Journal of Medicinal Chemistry

Article

pubs.acs.org/jmc

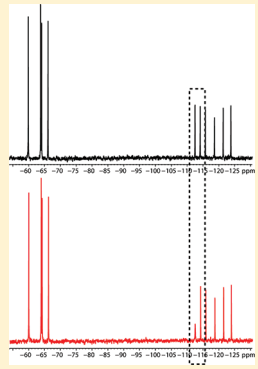
Fragment Based Drug Discovery: Practical Implementation Based on ^{19}F NMR Spectroscopy

John B. Jordan,*[,†] Leszek Poppe,*[,†] Xiaoyang Xia,[†] Alan C. Cheng,[†] Yax Sun,[†] Klaus Michelsen,[†] Heather Eastwood,[†] Paul D. Schnier,[†] Thomas Nixey,[‡] and Wenge Zhong[‡]

[†]Department of Molecular Structure, and [‡]Chemistry Research and Discovery, Amgen, Inc., One Amgen Center Drive, Thousand Oaks, California 91320, United States

Supporting Information

ABSTRACT: Fragment based drug discovery (FBDD) is a widely used tool for discovering novel therapeutics. NMR is a powerful means for implementing FBDD, and several approaches have been proposed utilizing ^{1}H−^{15}N heteronuclear single quantum coherence (HSQC) as well as one-dimensional ^{1}H and ^{19}F NMR to screen compound mixtures against a target of interest. While proton-based NMR methods of fragment screening (FBS) have been well documented and are widely used, the use of ^{19}F detection in FBS has been only recently introduced (Vulpetti et al. *J. Am. Chem. Soc.* **2009**, *131* (36), 12949−12959) with the aim of targeting "fluorophilic" sites in proteins. Here, we demonstrate a more general use of ^{19}F NMR-based fragment screening in several areas: as a key tool for rapid and sensitive detection of fragment hits, as a method for the rapid development of structure−activity relationship (SAR) on the hit-to-lead path using in-house libraries and/or commercially available compounds, and as a quick and efficient means of assessing target druggability.



INTRODUCTION

Fragment based drug discovery (FBDD) has become a major tool for discovering novel therapeutics.[1] In contrast to traditional screening approaches, fragment based screening (FBS) relies on the identification of small, weakly potent compounds with good binding efficiencies as a chemical starting point. There are several commonly cited benefits to FBDD: (1) sampling of higher chemical diversity, (2) higher hit rates, and (3) higher efficiency of binding.[2−5] Because of the small size and low chemical complexity, fragment libraries can sample a higher proportion of chemical space.[2] In addition, the decrease in complexity may result in less constrained, more optimal interactions with the target, thereby increasing ligand efficiency (LE).[6] Ligand efficiency is an important metric to assess the quality of fragments hit compounds and is characterized by the binding energy per heavy atom (HA) (eq 1).

$$\mathrm{LE} = -\frac{\Delta G}{\#\mathrm{HA}} \quad (1)$$

Another benefit of FBDD is that it is typically easier to elaborate fragments into larger, more potent compounds, while maintaining good LE (which typically decreases with increasing MW) compared with traditional high-throughput screening (HTS) hits. While HTS is still used abundantly, FBDD is gaining momentum in the drug discovery process by facilitating the identification of novel low molecular weight (<300 Da) chemical matter and, through structure-based drug design, facilitating the development of very potent molecules from initial fragments hits. Presently, there are numerous compounds in various phases of clinical development that originated from fragment-based screening campaigns.[7]

In addition to the widespread implementation of FBDD, there has been significant development in the analytical technologies used to perform FBS. The vast majority of fragment screening is performed using nuclear magnetic resonance (NMR) spectroscopy with proton detection and surface plasmon resonance (SPR) spectroscopy. In the present work, we will focus on NMR and, in particular, will discuss fragment screening methods based on ^{19}F NMR that offer several analytical advantages including speed, sensitivity, and selectivity.

Exploiting the fluorine nucleus for NMR screening has been extensively reviewed.[8] While the use of the ^{19}F nucleus has previously been proposed for NMR-based screening, these methods have been primarily focused on the screening of more drug-like molecules[9] and/or the use of ^{19}F-labeled molecules in reporter screening.[8] In the case of a ^{19}F library composed of drug-like compounds,[9] compound complexity and poor solubility rendered cocrystallization of the hits with the target highly problematic and posed a high barrier for chemistry follow-up. Recently, Vulpetti et al. proposed the construction of a fluorine fragment library based on the local environment of fluorine (LEF) principle.[10] This library, as described, was developed on the observation that the presence of a fluorine nucleus can be beneficial to some properties of small molecule drugs.[10−14] Thus, a novel library was constructed to search for the presence of "fluorophilic" protein environments and for identifying binders with fluorine in their pharmacophore.[10]

Received: July 13, 2011
Published: December 13, 2011

The prediction of target druggability, in drug discovery, is considered a measure of the likelihood of a target binding with high affinity to a small, drug-like molecule. The "druggable genome" analysis by Hopkins and Groom in 2002[15] highlighted that not all therapeutically relevant targets were "druggable", and suggested that only 5% of the human genome is both druggable and therapeutically relevant. This finding spurred attempts to assess target druggability in the drug discovery process. Now, because of the high rate of failure in small molecule drug discovery, the potential druggability of a target is generally a critical consideration in the target selection process.[16−18] Furthermore, the ability to accurately predict the probable success of a screening campaign may directly affect target attrition and can make the drug discovery process more efficient by allowing the prioritization of targets when they have similar or poor biological rationale and by allowing the effective allocation of resources around certain difficult targets. In recent years, several computational methods have been proposed to assess target druggability, most of which make use of geometric and energy-based algorithms that use available protein structures to search for and identify pockets.[19−23] These pockets can then be evaluated for their shape and volume, and then subsequently for their ability to bind a drug-like molecule.[24,25] Several methods, such as the maximal affinity model[26] have been shown to be valuable in rank ordering targets based on their computed druggability scores.[27] In addition to being used as a screening method to identify chemical matter, it has recently been proposed that hit rates from FBS may be a good indication of target "druggability".[25,27,28] Initially, Hajduk et al. suggested a strategy to evaluate target druggability by screening chemical libraries with two-dimensional (2D) heteronuclear-NMR.[27] The idea that hit rates from fragment screens directly correlate with the druggability of protein targets is now generally accepted.[16,7,25,28]

In this work, we demonstrate a general use of ^{19}F NMR for FBS in which the ^{19}F nucleus plays more of a sensor role rather than an affinity tag. Our library and methods are based on the simple fragment screening principles and have been extended to include hit validation, follow up, and target druggability assessment. Here, we demonstrate the straightforward construction of a ^{19}F fragment library based on the traditional fragment "rule of three"[29] with comparable size and chemical diversity to a typical fragment library (∼1k−2k compounds). We discuss the feasibility of constructing a diverse fragment library consisting of only fluorinated molecules and show the high sensitivity and speed of ^{19}F NMR in the detection of fragments that bind to the target protein. We demonstrate the performance of this library in a case study using β-secretase-1 (BACE-1) as a target molecule, which includes the ability to use ^{19}F NMR to rapidly perform the screen, identify hits, obtain the respective K_d values, and perform hit elaboration using ^{19}F NMR. In addition, we demonstrate how rapid structure−activity relationship (SAR) can be determined by searching small libraries of protonated analogues based on the initial ^{19}F fragment hits, and we introduce an efficient fragment linking strategy based on ^{19}F−^{19}F intermolecular NOE detection, which complements the original interligand nuclear Overhauser effect (ILNOE) experiment[30] commonly used in ^{1}H-based NMR fragment applications. Lastly, we discuss how the speed and sensitivity of ^{19}F-NMR based fragment screening may be leveraged to provide a rapid and reliable assessment of target druggability.

RESULTS AND DISCUSSION

^{19}F Fragment Library Construction and Screen of BACE-1. Using the criteria shown in Table 1, a ^{19}F fragment library consisting of 1200 diverse compounds was constructed (Figures 1 and 2) using both in-house and commercial sources.

Table 1. Selection Criteria for ^{19}F Fragment Library

- No metal or reactive compounds, no undesired chemotypes, no salt
- Number of atoms ≤ 18
- cLogP < 3 and AlogP < 3
- Molecular weight < 300 Da
- CF_3 groups = 1 or −F groups = 1
- H-bond acceptors =1−7
- H-bond donors =0−3
- Number of rings =1−4, rotatable bonds < 5
- Number of acid groups < 2
- Number of base groups < 2
- Prefer racemic over single enantiomers

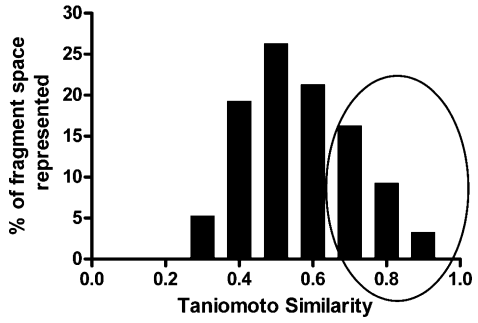


Figure 1. Creating "fragment space" and assessing diversity of the ^{19}F fragment library. Approximately 6 million compounds (ACD + ACDSC) were filtered using the methods discussed. The result was ∼67 000 compounds that were used to define "fragment space". The compounds selected for the fragment library were compared to "fragment space" using the nearest neighbor (NN) calculations and resulted in a similarity score. A similarity score of 0.7 or higher indicated that the reference compound is represented in the ^{19}F fragment library (circled area). In total, approximately 30% of the defined "fragment space" was represented by the compounds in the ^{19}F library.

We assessed the diversity of this library using a computational model of "fragment space" comprised of ∼67 000 reference compounds created from the Available Chemicals Directory (ACD) and Available Chemicals Directory Screening Compounds (ACDSC) databases (total of ∼6 million compounds) and using the selection filters depicted in Table 1 (excluding the requirement for fluorine atoms). In a similar manner, "^{19}F fragment space" was created using compounds selected for the fluorine library (Figures 1 and 2). Although the use of fluorine (or trifluoromethyl) as a selection filter in the library construction limited the coverage of "fragment space" to approximately 30% (Figure 1), we found that it had a negligible impact on the success of the screens using the targets presented here. To demonstrate the utility and robust nature of this library, we performed a ^{19}F-NMR fragment screen of BACE-1. The entire screen was completed in less than 24 h, and six hits were identified with five of these compounds being selected for subsequent follow up (Table 2 and Figure 3). Using the differential chemical shift perturbation (dCSP) method, we

dx.doi.org/10.1021/jm201441k | J. Med. Chem. 2012, 55, 678−687

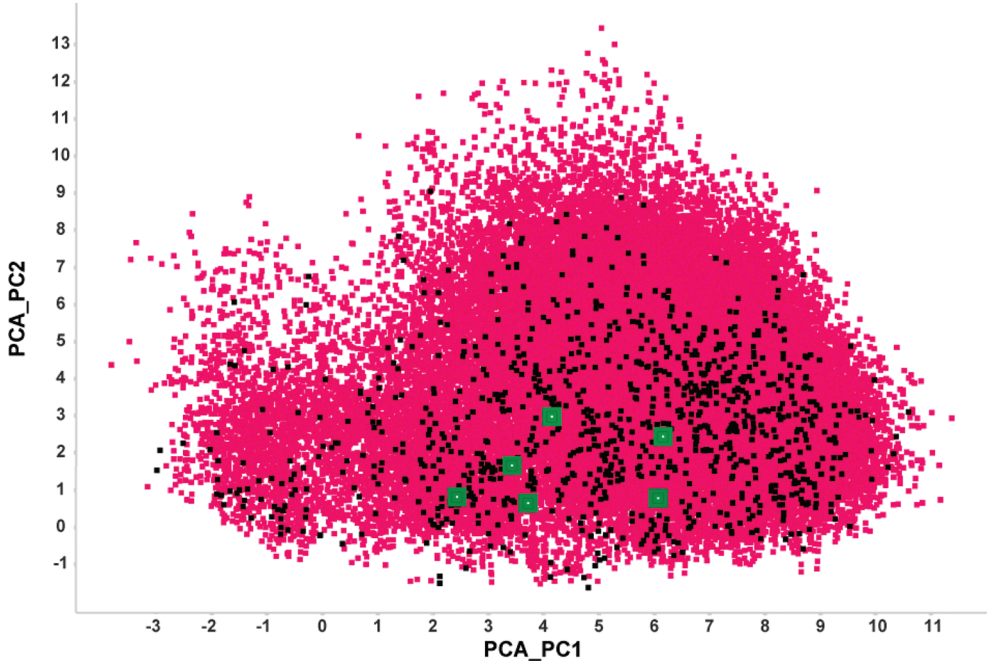
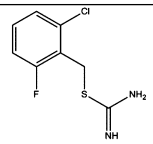
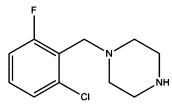
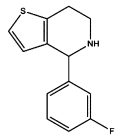
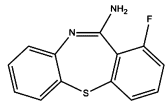
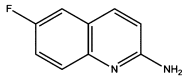


Figure 2. Graphical representation of "fragment space" coverage by ^{19}F compounds. Red dots represent compounds selected to represent "fragment space", while black dots represent compounds existing in the ^{19}F fragment library. Hit compounds from the screen of β-secretase discussed herein are shown as green squares.

Table 2. Fragment Compounds: Agreement of NMR and SPR Based Methods of K_d Determination

Compound	Structure	MW (Daltons)	$K_{d(NMR)}$ (μM)	$K_{d(SPR)}$ (μM)	Ligand Efficiency*	Crystal Structure
1		219	350	300	0.3	Yes
2		229	>1000	>2000	N/A	No
3		233	>1000	>2000	N/A	No
4		244	400	325	0.28	No
5		162	1000	1280	0.3	Yes

*Ligand Efficiency determined using K_d from SPR

determined K_d values for these hits using ^{19}F-NMR (see Material and Methods). The obtained K_d values are in excellent agreement with SPR data (Table 2 and Figure 4). In light of the abundance of previously published data on the binding of 2-aminoquinolines to BACE-1,[31] we elected to use the fluorinated version of this compound (hit number 5) to examine the effects of the fluorine nucleus on binding affinity. This was performed using a number of 2-aminoquinoline analogues, which were obtained from either our in-house sample collection (Table 3, compounds 5−10) or from chemical synthesis during the initial hit to lead campaign (Table 4, compounds 11−14). In the case of BACE-1, as well as other targets that we have screened, an improvement in the binding affinity of the small molecule was observed when replacing fluorine with hydrogen. This contrasts with the premise of the ^{19}F library design demonstrated by Vulpetti et al., where, using ^{19}F to search for "fluorophilic" environments, one would

dx.doi.org/10.1021/jm201441k | J. Med. Chem. 2012, 55, 678−687

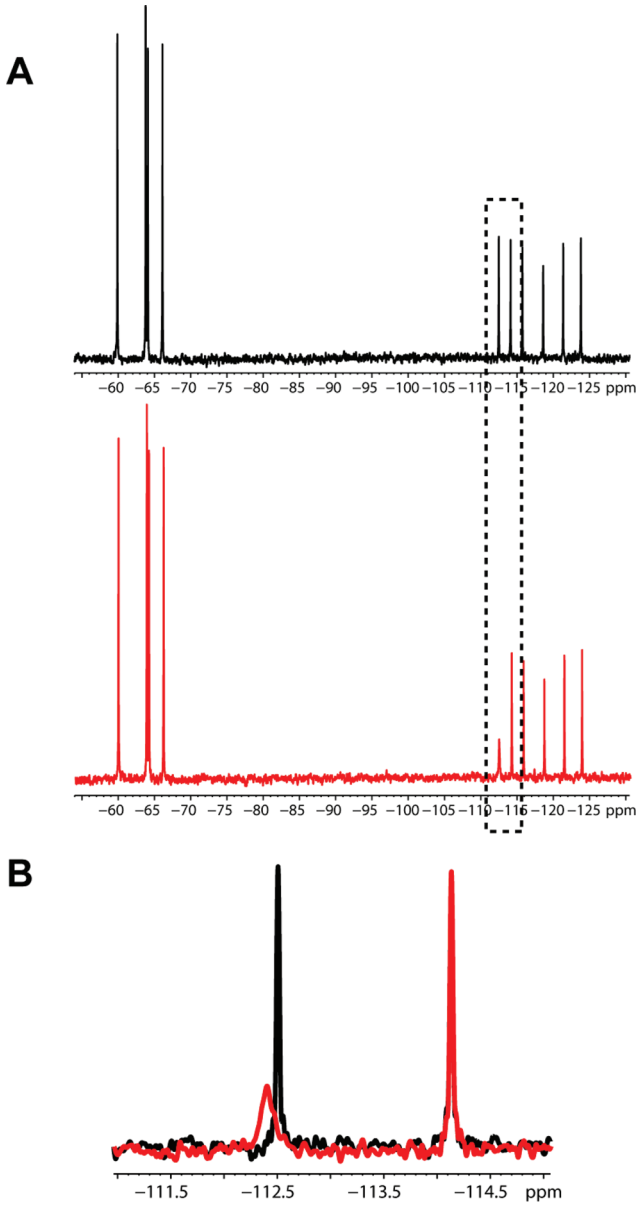


Figure 3. Typical ^{19}F NMR fragment screen spectra. (A) Two identical sets of spectra are acquired, one containing 10 μM protein (in red) and the other containing no protein (control plate, in black). (B) Expansion of box shown in (A) where spectra are compared in spectral overlay mode to determine hits. Compounds can be identified directly from the chemical shift.

expect the opposite trend.[10] We observed that substitution of −H by −F or by $-CH_3$ had a similar impact (~1.4- vs 1.7-fold increase) on the K_d values against BACE-1 as determined by SPR. However, substitution of a $-CF_3$ group resulted in a nearly 4-fold weaker interaction. This is likely due to the more hydrophobic nature of a trifluoromethyl group (versus a methyl) and/or the significant increase in the van der Waal's (VDW) volume of the trifluoromethyl group. The VDW radius of fluorine (1.47 Å) is only slightly larger than that of a proton (1.2 Å).[32,33] A $-CH_3$ to $-CF_3$ substitution, on the other hand, results in a significant increase in the bulk of the substituent, increasing the VDW volume from 21.6 Å^3 (for $-CH_3$) to 39.8 Å^3 (for $-CF_3$).[12]

It has been noted by others that substitutions at the 6-position of 2-aminoquinoline are not well tolerated.[31] However, we elected to pursue the effects of adding hydrophobic groups at this position.[34] As detailed in Cheng et al., it was found that hydrophobic substitutions of bromine at position 6 actually led to a 2−4 fold improvement in potency and a resulting increase in ligand efficiency.[34] Further substitutions at the 6-position using methyl-substituted aryl groups subsequently led to potencies in the low double- to single-digit micromolar.[34] At this point, we used the substitution of a ortho-trifluoromethyl-phenyl group at the 6-position of the 2-aminoquinoline core (a 200 μM binder, Table 4, compound **14**) as a seed to search for additional fragment binding partners.

We tested mixtures of ^{19}F containing compounds to identify potential candidates for fragment linking.[31] The detection of ternary complexes is facilitated by the increased spectral dispersion of the ^{19}F nucleus, allowing one to test more compounds in the mixture and without the presence of strong background resonances, which can easily interfere with weak ILNOEs in proton detected experiments.[30] Figure 5 illustrates the $^{19}F-^{19}F$ {^1H decoupled}- NOESY experiment for a mixture of six trifluoromethyl containing ligands. The presence of a $^{19}F-^{19}F$ NOE between two of the ligands (Figure 5, compounds **14** and **15**) indicates that the fragments bind close in space and suggests that these two compounds may be successfully combined. The hits typically found in fragment screens may occupy exclusive, partially exclusive, or nonexclusive sites within the binding space of a protein. In all these cases, if no special precautions are taken, the genuine $^{19}F-^{19}F$ ILNOE effect may be obscured by or even mistaken with the indirect magnetization transfer mediated by the protons belonging to the target protein. In the $^{19}F-^{19}F$ NOESY experiment, this effect may be easily eliminated by using either broadband proton decoupling or by a 180 degree refocusing ^1H pulse in the middle of the NOESY mixing time.[35] In the case shown here with BACE-1, we used a compound from the initial hit expansion as a seed and supplemented it with a set of targeted fluorinated compounds to demonstrate how the $^{19}F-^{19}F$ NOESY experiment can be used to search for proximity binders. The results of this search, supplemented by additional chemistry efforts, yielded a compound (compound **16**) with a more than 100-fold increase in potency compared to the original seed fragment.[34]

The major attributes of using ^{19}F as a detection tool have been reviewed extensively by Dalvit and co-workers.[8] In our case, the ability to detect ^{19}F with a cryogenic NMR probe drastically decreased the time necessary to perform a fragment screen on a small fragment library. Screening using fluorine is extremely sensitive to binding of the small molecule to the target protein, and the lack of a protonated background enables use of solvents, buffers, or detergents that would normally interfere with a ^1H-based NMR screen. These factors allow the detection of small quantities of compounds (~20 μM) in very short experiment times (2−3 min). In addition, the large chemical shift dispersion (~200 ppm) and narrow line width ($\Delta_{w1/2}$ ~ 1−2 Hz) for the ^{19}F signal of the free ligand allows the screening of a large number of compounds in a pool without the complication of signal overlap, and, in the case of this library, results in a unique chemical shift for every compound in the pool. Typically, ^1H-based NMR fragment screening is performed in pools of five compounds or less,[29] whereas our ^{19}F screening pools contained 12−13 compounds, with no need for any type of specific pooling strategy. Practically, this number can be safely extended to 20 or more compounds per pool, thereby essentially reducing the time of the screen to an overnight experiment (experimental observations). In comparison, some of the more commonly used ^1H-based NMR

dx.doi.org/10.1021/jm201441k | J. Med. Chem. 2012, 55, 678−687

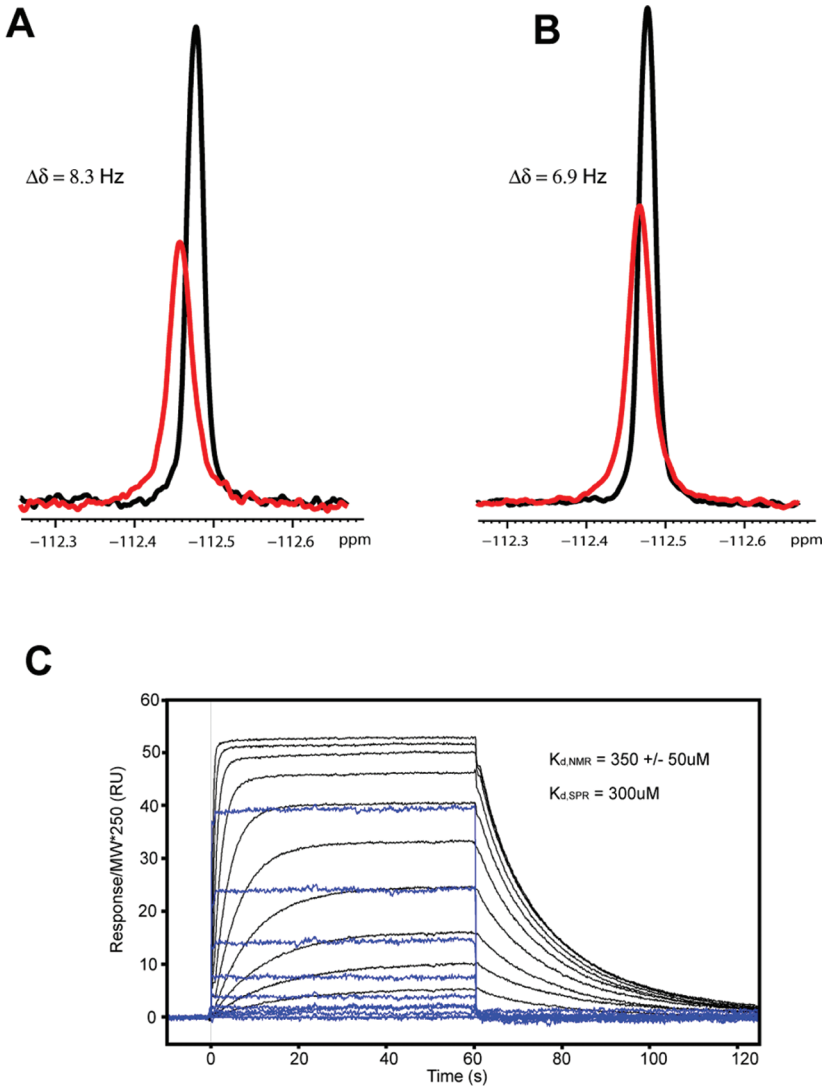


Figure 4. Fast and accurate determination of K_d values using dCSP. Using compound 1 from Table 1, ^{19}F spectra were acquired of samples prepared using 3 μM BACE-1 with (A) 100 μM and (B) 200 μM ligand (red). Traces are scaled for illustration. Chemical shifts were measured with respect to samples containing no protein (black) and K_d was determined using eq 4. (C) Results were confirmed using SPR. Blue traces represent raw sensorgrams for the fragment, and black traces represent raw sensorgrams for the control compound (K_d ~ 400 nM). Binding constants for both compounds were determined using steady-state analysis.

fragment screening methods (i.e., saturation transfer difference (STD) and WaterLOGSY) may require up to 30 min of experiment time per sample and are limited by the number of compounds that can be pooled due to signal overlap.[36] This type of approach, for a library of 1200 compounds using 4−5 compounds per pool, could potentially require more than 200 h of experiment time. In addition, ^{19}F NMR offers extremely high sensitivity to binding events and does not suffer from any of the drawbacks of traditional ^{1}H-detected NMR screening which is prone to false negatives (such as signal cancellation in WaterLOGSY experiments and reliance on fast protein−ligand exchange rates). The binding event can be easily detected by the pronounced changes in the intensity, line shape, and chemical shift of the ^{19}F resonances upon addition of a protein, as shown in Figure 6. These spectral changes can be easily quantified, and accurate K_d values can be quickly obtained using differential chemical shift perturbation (dCSP) experiments or line shape analyses. These methods can be reliably used for K_d determinations up to ~1 mM. For the weaker hits, spectral changes are too small for reliable quantification at ligand concentrations which are below the solubility limit of the monomer. It is worthwhile to point out that in this case, the solubility limit is defined when there is a detectable concentration-dependent ^{19}F resonance shift or line shape change above this limit in the absence of the protein.

One of the possible concerns with using a ^{19}F-based fragment library is the smaller fraction of chemical space that is covered due to the limitation of requiring fluorinated compounds. One would expect this to manifest in apparently lower overall hit rates. In the case of BACE-1, two previous publications have discussed fragment screens of this target where the libraries were not limited by ^{19}F selection.[31,37] In Folmer et al., a fragment screen was performed using solution ^{1}H NMR (WaterLOGSY) and resulted in a hit rate of about 0.5%.[37] Likewise, in 2007, Murray et al. reported the results of a fragment screen of BACE-1 using a focused library of 347 compounds with a hit rate of 0.6%.[31] In our experience using the ^{19}F fragment library, we obtained a hit rate of 0.5%. While the ^{19}F library may appear relatively more limited as far as its coverage of theoretical "fragment space", it appears to provide a sufficiently uniform coverage in accord with Figure 2.

dx.doi.org/10.1021/jm201441k | J. Med. Chem. 2012, 55, 678−687

Table 3. Influence of Fluorine on Affinities of 2-Aminoquinoline Fragments

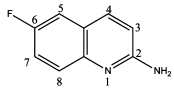
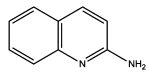
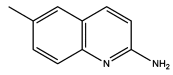
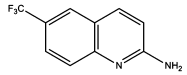
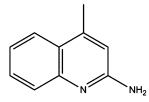
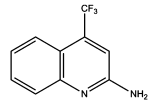
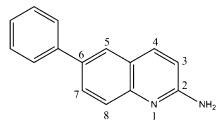
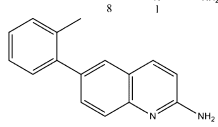
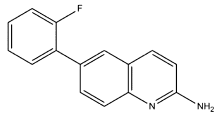
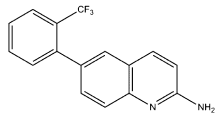
Compound	Structure	SPR K_d (mM)
5		1.2
6		0.9
7		1.6
8		3.9
9		1.3
10		2.9

Table 4. Influence of Fluorine on Affinities of Fragment Analogues

Compound	Structure	SPR K_d (uM)
11		34
12		25
13		39
14		200

We have observed comparable results with a number of other targets (vide infra).

Target Druggability Assessment Using ^{19}F Fragment Screening. Seven targets (Table 5) were chosen to assess the utility of ^{19}F-based fragment screening as a predictor of druggability. Briefly, two methods, MOE SiteFinder (Chemical Computing Group, v2010.10) and SiteMap (Schrodinger Inc., v2.4)[38] were used to identify pockets in each of the target crystal structures. Pocket features such as volume, hydrophobicity, and enclosure were then calculated in SiteMap. Finally, a druggability score, Dscore+, was calculated as Dscore+0.3*phobic, using terms provided by SiteMap. After target druggability was assessed by these computational methods, the results of ^{19}F-fragment screens performed on these targets using the library of 1200 fluorinated fragments were compiled and compared to the Dscore+ values (Table 5). We found that the fragment hit rates corresponded well to the Dscore+ druggability score, thus suggesting that the FBS hit rates were likely a good predictor of HTS success.[27,28]

The value of using FBS versus computational methods in predicting protein druggability lies in its ability to perform this function without the need for protein structures. Here, we extend the use of FBS in predicting protein druggability through the application of ^{19}F NMR spectroscopy. From a computational standpoint, targets with Dscore+ values of 1.2 and below are considered "difficult" targets for small molecule inhibitors.[34] We intentionally included targets of this type to act as negative controls. These were chosen based on available structural data as well as pocket volumes measured by Dscore+. In examining the results of these experiments, we found a clear correlation between the Dscore+ druggability index and the hit rates from the ^{19}F fragment screen (Table 5). In particular, the notably "difficult" targets (such as Target 1, Target 2, p27, and BACE-1)

dx.doi.org/10.1021/jm201441k | J. Med. Chem. 2012, 55, 678–687

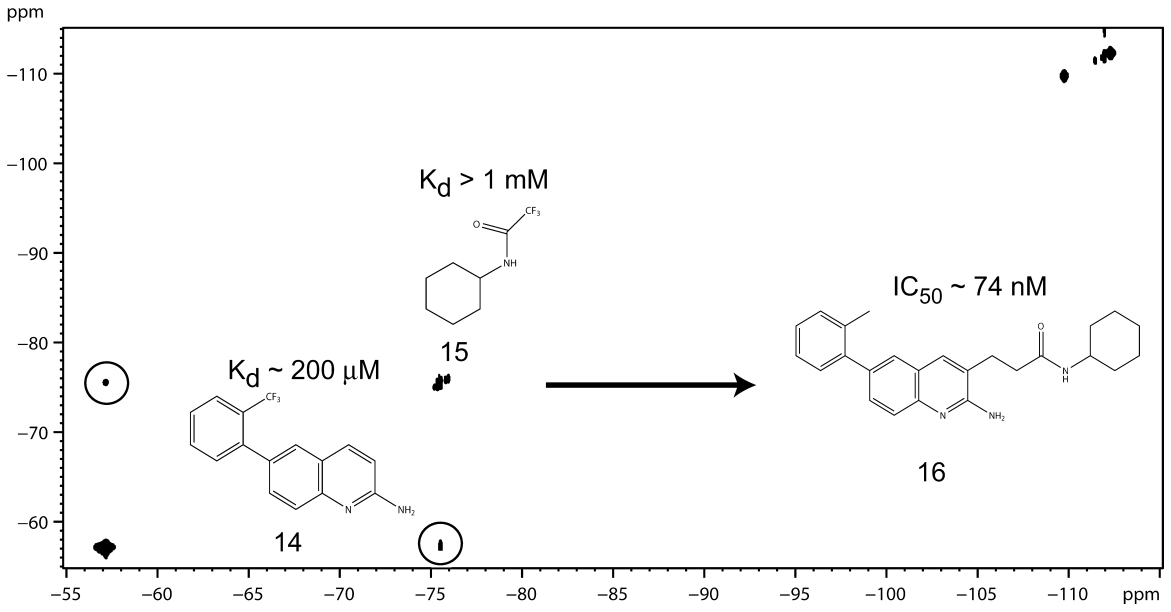


Figure 5. Fluorine−fluorine NOE experiment. ^{19}F-labeled fragments can be mixed in a manner similar to that with ^{1}H fragment libraries, and ^{19}F−^{19}F NOEs can be observed between fragments binding simultaneously to a target and close enough in space (∼5 Å) to yield an NOE. This experiment can give ideas as to how to pursue fragment linking approaches. The circled region of the spectrum shows the NOE between the core fragment and one of the additional six fragments. Protonated versions of the fragments were linked, and results of the linked fragments are shown in the inset (IC_{50} value determined from a FRET assay as described in Cheng et al.[34]).

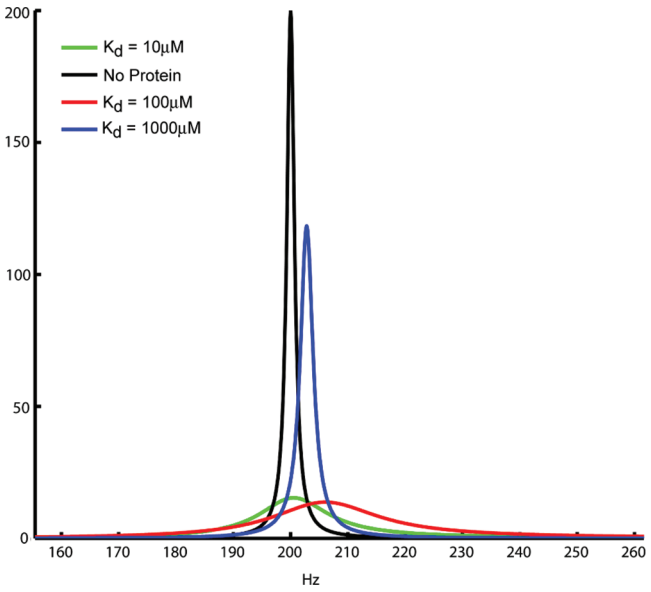


Figure 6. Ultrahigh sensitivity of ^{19}F line shape to the protein binding event. A simulation is shown for a theoretical compound with a chemical shift difference between bound and unbound ligand of 200 Hz, $k_{off} = |K_d\ (\mu M)|\ s^{-1}$, protein concentration, $[P] = 10\ \mu M$, ligand concentration, $[L] = 20\ \mu M$. The lines simulate the ^{19}F line shape of a fragment in the absence of protein (black), a fragment with $K_d = 1$ mM (blue), a fragment with $K_d = 100\ \mu M$ (red), and a fragment with $K_d = 10\ \mu M$ (green). All simulations were performed in MATLAB.

had extremely low hit rates, while targets more widely known to be tractable for small molecule therapeutics, such as Targets 3 and 4, had relatively high hit rates. This finding demonstrates the ability of ^{19}F-based fragment screening to rapidly provide information about a target's druggability when no protein structure is available.

Table 5. Target Druggability: Dscore+ Indices and Fragment Hit Rates

target[a]	Dscore+ index	pocket volume (Å^3)	fragment hit rate (%)
1	0.7	56	0.0
2	0.7	79	0.08
p27 KID	0.8	57	0.2
β-secretase	1.2	338	0.6
Derp7	1.3	170	2.2
3	1.7	176	7.3
4	1.9	334	7.8

[a]Target 1, cytokine receptor; Target 2, Wnt signaling protein; p27 KID, intrinsically disordered protein; β-secretase, aspartic acid protease; Derp 7, dust mite allergen protein; Target 3, apoptotic protein; Target 4, protein kinase.

CONCLUSION

In summary, we have described the construction and implementation of a generalized ^{19}F-based fragment library. In addition to demonstrating that the library assembly is relatively straightforward, we show that fragment screening with a simple one-dimensional ^{19}F NMR experiment (with ^{1}H decoupling) is significantly faster and, in many ways, more robust than traditional ^{1}H-based NMR screening. As shown here, ^{19}F-based fragment screening is sensitive enough to detect the most potent hits as well as hits with K_d values well over 1 mM (even approaching 5 mM, from experimental observations) and can be done using low concentrations of compound (∼20 μM). Low working concentrations and accurate solubility determination avoid, to a large extent, entropically driven compound aggregation in the presence of protein. Here, the sensitivity and speed of ^{19}F NMR allows one to efficiently control for this nonspecific phenomenon by screening fragment hits against ubiquitin. This is done using samples containing an equal mass of protein to that in the initial screen. If compounds appear to bind ubiquitin in the follow up experiments, their binding to

dx.doi.org/10.1021/jm201441k | J. Med. Chem. 2012, 55, 678−687

the target can be attributed to nonspecific binding, and these compounds can be excluded from further experiments. In using our ^{19}F fragment library, we have not observed any propensity toward aggregation, as all the compounds were prescreened for a minimum of 1 mM aqueous solubility using ^{19}F NMR. In addition, to test if the binding to a target has 1:1 stoichiometry, K_d values can be calculated from dCSP experiments recorded at different pairs of concentrations. If the K_d values from the two sets of concentrations are consistent, the binding is likely specific to a single site. If the K_d values are inconsistent, then the binding could be nonspecific or the compound could bind to multiple sites with different K_d values.

Although the fraction of chemical space covered by fluorine containing fragments may be limited compared to non-fluorinated fragments, our experience with multiple screens using this library suggests that similar hit rates are obtained using both methods. Our method focuses on the use of the fluorine nucleus as a detection tool, and subsequent to a ^{19}F fragment screen, protonated analogues of hit compounds can be mined and pursued in initial hit expansion. This approach can significantly add to the chemical space explored and can provide valuable information and early stage SAR to chemists. In addition, we propose the use of hit rates from screens with the ^{19}F fragment library as predictors of target druggability. Using targets from a wide range of biological processes, we demonstrate that ^{19}F FBS hit rates correspond very well with commonly accepted computational methods of assessing target druggability and provide an efficient tool for assessing protein topology in the absence of its structure. Taken together, our results suggest that the ability to rapidly perform fragment screens and obtain subsequent SAR using cryogenic NMR detection of the ^{19}F nucleus can provide significant benefits over traditional proton detected techniques.

MATERIALS AND METHODS

Construction of the ^{19}F-Fragment Library. The ^{19}F fragment library was constructed using both proprietary and commercially available compounds. A list of possible compounds was obtained by searching the internal sample bank as well as the ACD database (~675 000 compounds) using the criteria depicted in Table 1. After application of these filters, an additional diversity selection filter was applied using Daylight Fingerprint (Daylight Chemical Information Systems, Laguna Niguel, CA) as the descriptor. A Tanimoto distance[39] of 0.25 was used as the cutoff to ensure the selected compounds were sufficiently dissimilar.

After all computational filtering, approximately 1084 internal and 1200 external compounds were identified for possible inclusion in the library. Each of the 2284 compounds was then assessed for 1 mM aqueous solubility using ^{19}F NMR. The purpose of this exercise was 2-fold: First, though screens would be conducted at low concentrations (typically 10−20 μM), we wanted to ensure that soluble compounds made up the library in order to be able to pool a large number of compounds (10−20) and to facilitate X-ray crystallography follow up. Second, we used these experiments to build the database of ^{19}F chemical shift information which allowed for instant identification of screening hits. Briefly, stock solutions of each compound were prepared at 20 mM in nondeuterated DMSO followed by preparation of aqueous solutions (20 mM sodium phosphate, pH 7.0, 50 mM NaCl) consisting of 250 μM and 1 mM concentrations of each compound. The ^{19}F NMR spectrum of each compound was measured at each concentration, and, through automated processing, the intensities of each spectrum were compared. If the intensity of the 1 mM spectrum was not 4 times (±15%) that of the 250 μM sample, the compound was excluded from the final library. After the solubility assay, approximately 1200 compounds (~50%) were identified as suitable candidates for inclusion into this library.

We assessed the diversity of the final ^{19}F fragment library, using a computational model of "fragment space" created from the ACD and ACDSC databases (total of ~6 million compounds) and using the selection filters depicted in Table 1. This exercise yielded approximately 67 000 reference compounds. Using principal component analysis (PCA), each compound was described with two principal components, and a representative "^{19}F fragment space" was created (Figures 1 and 2). Principle component analysis (PCA) was carried out using the Pipeline Pilot (Accelrys, Inc.) "Learn Molecular PCA Model" component and using MDL public keys as the molecular properties for learning.[40] At least three components were required to explain a minimum of 75% variance. For each reference compound, the nearest neighbor in the ^{19}F fragment library was found, and a Tanimoto similarity score[39] based on Daylight Fingerprint was calculated. In general, a pair of compounds with 70% similarity was found to have reasonable structural similarity. Therefore, an arbitrary cutoff of 70% was used. If a nearest neighbor compound with >70% similarity was found, we assumed that the reference compound in "fragment space" was represented in the final library (Figures 1 and 2). Through these calculations, it was found that with as few as ~1200 compounds, approximately 30% of the defined "^{19}F fragment space" was covered. These compounds were acquired with the idea that we could continually acquire more compounds in order to more thoroughly cover the remaining 70% of "fragment space".

Execution of ^{19}F Fragment Screen. The initial fragment stock solutions were prepared at 20 mM stock concentration in nondeuterated DMSO and plated into 96-well plates. The plates containing the fragment library (~13 plates in all) were pooled by collapsing the plates into one final pool plate, yielding 12−13 compounds per pool. NMR samples of 600 μL in volume and prepared in 96, 1 mL well format were transferred into 5 mm diameter and 10 cm long NMR tubes using a Gilson GX-280 (Gilson, Middleton, WI) and subsequently placed into a refrigerated SampleJet (Bruker Biospin, Billerica, MA) with a temperature setting of 279 K. The reference plate (no protein) and the screening plate were both run under the same conditions for each screen. In principle, if the screening conditions are chosen to be identical for all screens, it should be sufficient to record the reference plate just once. However in practice, hit deconvolution is much easier by having recorded both plates prepared from the same DMSO stock solutions from each pool. This is mainly due to the high sensitivity of ^{19}F chemical shifts toward small mismatches of the DMSO content. Also, possible compound deterioration over time may complicate interpretation of spectra. Using this screening format (in 5 mm NMR tubes), the amount of protein used to perform the fragment screen was approximately 5.0×10^{-7} moles, which for BACE-1 (a ~47 kDa target) required approximately 20 mg of unlabeled protein.

Determination of K_d values. In the limit of $[L] \gg [P]$ (i.e., small fraction of the bound ligand ($p_b \ll 1$)), the difference between the observed frequency shift in the presence and in the absence of a protein is described by the following equation:[41]

$$\nu_{obs} - \nu_{free} = p_b \times \Delta\nu^{app} \quad (2)$$

where $\Delta\nu^{app}$ is a complex function of the nuclear and kinetic parameters. Importantly, this equation adequately quantifies frequency shift changes for both fast ($k_{off} \gg \sim\nu_{bound} - \nu_{free}$) and intermediate ($k_{off} \sim \nu_{bound} - \nu_{free}$) chemical exchange limits, where the second case is frequently encountered in ^{19}F detection due to the large chemical shift dispersion of this nucleus. The fraction of the bound ligand (p_b) for the $[L]_0 \gg [P]_0$ case can be expressed by Fielding as[42]

$$p_b = \frac{[P]_o}{K_d + [L]_o} \quad (3)$$

If one defines the differential frequency shift, γ, as

$$\gamma = \frac{\nu_{1,obs} - \nu_{free}}{\nu_{2,obs} - \nu_{free}} \quad (4)$$

Then the K_d value can be obtained using eqs 2 and 3 for two different ligand concentrations ($[L]_1$ and $[L]_2$) (see Supporting

dx.doi.org/10.1021/jm201441k | J. Med. Chem. 2012, 55, 678−687

Information for derivation):

$$K_d = \frac{\gamma[L]_1 - [L]_2}{1 - \gamma} \quad (5)$$

To obtain K_d values using this method, samples were usually prepared with L_1 = 100 and L_2 = 200 μM and with identical protein concentrations on the order of 1−5 μM. We refer to this method as dCSP.

In the case of tighter binding ligands ($k_{off} \ll \nu_{bound} - \nu_{free}$) (i.e., no chemical shift but line width changes of the observed resonance), the K_d values can be obtained in an analogous way from the differential line broadening of the observed free ligand signal at the two different concentrations. In this case, however, the changes of the line width are relatively much smaller and one needs to increase the protein concentration beyond the $[L]_0 \gg [P]_0$ limit. Substituting the frequency shift in eq 4 for the half-width of the resonance line, one can write the following relationship:

$$\gamma = \frac{p_{b1}(1 - p_{b2})}{p_{b2}(1 - p_{b1})} \quad (6)$$

where p_b is the bound fraction of the ligand:[43]

$$p_b = \{([L]_o + [P]_o + K_d) - \sqrt{([L]_o + [P]_o + K_d)^2 - 4[L]_o[P]_o}\} / (2[L]_o) \quad (7)$$

The K_d value is then obtained from the numerical solution of eq 6.

NMR Spectroscopy. All NMR experiments were performed on a Bruker Avance III NMR spectrometer (Bruker Biospin) operating at a ^{1}H frequency of 500.13 MHz using a SEF cryogenic probe equipped for direct ^{19}F detection. One dimensional ^{19}F spectra were acquired for each sample at 283 K using ^{1}H decoupling with a spectra width of 71 428 Hz, an acquisition time of 917 ms, and 128 scans with a relaxation delay of 1 s. This yielded experiment times on the order of 4 min each, including 2 min for initial temperature equilibration. This experimental set up allowed all reference and screen spectra to be acquired in less than 24 h. Typically, a standard buffer (20 mM sodium phosphate, 50 mM NaCl (± DTT), pH = 7) is used in the screening campaigns. This appeared suitable for all our measured targets and allowed the elimination of buffer-dependent variations in the ^{19}F chemical shifts while retaining good cryoprobe sensitivity at this moderate salt concentration.

All data were processed using Topspin 2.1 (Bruker Biospin, Billerica, MA) and were then compared visually to the reference spectra using the spectral overlay feature. Hits were identified by signal intensity and/or chemical shift changes. Since each compound in a pool had a unique chemical shift, hit identification was straightforward, and hit compounds could be identified by simply matching the chemical shift of the hit compound to that found in the compound database.

Surface Plasmon Resonance Spectroscopy. Dissociation constant (K_d) measurements were performed on Biacore S51 and T100 instruments (GE Healthcare). BACE-1 protein and inhibitors for Biacore measurements were generated in-house; all other reagents were purchased from GE Healthcare or Sigma-Aldrich. Glycosylated BACE-1 was reacted with sodium periodate to oxidize *cis*-diols groups on sugar chains to aldehydes. The oxidized BACE-1 was immobilized at high density (10000−12000 RU) onto CM5 chips using aldehyde coupling chemistry and resulted in surface activities close to 100% based on reference inhibitor binding. The immobilization running buffer consisted of 10 mM HEPES pH 7.4 with 150 mM NaCl and immobilization steps consisted of a 3−4 min EDC/NHS activation step [200 mM 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide hydrochloride, 50 mM *N*-hydroxysuccinimide], 7 min 5 mM carbohydrizide in water, 7 min 1 M ethanolamine hydrochloride pH 8.5, 10 min 20 μg/mL oxidized BACE-1 in 10 mM sodium acetate pH 4.0, 7 min 100 mM sodium cyanoborohydride in 100 mM sodium acetate pH 4.0, 3 × 30 s 50 mM glycine pH 9.5 and 3 × 30 s 1 M NaCl in 100 mM $NaHCO_3$ pH 9.5. The reference spot for all SPR experiments consisted of the carboxymethyl dextran surface of the CM5 chip treated in an identical manner as the other flow cell spot excluding the addition of protein.

For K_d measurements the buffer was replaced with 50 mM sodium acetate pH 5.0, 150 mM NaCl, 0.005 (v/v) Tween20, and 5% (v/v) DMSO. Compound stocks prepared in DMSO (typically 20 mM) were serially diluted in running buffer and injected over the immobilized BACE-1. Association and dissociation time were typically set to 45 s and 90 s, respectively. The runs were performed at 25 °C with a flow rate from 10 to 90 μL/min and data collection rate of 10 Hz.

The data were processed and analyzed using Scrubber-2 analysis software (BioLogic Software, Campbell, Australia). The sample response observed on the reference spot was subtracted from the sample response with immobilized BACE-1 to correct for systematic noise and baseline drift. Data was solvent corrected and the response from blank injections was used to double-reference the binding data. The data were molecular weight normalized and K_d values established using simple steady analysis with one global R_{max} (or individual R_{max} in cases where complete saturation was achieved).

Chemistry. All compounds were purified to ≥95% purity as determined by reverse phase HPLC. HPLC analysis was obtained on Agilent 1100, using the following methods: HPLC method (3.6 min LC-MS run): Zorbax analytical C18 column (50 mm × 3 mm, 3.5 μm, 40 °C); mobile phase, A = 0.1% TFA in water, B = 0.1% TFA in acetonitrile; gradient, 0.0−3.6 min, 5−95% B; flow rate = 1.5 mL/min; 254 nm; 0 min post time; 1.0 μL injection.

■ ASSOCIATED CONTENT

Supporting Information

Derivations of eqs 3 and 5 as used in the manuscript. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*Tel: 805-313-5647. Fax: 805-498-9057. E-mail: jbjordan@amgen.com (J.B.J.); lpoppe@amgen.com (L.P.).

■ ACKNOWLEDGMENTS

The authors would like to acknowledge Hanh Nguyen for synthesis of some of the compounds used in this work. We would also like to acknowledge Dr. Richard Kriwacki (St. Jude Children's Research Hospital) and Dr. Geoffrey Mueller (NIEHS) for providing protein with which to perform the druggability studies.

■ ABBREVIATIONS USED

FBDD, fragment based drug discovery; FBS, fragment based screening; LE, ligand efficiency; HA, heavy atoms; HTS, high throughput screening; NMR, nuclear magnetic resonance; SPR, surface plasmon resonance; LEF, local environment of fluorine; BACE-1, β-site APP cleaving enzyme-1; ILNOE, interligand nuclear Overhauser enhancement; ACD, Available Chemicals Directory; ACDSC, Available Chemicals Directory Screening Compounds; dCSP, differential chemical shift perturbation; NOESY, nuclear Overhauser enhancement spectroscopy; STD, saturation transfer difference; WaterLOGSY, water-ligand observe gradient spectroscopy; SAR, structure−activity relationship

■ REFERENCES

(1) Murray, C. W.; Rees, D. C. The rise of fragment-based drug discovery. *Nat. Chem.* **2009**, *1*, 187−192.

(2) Erlanson, D. A.; Hansen, S. K. Making drugs on proteins: site-directed ligand discovery for fragment-based lead assembly. *Curr. Opin. Chem. Biol.* **2004**, *8*, 399−406.
(3) Rees, D. C.; Congreve, M.; Murray, C. W.; Carr, R. Fragment-based lead discovery. *Nat. Rev. Drug Discovery* **2004**, *3*, 660−672.
(4) Carr, R.; Jhoti, H. Structure-based screening of low-affinity compounds. *Drug Discovery Today* **2002**, *7*, 522−527.
(5) Nienaber, V. L.; Richardson, P. L.; Klighofer, V.; Bouska, J. J.; Giranda, V. L.; Greer, J. Discovering novel ligands for macromolecules using X-ray crystallographic screening. *Nat. Biotechnol.* **2000**, *18*, 1105−1108.
(6) Hopkins, A. L.; Groom, C. R.; Alex, A. Ligand efficiency: a useful metric for lead selection. *Drug Discovery Today* **2004**, *9*, 430−431.
(7) Hajduk, P. J.; Greer, J. A decade of fragment-based drug design: strategic advances and lessons learned. *Nat. Rev. Drug Discovery* **2007**, *6*, 211−219.
(8) Dalvit, C. Ligand- and substrate-based ^{19}F NMR screening: Principles and applications to drug discovery. *Prog. Nucl. Magn. Reson. Spectrosc.* **2007**, *51*, 243−271.
(9) Poppe, L.; Harvey, T. S.; Mohr, C.; Zondlo, J.; Tegley, C. M.; Nuanmanee, O.; Cheetham, J. Discovery of ligands for Nurr1 by combined use of NMR screening with different isotopic and spin-labeling strategies. *J. Biomol. Screen.* **2007**, *12*, 301−311.
(10) Vulpetti, A.; Hommel, U.; Landrum, G.; Lewis, R.; Dalvit, C. Design and NMR-based screening of LEF, a library of chemical fragments with different local environment of fluorine. *J. Am. Chem. Soc.* **2009**, *131*, 12949−12959.
(11) Morgenthaler, M.; Schweizer, E.; Hoffmann-Roder, A.; Benini, F.; Martin, R. E.; Jaeschke, G.; Wagner, B.; Fischer, H.; Bendels, S.; Zimmerli, D.; Schneider, J.; Diederich, F.; Kansy, M.; Muller, K. Predicting and tuning physicochemical properties in lead optimization: amine basicities. *ChemMedChem* **2007**, *2*, 1100−1115.
(12) Leroux, F. Atropisomerism, biphenyls, and fluorine: a comparison of rotational barriers and twist angles. *ChemBiochem* **2004**, *5*, 644−649.
(13) Gerebtzoff, G.; Li-Blatter, X.; Fischer, H.; Frentzel, A.; Seelig, A. Halogenation of drugs enhances membrane binding and permeation. *ChemBiochem* **2004**, *5*, 676−684.
(14) Purser, S.; Moore, P. R.; Swallow, S.; Gouverneur, V. Fluorine in medicinal chemistry. *Chem. Soc. Rev.* **2008**, *37*, 320−330.
(15) Hopkins, A. L.; Groom, C. R. The druggable genome. *Nat. Rev. Drug Discovery* **2002**, *1*, 727−730.
(16) Barelier, S.; Krimm, I. Ligand specificity, privileged substructures and protein druggability from fragment-based screening. *Curr. Opin. Chem. Biol.* **2011**, *15*, 469−474.
(17) Kubinyi, H. Drug research: myths, hype and reality. *Nat. Rev. Drug Discovery* **2003**, *2*, 665−668.
(18) Brown, D.; Superti-Furga, G. Rediscovering the sweet spot in drug discovery. *Drug Discovery Today* **2003**, *8*, 1067−1077.
(19) Miranker, A.; Karplus, M. Functionality maps of binding sites: a multiple copy simultaneous search method. *Proteins* **1991**, *11*, 29−34.
(20) Laskowski, R. A. SURFNET: a program for visualizing molecular surfaces, cavities, and intermolecular interactions. *J. Mol. Graph.* **1995**, *13* (323−30), 307−308.
(21) Hendlich, M.; Rippmann, F.; Barnickel, G. LIGSITE: automatic and efficient detection of potential small molecule-binding sites in proteins. *J. Mol. Graph. Model.* **1997**, *15*, 359−363.
(22) Liang, J.; Edelsbrunner, H.; Fu, P.; Sudhakar, P. V.; Subramaniam, S. Analytical shape computation of macromolecules: I. Molecular area and volume through alpha shape. *Proteins* **1998**, *33*, 1−17.
(23) Liang, J.; Edelsbrunner, H.; Fu, P.; Sudhakar, P. V.; Subramaniam, S. Analytical shape computation of macromolecules: II. Inaccessible cavities in proteins. *Proteins* **1998**, *33*, 18−29.
(24) Perot, S.; Sperandio, O.; Miteva, M. A.; Camproux, A. C.; Villoutreix, B. O. Druggable pockets and binding site centric chemical space: a paradigm shift in drug discovery. *Drug Discovery Today* **2010**, *15*, 656−667.
(25) Chen, I. J.; Hubbard, R. E. Lessons for fragment library design: analysis of output from multiple screening campaigns. *J. Comput.-Aided Mol. Des.* **2009**, *23* (8), 603−620.
(26) Cheng, A. C.; Coleman, R. G.; Smyth, K. T.; Cao, Q.; Soulard, P.; Caffrey, D. R.; Salzberg, A. C.; Huang, E. S. Structure-based maximal affinity model predicts small-molecule druggability. *Nat. Biotechnol.* **2007**, *25*, 71−75.
(27) Hajduk, P. J.; Huth, J. R.; Fesik, S. W. Druggability indices for protein targets derived from NMR-based screening data. *J. Med. Chem.* **2005**, *48*, 2518−2525.
(28) Edfeldt, F. N.; Folmer, R. H.; Breeze, A. L. Fragment screening to predict druggability (ligandability) and lead discovery success. *Drug Discovery Today* **2011**, *16*, 284−287.
(29) Davies, T. G., Monfort, R. L. M., Williams, G., Jhoti, H. Pyramid: An integrated platform for fragment-based drug discovery. In *Fragment-based Approaches in Drug Discovery*; Erlanson, D. A., Jahnke, W., Ed.; Oxford University Press: Oxford, UK, 2006; pp 193−214.
(30) London, R. E. Theoretical analysis of the inter-ligand overhauser effect: a new approach for mapping structural relationships of macromolecular ligands. *J. Magn. Reson.* **1999**, *141*, 301−311.
(31) Murray, C. W.; Callaghan, O.; Chessari, G.; Cleasby, A.; Congreve, M.; Frederickson, M.; Hartshorn, M. J.; McMenamin, R.; Patel, S.; Wallis, N. Application of fragment screening by X-ray crystallography to beta-secretase. *J. Med. Chem.* **2007**, *50*, 1116−1123.
(32) Bondi, A. Van der Waals volumes and radii. *J. Phys. Chem.* **1964**, *68*, 441−451.
(33) Pauling, L. *The Nature of the Chemical Bond*, 6th ed.; Cornell University Press: Ithaca, NY, 1948.
(34) Cheng, Y.; Judd, T.; Bartberger, M.; Chen, K.; Fremeau, R.; Hickman, D.; Hitchcock, S.; Jordan, J.; Li, V.; Lopez, P.; Louie, S.; Luo, Y.; Michelsen, K.; Nixey, T.; Powers, T.; Rattan, C.; Sickmier, E. St.; Jean, D.; Wahl, R.; Wen, P.; Wood, S. From fragment screening to in vivo efficacy: optimization of a series of 2-aminoquinolines as potent inhibitors of BACE1. *J. Med. Chem.* **2011**, *54* (16), 5836−5857.
(35) Massefski, W.; Redfield, A. Elimination of multiple-step spin diffusion in two-dimensional NOE spectroscopy of nucleic acids. *J. Magn. Reson.* **1988**, *78*, 150−155.
(36) Dalvit, C.; Gossert, A. D.; Coutant, J.; Piotto, M. Rapid acquisition of ^{1}H and ^{19}F NMR experiments for direct and competition ligand-based screening. *Magn. Reson. Chem.* **2011**, *49*, 199−202.
(37) Geschwindner, S.; Olsson, L. L.; Albert, J. S.; Deinum, J.; Edwards, P. D.; de Beer, T.; Folmer, R. H. Discovery of a novel warhead against beta-secretase through fragment-based lead generation. *J. Med. Chem.* **2007**, *50*, 5903−5911.
(38) Halgren, T. A. Identifying and characterizing binding sites and assessing druggability. *J. Chem. Inf. Model.* **2009**, *49*, 377−389.
(39) Lipkus, A. A proof of the triangle inequality for the Tanimoto distance. *J. Math. Chem.* **1999**, *26*, 263−265.
(40) Durant, J. L.; Leland, B. A.; Henry, D. R.; Nourse, J. G. Reoptimization of MDL keys for use in drug discovery. *J. Chem. Inf. Comput. Sci.* **2002**, *42*, 1273−1280.
(41) Swift, T. J.; Connick, R. E. NMR-relaxation mechanisms of O^{17} in aqueous solutions of paramagnetic cations and the lifetime of water molecules in the first coordination sphere. *J. Chem. Phys.* **1962**, *37*, 307−320.
(42) Fielding, L. NMR methods for the determination of protein-ligand dissociation constants. *Prog. Nucl. Magn. Reson. Spectrosc.* **2007**, *51*, 219−242.
(43) Forsen, S.; Hoffman, R. Study of moderately rapid chemical exchange reactions by means of nuclear magnetic double resonance. *J. Chem. Phys.* **1963**, *39*, 2892.