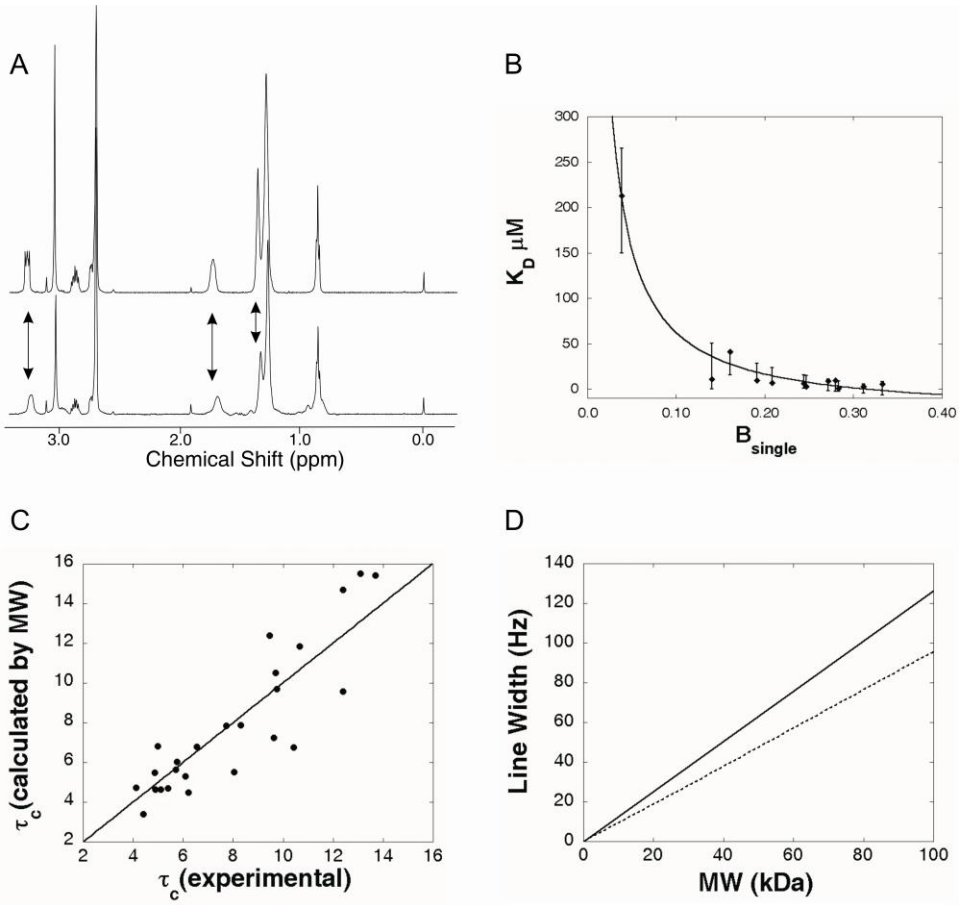

Figure 2. A) The increase in ligand line width (arrows) between free (top) and bound (bottom) is due to the bound ligand adopting the correlation time of the protein. B) Measurement of single point binding constants by NMR can qualitatively rank ligands based on binding affinity. C) A comparison of 27 experimental protein correlation times (τ_c) (ms) determined using NMR dynamics data with correlation times predicted from protein MW using eq 17 and a shape constant of 1.32. A best-fit line is shown with a slope of 1 and an R^2 of 0.81. D) A plot of line-width versus protein molecular-weight based on eq 17 for spherical proteins with ρ of 1 (solid line) and elliptical proteins with ρ of 1.32 (dashed) (B-D Reprinted with permission from reference [29], Copyright 2008 by the American Chemical Society).

viscosity, ligand-ligand interactions and non-specific binding were all ruled out. The results presented by Jadarzky *et al.* [116, 117] were critical in demonstrating the utility of NMR for monitoring specific binding interactions between proteins and small molecular-weight ligands. A binding interaction is simply identified by a decrease in the ligand's NMR signal (Figure 2A) in the presence of a protein relative to the ligand's free NMR spectrum due to peak broadening. The ligand signal line broadening is attributed to a difference in transverse relaxation rates (T_2) between the free and bound states of the ligand in the fast-exchange limit.

$$1/T_2 = (1/T_2)_{\text{free}} = B((1/T_2)_{\text{bound}} - (1/T_2)_{\text{free}}) \quad (4)$$

Assuming no other ligand or receptor dynamics, the decrease in ligand signal is due to the ligand adopting the larger rotational correlation time τ_c of the protein in the bound form. The primary relaxation mechanism in this type of ligand affinity screen is the dipole-dipole relaxation with the constant T_2^{-1} described by

$$T_2^{-1} = \frac{3}{20} b^2 \{3J(0) + 5J(\omega_0) + 2J(2\omega_0)\} \quad (5)$$

where

$$b = -\frac{\mu_o}{4\pi} \frac{\hbar\gamma^2}{r^3}, \; J(\omega) = \frac{\tau_c}{1+\omega^2\tau_c^2} \text{ and } \omega_0 = -\gamma B_0 \quad (6)$$

Here, $J(\omega)$ is the normalized spectral density function, μ_o is the vacuum permeability, γ is the magnetogyric ratio, ω is frequency (rad s^{-1}), $\hbar$ is Plank's constant, B_0 is the static magnetic field strength and r is the hydrodynamic radius of the protein. As shown, an increase in τ_c due to the ligand binding a protein has a dramatic effect on increasing the overall T_2^{-1} for the ligand. This makes the simple line broadening experiments exquisitely sensitive to binding events. Uniquely, the technique actually increases in sensitivity as the molecular-weight of the protein increases making the approach amenable to the vast majority of therapeutically interesting protein targets.

1D ^{1}H NMR line-broadening is a beneficial method for HTS-NMR because it is easy to interpret, experiment times are on the order of a few minutes, the NMR samples require minimal amounts of proteins, and relative binding constants can be rapidly identified [29, 118]. Cryoprobes coupled with high quality water suppression techniques [119] enable 1D ^{1}H line-broadening experiments to be acquired within 1-2 minutes with excellent signal-to-noise (S/N > 15000) for a 100 μM ligand sample. Also, the inclusion of sample changers with compound mixtures allows a chemical library composed of upwards of a 1000 compounds to be screened and analyzed in hours. An additional benefit to the line broadening experiment is the ability to measure qualitative binding affinities in a high-throughput manner [29]. Similar to traditional K_D measurements, NMR methods traditionally rely on the collection of multiple data points to accurately determine a dissociation equilibrium constant or binding affinity. This approach is usually impractical in a high–throughput mode that requires a rapid method for characterizing and ranking binding affinities. 1D ^{1}H line-broadening experiments are an integral component of the AIDA-NMR [60, 61], FAST-NMR [27, 28] FAXS [62], Multi-Step NMR [67], and TINS [74] assays. 1D ^{1}H line-broadening experiments have been used to identify a series of antagonists to estrogen receptor α [120], two (4- and 5-(2,5 dimethyl-pyrrol-1-yl)-2-hydroxybenzoic acid) antagonists of EphA4 ligand-binding domain [121], inhibitors of Nurr1 [122], and inhibitors of creatine kinase [123].

The saturation transfer difference (STD) experiment is also a popular means for detecting ligand binding that includes the SALMON [70] screening method. As an illustration, over 200 papers have been published within the last five years that utilizes STD-NMR. Recent examples include using STD-NMR to identify inhibitors for the innate immune system protein CD14 [124], carnitine acetyltransferase [125], β-

Ketoacyl-acyl carrier protein synthase III [126], MurD ligase [127], and yeast hexokinase [128]. In addition to identifying potential inhibitors, STD-NMR experiments have been used to understand the binding mode of α-conotoxins to acetylcholine binding protein (AChBP) from *Lymnea stagnalis* [129] and to determine the binding site for substrates for galactofuranosyltransferase GlfT2 [130].

Similar to the 1D ^{1}H line-broadening experiments, STD-NMR benefits from a low protein concentration, relative ease of implementation and data analysis, and the increase in sensitivity with large MW proteins [57, 131]. Additionally, STD-NMR generally doesn't require mixture deconvolution because only bound ligands produce an NMR signal (Figure 3). In the STD experiment, a binding interaction is identified by a transfer of magnetization from the protein target to the bound ligand. The transfer of magnetization within the protein occurs via ^{1}H-^{1}H cross relaxation and ligands that bind to the protein will experience a magnetization transfer through the binding interface. The ligand is usually in 20 to 30-fold excess relative to the protein. So for fast-exchange, a buildup of transferred magnetization occurs through multiple interactions between ligands and the protein target during the selective protein saturation pulse.

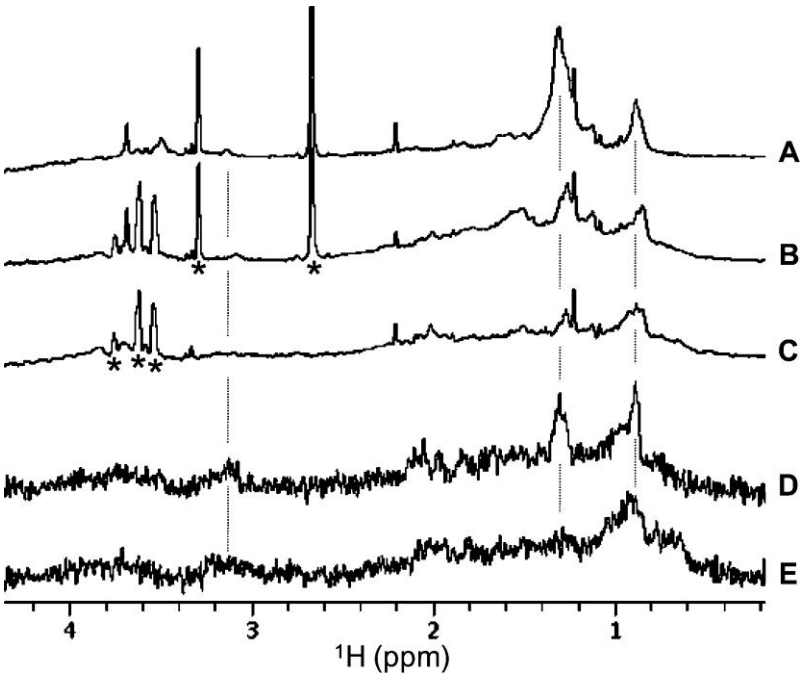

Figure 3. A comparison of line broadening and saturation transfer difference experiments performed on an inhibitor of the innate immune response protein CD14. The compound methyl 6-deoxy-6-amino-2,3-di-*O*-tetradecyl-α-d-glucopyranosidew was screened with sCD14. (A−C) 1D ^{1}H NMR spectra of (A) 19 μM free ligand, (B) 19 μM ligand with 0.5 μM sCD14, and (C) 0.5 μM sCD14 alone. (D) STD NMR spectrum obtained for sample B only. (E) STD NMR spectrum collected for control sample C. Comparison between D and E shows that the ligand binds to the sCD14. The stars denote ^{1}H peaks derived from solvents (e.g., DMSO at 2.66 ppm, methanol at 3.29 ppm) and/or buffer components (e.g., glycerol from sCD14 initial stock at 3.5−3.8 ppm). (Reprinted with permission from reference [124], Copyright 2009 by the American Chemical Society).

The STD experiment is collected by interleaving NMR spectra with and without protein saturation, where the spectra are alternatively subtracted using phase cycling. Selectively saturating the protein NMR spectrum occurs via frequency-selective radio-frequency pulses over a narrow, compound free frequency window (~ 0.0 ppm) for a time (T_{sat}) of about 1 to 3 seconds. A second off-resonance saturation experiment is acquired by simply shifting the frequency-selective radio-frequency pulse to a frequency distant from any protein or ligand NMR resonances. The final STD spectrum will resemble the free ligand's 1D ^{1}H spectrum for a compound that binds the protein target because the intensity of the ligand's NMR spectrum with protein saturation is greater than without due to the saturation transfer. Conversely, a non-binding ligand will produce a null spectrum. The protein NMR resonances are suppressed due to the very low concentrations and relaxation filtering. Comparable to HTS-NMR screens that use 1D ^{1}H line-broadening experiments, it is also desirable to measure dissociation constant from the STD experiments. But, collecting a complete binding isotherm to measure a K_D is not practical for a high-throughput screen. Additionally, STD experiments are overly sensitive to weak, non-specific binders leading to a significant number of false positives, further undermining the value of directly measuring K_D values as part of an HTS screen.

1.3 *Target Based HTS-NMR Screening Methods*

The physiochemical environment of a protein's surface is an extremely sensitive probe for ligand interactions. The addition of a compound to a uniformly ^{15}N and/or ^{13}C labeled protein target can easily monitor a binding interaction based on chemical shift changes in the protein's NMR spectrum. Therefore, target based screening methods by NMR predominately use chemical shift perturbations (CSPs) to identify ligand binding and to characterize the ligand binding site. The CSPs are mapped to the surface of a protein, where a cluster of CSPs provide a visual identification of the ligand binding site. Conversely, the lack of a clear clustering pattern would suggest non-specific binding. CSPs are typically measured by overlaying 2D ^{1}H-^{15}N HSQC/HMQC, 2D ^{1}H-^{13}C HSQC/HMQC or 2D ^{1}H-^{15}N TROSY [132] spectra for the free protein and the protein ligand complex.

These NMR methods are incredibly informative and beneficial for the analysis of protein-ligand interactions, however; a major drawback to target based screening is the increase in experimental time required for data collection. Additionally, significantly higher protein sample concentrations are necessary to collect reliable HSQC, HMQC or TROSY spectra. Recent advances in NMR pulse programs focusing on better water suppression and rapid data acquisition have dramatically increased the throughput, improving the application to HTS-NMR. Spectra can be rapidly acquired in minutes using FHSQC [133] or SOFAST-HMQC [134]. The FHSQC and SOFAST-HMQC pulse scheme permits the collection of the NMR spectrum in a few seconds by dramatically reducing the recycle time and allowing for high repetition rates. The analysis of the NMR spectra can be further improved and simplified by incorporating deuterium labeling [135], selective residue labeling [136], or selective methyl labeling [137]. Basically, the spectrum is reduced to only contain NMR resonances for residues or methyl-groups associated with the protein's ligand binding site.

The SAR by NMR [44, 55] methodology effectively established HTS-NMR and is based on CSPs from 2D ^{1}H-^{15}N HSQC. Subsequent HTS-NMR assays like FAST-NMR [27, 28], MS/NMR [68], Multi-Step NMR [67], RAMPED-UP NMR [56],

SHAPES [138], and SMILI-NMR [71] similarly utilize CSPs from 2D ^{1}H-^{15}N HSQC. HTS-NMR based on CSPs and fragment-based libraries have been an extremely successful approach for drug discovery resulting in a number of compounds proceeding to clinical trials that includes inhibitors to matrix metalloproteinase [139], aurora kinase, cyclin-dependent kinase, and peroxisome proliferator-activated receptor [30, 140].

1.4 In-cell HTS NMR Methods

HTS-NMR techniques have established themselves over the past decade as invaluable and robust *in vitro* screening methods for drug discovery. However, the results of *in vitro* experiments may fail when transferred to the physiological constraints of a cell or organism. This failure is not attributed to the lack of identifying promising chemical leads from HTS-NMR, but a result of the *in vitro* hits being unable to pass through the cell membrane, being rapidly metabolized by other cellular components or being actively pumped out of the cell. To evolve a chemical lead to a drug, it is fundamentally essential for a compound to overcome these obstacles in order to demonstrate *in vivo* efficacy. Monitoring the activity of a chemical lead within the cell is one approach to address this issue. Cell-based HTS assays are a common method of screening compound libraries [141-143], but are prone to off-target side effects since the mechanism of *in vivo* activity is unknown. These problems may occur even for high affinity ligand-protein complexes [144]. Despite these difficulties, a number of drug discovery successes have been reported [145-147].

Recent advances with in-cell NMR technology provide an alternative approach to increase *in vivo* drug efficacy and are a valuable HTS-NMR advancement [148-151]. Due to the complexity of the metabolome and the large number of metabolites, in-cell NMR methods primarily use multi-dimensional data collection of uniformly labeled proteins. However methods for 1D *in vivo* screening have been reported [152]. In-cell NMR methods have been used to investigate protein conformational changes [153], protein post-translational modifications [154], or protein-protein interactions [72, 73] using 2D HSQC experiments. Since the structure of a protein can differ under *in vitro* and *in vivo* conditions, the resulting changes in the HSQC spectra require the complete re-assignment of the backbone resonances for in-cell proteins. Since traditional multi-dimensional NMR experiments for complete backbone assignment [155, 156] can take weeks to complete, keeping cells alive in an NMR spectrometer for that long is challenging at best. Advancements in NMR techniques for rapid data collection have shortened the time required to collect a complete set of NMR spectra for backbone assignments and structure determination [149, 157] that are applicable to live cells [158-163]. These methods include automated projection spectroscopy (APSY) [164], G-matrix Fourier transform NMR (GFT-NMR) [31, 165], high-resolution iterative frequency identification (HIFI-NMR) [166], projection-reconstruction NMR (PR-NMR) [167], and reduced-dimensionality NMR (RD-NMR) [168, 169].

In-cell NMR studies have primarily focused on addressing specific biological questions about proteins and related cellular processes, but the techniques are also applicable to ligand-affinity screens. The SMILI-NMR (Small Molecule Interactor

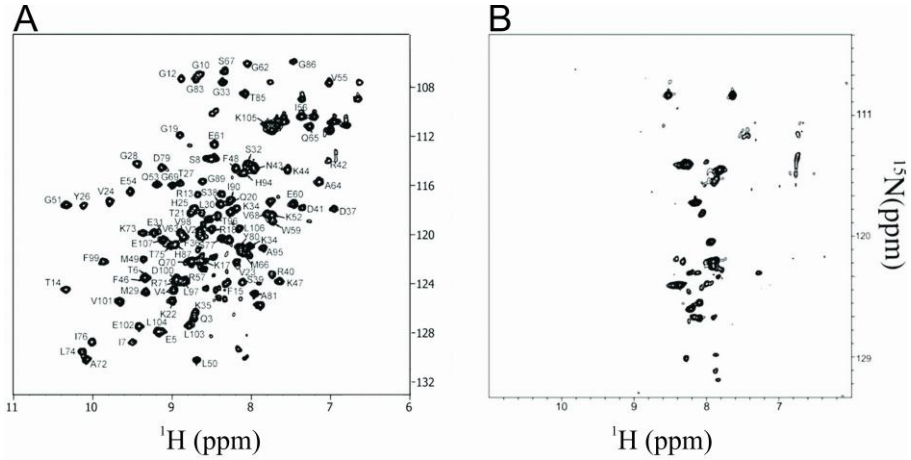

Figure 4. In-cell NMR methods provide a new avenue for NMR screening to help drug discovery efforts. A) A biocomplex is formed between FKBP and FRB upon sequential expression of uniformly ^{15}N labeled FKBP and unlabeled FRB. The biocomplex makes the FKBP visible to NMR by forming the lower molecular weight FKBP-FRB complex. B) Screening the FKBP-FRB complex with a peptide library identified a hit which disrupted the FKBP-FRB complex making the FKBP invisible to NMR. (Reprinted with permission from reference [71], Copyright 2009 by the American Chemical Society).

Library by In-cell NMR) method has been reported for screening ligands in whole cells [71]. SMILI-NMR is based on the formation of a complex between two interacting proteins similar to the STINT-NMR method [149, 170, 171]. The biomolecular complex contains one uniformly ^{15}N labeled protein visible by NMR, where CSPs indicate ligand binding and complex formation or disruption. The SMILI-NMR protocol was demonstrated using the FKBP-FRB interaction, which forms in the presence of rapamycin. The FKBP-FRB interaction is an important regulator of the mTOR (mammalian target of rapamycin) signaling pathway. A biomolecular complex was formed by the over-expression of uniformly ^{15}N labeled FKBP and unlabeled FRB in *E. coli.* The complex was only formed at sufficiently high concentrations of FRB, suggesting the FKBP was involved in a large molecular weight complex and therefore invisible to NMR. The sequential over-expression of FRB resulted in a detectable FKBP NMR spectrum due to the significantly lower molecular weight of the FKBP-FRB complex. Correspondingly, a ligand that disrupts the biomolecular complex causes resonances in the FKBP 2D ^{1}H-^{15}N HSQC spectrum to disappear (Figure 4).

1.5 Dissociation Constants from High-Throughput NMR Screening

As with any biochemical assay that detects ligand binding, NMR methods are governed by the same equilibrium and kinetic parameters, which are, in turn, dependent on the concentration of the free ligand $[L]_F$, free receptor $[P]_F$, and the receptor-ligand complex [PL]. For single-site binding, the relative ratios of these concentrations are governed by the on (k_{on}) and off (k_{off}) rates between the free and bound forms as described in eqn 7.

$$[L]_F + [P]_F \underset{k_{off}}{\overset{k_{on}}{\rightleftharpoons}} [PL] \quad (7)$$

The relative strength of a ligand's binding affinity is quantified by the dissociation constant (K_D):

$$K_D = \frac{k_{off}}{k_{on}} = \frac{[L]_F[P]_F}{[PL]} \quad (8)$$

Correspondingly, the bound receptor fraction (fractional occupancy, f_B) is described by $f_B = [PL]/([P]_F + [L]_F)$. When the ligand is in large excess of receptor, the total ligand concentration ($[L]_T$) is approximately equal to the free ligand concentration. Combining the definition of K_D with the expression for the bound receptor fraction leads to eq. 9.

$$f_b = \frac{[PL]}{[L]_T} = \frac{1}{1 + \frac{K_D}{[P]_F}} \quad (9)$$

Typically, NMR ligand affinity screens are not conducted under conditions where the total ligand concentration is in excess of the maximum complex concentration. Also, the direct measurement of the free protein concentration is generally not possible. Therefore, eq. 9 can be rearranged in terms of total protein ($[P]_T$) and total ligand concentration.

$$f_B = \frac{[PL]}{[L]_T} = \frac{1}{1 + \frac{2K_D}{([P]_T - [L]_T - K_D) + \sqrt{([P]_T - [L]_T + K_D)^2 + 4K_D[L]_T}}} \quad (10)$$

The ultimate goal of HTS-NMR is to identify ligands with the highest affinity for the protein of interest. Effectively, the NMR screen attempts to minimize K_D. Various NMR techniques allow for the direct measurement of K_D by monitoring changes in chemical shifts [172], line-broadening [29], diffusion rates [173], and a saturation transfer [174]. The most routine approach for measuring dissociation constants by NMR is the observation of protein chemical shift changes in a 2D ^{1}H-^{15}N HSQC spectrum as the ligand is slowly titrated into a constant protein concentration. Some recent examples include: suramin bound to *Pseudomonas aeroginosa* protein PA1324 (K_D ~ 51 μM) [175], 5′-CUCUCU-3′ bound to polypyrimidine tract-binding protein (PTB) (K_D ~ 0.93 to 3.2 μM) [176], *N*-phenyl-naphthylamine bound to mouse major urinary protein (MUP) (K_D ~ 0.032 μM) [177], and indole derivatives bound to $GABA_A$ receptor associated protein (GABARAP) (K_D ~ 6-208 mM) [178].

Comparable approaches have been developed using STD experiments, but require the introduction of an amplification factor (A_{STD}) [179]. The STD amplification factor is determined by normalizing each peak in the STD spectrum to the corresponding peak

in the off-resonance NMR spectrum followed by multiplying by the excess ligand concentration relative to the protein concentration. Because only bound ligands experience a magnetization transfer, the intensity of an STD response (I_{STD}) is directly proportional to the ligand-receptor complex [PL] concentration:

$$I_{STD} = C\alpha_{STD}[PL] \tag{11}$$

where C is a constant and α_{STD} is a scaling factor based on the STD enhancement. The reference intensity (I_o) is proportional to the total ligand concentration and the ratio of bond and free ligand signal (η_{STD}) is:

$$\eta_{STD} = \frac{I_{STD}}{I_o} = \frac{\alpha_{STD}[PL]}{[L]_T} = \alpha_{STD} f_B \tag{12}$$

where f_B is the fractional occupancy from eq. 9. A K_D can be measured by simply fitting the STD amplification factor as a function of ligand concentration to eq. 13, which is similar to the standard Henri-Michaelis-Menten equation.

$$A_{STD} = \left(\frac{\alpha_{STD}[L]}{[L] + K_D} \right) \tag{13}$$

Nevertheless, collecting a series of STD or 2D ^{1}H-^{15}N HSQC spectra to obtain a complete binding isotherm in the context of a high-throughput screen for hundreds to thousands of compounds is clearly not practical. As a result, a number of ligand-detected 1D ^{1}H NMR experiments have been developed to simultaneously identify binders and estimate dissociation constants from an HTS-NMR assay. Dalvit *et al.* [180] describe a single-point measurement using a 1D ^{1}H line-broadening or selective longitudinal relaxation (T_1) competition experiment. Basically, a compound with a known K_D and complete NMR titration curve is used as a reference or spy compound and is included within each mixture screened by NMR. The presence of a second compound in the mixture that competes with the reference compound will result in a proportional change in the reference compound's line-width or peak intensity. This enables the determination of $[PL]/[L]_T$ by simply reading the value off of the reference compound's original NMR titration curve from the observed intensity change. Given that $[L]_T$ and $[P]_T$ are known, an apparent K_D for the reference compound and a binding constant (K_I) for the new compound can be calculated:

$$K_D^{app} = \frac{[P]_T[L]_T - [P]_T[PL] + [PL]^2 - [L]_T[PL]}{[PL]} \qquad K_I = \frac{[I]K_D}{K_D^{app} - K_D} \tag{14}$$

Other HTS-NMR competition screens that also simultaneously measure K_Ds have been proposed using STD [181], c-WaterLOGSY [182], line-broadening changes [183], fluorine NMR [184], multi-selective NMR experiments [185], and spin-labels [186]. Of course, all these methods require a known binder with a defined dissociation constant that is not always available for a specific protein, especially a new drug

discovery target. Recently, Shortridge *et al.* [29] described a single-point measurement using 1D ^{1}H-line broadening that does not require a reference or spy molecule (Figure 2B). Line broadening experiments can rapidly measure a qualitative binding dissociation constant (K_D) of a protein-ligand complex in a single comparison between the free and bound states of the ligand [29]:

$$K_D = \left[\left(\frac{c[P]_T}{B_{single}} - c[P]_T \right) - [L]_T \right] \quad (15)$$

where B_{single} is the ratio of ligand intensities in the free (I_F) and bound (I_B) states and the constant *c* depends on the line widths of the free (ν_F) and bound (ν_B) ligand.

$$B_{single} = 1 - \frac{I_B}{I_F} \quad \text{with,} \quad c = \frac{\nu_B}{\nu_F} - 1 \quad (16)$$

The bound ligand line width can be approximated by the protein line width, which is dependent on the molecular weight of the protein. The Stokes-Einstein equation relates τ_c to molecular weight (MW) for a globular protein (Figures 2C,D):

$$\tau_c = \frac{4\pi\rho\eta r^3}{3kT} \quad \text{with,} \quad \tau_c \approx \rho * \frac{MW}{2400}(ns) \quad (17)$$

where T is the temperature, k is the Boltzmann constant, η is the viscosity of the solvent, r is the radius, and ρ is the shape constant.

2. Ligand Bound Co-structures from NMR Data

Knowing that a ligand binds to a particular protein target is invaluable, but not sufficient. It is also critical to understand the structural details of the ligand protein interaction for a successful structure-based design approach to drug discovery [187, 188]. Generating an NMR structure for a protein-ligand complex is relatively routine with 1,466 NMR ligand bound structure deposited in the PDB as of December 9, 2009 [189]. Standard NMR methods used to generate protein-ligand co-structures are generally limited to tightly bound ligands and require complete NMR assignments. Applying these approaches as a component of HTS-NMR is impractical because of the large time commitment per structure. Furthermore, ligand bound co-structures should lead the optimization process to evolve the chemical leads into drug candidates. In this section, methods for generating rapid protein-ligand co-structures will be discussed, with a particular emphasis on using NMR data to drive the docking of weak binding ligands that are typically identified by HTS-NMR.

2.1 Co-structures of Weakly Bound Protein-Ligand Complexes Using NMR

The mapping of CSPs on a protein surface, as described in the target focused HTS-

NMR section, can only approximate a ligand binding site. An NMR method that uses CSPs to improve the description of the ligand binding site to generate a modeled co-structures was described by Medek *et al.* [190]. A series of high affinity ligands are used to generate a differential chemical shift map to dock additional chemically related ligands to an existing NMR or x-ray protein structure. Basically, chemical alterations (different substituents or chemical moieties) between the ligand structures are correlated with the different CSPs observed between ligands to orient the new compound within the protein's active-site. Simply, an additional substituent is positioned proximal to the amino-acid that uniquely incurred a CSP with the new ligand. The NMR-DOC [191] approach simplifies the analysis of differential CSPs developed by Medek *et al.* [190] by using a series of residue-specific ^{13}C-methly labeled protein samples. The orientation of the ligand is then simply based on which of these specifically-labeled residues incurs a CSP upon the addition of the ligand. The SOS-NMR [192] technique is comparable to NMR-DOC, but differential STDs instead of CSPs are observed between the ligand and specifically deuterated labeled residues. Every residue type except one is deuterium labeled in a series of protein samples. If an unlabeled residue is proximal to the ligand then an STD is still observed. The ligand's orientation is based on which deuterated residues are necessary to observe an STD spectrum. The NMR-SOLVE [193] method uses a reference ligand (natural cofactor) to identify chemical shifts and NOEs associated with a second ligand binding site. This information is then used to map the location and orientation of novel compounds relative to the reference ligand.

McCoy and Wyss [194] further developed the CSP approach by comparing experimental and predicted CSPs using a three step procedure that first measures experimental proton chemical shifts between the free and bound states of an ^{15}N labeled protein. The program SHIFTS [195] is then used to predict the protein chemical shift for both the ligand-free and a ligand-bound structure, where the simulated CSPs are simply the difference between these two calculated proton chemical shifts. Finally, an iterative alignment of the ligand is performed to minimize the difference between the experimental and predicted CSPs. The method was applied to successfully generate a structure for the Ca^{2+}-CaM-W-7 complex. McCoy and Wyss extended their method by using the electron current density from the aromatic rings of the interacting ligand to dock the compound [196, 197]. The aromatic ring current effect is approximated as a single point dipole to estimate potential ligand locations. Thus, a dot-density plot for each observed CSP is described by a volume of randomly distributed dots. Overlapping dot-density maps are plotted using a GRASP surface [198] and the highest densities are predicted to be the most likely ligand binding site (Figure 5).

To extend the utility of chemical shift mapping to generate protein-ligand co-structures, a number of NMR-based methods have incorporated computer aided docking. Docking programs such as AutoDock [199], FRED [200], and DOCK [201] can rapidly generate thousands of structures in a high-throughput fashion. The selection or ranking of the best-structures are based on an empirical energy scoring function. These scoring functions are robust, but are also often dependent on a particular protein target and virtual library [202], and have a high false-positive rate [78]. This can provide misleading information about a "correct" binding site or pose for a ligand unless *a priori* knowledge of a binding site is used to help assist and evaluate the docked structures. In this manner, NMR CSP data can be used to guide and filter the molecular docking of weakly bound ligands.

One of the first methods to use a combination of ligand induced chemical shift

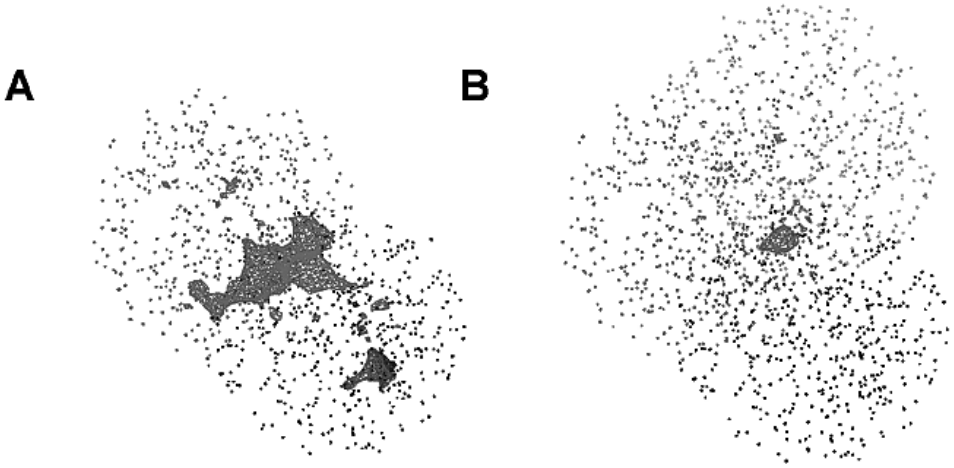

Figure 5. Identification of active sites via chemical shift mapping. The changes in chemical shifts are represented as a dot-density distribution (I x N) with overlapping dots approximating the binding site of the ligand. The spheres (I) of ring current are centered on each perturbed HN with each sphere having (N) dots. Surfaces created from dots with high densities localize the ligand ring position (Reprinted with permission from reference [203], Copyright 1990 by the American Chemical Society).

mapping and computer aided docking was described by Lugovskoy *et al*. [204]. The method was used to analyze BH3Is ligand bound to Bcl-xL. CSPs from backbone amides were used to map and localize the ligand binding site. The program TreeDock [205] was then used to derive a ligand bound co-structure. The best representative structure was predicted to be the co-structure with the lowest Lennard-Jones potential energy within the CSP localized binding site. While the method successfully generated structure activity relationships for the series of Bcl-xL binding ligands, the use of the lowest energy potential to define the ligand pose can be problematic. Newer methods attempt to solve this limitation by using docking filtering techniques [24] or combination of chemical shift density restraints and docking [206].

The method described by Stark and Powers (Figure 6) uses a pseudo-distance (d_{CSP}) based on CSPs from 2D ^{1}H-^{15}N HSQC spectra [24]. The CSPs are used to minimize the structure search space by using a significantly reduced AutoDock 3D grid during the docking calculation that conforms to the protein's surface defined by the CSPs. AutoDock 4.0 is then used to generate 100 docked protein-ligand co-structures using the Lamarckian search algorithm with a population size of 300 and 500,000 energy evaluations [207]. The AutoDockFilter (ADF) program then uses an NMR energy function based on the magnitude of CSPs to select the best ligand conformation or pose:

$$E_{NMR} = k\sum_{i=1}^{n}(\Delta_{Dist})^2 \qquad \Delta_{Dist} = \begin{cases} d_{CSP} & d_{CSP} < d_S \\ 0 & d_S \le d_{CSP} \end{cases} \tag{18}$$

where ADF calculates a pseudo-distance (d_{CSP}) based on the magnitude of the NH CSP, which is then compared to the shortest distance (d_S) between any atom in the residue

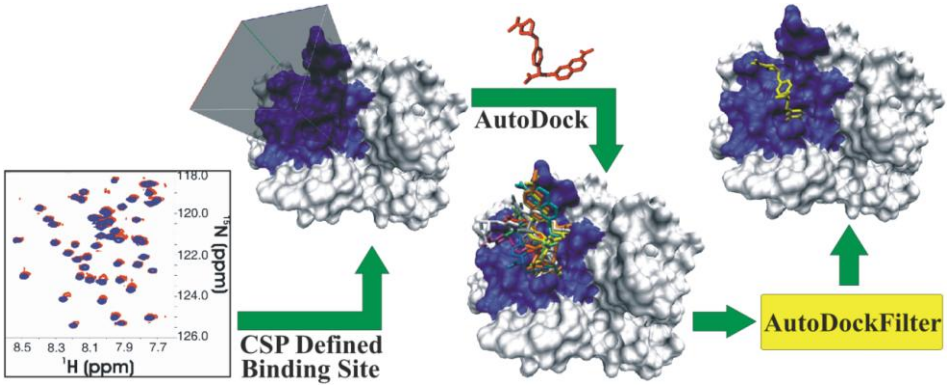

Figure 6. Flow diagram illustrating the application of NMR chemical shift perturbations and computer aided docking to rapidly generate a protein-ligand co-structure. The CSPs are used to map the ligand binding site and guide the AutoDock 3D search grid. A total of 100 conformers are generated bound in the NMR defined ligand binding site. The ADF program selects the best pose based on a consistency with the magnitude of observed CSPs. (Reprinted with permission from reference [24], Copyright 2008 by the American Chemical Society).

that incurred an NH CSP and any atom in each docked ligand conformer. Comparison of these CSP-directed and selected ligand-docked structures with experimental x-ray and NMR structures has yielded an overall average rmsd of 1.17 ± 0.74 Å [24]. The determination of PrgI-ligand co-structures is a recent example of using CSPs to guide and filter a protein-ligand complex [208].

The HADDOCK program [209] has also successfully used CSPs to generate protein-ligand complexes [31]. CSPs are used to define ambiguous interaction constraints (AIR), an intermolecular distance (≤ 3Å) between all sets of residues that incurred a CSP to the ligand [210]. Thus, a co-structure is directly refined against the experimental CSPs, which requires three steps. First, the ligand is positioned in the binding site with a rigid body docking and energy minimization, where the protein structure remains fixed. This is followed by a semi-rigid simulated annealing in torsional angle space to enable re-orientation of both the protein and ligand to further optimize the interaction. A final refinement with explicit solvent removes clashes, other structural problems and optimizes hydrogen-bond interactions.

3. NMR Screening and Functional Annotation

Over the last decade, the pharmaceutical industry has experienced an unprecedented decline in creativity and productivity leading to a dramatic reduction in new drugs [211-215]. Drug discovery is a complex and challenging endeavor and a number of factors are contributing to this decline. But one key issue is the concentrated effort in a limited number of research areas [216]. This results in multiple companies simultaneously pursuing therapeutics against the same small set of protein targets. This inevitably leads to a high-level of competition, a limited number of new therapies and a number of redundant drugs [217] serving the same market [218]. Drug discovery would benefit from new methodologies that enhance the identification of novel therapeutic targets.

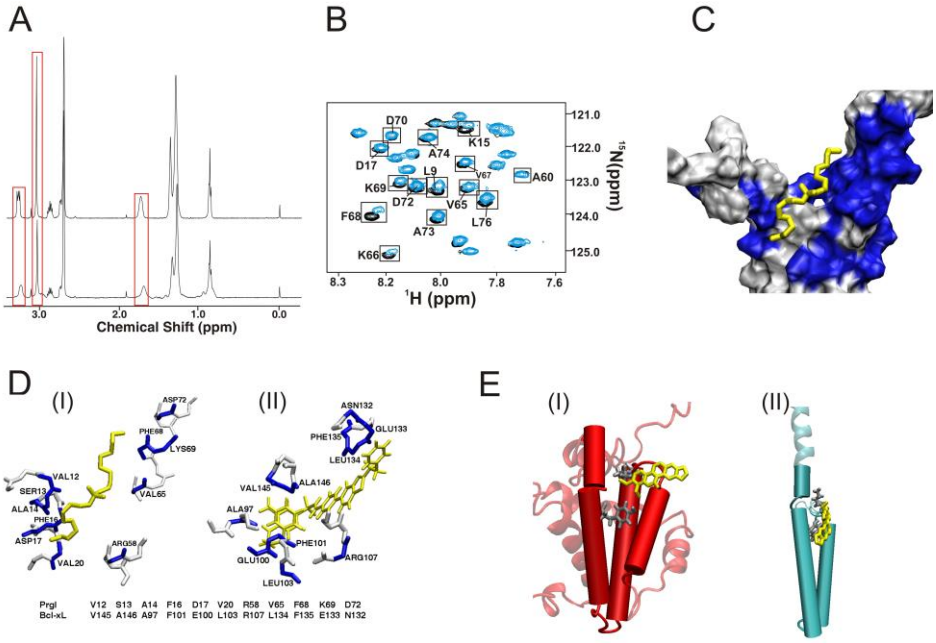

Figure 7. FAST-NMR approach to screening with the target protein PrgI from *S. typhirium.* A) 1D-^{1}H line broadening experiments identified 5 compounds that bound PrgI from a chemical library of ~450 compounds. B) Secondary screening using ^{1}H-^{15}N HSQC of the PrgI protein showed that only one compound of the 5 initial hits, didecyldimethylammonium bromide, specifically bound PrgI with an approximate K_D of 553 μM. C) A rapid ligand bound co-structures was generated and used for functional annotation by active site similarity searching. The protein with the highest similarity to PrgI active site D(I) was the Bcl-2 family protein, Bcl-xL bound to ABT-737 D(II). The similarity between PrgI and Bcl-xL binding sites was used to predict the known Bcl-xL inhibitor, chelerythrine E(I), was also a binder of PrgI Bcl-xL in a similar manner to PrgI E(II). (Reprinted with permission from reference [208], Copyright 2009 by the Public Library of Science).

With the rapid increase in complete genome sequences of various organisms [219-223], there is an ever expanding data set available to identify new drug discovery targets. Bioinformatic methods attempt to annotate function using protein sequence and structure similarities [224, 225] or a combination of multiple experimental and computational approaches [226]. Nevertheless, it is still challenging and error prone to predict function using global sequence and structure similarity alone [227, 228]. A growing trend in protein function prediction is using active-site similarity [27, 229, 230] to identify regions on a protein that interact with biologically important compounds. Active-sites or functional epitopes with similar sequence and structure, and that bind similar ligands are considered to be related by function. Protein active-sites tend to be evolutionary more stable relative to the remainder of a protein structure [231]. A major obstacle to leveraging active-site similarity to infer biological function is the accurate identification of ligand binding sites and the corresponding ligands.

As described in detail above, NMR is uniquely capable of characterizing ligand binding sites and screening chemical libraries to identify hits. Correspondingly, FAST-NMR [27, 28] employs a multi-step NMR screening protocol [67] and a biologically active ligand library to identify binding ligands and their representative binding site to

annotate proteins of unknown function. The experimentally determined ligand, ligand-binding site and protein-ligand co-structure provide valuable resources for active site similarity searches (CPASS) [232]. The multi-step NMR screening technique combines a 1D ^{1}H line-broadening screen to rapidly identify binding ligands, followed by a 2D ^{1}H-^{15}N HSQC NMR screen using only the hits from the 1D ^{1}H line-broadening screen to map the ligand-binding site and generate a co-structure with CSPs. The multi-step NMR approach reduces both the total time required to complete the screen along with minimizing sample requirements. As an example, screening a library of about 450 compounds [107] using the FAST-NMR multi-step screening method takes ~11 hrs with an average hit rate of ~20 compounds. Coupled with rapid active site similarity search tools [232] and other bioinformatics techniques, the functional annotation of a protein of unknown function can be reduced to just days. The FAST-NMR method has been used to functionally annotate a number of novel proteins and build supporting evidence for the functional similarity between the type 3 secretion system proteins PrgI and the Bcl-2 family of apoptosis regulators [27, 28, 175, 208] (Figure 7).

4. Conclusions

The NMR has proven to be a valuable and versatile tool for the drug discovery community. A variety of HTS-NMR techniques that include chemical shift perturbations from HSQC experiments, 1D ^{1}H line-broadening experiments and saturation transfer difference experiments are routinely used to screen chemical libraries to identify chemical leads. Additionally, various single-point and competition techniques have been developed to obtain dissociation constants from NMR ligand affinity screens. Efficient library designs coupled with fragment-based approaches have increased the efficiency and success rate of HTS-NMR screens and the drug discovery process. Furthermore, the CSPs, STDs and NOEs from HTS-NMR screens have been combined with ligand-docking software to generate rapid protein-ligand co-structures to aid in the further optimization of chemical leads. Finally, the NMR techniques developed for drug discovery have also been used to develop the FAST-NMR methodology for protein functional annotation. FAST-NMR further benefits drug discovery by aiding in the identification of novel therapeutic targets.

References

[1] F. Sams-Dodd, Drug discovery: selecting the optimal approach. *Drug Discov. Today*, **11** (2006), 465-472.

[2] G.C. Terstappen, C. Schluepen, R. Raggiaschi, and G. Gaviraghi, Target deconvolution strategies in drug discovery. *Nat. Rev. Drug Discovery*, **6** (2007), 891-903.

[3] J.P.E. Karlberg, Trends in disease focus of drug development. *Nat. Rev. Drug Discovery*, **7** (2008), 639-640.

[4] A.C. Anderson, The Process of Structure-Based Drug Design. *Chem. Biol.*, **10** (2003), 787-797.

[5] G.M. Lee and C.S. Craik, Trapping Moving Targets with Small Molecules. *Science*, **324** (2009), 213-215.

[6] O. Keskin, A. Gursoy, B. Ma, and R. Nussinov, Towards drugs targeting multiple proteins in a systems biology approach. *Curr. Top. Med. Chem. (Sharjah, United Arab Emirates)*, **7** (2007), 943-951.

[7] J.W. Armstrong, A review of high-throughput screening approaches for drug discovery. *Am. Biotechnol. Lab.*, **17** (1999), 26-28.

[8] J.E. Gonzalez and P.A. Negulescu, Intracellular detection assays for high-throughput screening. *Curr. Opin. Biotechnol.*, **9** (1998), 624-631.
[9] K.R. Oldenburg, Current and future trends in high throughput screening for drug discovery. *Annu. Rep. Med. Chem*, **33** (1998), 301-311.
[10] P.B. Fernandes, Technological advances in high-throughput screening. *Curr. Opin. Chem. Biol.*, **2** (1998), 597-603.
[11] B.A. Kenny, M. Bushfield, D.J. Parry-Smith, S. Fogarty, and J.M. Treherne, The application of high-throughput screening to novel lead discovery. *Prog. Drug Res.*, **51** (1998), 245-269.
[12] J.T. Koh, Making virtual screening a reality. *Proc. Natl. Acad. Sci. U. S. A.*, **100** (2003), 6902-6903.
[13] S.S. Phatak, C.C. Stephan, and C.N. Cavasotto, High-throughput and in silico screenings in drug discovery. *Expert Opin. Drug Discovery*, **4** (2009), 947-959.
[14] M.J. McGregor, Z. Luo, and X. Jiang, Virtual screening in drug discovery. *Drug Discovery Res.*, 2007), 63-88, 2 plates.
[15] G.M. Rishton, Reactive compounds and in vitro false positives in HTS. *Drug Discov. Today*, **2** (1997), 382-384.
[16] S.L. McGovern, E. Caselli, N. Grigorieff, and B.K. Shoichet, A Common Mechanism Underlying Promiscuous Inhibitors from Virtual and High-Throughput Screening. *J. Med. Chem.*, **45** (2002), 1712-1722.
[17] S.L. McGovern, B.T. Helfand, B. Feng, and B.K. Shoichet, A Specific Mechanism of Nonspecific Inhibition. *J. Med. Chem.*, **46** (2003), 4265-4272.
[18] J. Seidler, S.L. McGovern, T.N. Doman, and B.K. Shoichet, Identification and Prediction of Promiscuous Aggregating Inhibitors among Known Drugs. *J. Med. Chem.*, **46** (2003), 4477-4486.
[19] D.R. Goode, R.K. Totten, J.T. Heeres, and P.J. Hergenrother, Identification of Promiscuous Small Molecule Activators in High-Throughput Enzyme Activation Screens. *J. Med. Chem.*, **51** (2008), 2346-2349.
[20] T. Polgar, A. Baki, G.I. Szendrei, and G.M. Keseru, Comparative Virtual and Experimental High-Throughput Screening for Glycogen Synthase Kinase-3β Inhibitors. *J. Med. Chem.*, **48** (2005), 7946-7959.
[21] T.N. Doman, S.L. McGovern, B.J. Witherbee, T.P. Kasten, R. Kurumbail, W.C. Stallings, D.T. Connolly, and B.K. Shoichet, Molecular docking and high-throughput screening for novel inhibitors of protein tyrosine phosphatase-1B. *J. Med. Chem.*, **45** (2002), 2213-2221.
[22] A.M. Paiva, D.E. Vanderwall, J.S. Blanchard, J.W. Kozarich, J.M. Williamson, and T.M. Kelly, Inhibitors of dihydrodipicolinate reductase, a key enzyme of the diaminopimelate pathway of Mycobacterium tuberculosis. *Biochim. Biophys. Acta, Protein Struct. Mol. Enzymol.*, **1545** (2001), 67-77.
[23] A. Dove and A. Marshall, Drug screening-beyond the bottleneck. *Nat. Biotechnol.*, **17** (1999), 859-863.
[24] J. Stark and R. Powers, Rapid protein-ligand costructures using chemical shift perturbations. *J. Am. Chem. Soc.*, **130** (2008), 535-45.
[25] M. Pellecchia, I. Bertini, D. Cowburn, C. Dalvit, E. Giralt, W. Jahnke, T.L. James, S.W. Homans, H. Kessler, C. Luchinat, B. Meyer, H. Oschkinat, J. Peng, H. Schwalbe, and G. Siegal, Perspectives on NMR in drug discovery: a technique comes of age. *Nat Rev Drug Discov*, **7** (2008), 738-45.
[26] R. Powers, Advances in nuclear magnetic resonance for drug discovery. *Expert Opin. Drug Discovery*, **4** (2009), 1077-1098.
[27] K.A. Mercier, M. Baran, V. Ramanathan, P. Revesz, R. Xiao, G.T. Montelione, and R. Powers, FAST-NMR: functional annotation screening technology using NMR spectroscopy. *J. Am. Chem. Soc.*, **128** (2006), 15292-9.
[28] R. Powers, K.A. Mercier, and J.C. Copeland, The application of FAST-NMR for the identification of novel drug discovery targets. *Drug Discov. Today*, **13** (2008), 172-9.
[29] M.D. Shortridge, D.S. Hage, G.S. Harbison, and R. Powers, Estimating protein-ligand binding affinity using high-throughput screening by NMR. *J. Comb. Chem.*, **10** (2008), 948-58.
[30] P.J. Hajduk and J. Greer, A decade of fragment-based drug design: strategic advances and lessons learned. *Nat. Rev. Drug Discovery*, **6** (2007), 211-219.
[31] G. Liu, Y. Shen, H.S. Atreya, D. Parish, Y. Shao, D.K. Sukumaran, R. Xiao, A. Yee, A. Lemak, A. Bhattacharya, T.A. Acton, C.H. Arrowsmith, G.T. Montelione, and T. Szyperski, NMR data collection and analysis protocol for high-throughput protein structure determination. *Proc. Natl. Acad. Sci. U. S. A.*, **102** (2005), 10487-10492.
[32] S. Tomaselli, L. Ragona, L. Zetta, M. Assfalg, P. Ferranti, R. Longhi, A.M. Bonvin, and H. Molinari, NMR-based modeling and binding studies of a ternary complex between chicken liver bile acid binding protein and bile acids. *Proteins*, **69** (2007), 177-91.
[33] R. Powers, NMR metabolomics and drug discovery. *Magn. Reson. Chem.*, **47** (2009), S2-S11.

[34] D. Ekonomiuk, X.C. Su, K. Ozawa, C. Bodenreider, S.P. Lim, Z. Yin, T.H. Keller, D. Beer, V. Patel, G. Otting, A. Caflisch, and D. Huang, Discovery of a non-peptidic inhibitor of west nile virus NS3 protease by high-throughput docking. *PLoS Negl. Trop. Dis.*, **3** (2009), e356.
[35] J. Liu, D. Begley, D.D. Mitchell, C.L. Verlinde, G. Varani, and E. Fan, Multivalent drug design and inhibition of cholera toxin by specific and transient protein-ligand interactions. *Chem. Biol. Drug Des.*, **71** (2008), 408-19.
[36] Z. Shi, S. Tabassum, W. Jiang, J. Zhang, S. Mathur, J. Wu, and Y. Shi, Identification of a potent inhibitor of human dual-specific phosphatase, VHR, from computer-aided and NMR-based screening to cellular effects. *ChemBioChem*, **8** (2007), 2092-9.
[37] U. Rothweiler, A. Czarna, M. Krajewski, J. Ciombor, C. Kalinski, V. Khazak, G. Ross, N. Skobeleva, L. Weber, and T.A. Holak, Isoquinolin-1-one inhibitors of the MDM2-p53 interaction. *ChemMedChem*, **3** (2008), 1118-28.
[38] D.H. Tsao, A.G. Sutherland, L.D. Jennings, Y. Li, T.S. Rush, 3rd, J.C. Alvarez, W. Ding, E.G. Dushin, R.G. Dushin, S.A. Haney, C.H. Kenny, A.K. Malakian, R. Nilakantan, and L. Mosyak, Discovery of novel inhibitors of the ZipA/FtsZ complex by NMR fragment screening coupled with structure-based design. *Bioorg. Med. Chem.*, **14** (2006), 7953-61.
[39] D. Grandy, J. Shan, X. Zhang, S. Rao, S. Akunuru, H. Li, Y. Zhang, I. Alpatov, X.A. Zhang, R.A. Lang, D.L. Shi, and J.J. Zheng, Discovery and characterization of a small molecule inhibitor of the PDZ domain of dishevelled. *J. Biol. Chem.*, **284** (2009), 16256-63.
[40] M. Brewer, O. Ichihara, C. Kirchhoff, M. Schade, and M. Whittaker, *Assembling a Fragment Library*, in *Fragment-Based Drug Discovery*, M.J.S. Edward R. Zartler, Editor. 2008, John Wiley & Sons, Ltd: Hoboken, NJ 39-62.
[41] A.L. Hopkins, C.R. Groom, and A. Alex, Ligand efficiency: a useful metric for lead selection. *Drug Discov. Today*, **9** (2004), 430-1.
[42] A. Schuffenhauer, S. Ruedisser, A.L. Marzinzik, W. Jahnke, M. Blommers, P. Selzer, and E. Jacoby, Library design for fragment based screening. *Curr. Top. Med. Chem. (Sharjah, United Arab Emirates)*, **5** (2005), 751-762.
[43] T. Fink, H. Bruggesser, and J.-L. Reymond, Virtual exploration of the small-molecule chemical universe below 160 D. *Angew. Chem., Int. Ed.*, **44** (2005), 1504-1508.
[44] P.J. Hajduk, SAR by NMR: putting the pieces together. *Mol. Interventions*, **6** (2006), 266-272.
[45] A. Smith, Screening for drug discovery: the leading question. *Nature*, **418** (2002), 453-9.
[46] J. Drews, Drug discovery: A historical perspective. *Science*, **287** (2000), 1960-1963.
[47] J. Bajorath, Rational drug discovery revisited: interfacing experimental programs with bio- and chemo-informatics. *Drug Discov. Today*, **6** (2001), 989-995.
[48] R. Lahana, How many leads from HTS? *Drug Discov. Today*, **4** (1999), 447-448.
[49] J. Fejzo, C.A. Lepre, J.W. Peng, G.W. Bemis, Ajay, M.A. Murcko, and J.M. Moore, The SHAPES strategy: and NMR-based approach for lead generation in drug discovery. *Chem. Biol.*, **6** (1999), 755-769.
[50] P.J. Hajduk, E.T. Olejniczak, and S.W. Fesik, One-Dimensional Relaxation- and Diffusion-Edited NMR Methods for Screening Compounds That Bind to Macromolecules. *J. Am. Chem. Soc.*, **119** (1997), 12257-12261.
[51] S.B. Shuker, P.J. Hajduk, R.P. Meadow, and S.W. Feisk, Discovering High-Affinity Ligands for Proteins: SAR by NMR. *Science*, **274** (1996), 1531-1534.
[52] G. Wagner, S.G. Hyberts, and T.F. Havel, NMR Structure Determination in Solution: A Critique and Comparison with X-Ray Crystallography. *Annu. Rev. Biophys. Biomol. Struct*, **21** (1992), 167-198.
[53] S. Vanwetswinkel, R.J. Heetebrij, J.v. Duynhoven, J.G. Hollander, D.V. Filippov, P. Hajduk, and G. Siegal, TINS, Target Immobilized NMR Screening. *Chem. Biol.*, **12** (2005), 207-216.
[54] J.W. Peng, J. Moore, and N. Abdul-Manan, NMR experiments for lead generation in drug discovery. *Prog. Nucl. Magn. Reson. Spectrosc.*, **44** (2004), 225-256.
[55] S.B. Shuker, P.J. Hajduk, R.P. Meadows, and S.W. Fesik, Discovering high-affinity ligands for proteins: SAR by NMR. *Science*, **274** (1996), 1531-4.
[56] E.R. Zartler, J. Hanson, B.E. Jones, A.D. Kline, G. Martin, H. Mo, M.J. Shapiro, R. Wang, H. Wu, and J. Yan, RAMPED-UP NMR: multiplexed NMR-based screening for drug discovery. *J. Am. Chem. Soc.*, **125** (2003), 10941-6.
[57] M. Mayer and B. Meyer, Characterization of ligand binding by saturation transfer difference NMR spectroscopy. *Angew. Chem., Int. Ed.*, **38** (1999), 1784-1788.
[58] D.S. Sem, B. Bertolaet, B. Baker, E. Chang, A.D. Costache, S. Coutts, Q. Dong, M. Hansen, V. Hong, X. Huang, R.M. Jack, R. Kho, H. Lang, C.-T. Ma, D. Meininger, M. Pellecchia, F. Pierre, H. Villar, and L. Yu, Systems-Based Design of Bi-Ligand Inhibitors of Oxidoreductases Filling the Chemical Proteomic Toolbox. *Chem. Biol.*, **11** (2004), 185-194.

[59] C. Dalvit, P. Pevarello, M. Tato, M. Veronesi, A. Vulpetti, and M. Sundstrom, Identification of compounds with binding affinity to proteins via magnetization transfer from bulk water. *J. Biomol. NMR*, **18** (2000), 65-68.

[60] M. Bista, K. Kowalska, W. Janczyk, A. Doemling, and T.A. Holak, Robust NMR Screening for Lead Compounds Using Tryptophan-Containing Proteins. *J. Am. Chem. Soc*., **131** (2009), 7500-7501.

[61] L. D'Silva, P. Ozdowy, M. Krajewski, U. Rothweiler, M. Singh, and T.A. Holak, Monitoring the Effects of Antagonists on Protein-Protein Interactions with NMR Spectroscopy. *J. Am. Chem. Soc*., **127** (2005), 13220-13226.

[62] C. Dalvit, P.E. Fagerness, D.T.A. Hadden, R.W. Sarver, and B.J. Stockman, Fluorine-NMR experiments for high-throughput screening: Theoretical aspects, practical considerations, and range of applicability. *J. Am. Chem. Soc.*, **125** (2003), 7696-7703.

[63] W. Jahnke, S. Ruedisser, and M. Zurini, Spin label enhanced NMR screening. *J. Am. Chem. Soc*., **123** (2001), 3149-3150.

[64] M. Lin, M.J. Shapiro, and J.R. Wareing, Diffusion-Edited NMR-Affinity NMR for Direct Observation of Molecular Interactions. *J. Am. Chem. Soc.*, **119** (1997), 5249-5250.

[65] C. Dalvit, E. Ardini, G.P. Fogliatto, N. Mongelli, and M. Veronesi, Reliable high-throughput functional screening with 3-FABS. *Drug Discov. Today*, **9** (2004), 595-602.

[66] A. Chen and M.J. Shapiro, NOE Pumping: A Novel NMR Technique for Identification of Compounds with Binding Affinity to Macromolecules. *J. Am. Chem. Soc.*, **120** (1998), 10258-10259.

[67] K.A. Mercier, M.D. Shortridge, and R. Powers, A multi-step NMR screen for the identification and evaluation of chemical leads for drug discovery. *Comb. Chem. High Throughput Screen*, **12** (2009), 285-95.

[68] F.J. Moy, K. Haraki, D. Mobilio, G. Walker, R. Powers, K. Tabei, H. Tong, and M.M. Siegel, MS/NMR: a structure-based approach for discovering protein ligands and for drug design by coupling size exclusion chromatography, mass spectrometry, and nuclear magnetic resonance spectroscopy. *Anal. Chem.*, **73** (2001), 571-81.

[69] V.M. Sanchez-Pedregal, M. Reese, J. Meiler, M.J.J. Blommers, C. Griesinger, and T. Carlomagno, The INPHARMA method: Protein-mediated interligand NOEs for pharmacophore mapping. *Angew. Chem., Int. Ed.*, **44** (2005), 4172-4175.

[70] C. Ludwig, P.J.A. Michiels, X. Wu, K.L. Kavanagh, E. Pilka, A. Jansson, U. Oppermann, and U.L. Guenther, SALMON: Solvent Accessibility, Ligand binding, and Mapping of ligand Orientation by NMR Spectroscopy. *J. Med. Chem.*, **51** (2008), 1-3.

[71] J. Xie, R. Thapa, S. Reverdatto, D.S. Burz, and A. Shekhtman, Screening of small molecule interactor library by using in-cell NMR spectroscopy (SMILI-NMR). *J. Med. Chem.*, **52** (2009), 3516-22.

[72] D.S. Burz, K. Dutta, D. Cowburn, and A. Shekhtman, In-cell NMR for protein-protein interactions (STINT-NMR). *Nat. Protoc.*, **1** (2006), 146-152.

[73] D.S. Burz, K. Dutta, D. Cowburn, and A. Shekhtman, Mapping structural interactions using in-cell NMR spectroscopy (STINT-NMR). *Nat. Methods*, **3** (2006), 91-93.

[74] S. Vanwetswinkel, R.J. Heetebrij, J. van Duynhoven, J.G. Hollander, D.V. Filippov, P.J. Hajduk, and G. Siegal, TINS, Target Immobilized NMR Screening: An Efficient and Sensitive Method for Ligand Discovery. *Chem. Biol.*, **12** (2005), 207-216.

[75] R.S. Bohacek, C. McMartin, and W.C. Guida, The art and practice of structure-based drug design: a molecular modeling perspective. *Med. Res. Rev.*, **16** (1996), 3-50.

[76] O. Koji, I. Tetsu, Y. Hiroshi, and K. Hideo, A Quantitative Approach to the Estimation of Chemical Space from a Given Geometry by the Combination of Atomic Species. *QSAR Comb. Sci.*, **26** (2007), 596-607.

[77] R.S. Bohacek, C. McMartin, and W.C. Guida, The art and practice of structure-based drug design: a molecular modeling perspective. *Med. Res. Rev.*, **16** (1996), 3-50.

[78] H.J. Seifert Markus, Targeted scoring functions for virtual screening. *Drug Discov. Today*, **14** (2009), 562-9.

[79] J. Xu and J. Stevenson, Drug-like Index: A New Approach To Measure Drug-like Compounds and Their Diversity. *J. Chem. Inf. Comput. Sci.*, **40** (2000), 1177-1187.

[80] L. Xue and J. Bajorath, Molecular descriptors in chemoinformatics, computational combinatorial chemistry, and virtual screening. *Comb. Chem. High Throughput Screening*, **3** (2000), 363-372.

[81] R.A. Lewis, S.D. Pickett, and D.E. Clark, Computer-aided molecular diversity analysis and combinatorial library design. *Rev. Comput. Chem*, **16** (2000), 1-51.

[82] P. Willett, Chemoinformatics - similarity and diversity in chemical libraries. *Curr. Opin. Biotechnol.*, **11** (2000), 85-88.

[83] D.C. Spellmeyer and P.D.J. Grootenhuis, Recent developments in molecular diversity. Computational approaches to combinatorial chemistry. *Annu. Rep. Med. Chem*, **34** (1999), 287-296.

[84] D. Gorse and R. Lahana, Functional diversity of compound libraries. *Curr. Opin. Chem. Biol.*, **4** (2000), 287-294.
[85] P.D. Leeson and B. Springthorpe, The influence of drug-like concepts on decision-making in medicinal chemistry. *Nat. Rev. Drug Discovery*, **6** (2007), 881-890.
[86] C.A. Lipinski, F. Lombardo, B.W. Dominy, and P.J. Feeney, Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Delivery Rev.*, **46** (2001), 3-26.
[87] C.A. Lipinski, Lead- and drug-like compounds: the rule-of-five revolution. *Drug Discov. Today Technol.*, **1** (2004), 337-341.
[88] C.A. Lepre, Library design for NMR-based screening. *Drug Discov. Today*, **6** (2001), 133-140.
[89] K.A. Mercier and R. Powers, Determining the optimal size of small molecule mixtures for high throughput NMR screening. *J. Biomol. NMR*, **31** (2005), 243-58.
[90] D.A. Erlanson, Fragment-based lead discovery: A chemical update. *Curr. Opin. Biotechnol.*, **17** (2006), 643-652.
[91] M. Fischer and R.E. Hubbard, Fragment-based ligand discovery. *Mol. Interventions*, **9** (2009), 22-30.
[92] M. Congreve, C.W. Murray, R. Carr, and D.C. Rees, Fragment-based lead discovery. *Annu. Rep. Med. Chem.*, **42** (2007), 431-448.
[93] G. Siegal, E. Ab, and J. Schultz, Integration of fragment screening and library design. *Drug Discov. Today*, **12** (2007), 1032-1039.
[94] I.V. Tetko, J. Gasteiger, R. Todeschini, A. Mauri, D. Livingstone, P. Ertl, V.A. Palyulin, E.V. Radchenko, N.S. Zefirov, A.S. Makarenko, V.Y. Tanchuk, and V.V. Prokopenko, Virtual computational chemistry laboratory - design and description. *J. Comput.-Aided Mol. Des.*, **19** (2005), 453-463.
[95] O. Ivanciuc and A.T. Balaban, The graph description of chemical structures. in *Topological Indices and Related Descriptors in QSAR and QSPR*, J. Devillers and A.T. Balaban, Editors.1999, Gordon and Breach Science Publishers, The Netherlands, 59-167.
[96] Y.S. Prabhakar and M.K. Gupta, Chemical structure indices in *in silico* molecular design. *Sci. Pharm.*, **76** (2008), 101-132.
[97] A.T. Balaban, D. Ciubotariu, and M. Medeleanu, Topological indices and real number vertex invariants based on graph eigenvalues or eigenvectors. *J. Chem. Inf. Comput. Sci.*, **31** (1991), 517-23.
[98] J. Auer and J. Bajorath, Molecular similarity concepts and search calculations. *Methods Mol. Biol.*, **453** (2008), 327-347.
[99] P. Willett, Searching techniques for databases of two- and three-dimensional chemical structures. *J. Med. Chem.*, **48** (2005), 4183-99.
[100] D.E. Patterson, R.D. Cramer, A.M. Ferguson, R.D. Clark, and L.E. Weinberger, Neighborhood behavior: a useful concept for validation of "molecular diversity" descriptors. *J. Med. Chem.*, **39** (1996), 3049-59.
[101] E. Lasser Karen, D. Allen Paul, J. Woolhandler Steffie, U. Himmelstein David, M. Wolfe Sidney, and H. Bor David, Timing of new black box warnings and withdrawals for prescription medications. *JAMA*, **287** (2002), 2215-20.
[102] J.A. Kramer, J.E. Sagartz, and D.L. Morris, The application of discovery toxicology and pathology towards the design of safer pharmaceutical lead candidates. *Nat. Rev. Drug Discovery*, **6** (2007), 636-649.
[103] C.W. Andrews, L. Bennett, and L.X. Yu, Predicting human oral bioavailability of a compound: development of a novel quantitative structure-bioavailability relationship. *Pharm. Res.*, **17** (2000), 639-644.
[104] D.E. Clark and S.D. Pickett, Computational methods for the prediction of `drug-likeness'. **5** (2000), 49-58.
[105] J.C. Adams, M.J. Keiser, L. Basuino, H.F. Chambers, D.S. Lee, O.G. Wiest, and P.C. Babbitt, A mapping of drug space from the viewpoint of small molecule metabolism. *PLoS Comput. Biol.*, **5** (2009), e1000474.
[106] H. Mishra, N. Singh, T. Lahiri, and K. Misra, A comparative study on the molecular descriptors for predicting drug-likeness of small molecules. *Bioinformation*, **3** (2009), 384-8.
[107] G.M. Rishton, Reactive compounds and in vitro false positives in HTS. *Drug Discov. Today*, **2** (1997), 382-384.
[108] G. Chessari and A.J. Woodhead, From fragment to clinical candidate--a historical perspective. *Drug Discov. Today*, **14** (2009), 668-75.
[109] K. Eitner and U. Koch, From fragment screening to potent binders: strategies for fragment-to-lead evolution. *Mini Rev. Med. Chem.*, **9** (2009), 956-61.
[110] M. Pellecchia, Fragment-based drug discovery takes a virtual turn. *Nat. Chem. Biol.*, **5** (2009), 274-5.

[111] M.F. Schmidt and J. Rademann, Dynamic template-assisted strategies in fragment-based drug discovery. *Trends Biotechnol.*, **27** (2009), 512-21.
[112] K.A. Mercier, K. Germer, and R. Powers, Design and characterization of a functional library for NMR screening against novel protein targets. *Comb. Chem. High Throughput Screen*, **9** (2006), 515-34.
[113] A. Dove, Drug screening--beyond the bottleneck. *Nat. Biotechnol.*, **17** (1999), 859-63.
[114] C.A. Lepre, J. Peng, J. Fejzo, N. Abdul-Manan, J. Pocas, M. Jacobs, X. Xie, and J.M. Moore, Applications of SHAPES screening in drug discovery. *Comb. Chem. High Throughput Screen*, **5** (2002), 583-90.
[115] H. Widmer and W. Jahnke, Protein NMR in biomedical research. *Cell Mol. Life Sci.*, **61** (2004), 580-99.
[116] O. Jardetzky and N.G. Wade-Jardetzky, On the mechanism of the binding of sulfonamides to bovine serum albumin. *Mol. Pharmacol.*, **1** (1965), 214-30.
[117] J.J. Fischer and O. Jardetzky, Nuclear Magnetic Relaxation Study of Intermolecular Complexes. The Mechanism of Penicillin Binding to Serum Albumin. *J. Am. Chem. Soc.*, **87** (1965), 3237-44.
[118] M.J. van Dongen, J. Uppenberg, S. Svensson, T. Lundback, T. Akerud, M. Wikstrom, and J. Schultz, Structure-based screening as applied to human FABP4: a highly efficient alternative to HTS for hit generation. *J. Am. Chem. Soc.*, **124** (2002), 11874-80.
[119] R.T. McKay and A.W. Graham, *Chapter 2 Recent Advances in Solvent Suppression for Solution NMR: A Practical Reference*, in *Annual Reports on NMR Spectroscopy*. 2009, Academic Press, 33-76.
[120] D. Shao, T.J. Berrodin, E. Manas, D. Hauze, R. Powers, A. Bapat, D. Gonder, R.C. Winneker, and D.E. Frail, Identification of novel estrogen receptor a antagonists. *J. Steroid. Biochem. Mol. Biol.*, **88** (2004), 351-360.
[121] H. Qin, J. Shi, R. Noberini, E.B. Pasquale, and J. Song, Crystal Structure and NMR Binding Reveal That Two Small Molecule Antagonists Target the High Affinity Ephrin-binding Channel of the EphA4 Receptor. *J. Biol. Chem.*, **283** (2008), 29473-29484.
[122] L. Poppe, T.S. Harvey, C. Mohr, J. Zondlo, C.M. Tegley, O. Nuanmanee, and J. Cheetham, Discovery of ligands for Nurr1 by combined use of NMR screening with different isotopic and spin-labeling strategies. *J. Biomol. Screening*, **12** (2007), 301-311.
[123] A.-S. Bretonnet, A. Jochum, O. Walker, I. Krimm, P. Goekjian, O. Marcillat, and J.-M. Lancelin, NMR Screening Applied to the Fragment-Based Generation of Inhibitors of Creatine Kinase Exploiting a New Interaction Proximate to the ATP Binding Site. *J. Med. Chem.*, **50** (2007), 1865-1875.
[124] M. Piazza, L. Yu, A. Teghanemt, T. Gioannini, J. Weiss, and F. Peri, Evidence of a Specific Interaction between New Synthetic Antisepsis Agents and CD14. *Biochemistry*, **48** (2009), 12337-12344.
[125] K. Jaudzems, J. Kuka, A. Gutsaits, K. Zinovjevs, I. Kalvinsh, E. Liepinsh, E. Liepinsh, and M. Dambrova, Inhibition of carnitine acetyltransferase by mildronate, a regulator of energy metabolism. *J. Enzyme Inhib. Med. Chem.*, **24** (2009), 1269-1275.
[126] J.-Y. Lee, K.-W. Jeong, J.-U. Lee, D.-I. Kang, and Y. Kim, Novel E. coli β-ketoacyl-acyl carrier protein synthase III inhibitors as targeted antibiotics. *Bioorg. Med. Chem.*, **17** (2009), 1506-1513.
[127] M. Simcic, M. Hodoscek, J. Humljan, K. Kristan, U. Urleb, D. Kocjan, and S.G. Grdadolnik, NMR and Molecular Dynamics Study of the Binding Mode of Naphthalene-N-sulfonyl-D-glutamic Acid Derivatives: Novel MurD Ligase Inhibitors. *J. Med. Chem.*, **52** (2009), 2899-2908.
[128] A. Blume, M. Fitzen, A.J. Benie, and T. Peters, Specificity of ligand binding to yeast hexokinase PII studied by STD-NMR. *Carbohydr. Res.*, **344** (2009), 1567-1574.
[129] J.-C. Westermann, R.J. Clark, and D.J. Craik, Binding mode of α-conotoxins to an acetylcholine binding protein determined by saturation transfer difference NMR. *Protein Pept. Lett.*, **15** (2008), 910-914.
[130] M.G. Szczepina, R.B. Zheng, G.C. Completo, T.L. Lowary, and B.M. Pinto, STD-NMR Studies Suggest that Two Acceptor Substrates for GlfT2, a Bifunctional Galactofuranosyltransferase Required for the Biosynthesis of Mycobacterium tuberculosis Arabinogalactan, Compete for the Same Binding Site. *ChemBioChem*, **10** (2009), 2052-2059.
[131] V.V. Krishnan, Ligand screening by saturation-transfer difference (STD) NMR spectroscopy. *Curr. Anal. Chem.*, **1** (2005), 307-320.
[132] G. Zhu and X. Yao, TROSY-based NMR experiments for NMR studies of large biomolecules. *Prog. Nucl. Magn. Reson. Spectrosc.*, **52** (2008), 49-68.
[133] S. Mori, C. Abeygunawardana, M.O.N. Johnson, and P.C.M. van Zijl, Improved sensitivity of HSQC spectra of exchanging protons at short interscan delays using a new fast HSQC (FHSQC) detection scheme that avoids water saturation. *J. Magn. Reson., Ser. B*, **108** (1995), 94-8.

[134] P. Schanda, E. Kupce, and B. Brutscher, SOFAST-HMQC Experiments for Recording Two-dimensional Heteronuclear Correlation Spectra of Proteins within a Few Seconds. *J. Biomol. NMR*, **33** (2005), 199-211.
[135] D.M. LeMaster, Deuterium labeling in NMR structural analysis of larger proteins. *Q. Rev. Biophys.*, **23** (1990), 133-74.
[136] T. Kigawa, Y. Muto, and S. Yokoyama, Cell-free synthesis and amino acid-selective stable isotope labeling of proteins for NMR analysis. *J. Biomol. NMR*, **6** (1995), 129-34.
[137] N.K. Goto and L.E. Kay, New developments in isotope labeling strategies for protein solution NMR spectroscopy. *Curr. Opin. Struct. Biol.*, **10** (2000), 585-592.
[138] J. Fejzo, C.A. Lepre, J.W. Peng, G.W. Bemis, Ajay, M.A. Murcko, and J.M. Moore, The SHAPES strategy: an NMR-based approach for lead generation in drug discovery. *Chem. Biol.*, **6** (1999), 755-69.
[139] C.K. Wada, J.H. Holms, M.L. Curtin, Y. Dai, A.S. Florjancic, R.B. Garland, Y. Guo, H.R. Heyman, J.R. Stacey, D.H. Steinman, D.H. Albert, J.J. Bouska, I.N. Elmore, C.L. Goodfellow, P.A. Marcotte, P. Tapang, D.W. Morgan, M.R. Michaelides, and S.K. Davidsen, Phenoxyphenyl Sulfone N-Formylhydroxylamines (Retrohydroxamates) as Potent, Selective, Orally Bioavailable Matrix Metalloproteinase Inhibitors. *J. Med. Chem.*, **45** (2002), 219-232.
[140] M.N. Schulz and R.E. Hubbard, Recent progress in fragment-based lead discovery. *Curr. Opin. Pharmacol.*, **9** (2009), 615-621.
[141] T.J. Mitchison, Towards a pharmacological genetics. *Chem. Biol.*, **1** (1994), 3-6.
[142] B.R. Stockwell, S.J. Haggarty, and S.L. Schreiber, High-throughput screening of small molecules in miniaturized mammalian cell-based assays involving post-translational modifications. *Chem. Biol.*, **6** (1999), 71-83.
[143] C.M. Crews and U. Splittgerber, Chemical genetics: exploring and controlling cellular processes with chemical probes. *Trends Biochem Sci*, **24** (1999), 317-20.
[144] L. Burdine and T. Kodadek, Target identification in chemical genetics: the (often) missing link. *Chem. Biol.*, **11** (2004), 593-7.
[145] T.U. Mayer, T.M. Kapoor, S.J. Haggarty, R.W. King, S.L. Schreiber, and T.J. Mitchison, Small molecule inhibitor of mitotic spindle bipolarity identified in a phenotype-based screen. *Science*, **286** (1999), 971-4.
[146] J.R. Peterson, R.S. Lokey, T.J. Mitchison, and M.W. Kirschner, A chemical inhibitor of N-WASP reveals a new mechanism for targeting protein interactions. *Proc. Natl. Acad. Sci. U. S. A.*, **98** (2001), 10624-9.
[147] K.L. Carey, N.J. Westwood, T.J. Mitchison, and G.E. Ward, A small-molecule approach to studying invasive mechanisms of Toxoplasma gondii. *Proc. Natl. Acad. Sci. U. S. A.*, **101** (2004), 7433-8.
[148] A.H. Julia, K.M. Lesley, W.K. Gavin, J.J. Joanna, and P.F. Andrew, Nuclear magnetic resonance spectroscopy reveals the functional state of the signalling protein CheY *in vivo* in *Escherichia coli*. *Mol. Microbiol.*, **49** (2003), 1191-1200.
[149] D.S. Burz and A. Shekhtman, Structural biology: Inside the living cell. *Nature*, **458** (2009), 37-8.
[150] S. Reckel, R. Haensel, F. Loehr, and V. Doetsch, In-cell NMR spectroscopy. *Prog. Nucl. Magn. Reson. Spectrosc.*, **51** (2007), 91-101.
[151] T. Sakai, H. Tochio, T. Tenno, Y. Ito, T. Kokubo, H. Hiroaki, and M. Shirakawa, In-cell NMR spectroscopy of proteins inside Xenopus laevis oocytes. *J. Biomol. NMR*, **36** (2006), 179-188.
[152] S. Mari, D. Serrano-Gomez, F.J. Canada, A.L. Corbi, and J. Jimenez-Barbero, 1D saturation transfer difference NMR experiments on living cells: the DC-SIGN/oligomannose interaction. *Angew. Chem. Int. Ed. Engl.*, **44** (2004), 296-8.
[153] V. Doetsch, Investigation of proteins in living bacteria with in-cell NMR experiments. *Top. Curr. Chem.*, **273** (2008), 203-229.
[154] S. Ogino, S. Kubo, R. Umemoto, S. Huang, N. Nishida, and I. Shimada, Observation of NMR Signals from Proteins Introduced into Living Mammalian Cells by Reversible Membrane Permeabilization Using a Pore-Forming Toxin, Streptolysin O. *J. Am. Chem. Soc.*, **131** (2009), 10834-10835.
[155] A.E. Ferentz and G. Wagner, NMR spectroscopy: a multifaceted approach to macromolecular structure. *Q. Rev. Biophys.*, **33** (2000), 29-65.
[156] G.M. Clore and A.M. Gronenborn, Two-, three-, and four-dimensional NMR methods for obtaining larger and more precise three-dimensional structures of proteins in solution. *Annu. Rev. Biophys. Biophys. Chem*, **20** (1991), 29-63.
[157] P.N. Reardon and L.D. Spicer, Multidimensional NMR spectroscopy for protein characterization and assignment inside cells. *J. Am. Chem. Soc.*, **127** (2005), 10848-9.
[158] Y. Shrot and L. Frydman, Single-scan NMR spectroscopy at arbitrary dimensions. *J. Am. Chem. Soc.*, **125** (2003), 11385-96.

[159] M. Gal, M. Mishkovsky, and L. Frydman, Real-time monitoring of chemical transformations by ultrafast 2D NMR spectroscopy. *J. Am. Chem. Soc.*, **128** (2006), 951-6.
[160] T. Szyperski and H.S. Atreya, Principles and applications of GFT projection NMR spectroscopy. *Magn Reson Chem*, **44 Spec No** (2006), S51-60.
[161] H.S. Atreya and T. Szyperski, Rapid NMR data collection. *Methods Enzymol*, **394** (2005), 78-108.
[162] P. Schanda, V. Forge, and B. Brutscher, HET-SOFAST NMR for fast detection of structural compactness and heterogeneity along polypeptide chains. *Magn Reson Chem*, **44 Spec No** (2006), S177-84.
[163] P. Schanda, H. Van Melckebeke, and B. Brutscher, Speeding up three-dimensional protein NMR experiments to a few minutes. *J. Am. Chem. Soc.*, **128** (2006), 9042-3.
[164] S. Hiller, F. Fiorito, K. Wuethrich, and G. Wider, Automated projection spectroscopy (APSY). *Proc. Natl. Acad. Sci. U. S. A.*, **102** (2005), 10876-10881.
[165] H.S. Atreya and T. Szyperski, G-matrix Fourier transform NMR spectroscopy for complete protein resonance assignment. *Proc. Natl. Acad. Sci. U. S. A.*, **101** (2004), 9642-9647.
[166] H.R. Eghbalnia, A. Bahrami, M. Tonelli, K. Hallenga, and J.L. Markley, High-Resolution Iterative Frequency Identification for NMR as a General Strategy for Multidimensional Data Collection. *J. Am. Chem. Soc.*, **127** (2005), 12528-12536.
[167] B.E. Coggins, R.A. Venters, and P. Zhou, Filtered Backprojection for the Reconstruction of a High-Resolution (4,2)D CH3-NH NOESY Spectrum on a 29 kDa Protein. *J. Am. Chem. Soc.*, **127** (2005), 11562-11563.
[168] D.K. Staykova, J. Fredriksson, W. Bermel, and M. Billeter, Assignment of protein NMR spectra based on projections, multi-way decomposition and a fast correlation approach. *J. Biomol. NMR*, **42** (2008), 87-97.
[169] T. Szyperski, D.C. Yeh, D.K. Sukumaran, H.N.B. Moseley, and G.T. Montelione, Reduced-dimensionality NMR spectroscopy for high-throughput protein resonance assignment. *Proc. Natl. Acad. Sci. U. S. A.*, **99** (2002), 8009-8014.
[170] D.S. Burz, K. Dutta, D. Cowburn, and A. Shekhtman, In-cell NMR for protein-protein interactions (STINT-NMR). *Nat Protoc*, **1** (2006), 146-52.
[171] D.S. Burz, K. Dutta, D. Cowburn, and A. Shekhtman, Mapping structural interactions using in-cell NMR spectroscopy (STINT-NMR). *Nat Methods*, **3** (2006), 91-3.
[172] L. Fielding, NMR methods for the determination of protein-ligand dissociation constants. *Prog. Nucl. Magn. Reson. Spectrosc.*, **51** (2007), 219-242.
[173] L.H. Lucas and C.K. Larive, Measuring ligand-protein binding using NMR diffusion experiments. *Concepts Magn. Reson., Part A*, **20A** (2004), 24-41.
[174] Z. Ji, Z. Yao, and M. Liu, Saturation transfer difference nuclear magnetic resonance study on the specific binding of ligand to protein. *Anal. Biochem.*, **385** (2009), 380-382.
[175] K.A. Mercier, J.R. Cort, M.A. Kennedy, E.E. Lockert, S. Ni, M.D. Shortridge, and R. Powers, Structure and function of Pseudomonas aeruginosa protein PA1324 (21-170). *Protein Sci*, **18** (2009), 606-18.
[176] S.D. Auweter, F.C. Oberstrass, and F.H.T. Allain, Solving the Structure of PTB in Complex with Pyrimidine Tracts: An NMR Study of Protein-RNA Complexes of Weak Affinities+. *J. Mol. Biol.*, **367** (2007), 174-186.
[177] T.A. Pertinhez, E. Ferrari, E. Casali, J.A. Patel, A. Spisni, and L.J. Smith, The binding cavity of mouse major urinary protein is optimised for a variety of ligand binding modes. *Biochem. Biophys. Res. Commun.*, **390** (2009), 1266-1271.
[178] Y. Thielmann, J. Mohrlueder, B.W. Koenig, T. Stangler, R. Hartmann, K. Becker, H.-D. Hoeltje, and D. Willbold, An indole-binding site is a major determinant of the ligand specificity of the GABA type A receptor-associated protein GABARAP. *ChemBioChem*, **9** (2008), 1767-1775.
[179] M. Mayer and B. Meyer, Group Epitope Mapping by Saturation Transfer Difference NMR To Identify Segments of a Ligand in Direct Contact with a Protein Receptor. *J. Am. Chem. Soc.*, **123** (2001), 6108-6117.
[180] C. Dalvit, M. Flocco, S. Knapp, M. Mostardini, R. Perego, B.J. Stockman, M. Veronesi, and M. Varasi, High-throughput NMR-based screening with competition binding experiments. *J. Am. Chem. Soc.*, **124** (2002), 7702-7709.
[181] W. Jahnke, P. Floersheim, C. Ostermeier, X. Zhang, R. Hemmig, K. Hurth, and D.P. Uzunov, NMR reporter screening for the detection of high-affinity ligands. *Angew. Chem., Int. Ed.*, **41** (2002), 3420-3423.
[182] C. Dalvit, M. Fasolini, M. Flocco, S. Knapp, P. Pevarello, and M. Veronesi, NMR-Based Screening with Competition Water-Ligand Observed via Gradient Spectroscopy Experiments: Detection of High-Affinity Ligands. *J. Med. Chem.*, **45** (2002), 2610-2614.

[183] A.H. Siriwardena, F. Tian, S. Noble, and J.H. Prestegard, A straightforward NMR-spectroscopy-based method for rapid library screening. *Angew. Chem., Int. Ed.*, **41** (2002), 3454-3457.
[184] C. Dalvit, M. Flocco, M. Veronesi, and B.J. Stockman, Fluorine-NMR competition binding experiments for high-throughput screening of large compound mixtures. *Comb. Chem. High Throughput Screening*, **5** (2002), 605-611.
[185] C. Dalvit, D.T.A. Hadden, R.W. Sarver, A.M. Ho, and B.J. Stockman, Multi-selective one dimensional proton NMR experiments for rapid screening and binding affinity measurements. *Comb. Chem. High Throughput Screening*, **6** (2003), 445-453.
[186] W. Jahnke, Spin labels as a tool to identify and characterize protein - ligand interactions by NMR spectroscopy. *ChemBioChem*, **3** (2002), 167-173.
[187] R.V.C. Guido, M.S. Castilho, G. Oliva, and A.D. Andricopulo, Structure-based pharmacophore strategies in drug design. *Curr. Methods Med. Chem. Biol. Phys.*, **2** (2008), 93-112.
[188] K.M. Honorio, C.A. Montanari, and A.D. Andricopulo, Advances and applications of structure-based approaches in drug discovery. *Curr. Methods Med. Chem. Biol. Phys.*, **2** (2008), 75-92.
[189] A.M. Petros and S.W. Fesik, Nuclear magnetic resonance methods for studying protein-ligand complexes. *Methods Enzymol.*, **239** (1994), 717-39.
[190] A. Medek, P.J. Hajduk, J. Mack, and S.W. Fesik, The Use of Differential Chemical Shifts for Determining the Binding Site Location and Orientation of Protein-Bound Ligands. *J. Am. Chem. Soc.* **122** (2000), 1241-1242.
[191] M. Pellecchia, D. Meininger, Q. Dong, E. Chang, R. Jack, and D.S. Sem, NMR-based structural characterization of large protein-ligand interactions. *J. Biomol. NMR*, **22** (2002), 165-173.
[192] P.J. Hajduk, J.C. Mack, E.T. Olejniczak, C. Park, P.J. Dandliker, and B.A. Beutel, SOS-NMR: A Saturation Transfer NMR-Based Method for Determining the Structures of Protein-Ligand Complexes. *J. Am. Chem. Soc.*, **126** (2004), 2390-2398.
[193] D.S. Sem, L. Yu, S.M. Coutts, and R. Jack, Object-oriented approach to drug design enabled by NMR SOLVE: first real-time structural tool for characterizing protein-ligand interactions. *J. Cell. Biochem.*, 2001), 99-105.
[194] M.A. McCoy and D.F. Wyss, Alignment of weakly interacting molecules to protein surfaces using simulations of chemical shift perturbations. *J. Biomol. NMR*, **18** (2000), 189-98.
[195] D. Sitkoff and D.A. Case, Theories of chemical shift anisotropies in proteins and nucleic acids. *Prog. Nucl. Magn. Reson. Spectrosc.*, **32** (1998), 165-190.
[196] M.A. McCoy and D.F. Wyss, Spatial localization of ligand binding sites from electron current density surfaces calculated from NMR chemical shift perturbations. *J. Am. Chem. Soc.*, **124** (2002), 11758-11763.
[197] K. Osapay and D.A. Case, A new analysis of proton chemical shifts in proteins. *J. Am. Chem. Soc.*, **113** (1991), 9436-44.
[198] A. Nicholls, K.A. Sharp, and B. Honig, Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins Struct., Funct., Genet.*, **11** (1991), 281-96.
[199] D.S. Goodsell, G.M. Morris, and A.J. Olson, Automated docking of flexible ligands: applications of AutoDock. *J. Mol. Recognit.*, **9** (1996), 1-5.
[200] M.R. McGann, H.R. Almond, A. Nicholls, J.A. Grant, and F.K. Brown, Gaussian docking functions. *Biopolymers*, **68** (2003), 76-90.
[201] D.T. Moustakas, P.T. Lang, S. Pegg, E. Pettersen, I.D. Kuntz, N. Brooijmans, and R.C. Rizzo, Development and validation of a modular, extensible docking program: DOCK 5. *J. Comput. Aided Mol. Des.*, **20** (2006), 601-19.
[202] G.B. McGaughey, R.P. Sheridan, C.I. Bayly, J.C. Culberson, C. Kreatsoulas, S. Lindsley, V. Maiorov, J.F. Truchon, and W.D. Cornell, Comparison of topological, shape, and docking methods in virtual screening. *J. Chem. Inf. Model*, **47** (2007), 1504-19.
[203] M.A. McCoy and D.F. Wyss, Spatial localization of ligand binding sites from electron current density surfaces calculated from NMR chemical shift perturbations. *J. Am. Chem. Soc.*, **124** (2002), 11758-63.
[204] A.A. Lugovskoy, A.I. Degterev, A.F. Fahmy, P. Zhou, J.D. Gross, J. Yuan, and G. Wagner, A Novel Approach for Characterizing Protein Ligand Complexes: Molecular Basis for Specificity of Small-Molecule Bcl-2 Inhibitors. *J. Am. Chem. Soc.*, **124** (2002), 1234-1240.
[205] A. Fahmy and G. Wagner, TreeDock: A tool for protein docking based on minimizing van der Waals energies. *J. Am. Chem. Soc.*, **124** (2002), 1241-1250.
[206] D. Gonzalez-Ruiz and H. Gohlke, Steering protein-ligand docking with quantitative NMR chemical shift perturbations. *J. Chem. Inf. Model*, **49** (2009), 2260-71.
[207] G.M. Morris, D.S. Goodsell, R.S. Halliday, R. Huey, W.E. Hart, R.K. Belew, and A.J. Olson, Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *J. Comput. Chem.*, **19** (1998), 1639-1662.

[208] M.D. Shortridge and R. Powers, Structural and Functional Similarity between the Bacterial Type III Secretion System Needle Protein PrgI and the Eukaryotic Apoptosis Bcl-2 Proteins *PLoS ONE*, **4** (2009), 1-10.
[209] C. Dominguez, R. Boelens, and A.M.J.J. Bonvin, HADDOCK: A protein-protein docking approach based on biochemical or biophysical information. *J. Am. Chem. Soc.*, **125** (2003), 1731-1737.
[210] M. Nilges, Ambiguous distance data in the calculation of NMR structures. *Folding Des.*, **2** (1997), S53-S57.
[211] F. Sams-Dodd, Target-based drug discovery: is something wrong? *Drug Discov. Today*, **10** (2005), 139-147.
[212] J.-P. Garnier, Rebuilding the R&D engine in big pharma. *Harv. Bus. Rev.*, **86** (2008), 68-70, 72-6, 128.
[213] J.A. Weisbach and W.H. Moos, Diagnosing the decline of major Pharm. Res. laboratories: a prescription for drug companies. *Drug Dev. Res.*, **34** (1995), 243-59.
[214] D.F. Horrobin, Realism in drug discovery-could Cassandra be right? *Nat. Biotechnol.*, **19** (2001), 1099-1100.
[215] P. Cuatrecasas, Drug discovery in jeopardy. *J. Clin. Invest.*, **116** (2006), 2837-2842.
[216] J. Owens, Big pharma slims down to bolster productivity. *Nat. Rev. Drug Discovery*, **6** (2007), 173-174.
[217] C.G. Wermuth, Similarity in drugs: reflections on analogue design. *Drug Discov. Today*, **11** (2006), 348-354.
[218] B. Booth and R. Zemmel, Opinion: Quest for the best. *Nat. Rev. Drug Discovery*, **2** (2003), 838-841.
[219] D. Frishman, M. Mokrejs, D. Kosykh, G. Kastenmuller, G. Kolesov, I. Zubrzycki, C. Gruber, B. Geier, A. Kaps, K. Albermann, A. Volz, C. Wagner, M. Fellenberg, K. Heumann, and H.-W. Mewes, The PEDANT genome database. *Nucleic Acids Res.*, **31** (2003), 207-211.
[220] H.W. Mewes, D. Frishman, U. Gueldener, G. Mannhaupt, K. Mayer, M. Mokrejs, B. Morgenstern, M. Muensterkoetter, S. Rudd, and B. Weil, MIPS: a database for genomes and protein sequences. *Nucleic Acids Res.*, **30** (2002), 31-34.
[221] P.D. Thomas, A. Kejariwal, M.J. Campbell, H. Mi, K. Diemer, N. Guo, I. Ladunga, B. Ulitsky-Lazareva, A. Muruganujan, S. Rabkin, J.A. Vandergriff, and O. Doremieux, PANTHER: A browsable database of gene products organized by biological function, using curated protein family and subfamily classification. *Nucleic Acids Res.*, **31** (2003), 334-341.
[222] A. Bernal, U. Ear, and N. Kyrpides, Genomes OnLine Database (GOLD): a monitor of genome projects world-wide. *Nucleic Acids Res.*, **29** (2001), 126-127.
[223] M.T. Black and J. Hodgson, Novel target sites in bacteria for overcoming antibiotic resistance. *Adv. Drug Delivery Rev.*, **57** (2005), 1528-1538.
[224] D. Lee, O. Redfern, and C. Orengo, Predicting protein function from sequence and structure. *Nat. Rev. Mol. Cell Biol.*, **8** (2007), 995-1005.
[225] A. Valencia, Automatic annotation of protein function. *Curr. Opin. Struct. Biol.*, **15** (2005), 267-274.
[226] R. Rentzsch and C.A. Orengo, Protein function prediction--the power of multiplicity. *Trends Biotechnol.*, **27** (2009), 210-9.
[227] J.A. Gerlt and P.C. Babbitt, Can sequence determine function? *GenomeBiology*, **1** (2000), Epub
[228] R.H. White, The difficult road from sequence to function. *J. Bacteriol.*, **188** (2006), 3431-3432.
[229] S.L. Kinnings and R.M. Jackson, Binding site similarity analysis for the functional classification of the protein kinase family. *J. Chem. Inf. Model*, **49** (2009), 318-29.
[230] R.J. Najmanovich, A. Allali-Hassani, R.J. Morris, L. Dombrovsky, P.W. Pan, M. Vedadi, A.N. Plotnikov, A. Edwards, C. Arrowsmith, and J.M. Thornton, Analysis of binding site similarity, small-molecule similarity and experimental binding profiles in the human cytosolic sulfotransferase family. *Bioinformatics*, **23** (2007), e104-9.
[231] L.A. Mirny and E.I. Shakhnovich, Universally conserved positions in protein folds: reading evolutionary signals about stability, folding kinetics and function. *J. Mol. Biol.*, **291** (1999), 177-96.
[232] R. Powers, J.C. Copeland, K. Germer, K.A. Mercier, V. Ramanathan, and P. Revesz, Comparison of protein active site structures for functional annotation of proteins and drug design. *Proteins*, **65** (2006), 124-35.