

New Frontiers in Druggability

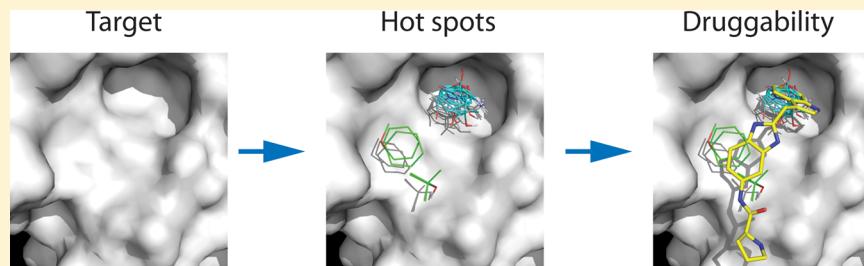
Dima Kozakov,^{*,†} David R. Hall,[‡] Raeanne L. Napoleon,[§] Christine Yueh,^{||} Adrian Whitty,^{*,§} and Sandor Vajda^{*,§,||}

[†]Department of Applied Mathematics & Statistics, Stony Brook University, Stony Brook, New York 11794, United States

[‡]Acpharis Inc., Holliston, Massachusetts 01746, United States

[§]Department of Chemistry and ^{||}Department of Biomedical Engineering, Boston University, Boston, Massachusetts 02215, United States

Supporting Information



ABSTRACT: A powerful early approach to evaluating the druggability of proteins involved determining the hit rate in NMR-based screening of a library of small compounds. Here, we show that a computational analog of this method, based on mapping proteins using small molecules as probes, can reliably reproduce druggability results from NMR-based screening and can provide a more meaningful assessment in cases where the two approaches disagree. We apply the method to a large set of proteins. The results show that, because the method is based on the biophysics of binding rather than on empirical parametrization, meaningful information can be gained about classes of proteins and classes of compounds beyond those resembling validated targets and conventionally druglike ligands. In particular, the method identifies targets that, while not druggable by druglike compounds, may become druggable using compound classes such as macrocycles or other large molecules beyond the rule-of-five limit.

INTRODUCTION

Among the distinct proteins encoded by the human proteome, it had been estimated that only about 3000 belong to structural families for which small molecule (i.e., synthetic organic) drugs have been developed.¹ Many biologically compelling drug targets belong to other protein families that lack such empirical proof of principle that they can be inhibited by small molecule drugs. Consequently, approaches to assessing the druggability of potential drug targets, i.e., the likelihood of being able to identify a druglike small ligand that can modulate the activity of the target, have emerged as an important tool for target-based drug discovery. We note that throughout this perspective, we call a small compound druglike if it satisfies Lipinski's rule of five (no more than five hydrogen-bond donors, no more than 10 hydrogen-bond acceptors, a molecular mass less than 500 g/mol, and an octanol–water partition coefficient not greater than 5).² There are two major approaches to determining the druggability of a protein. One is based on empirical analysis of known target/drug associations,^{1,3,4} which can be applied only when the protein in question belongs to a well-studied family that includes close homologues for which reliable druggability information is available. The other is based solely on analysis of the structure of the target protein.^{5–17} For novel targets, from families for which no drugs or advanced drug candidates are known, one has to rely on the latter method of structure-based

predictive models. This approach was pioneered by Hajduk et al.,⁷ who re-examined the results of NMR-based fragment screening for a diverse set of 23 proteins and observed a high correlation between the fragment hit rate and the ability to identify high-affinity ($K_d < 300$ nM) druglike ligands. They used these data to derive a regression-based model to predict the likely fragment hit rate for untested targets, based on structural properties of the protein binding site such as total and apolar surface area and pocket compactness, generating for each a predicted druggability score, defined as the logarithm of the predicted NMR hit rate. The regression formula provided a good approximation of the original hit rates and, as we will show, resulted in fairly accurate prediction of druggability. Another important contribution came from Cheng et al.,⁸ who used the curvature and the lipophilic surface area of pockets to predict the maximal affinity achievable by passively absorbed drugs. Despite its simplicity, the model also provided good discrimination of druggable binding sites from those labeled as difficult or undruggable, but it was tested only on a limited number of proteins.

To further improve druggability prediction, a variety of methods have been published in the last 6 years, with most of

Received: April 15, 2015

Published: July 31, 2015



them trained on the growing number of protein targets with some level of druggability information available.^{5,10–12,14,16} For method validation, Schmidtke and Barril compiled a large set of protein targets with associated degrees of druggability.¹¹ Krasowski et al.¹² added further proteins to the training set and developed the DrugPred druggability score. All of these methods are based on the analysis of various pocket properties, including hydrophobicity, size, compactness, hydrogen-bond donor and -acceptor surface areas, and amino acid composition in pockets. Thus, evaluating the druggability of proteins with respect to conventional druglike compounds is now a mature technology that can perform such assessments with some reliability. However, this technology has considerable limitations. Apart from the NMR-based fragment screening by Hajduk et al.,⁷ all methods correlate pocket properties with druggability data available for traditional targets and therefore, in essence, simply evaluate whether the novel target in question contains a binding pocket that in some sense resembles pockets on known druggable targets that bind conventional druglike ligands. Moreover, the methods provide a single druggability scale that attempts to capture both the expected affinity that can be achieved as well as the druglikeness, and, implicitly, the potential for oral bioavailability, of ligands that will potentially bind to the site. A general shortcoming of all of these methods, therefore, is that they are poorly equipped to evaluate the druggability of proteins with surface topologies that substantially differ from the relatively small number of protein families for which drugs or high-affinity druglike ligands exist, such as protein–protein interaction (PPI) targets, or to assess the potential of any target to bind ligands that do not conform to conventional definitions of druglikeness. Moreover, because these methods are parametrized rather than being based on biophysical principles and return only a monolithic measure of druggability, they provide little insight into why a particular target is or is not druggable and therefore into the prospects that strong binding might be achievable by an appropriately chosen nontraditional drug chemotype such as a natural product, a macrocyclic compound,^{18–21} a covalent inhibitor,^{22–24} or a peptide-derived foldamer.^{25–28}

It is additionally noteworthy that, among the above methods, only Seco et al.⁹ explicitly take into account the potential for conformational flexibility of the protein to affect the druggability of a binding site. The other methods base their analysis on static X-ray crystal structures. Several of the studies did, however, test the robustness of their results to variations in protein conformation, by comparing the results obtained after computationally perturbing the structure¹⁰ or by comparing the results obtained for different crystal structures of the same target protein,^{5,8,11,12} and generally reported that the ability to identify druggable sites was not greatly sensitive to such conformational effects. Other work has suggested, however, that conformational adaptivity at the binding site can substantially affect druggability, in cases where the alternative conformations that are sampled through thermal motions persist for a sufficient period of time.²⁹

In this perspective, we first discuss the advantages of the fragment screening approach used by Hajduk et al.⁷ and introduce an analog of the method that uses a computational algorithm rather than experimental NMR data for fragment screening. The computational approach was used to evaluate some of the historical druggability results published during the past decade, including the 23 proteins considered in the NMR-based screening study^{6,7} and the additional set of 72 proteins

previously studied by the regression model, based on the NMR results and developed by the same authors.⁷ We also studied the protein set selected by Cheng et al.⁸ and some of the more contentious results from the DrugPred algorithm of Krasowski et al.¹² As will be shown, the method can reliably reproduce earlier druggability results. Moreover, evaluation of the small number of cases in which the original assessment was uncertain or was subsequently proved to be incorrect shows that the method presented here can, in fact, provide an even more meaningful assessment of druggability. We additionally show that biophysics-based computational approaches (as opposed to empirical approaches that are parametrized against a particular training set of protein–ligand complexes) are now sufficiently reliable that they can be used to address the potential for highly challenging or atypical drug targets, such as protein–protein interfaces, to be addressed using non-canonical drug chemotypes such as synthetic macrocycles. These methods, therefore, point the way to expanding the range of drug targets that can be addressed by using chemotypes that are known to be pharmaceutically relevant but that fall outside of conventional druglike chemical space. We also provide detailed discussion of the druggability of six targets of high pharmaceutical interest for which the prospects of success have been controversial.

Experimental Fragment-Based Analysis of Druggability. As already mentioned, the structure-based analysis of druggability was introduced in a landmark paper by Hajduk and co-workers,⁷ who re-examined the results of NMR-based fragment screening experiments and observed a high correlation between the fragment hit rate and the ability to identify high-affinity druglike inhibitors. This early method has two major advantages over more recent approaches. First, by determining the binding locations of fragment-sized compounds, the screening based method automatically identifies the binding site. In contrast, the prediction methods described in the [Introduction](#) are based on the analysis of pocket properties, primarily size, shape, complexity, and hydrophobicity, and thus have first to select the binding site residues. As noticed by Rarey and co-workers,¹⁷ pocket descriptors depend heavily on correctly selecting the region of the protein surface that is most likely to form a binding site. This is a nontrivial task for a ligand-free protein and is also complicated by potential conformational differences between ligand-bound and -free structures. Second, as Hajduk et al.⁷ were aware, their method actually identifies binding hot spots, i.e., regions of the binding site that contribute a disproportionate amount of the binding free energy^{30,31} and hence that are primarily responsible for ligand binding, in contrast to other methods that base their predictions on global pocket properties without regard for the extent to which sites within the pocket can contribute significant binding energy.

The concept of hot spots was originally introduced in the context of mutating interface residues to alanine in protein–protein or protein–peptide interfaces.^{30,32–34} On the basis of this method, a residue is considered to be a hot spot if its mutation to alanine gives rise to a substantial drop in binding affinity. The concept was extended to the binding of small molecules by Ringe and co-workers, who used X-ray crystallography to determine the structure of the target protein soaked in aqueous solutions of 6–8 organic solvents.^{35,36} By superimposing the structures in these so-called multiple solvent crystal structure (MSCS) experiments, regions that bind multiple different probes were detected.^{35,36} While individual probes were able to bind at a number of locations, their clusters

indicated binding hot spots. In their NMR-based experiments, Hajduk et al.⁷ incubated the target protein with a series of small organic probes, generically referred to as fragments, and used the perturbations in residue chemical shifts to identify residues that participate in small molecule binding, thereby also finding the binding hot spots. While results of the MSCS method requiring the determination of X-ray structures in multiple solvents were reported only for a few proteins,^{36–40} Hajduk et al.⁷ analyzed the results of NMR-based fragment screening for 23 proteins. These results convincingly demonstrated that the small fragment probe ligands cluster at hot spots and that the hit rate predicts the druggability of the site,^{7,35} thereby initiating the development of structure-based druggability conditions.

Computational Fragment Screening by FTMap. On the basis of the above results from the Hajduk⁷ and Ringe³⁵ groups, NMR and X-ray based fragment screening provides information on druggability and is now broadly used in pharmaceutical research.^{41,42} However, the experiments are expensive and may be hampered by the requirement that the target protein be expressed and purified in relatively large quantities for experimental analysis.¹⁷ Thus, it is desirable to replace these experimental methods with *in silico* approaches to evaluating druggability, particularly when exploring novel targets. Hajduk et al.⁷ accomplished this goal by constructing a regression equation to express the NMR hit rates in terms of pocket properties. In this perspective, we describe druggability conditions based on direct simulation of the X-ray and NMR-based screening experiments. In fact, modeling the weakly specific binding of small and rigid ligands to proteins is not very difficult and hence computational methods offer attractive alternatives.^{43,44} We use computational solvent mapping, implemented as the FTMap algorithm.⁴⁵ The method places small molecular probes of various sizes and shapes on a dense grid around the protein, finds favorable positions using empirical energy functions, refines the bound positions while accounting for side chain and probe flexibility, clusters the conformations, and ranks the clusters on the basis of the average energy. All ligands and crystallographic water molecules are removed prior to mapping, and the probes are initially distributed over the entire protein surface without any assumption as to the location of the binding site(s). As has been extensively validated, the regions that bind multiple low-energy probe clusters, called consensus cluster (CC) sites, identify the locations of the binding energy hot spots.^{40,45–50} The CCs are ranked on the basis of the number of probe clusters contained, which we have shown corresponds to relative energetic importance,⁵¹ and are denoted as CC1, CC2, and so on. The CC with the highest number of probe clusters in the binding site of interest is defined as the primary hot spot, whereas other CCs with fewer probe clusters are secondary hot spots.

FTMap has been implemented as a server (<http://ftmap.bu.edu/>), which currently uses 16 small molecules as probes (ethanol, isopropanol, isobutanol, acetone, acetaldehyde, dimethyl ether, cyclohexane, ethane, acetonitrile, urea, methylamine, phenol, benzaldehyde, benzene, acetamide, and *N,N*-dimethylformamide). The server and its potential applications were recently discussed in detail.⁵² We explored the use of substantially larger probe libraries, but, since this did not affect druggability predictions, we returned to the 16 compounds whose aqueous solutions have been used previously in mapping experiments based on NMR⁵³ or X-ray techniques.^{35–38,54} Since most probes include both polar and nonpolar moieties,

our mapping results show that the highest number of probes cluster in pockets that have a mosaic-like pattern of hydrophobic and polar regions, enabling the binding of many different compounds in a variety of poses.⁴⁸

Druggability Analysis by FTMap. As will be described, we have mapped and analyzed over 150 ligand–protein complexes as well as the corresponding ligand-free proteins. To establish objective FTMap-based criteria for evaluating the druggability of proteins, we set out to benchmark our results by using the 16 probe types used in the FTMap server. For most target proteins, we mapped both the highest resolution ligand-free and -bound structures available in the Protein Data Bank (PDB).⁵⁵ As explained below, generally, druggability can best be determined by mapping the ligand-free structure; although analysis of ligand-bound structure provides additional information, in most cases it is used only to indicate the binding location of a known ligand. On the basis of the analysis of this large set of proteins, we have observed that the hot spots of druggable proteins satisfy conditions on (1) strength, (2) connectivity or compactness, and (3) the maximum dimension of the hot spot region. We describe the origin of these conditions in turn.

As already mentioned, Hajduk et al. observed that the hit rate in fragment screening is a predictor of druggability.⁷ In agreement with this observation, we have previously established that, when using 16 probes for the mapping, the sites that are known to be druggable invariably contain a strong hot spot containing 16 or more probe clusters.⁴⁸ Indeed, among the 121 drug targets considered in earlier studies by Hajduk et al.⁷ and Cheng et al.,⁸ and also examined by FTMap here, only in a single case was there a druggable target with known, potent ligands that had fewer than 16 probe clusters in the strongest hot spot at the drug binding site (Figure S1), whereas we found 21 such targets among those that experimental evidence suggests are not druggable (Figure S2). Since FTMap clusters probe positions using pairwise 4 Å root-mean-square deviation (RMSD) as the clustering radius, the probe clusters are relatively small; thus, a consensus cluster (CC) can include several clusters of the same probe type. Thus, the requirement of 16 probe clusters does not imply that a druggable site has to bind all 16 different probes, although very frequently this is the case. FTMap analysis of the targets considered herein indicates that the requirement for a hot spot of at least 16 probe clusters is the most important condition for druggability: without it, high-affinity binding is essentially impossible (Figure S1).

The second condition, requiring connectivity or compactness, is based on the observation that a druggable site should also include one or more secondary hot spots that are close enough to the primary hot spot to be reached by a drug-sized molecule.⁴⁸ For the current work, we attempted to establish a quantitative threshold for how close these additional hot spots must be to support the strong binding of a druglike ligand so that we could include this measure as part of an objective metric for site druggability. We therefore analyzed the separation between hot spots seen for a large set of druggable sites that bind druglike ligands with molecular weight not exceeding 500 g/mol, as shown in Figure S3. The results show that the mean distance between the centers of CCs in druggable sites (i.e., the points with the highest density of probe atoms within the consensus clusters) is between 5 and 6 Å. However, to avoid missing any druggable sites, we designated 8 Å as the maximum allowed distance between connected hot spots, which captures over 99% of the druggable

Table 1. Classification of Protein Druggability Based on FTMap Results

druggability class	strength (S) (no. of probe clusters)	center-to-center distance (CD), Å	maximum dimension (MD), Å	druggability subclass	notation
druggable	$S \geq 16$	$CD < 8$	$MD \geq 10$	druggable using druglike compounds	D
not druggable	$S < 13$	any CD	any MD	not druggable due to weak hot spots	N
not druggable	any S	any CD	$MD < 7$	not druggable due to small hot spot ensemble	
non-canonically druggable—large	$S \geq 16$	$CD \geq 8$	$MD \geq 10$	druggable only by large chemotype such as macrocycle or foldamer	D^*_L
non-canonically druggable—small	$S \geq 16$	$CD < 8$	$7 \leq MD < 10$	druggable only by peptide, macrocycle, or charged compound	D^*_S
borderline druggable	$13 \leq S < 16$	$CD < 8$	$MD \geq 10$	at most micromolar affinity using druglike compounds	B
borderline non-canonically druggable—large	$13 \leq S < 16$	$CD \geq 8$	$MD \geq 10$	micromolar affinity by large chemotype such as macrocycle or foldamer	B^*_L
borderline non-canonically druggable—small	$13 \leq S < 16$	$CD < 8$	$7 \leq MD < 10$	micromolar affinity by peptide, macrocycle, or charged compound	B^*_S

sites. In evaluating druggability, we used this criterion for hot spot proximity to characterize the ensemble of hot spots that constitute a particular binding site. Specifically, we first selected the primary hot spot in a given site (not necessarily CC1) to form the kernel of the ensemble. The ensemble was then expanded by iteratively adding all secondary hot spots within 8 Å from any hot spot already in the ensemble until no further expansion was possible. In this process, secondary hot spots with less than three probe clusters were ignored.

The third criterion involving the size of the hot spot region is based on the observation that a binding site can be too small to be easily druggable. It is well-known that very small binding sites, and the fragment-sized ligands that bind to them, tend to be promiscuous.^{56–60} Indeed, this fact is implicit in the observations that experimental fragment libraries containing only 1000–2000 compounds typically provide multiple hits when screened against essentially any druggable target and that a druggable site is characterized by the observation of a high fragment hit rate, indicating that each site can bind many different fragment structures.^{56,57} Thus, while a small site containing one or more strong CCs can be sufficient to enable strong ligand binding, for the binding to be selective, it appears to be also necessary that the site be sufficiently large and topologically complex that a ligand that is complementary to this site is unlikely to bind as strongly to unrelated sites.⁴² As the simplest way to capture the extent of the binding site, or at least that portion of the binding site that is relevant for the generation of binding energy with the ligand, we simply report the longest dimension of the CC ensemble, which we defined as the distance between the two most widely separated probe atoms across the ensemble of probe clusters. In the following analysis, we found that the maximum extent of the binding site should exceed 10 Å for a site to be considered druggable. Indeed, among the druggable targets that have known druglike ligands, we found only three violations of this rule (Figure S4).

On the basis of the three properties described above, we classify target sites on proteins into four broad classes of druggability (Table 1).

(1). *Druggable*. A binding site is categorized as druggable with respect to conventional, druglike (e.g., rule-of-five compliant)² compounds if it meets the following criteria: (a) it includes a strong primary hot spot with at least 16 probe clusters, denoted $S \geq 16$. (b) One or more secondary hot spots are connected to the primary hot spot with a center-to-center distance (CD) < 8 Å. (c) The maximum dimension (MD) of the connected hot spot ensemble is at least 10 Å, i.e., $MD \geq 10$ Å (Table 1).

(2). *Not Druggable*. If the primary hot spot is very weak ($S < 13$) or if the hot spot ensemble is very small ($MD < 7$ Å), then a site is not druggable irrespective of the compounds used.

(3). *Non-canonically Druggable—Large*. If a sufficient number of sufficiently strong hot spots are present ($S \geq 16$) but the hot spot ensemble is too sparse, requiring a center-to-center distance exceeding 8 Å to reach a secondary hot spot from the primary one, then small druglike compounds cannot bind with sufficient affinity. However, the evidence indicates that such sites generally can bind strongly to larger but still pharmaceutically relevant compounds such as macrocycles^{61,62} or peptide/peptidomimetic foldamers.^{63,64} As will be discussed, it is possible that a subset of hot spots can satisfy the conditions $CD < 8$ Å, $S \geq 16$, and $MD \geq 10$ Å but that there exist additional strong hot spots beyond the ensemble. In such cases, the target is druggable using druglike molecules, but there is the potential for binding affinity to be improved if the size of the compounds is increased to reach all hot spots, although doing so will likely require using compounds that are large and so lie outside of conventional druglike chemical space.

(4). *Non-canonically Druggable—Small*. Another class of sites is encountered when the hot spot ensemble is small but not very small ($7 \text{ } \text{\AA} \leq MD < 10$ Å). Such sites may be druggable using ligands with a charged group that is not fully desolvated upon binding. Alternatively, druggability can be achieved by large and thus not druglike compounds (peptides,⁶⁵ peptidomimetics,⁶⁴ or macrocycles⁶¹) that extend beyond the small hot spot ensemble to interact with large surface regions without any strong hot spot. As will be demonstrated, while such interactions are weak, they still contribute to the binding free energy.

It is important to note that druggability is not an on–off property. In particular, for categories (1), (3), and (4), a site where the primary hot spot is somewhat weaker than the threshold ($S \leq 13 < 16$), we use the term borderline druggable (Table 1). For such sites, ligands can be found that bind with micromolar affinity, but attempts to increase potency usually fail. More generally, the strengths and arrangement of hot spots provide information beyond the simple four categories above. As an example, we discuss here the analysis of druggability for the FK506 binding protein (FKBP), which binds 8-deethyl-8-[but-3-enyl]-ascomycin (FK506), a macrocyclic compound with a molecular weight (MW) of 804 g/mol (Figure 1A).⁶⁶ Later, we will also study ABT737, a well-known regulator of apoptosis by inhibiting the interaction between the Bcl-xL protein and BAK peptide (Figure 1B).⁶⁷ Figure 1C shows the results of mapping a ligand-free structure of FKBP (PDB ID

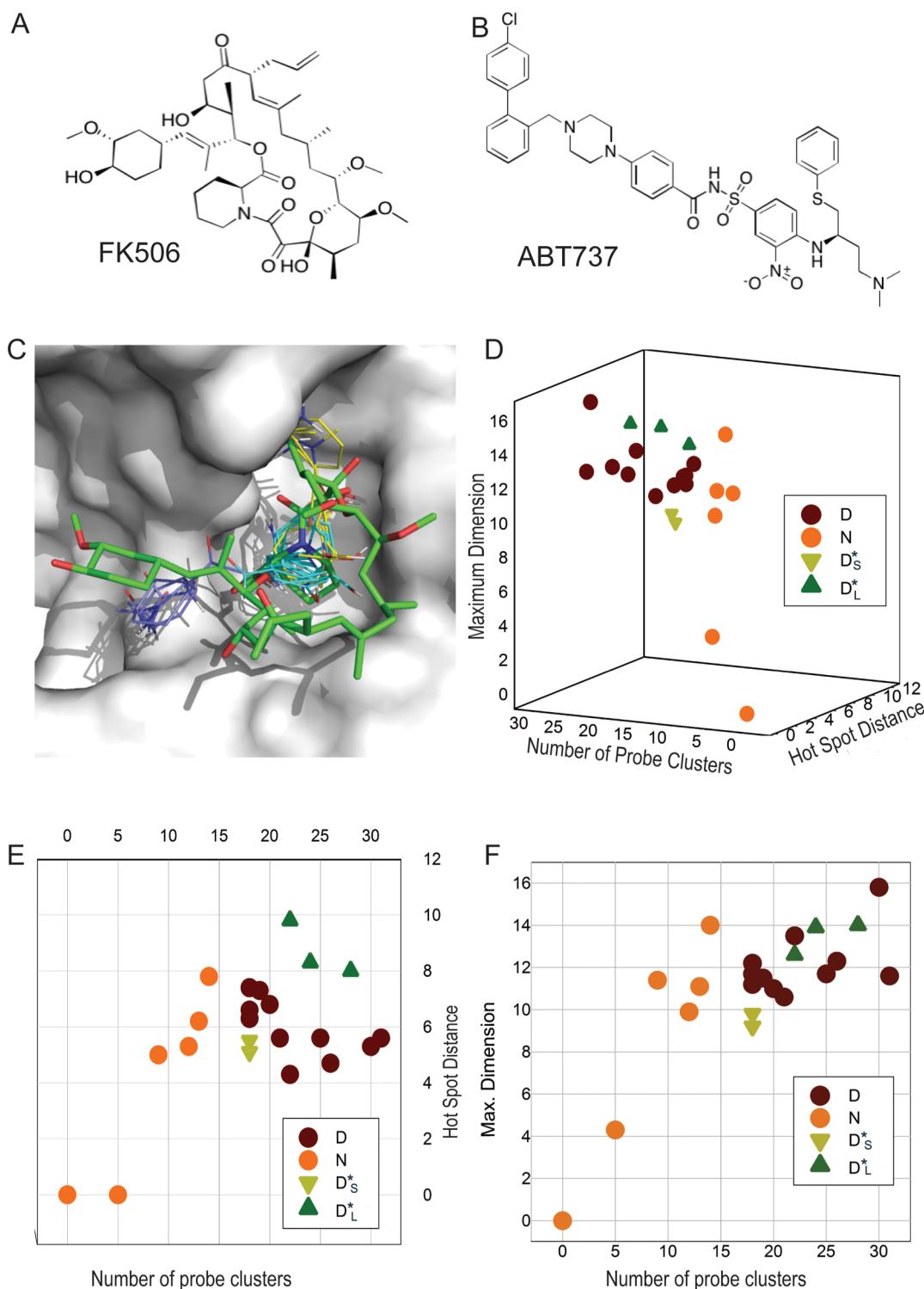


Figure 1. Using hot spot properties to assess druggability. (A) Structure of FK506, considered in the example of druggability analysis by FTMap. (B) Structure of ABT737, a well-known example of protein–protein interaction inhibitors. (C) Mapping the ligand-free FK506 binding protein (PDB ID 2ppn, shown as surface), with bound FK506 (green sticks, from the structure with PDB ID 1fkj) superimposed for reference. The consensus clusters shown are CC1 (cyan, 18 probe clusters), CC3 (yellow, 14 probe clusters), and CC6 (blue, 7 probe clusters). The center-to-center distance (CD) between CC1 and CC3 is 6.3 Å, and the maximum dimension of the ensemble of the two hot spots is 10.6 Å. Thus, the subset of these two hot spots is druggable on its own but would bind small ligands with only moderate potency. Using all three hot spots increases MD to 12.3 Å and allows for high-affinity binding. (D) Distribution of the parameters used for assessing druggability of the 23 protein targets in the Hajduk et al.⁷ data set. (E) Top view of the distribution shown in panel D to show the separation of different druggability classes. Note that the non-canonically druggable—large (D^*_L) category is defined by the large ($CD \geq 8$ Å) distance between the centers of the hot spots. (F) Right view of the distribution shown in panel D. The non-canonically druggable—small (D^*_S) category is defined by the small ($MD < 10$ Å) dimension of the entire hot spot ensemble.

Table 2. Druggability Results from NMR-Based Screening versus Those Predicted by FTMap

	target	site	structure ⁷	PDB		affinity (nM)	consensus sites	D ^a	druggability score (DS) ⁷		distances, Å	
				U	B				exp.	pred.	to further CCs	max. dim.
1	AK	Adenosine	Internal	1bx4	2i6b	<10 68	1bx4: 1(19), 8(4) 2a6b: 1(21), 6(6), 7(5)	D	−0.66 −0.42	−0.42 2a6b: 6.5, 6.8	1bx4: 7.3 2a6b: 11.5 11.7	1bx4: 11.5 2a6b: 11.7
2	Akt-PH	IP3	1h10	1unr	1h10	>1000 6300	1unr: 3(14), 4(10) 1h10: 5(9), 6(8)	B	−1.91 −1.98	−1.98 1h10: 7.0	1unr: 7.8 1h10: 11.0	1unr: 14.0 1h10: 11.0
3	Bcl-xL	Bak	Internal	1r2d	2yxj	<200 60	1r2d: 1(18), 2(17), 5(9) 2yxj: 1(28), 2(25), 4(9), 5(8)	D	−0.11 −0.64	1r2d: 7.4, 5.8, 2yxj: 6.3, 7.7, 7.0	1r2d: 12.2 2yxj: 16.5	
4	XIAP Bir3	Pept.	1g3f		2opz	28	2opz: 1(23), 7(5)	D	−1.03 −0.72	2opz: 3.1	2opz: 10.6	
5	CMPK	CMP	Internal	1cke	2cmk		1cke: 1(20), 3(10) 2cmk: 1(19), 4(10)	D	−1.13 −0.81	1cke: 5.3 2cmk: 4.7	1cke: 14.1 2cmk: 11.0	
6	E2-31	DNA	Internal	1dhm	1jj4	>1000 60 000	1dhm: 5(12), 6(7) 1jj4: 4(11)	N	−0.71 −0.72	1dhm: 5.4 1jj4: −	1dhm: 9.9 1jj4: 4.8	
7	ErmAM	SAH	1qam	1qam	1qan	>1000 1100	1qam: 1(25), 3(16) 1qan: 2(15), 3(12), 5(8)	D	−1.01 −0.87	1qam: 5.6 1qan: 5.3, 7.8	1qam: 11.7 1qan: 15.8	
8	FBP	DNA	1j4w	1fkj	2ppn	1j4w <10 2.3	1j4w: 1(18), 2(15) 2ppn: 1(18), 3(14), 6(7) 1fkj: 1(34), 2(16)	D ^{*s} D	−1.61 −0.03	−1.04 −0.24	1j4w: 5.1 2ppn: 6.3, 7.2 1fkj: 7.9	1j4w: 9.8 2ppn: 11.7 1fkj: 10.7
10	HI-0065	ADP	1fl9	1fl9	1htw	>1000 10 000	1fl9: 3(13), 5(8) 1htw: 2(19)	B	−0.82 −1.28	1fl9: 6.2 1htw: −	1fl9: 11.1 1htw: 7.9	
11	LCK	PTyr	1lkl	3hck	1lkl	300	3hck: 3(18), 5(9), 6(7)	D	−0.21 −1.07	3hck: 6.6, 7.6	3hck: 11.2	
12	LFA	IDAS	1rd4	1lfa	1rd4	<50 18.3	1lfa: 1(31), 4(12) 1rd4: 1(18), 2(13), 3(12), 7(6), 8(6)	D	−0.40 −0.35	1lfa: 5.6 1rd4: 5.7, 6.0, 4.7, 3.7	1lfa: 11.6 1rd4: 13.2	
13	MDM2	p53	1ycr	1z1m	1rv1	<100	1z1m: 1(21), 2(21) 1rv1: 1(21), 2(20), 3(19)	D	−0.49 −0.35	1z1m: 5.6 1rv1: 5.7, 4.7	1z1m: 10.6 1rv1: 14.8	
14	MurA	UDPNAG	1uae	1ejd	1uae	250	1ejd: 2(22), 4(10), 5(8) 1uae: 1(21), 2(12), 3(11), 4(8)	D	−1.38 −1.44	1ejd: 4.3, 0.8, 5.9, 7.2	1uae: 16.0	
15	MurI	Glu	Internal	1b73	1b74	<20 1b74: 1(20), 3(14), 4(12), 6(9)	1b73: 1(30), 2(14), 3(12), 4(11), 5(9), 6(8), 7(7) 1b74: 1(20), 3(14), 4(12), 6(9)	D	−1.93 −2.00	1b73: 5.3, 7.5, 5.3, 7.1, 3.0, 4.9 1b74: 4.0, 7.6, 3.8,	1b73: 15.8 1b74: 14.4	
16	PAK4	ATP	Inter.	1i8h	1i6c	1i8g 1i8h	<200	1i6c: 1(28), 2(18), 4(13)	D ^{*L}	−0.78 −0.94	−0.63 −1.49	1i6c: 8.0, 4.6 1i8h: 6.3, 5.6
17	Pin1 WW domain	Pept.		1ph0	3a5j	1pty	<10	3a5j: 1(20), 2(16)	D	−0.68 −0.68	−1.15 3a5j: 6.8	3a5j: 11.0
18	PDZ-PSD95	Pept.	1iu0	1iu2	1rgr	<300	1iu2: 1(22), 5(11) 1rgr: 1(22), 6(7)	D ^{*L}	−2.00 −1.99	1iu2: 9.8 1rgr: 5.9	1iu2: 12.6 1rgr: 8.7	
19	PTP1B	Cat. pY	1ph0	3a5j	1pty	>1000	3a5j: 4(9), 7(7), 8(7)	D	−0.68 −1.77	−1.15 3a5j: 5.0, 6.2	3a5j: 11.4	

Table 2. continued

	target	site	structure ⁷	PDB			affinity (nM)	consensus sites	D ^a	druggability score (DS) ⁷		distances, Å	
				U	B	exp.				pred.	to further CCs	max. dim.	
								1pty: 8(7), 9(7)			1pty: 5.9	1pty: 10.2	
20	SARS N-term	RNA	Internal	1ssk		>1000 <i>~ 10⁶</i>	none		N	-1.93	-1.92		
21	SCD	Subst.	1g4k	1cqz	1g4k	<10	1cqz: 1(26), 5(7)		D	-0.09	-0.55	1cqz: 4.7	1cqz: 12.3
							1g4k: 1(20), 2(17), 5(9), 7(4)				1g4k: 4.6, 4.8, 3.2	1g4k: 13.0	
22	Survivin	Bir3	1e31	1e31	3uih	>1000	1e31: 7(5) 3uih: 8(5)		N	-1.97	-1.99	1e31: – 3uih: –	1e31: 4.3 3uih: 4.1
22	Survivin	other	1e31	1e31		<300	1(24), 2(24)		D [*] _L	-0.45	-0.35	1e31: 8.3	1e31: 13.9
23	UK	Pept.	1fv9	2o8t	1fv9	<50	2o8t: 1(29), 2(26) 1fv9: 1(18), 2(15), 5(9)		D	-0.40	-0.81	2o8t: 6.0 1fv9: 4.3, 2.1	2o8t: 17.8 1fv9: 13.0

^aDruggability class based on FTMap analysis, as described in Table 1.

2ppn). As mentioned, each hot spot is identified by a consensus cluster (CC). In all of our figures, each probe cluster contained in a CC is represented by the single lowest energy probe structure to avoid the overcrowding of the figure. All probes in a CC have the same color, using the color code as follows: CC1, cyan; CC2, magenta; CC3, yellow; CC4, salmon; CC5, white; CC6, blue; CC7, orange; and CC8, green. Figure 1C also shows the location of the bound FKS06 (taken from the structure 1fkj) superimposed for reference. The ligand binding site contains three connected hot spots. These were the top-ranked hot spot, CC1, which contained 18 probe clusters, plus two secondary CCs, CC3 (14 probe clusters) and CC6 (7 probe clusters). The center-to-center distance (CD) between CC1 and CC3 is 6.3 Å, and the maximum dimension of the ensemble with the two hot spots is 10.6 Å. Thus, according to our conditions, the site would be druggable when using a drug that binds only to these two hot spots. However, since CC1 is not much stronger than the required S = 16, and the MD value is close to the required threshold of 10 Å, the affinity would be moderate. Since the CD distance between CC1 and CC6 is 7.2 Å, CC6 can be also added to the hot spot example used for the binding. Adding the third hot spot CC6 (7 probe clusters) increases MD to 12.3 Å. According to Table 1, all three hot spots in this extended ensemble still can be reached using druglike compounds. In fact, FKBP has been cocrystallized with smaller but still potent inhibitors. Examples are PDB ID 1fkh (MW = 453, K_i = 7 nM)⁶⁸ and PDB ID 1f40 (MW = 360.5, K_i = 7.5 nM).⁶⁹ However, the inhibitors that reach the three hot spots are fairly long, and connecting the two far ends by forming a macrocycle in FKS06 improves the affinity to K_i = 0.2 nM.⁶⁶ This example also demonstrates that, while the analysis of druggability could be simplified by integrating the three different requirements into a single druggability measure, considering the three requirements separately can more explicitly show problems that may lead to the lack of druggability, as well as explore opportunities for using non-canonical compounds as drugs.

If several structures of a target are available, then an important question is which structure to use for the analysis. Frequently, druggability needs to be established before discovering any ligand, motivating preference for unbound

structures. If multiple unbound structures are available, then it is useful to map all and to consider the one that results in the highest level of druggability, assuming that ligands can be optimized to take advantage of the available hot spots. While ligand binding may affect their location and relative strength,^{70,71} hot spots are generally less sensitive to conformational changes than binding sites are,⁴⁸ and the analysis of the ligand-free structures provides valid druggability assessment for most targets. However, the binding sites in the unliganded structures of some proteins may be completely closed, or ligand binding may create new hot spots; hence, it is generally advisable to also map ligand-bound structures when they become available in the process of drug discovery. This additional information frequently shows if there is a choice between a large and hence nondruglike ligand that reaches all major hot spots or a smaller and less potent druglike ligand that can take advantage of only the main hot spot and some of the weaker hot spots around it.

Druggability Results from NMR-Based Screening versus Predicted by FTMap. Table 2 shows the analysis of the 23 proteins in the test set from Hajduk et al.⁷ using the three FTMap criteria. In two cases, PTP1B (T19 in Table 2) and survivin (T22), the protein was found to contain two distinct binding sites that were characterized separately, resulting in a total of 25 binding sites. Druggability results are also shown graphically in Figure 1D–F. Table 2 lists the short names of the targets and specifies their binding sites, as given by Hajduk et al.⁷ While we quote the Protein Data Bank (PDB) entry as given in the original publication,⁷ for our analysis, we mapped both the highest resolution ligand-free and -bound structures currently available in the PDB.⁵⁵ As discussed, in most cases, druggability can be determined by mapping the ligand-free structure alone, and the ligand-bound structure is needed only to judge whether the hot spots are located in the appropriate binding site for the purpose of retrospective validation with respect to experimentally identified ligands. For the protein PAK4 (T16 in Table 2), no structure was found in the PDB; hence, the druggability of this target is not discussed. The affinities shown in Table 2 are the ones given by Hajduk et al.,⁷ but we also list (in italics) if more potent compounds have subsequently been reported. Table S1

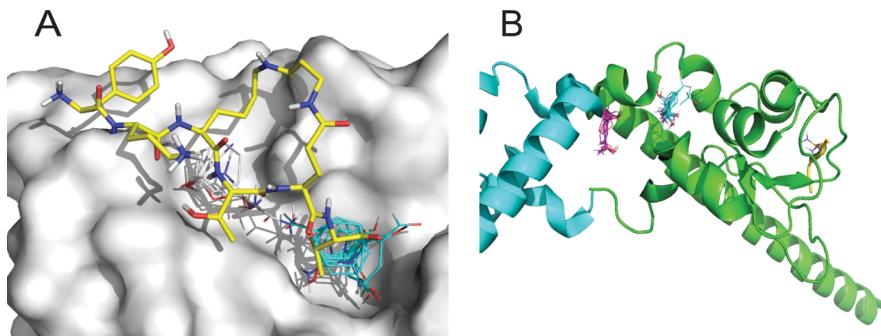


Figure 2. Hot spots in two protein–protein interaction targets, belonging to the druggability class D^*_L . (A) Mapping of the ligand free PDZ-1 domain of PSD-95 (PDB ID 1iu2, shown as surface), with a cyclic peptide (from PDB ID 1rgr, shown as yellow stick model) superimposed for reference. The two hot spots in the interface, CC1 (cyan, 22 probe clusters) and CCS (white, 11 probe clusters), are 9.8 Å from each other. (B) Mapping of the ligand-free survivin dimer (PDB ID 1e31, shown as cyan and green cartoon), with the Smac/DIABLO peptide (from PDB ID 3uih, shown as yellow cartoon) superimposed for reference. The only hot spot in the survivin/Smac interface is CC7 (orange, 5 probe clusters); thus, the peptide binding site is not druggable. A novel site is defined by the strong hot spots CC1 (cyan, 24 probe clusters) and CC2 (magenta, also 24 probe clusters) at the dimerization interface. Since the center-to-center distance between CC1 and CC2 is 8.3 Å, ligands that bind to both are over the 500 g/mol MW limit.

is a copy of Table 2, but its footnotes provide detailed information on targets and their affinities, with references. Additional references may not be listed if the value is given in PDBbind,⁷² BindingDB,⁷³ or BindingMOAD.⁷⁴ Table 2 also shows the experimental and predicted druggability scores (DS) calculated by Hajduk et al.,⁷ which represent the logarithm of the actual or predicted hit rate in an NMR-based fragment screen. These authors observed that most of the targets for which high-affinity, druglike leads could be identified had DS values greater than -1.0 , whereas DS values below -1.5 indicated proteins without such ligands. Consequently, a protein pocket with DS > -1.0 was considered to have high druggability, whereas low druggability was assigned for targets with DS < -1.5 . The DS value may be close to -2.0 for very undruggable proteins. Protein pockets with intermediate values ($-1.5 < \text{DS} < -1.0$) were assigned as having moderate druggability.

According to the FTMap results, 14 of the 25 distinct binding sites in this test set satisfy all three criteria for druggability: i.e., the top-ranking CC in the cluster contains at least 16 probe clusters, the ensemble contains at least two CCs with a center-to-center distance of less than 8 Å, and the maximum dimension of the ensemble is at least 10 Å. These 14 sites are therefore classified as druggable in our analysis, indicating a strong likelihood of finding a druglike small molecule ligand that binds strongly to the site. Such ligands are known for 12 of these 14 targets (see Table S1). The exceptions are cytidine monophosphate kinase (CMPK, T5), with only micromolar ligands identified so far,⁷⁵ and rRNA methyltransferase ErmC (ErmAM, T7), for which the best known inhibitor has $\text{IC}_{50} \sim 1100$ nM.⁷⁶ Two targets, Akt-PH (T2) and HI-0065 (T10), are predicted as being borderline druggable and have micromolar ligands.^{77,78} Four sites (E2-31, the noncatalytic phosphotyrosyl binding site of PTP1B, the N-terminal domain of SARS, and the Bir3 domain of survivin) are classified as not druggable, in each case because of a weak main hot spot ($S < 13$). According to FTMap, three targets (the WW domain of Pin1, PDZ-PSD95, and the “other” site of survivin) belong to the D^*_L category; thus, achieving substantial potency requires peptide-like or macrocyclic ligands. PZD-PSD95 (T18) was considered not to be druggable by Hajduk et al.⁷ (DS = -2.00), but FTMap identifies two hot spots in the site

that are 9.8 Å apart (Figure 2A); hence, based on the FTMap results, this site can be targeted by peptides or macrocycles that are large enough to reach the two hot spots.⁷⁹ In fact, PZD-PSD95 binds cyclic peptides with $\text{IC}_{50} \sim 1000$ nM in a solvent-exposed shallow pocket.⁷⁹ The two main hot spots are also too far apart in the peptide binding site of Pin1 WW domain (T17). We note that, based on the PDB code (1ih8) given by Hajduk et al.,⁷ we consider here the WW domain of Pin1 that has only peptides as known ligands⁸⁰ rather than the peptidyl prolyl cis/trans isomerase (PPIase) domain that is the target of more recent studies.^{81–83} Survivin (T22 in Table 2) has two binding sites of interest. As shown in Figure 2B, in contrast to XIAP Bir3, the Smac peptide binding site of this Bir3 domain includes only CC7 with 5 probe clusters and hence is not druggable. However, the “other” site in Table 2, close to the dimer interface, includes two strong hot spots that are 8.3 Å apart, almost within the reach of druglike ligands. Inhibitors that binds to both hot spots have K_d values as low as 37 nM.⁸⁴ Finally, based on the FTMap results, one target, FUSE-binding protein (FBP, T8 in Table 2), is in the D^*_S category due to the small size of the hot spot ensemble. In fact, the only known inhibitors of DNA binding to FBP are charged and of high micromolar affinity.⁸⁵

The FTMap results generally agree well with the assessment based on experimental and predicted druggability scores given by Hajduk et al.⁷ as well as with the actual success or failure to identify a high-affinity druglike ligand that binds to the site in question. We have substantive differences with the DS value from Hajduk et al. only for 3 of the 25 sites. Moreover, the available experimental data suggest that, in these three cases, the druggability assignment based on FTMap is, in fact, the correct one. Specifically, the DNA binding site of the human papillomavirus (HPV) E2-31 protein (Target T6 in Table 1) is considered to be druggable by Hajduk et al.⁷ but not by FTMap. The best inhibitor of the HPV E2 DNA binding site in the literature has an IC_{50} value of approximately 10 μM,⁸⁶ and due to the consistent lack of success in finding more potent ligands for this site, efforts have been largely redirected to target the DNA rather than the protein.^{87,88} Moreover, the majority of DNA binding sites of proteins have limited druggability.⁶ FTMap also differs from the DS method in predicting that the adenosine diphosphate (ADP) binding site of the hypothetical

bacterial ATPase HI-0065, also known as YjeE (T10 in Table 1), is at most borderline druggable. HI-0065 was targeted by the Abbott group,⁷⁸ but the best compound identified by NMR-based screening and optimization had a K_d value of about 5 μM . We note that both HPV E2-31 and HI-0065 have strong binding hot spots close to, but not overlapping with, the ligand binding site, which may have resulted in a misleadingly high NMR fragment hit rates. The third deviation between FTMap and the earlier results occurs for the glutamate racemase MurI that was rated as not druggable by Hajduk et al.,⁷ although they reported that 4-substituted D-glutamic acid analogues (not considered druglike) bind with nanomolar affinity.⁸⁹ It is not clear why the NMR-based screening did not result in hits. According to FTMap, this target is druggable, emphasizing that our condition does not directly account for the lack of druglikeness due to the need for a charged ligand, unless the hot spot region is too small. In this particular case, we note that a close homologue, glutamate racemase from *Helicobacter pylori*, has potent and druglike noncompetitive inhibitors that bind to the corresponding site.^{90,91}

Comparison of FTMap results obtained using the ligand-bound structure rather than the unbound structure of the protein (Table 2) shows that small differences between the structures can affect the strength and location of hot spots but that such changes impact druggability assessment only for three targets, HI-0065, LCK, and PTP1B, all having strongly charged ligands with a phosphate group (see footnotes to Table S1). It appears that the binding of such ligands tightens the binding pocket, which then binds fewer probe clusters. Thus, the mapping of ligand-free proteins generally provides valid druggability information, emphasizing that druggability is an inherent property of a protein. Nevertheless, as will be further discussed, it may be useful to also map ligand-bound protein structures if available, since, in some cases, ligand binding may open additional subsites and create additional hot spots, favoring druggability.

Validation of FTMap Druggability Assessment against Extended 72 Protein Set of Hajduk et al.⁷ In addition to analyzing the 23 drug targets for which experimental NMR fragment screening results were available, Hajduk et al.⁷ calculated druggability indices for a set of 72 proteins based on cocrystal structures with ligands. They did this by constructing a regression equation to calculate a druggability score based on geometric features of the binding site, including the total and apolar surface area, compactness, and the first and third principal moments of the pocket, which they parametrized using the results of NMR fragment screening of the original 23 protein test set. The 72 protein set includes 35 proteins for which high-affinity, druglike molecules have been reported plus 37 proteins for which the known ligands were either not high affinity or not druglike. Table S2 compares the results of FTMap analysis of this extended 72 protein set to the druggability scores (DS) reported by Hajduk et al.⁷

For each of these 72 targets, we mapped the highest resolution ligand-free structure available in the PDB. Bound structures are listed only to specify the binding site of interest. Among the 35 proteins that were considered by Hajduk et al. to have druggability, the FTMap results disagree with the high DS value only for two targets, cytochrome c peroxidase (Target T17) and leghemoglobin (Target T32). Target T17, given by PDB ID 1ccg,⁷ is actually the His175Gly mutant of cytochrome c peroxidase, which binds imidazole in a cavity created by the removal of His175 side chain, but it does so with a K_d value of

only 2.7 mM.⁹² FTMap finds this imidazole binding site to be not druggable. Although a number of strong hot spots exist in the large cavity of the protein, these do not overlap with the bound imidazole position. The problem with leghemoglobin (T32) is that the binding site is formed in the process of ligand binding; hence, only the bound structure, which was studied by Hajduk et al.,⁷ reveals the druggability of this target. Since we generally consider the mapping of the ligand-free structure to be more predictive of druggability, this may appear to be a shortcoming of our approach. However, leghemoglobin is not a drug target, and its binding pocket is extremely small, accommodating just a nicotinate with molecular weight 122 g/mol.⁹³ As we have discussed, such small ligands are unlikely to be drugs, and, in fact, the binding of nicotinate is weak ($K_d = 1400 \text{ nM}$).⁹³

Among the remaining 37 structures that did not have high-affinity druglike ligands, the agreement between FTMap and DS druggability predictions was also generally good; however, we again see some differences. Phosphoserine phosphatase (T36) is rated as being druggable by Hajduk et al.⁷ (DS = -0.808), but it is only borderline druggable by FTMap. No potent inhibitor is indicated by Hajduk et al.,⁷ and such a compound has not been reported in the subsequent literature. In the apo structure of flavodoxin (T37, DS = -0.857), the side chains of W57 and Y94 completely close down the flavin mononucleotide binding site;⁹⁴ hence, FTMap does not find a hot spot there. As mentioned, proteins for which the druggability of the bound conformation is so greatly different from that of the apoprotein are very rare; we have found only two cases (leghemoglobin and flavodoxin) among the more than 150 proteins we have studied, and these two are not drug targets. FTMap confirms druggability for most other targets with DS > -1.2 (Targets T38 through T45), and, apart from inositol monophosphatase (T44), these proteins indeed have known potent ligands, as does sialidase (T46), with DS = -1.201 (Table S2). Targets PI phospholipase (T47) and enterotoxin (T48), with lower DS values, we rate as borderline druggable, and phosphofructokinase (T49), as not druggable. Consistent with these ratings, none of these proteins has any known potent ligand. The druggability of targets with DS values between -1.334 and -1.5 is somewhat undefined by Hajduk et al.⁷ FTMap shows substantial variation in this range, and for several proteins, our results are supported by binding data. In particular, protein kinase C δ (PKC δ , T50 in Table S2) is in the D*_S category, and it binds phorbol-12,13-dibutyrate, a nondruglike compound with MW = 504.6 g/mol, with nanomolar K_d .⁹⁵ FTMap confirms the druggability of the GTP binding site of Ras p21 (T52 in Table S2). Guanylate kinase (T53) is not druggable, whereas based on FTMap, chloramphenicol acetyltransferase (T54) is in the D*_S category and indeed is known to bind fusidic acid with $K_i = 5400 \text{ nM}$.⁹⁶ FTMap predicts seven targets to be druggable (aldehyde ferredoxin oxidoreductase, galactose binding protein, cytidine deaminase, trichosanthin, achromobacter protease, diaminopimelic acid dehydrogenase, and glutamine methyltransferase) that were assigned a low druggability index (DS < -1.5) by Hajduk et al.⁷ In three cases, the FTMap assignment is supported by known affinity data. These three are the galactose binding protein (T61), which binds galactose with $K_d = 400 \text{ nM}$,⁹⁷ cytidine deaminase (T62), which has potent inhibitors,⁹⁸ and achromobacter protease (T64), which has inhibitors with $K_i = 6.5 \text{ nM}$.⁹⁹ We could not find data confirming druggability of the remaining four targets, aldehyde ferredoxin oxidore-

eductase (T60), trichosanthin (T63), diaminopimelic acid dehydrogenase (T66), and glutamine methyltransferase (T71), but in three cases, T60, T63, and T66, the proteins are known to bind (nondruglike) small substrates or cofactors with relatively high affinity (see Table S2). In the same range (i.e., with DS < -1.5), we predict four targets in the D*_S category. Two of these, periplasmic glucose/galactose receptor (T69) and IP phosphatase (T70), are known to bind charged ligands with nanomolar affinity (Table S2). We have no information on the other two, deoxynucleoside monophosphate kinase (T56) and glucoamylase (T58). Finally, Table S2 has one target, ^{PP60}Src SH2 domain (T57), with DS < -1.5, that FTMap places in the D*_L category. This domain is known to bind a peptidomimetic ligand with IC₅₀ = 1800 nM.¹⁰⁰

As shown, we were able to find binding data for a substantial fraction of targets that were predicted to be druggable by FTMap but that had DS < -1.5 assigned by Hajduk et al.⁷ This poses the question of why such low DS values were predicted. The most likely reason is that the regression formula for DS is based on the properties of pockets rather than of the hot spots, and this may underestimate potential binding capacity for pockets that have a somewhat unusual shape. Indeed, targets T55, T60, and T61 all have completely buried ligand binding sites with no solvent accessibility in the ligand-bound structures considered by Hajduk et al.,⁷ which likely contributed to their low DS values.⁷ However, FTMap can still place probes in such internal cavities, and the proteins are likely to have alternative conformations with more open pockets. Trichosanthin (T63) and glutamine methyltransferase (T71) have deep and narrow binding pockets, whereas the binding site in diaminopimelic acid dehydrogenase (T66) is largely solvent-exposed, and it seems that the simple regression formula by Hajduk et al.⁷ underestimates the druggability for such unusually shaped pockets.

Validation of FTMap Druggability Assessment against the Test Set of Cheng et al.⁸ A biophysical model developed by Cheng et al.⁸ accounts for lipophilic surface area and curvature of binding sites to compute a score that corresponds to the maximal affinity predicted for a passively absorbed oral drug (MAP_{POD}). MAP_{POD} is based on the hypothesis that the maximal achievable binding free energy is largely due to desolvation, and it models the desolvation contributions from target and ligand in terms of nonpolar surface areas with coefficients that depend on surface curvature. The expression does not have an explicit polar desolvation term, since it is assumed that the electrostatic interactions and the desolvation energies largely cancel each other.⁸ This model was used to assess the druggability of 27 binding sites. The value of this test set is that all sites are of pharmaceutical interest, in contrast to later test sets that included many proteins without this property. The 27 target sites were classified as druggable (17), difficult (6), or undruggable (4). Nearly all targets classified as druggable or difficult had associated orally bioavailable, marketed drugs, whereas the undruggable targets did not, despite considerable efforts within the industry.

In Table S3, we compare the FTMap results to the maximal affinity (MAP_{POD}) values predicted by Cheng et al.⁸ that were recently reviewed.¹⁰¹ Both ligand-free and -bound structures were mapped, but, as mentioned, we consider the unbound results to be more relevant for assessing druggability. The agreement with the results of Cheng et al.⁸ is generally good. For the 23 druggable or difficult targets, we have different results only for the binding site of COX-2 (Target T6 in Table

S3), which FTMap places into the D*_S category due to the small size of the hot spot ensemble, predicting that potent ligands binding to this site must have a charged or highly polar group not fully desolvated upon binding. Indeed, all strong COX-2 inhibitors seem to have either a carboxyl or a terminal sulfonamide group.¹⁰² Cheng et al.⁸ examined six targets that they considered to be difficult, namely, ACE-1 (T1), HIV RT nucleotide site (T13), IMPDH (T16), neuraminidase (T18), penicillin binding protein (T22), and thrombin (T23), because they required highly polar molecules that were either processed from prodrugs or relied on active transport mechanisms. Thrombin (T23) is predicted to be druggable both by Cheng et al.⁸ and by FTMap and now has direct (i.e., not prodrug) oral inhibitors.^{103,104} With 15 probe clusters, the penicillin binding protein (T22) is borderline druggable by FTMap, and ligand binding substantially strengthens the hot spot. This protein binds piperacillin,¹⁰⁵ an approved but not orally bioavailable drug (MW= 517.5 g/mol), with an IC₅₀ of 166 nM.¹⁰⁶ FTMap predicts that ACE-1 (T1), HIV RT nucleotide site (T13), IMPDH (T16), and neuraminidase (T18) are all druggable, although the primary hot spot in IMPDH is relatively weak, with only 17 probe clusters, and ACE-1 and neuraminidase have strong but small hot spots (MD < 11.0). Both ACE-1 and neuraminidase have approved drugs that include a carboxyl group but are not prodrugs, for example, N2-[(1S)-1-carboxy-3-phenylpropyl]-L-lysyl-L-proline (lisinopril)¹⁰⁷ for ACE-1 and (2R,3R,4S)-3-acetamido-4-(diaminomethylideneamino)-2-[(1R,2R)-1,2,3-trihydroxypropyl]-3,4-dihydro-2H-pyran-6-carboxylic acid (zanamivir)¹⁰⁸ for influenza neuraminidase. These results confirm that ACE-1 and neuraminidase are correctly classified as druggable, but they also show that a small hot spot is likely to require a charged ligand. Two targets, Factor Xa (T10) and HMG CoA reductase (T15), were placed on the border of the druggable region by Cheng et al.¹³ Factor Xa is categorized by FTMap as D*_L, and it has druglike inhibitors close to the upper limit on size. For example, the inhibitor methyl(2R,3R)-3-((3'-(aminomethyl)biphenyl-4-yl]carbonyl)-amino)-2-(3-carbamimidoylbenzyl)butanoate (RPR128515),¹⁰⁹ bound in the structure 1ezq, has a molecular weight of 458.6 g/mol.¹⁰⁹ Both Cheng et al.⁸ and FTMap underestimate the druggability of HMG CoA reductase, which has potent statin inhibitors,¹¹⁰ but is considered to be only borderline druggable by FTMap (Table S3). This is due to the fact that statins occupy only a portion of the large ligand binding site. While this is enough to block access of the substrate to the active site of the enzyme, the unoccupied region also includes a number of hot spots, thus resulting in fewer probe clusters overlapping with the inhibitor.¹¹⁰

Cheng et al.⁸ classify four targets as being not druggable: cathepsin K (T24), Caspase 1 (T25), the active site of HIV integrase (T26), and PTP1B. Cathepsin K (T24) and the catalytic site of the HIV integrase are predicted to be druggable by FTMap. This finding is supported by data that became available after the publication of the analysis by Cheng et al.⁸ Cathepsin K has acrylic cyanamide-based inhibitors with *K*_i values as low as 0.003 nM.¹¹¹ The osteoporosis drug *N*-(1-cyanocyclopropyl)-4-fluoro-N2-[(1S)-2,2,2-trifluoro-1-[4'-(methylsulfonyl)biphenyl-4-yl]ethyl]-L-leucinamide (Odanacatib),¹¹² in Phase III clinical trial by Merck, is a potent cathepsin K inhibitor, and current studies focus on safety issues rather than affinity or oral bioavailability.¹¹³ For HIV integrase, there exist several FDA approved orally bioavailable inhibitors, such as *N*-(4-fluorobenzyl)-5-hydroxy-1-methyl-2-(2-[(5-methyl-1,3,4-

Table 3. Druggability of Proteins Binding Macroyclic Compounds

protein name	PDB	U		B		hot spots	D ^a	comment
		U	B	U	B			
1 Cyclophilin A	2cpl	1cw3, 1sfa	1(23), 3(15), 5(6)	D	CC1 and CC5 are overlapping, CC5–CC3 = 5.8 Å, MD = 14.0 Å, druggable by druglike compounds			
2 FKBP	2ppn	2tke, 2ag3	1(18), 3(14), 6(7), 7(6)	D	CC1, CC3, and CC7 are close and make it already druggable by druglike compounds (MD = 10.5 Å); CCl–CC6 = 7.2 Å			
3 Pancreatic elastase	1lvx	1okx	2(19), 5(8), 6(6), 7(5)	D	CC2 is elongated, MD = 9.5 Å (borderline), but CC2–CC7 = 6 Å, CC6–CC4= 5.3 Å, CC2–CC6 = 7.2 Å, thus all are connected			
4 Elongation factor TU	1efc	1dt8	1(21), 4(12), 6(5)	D ^{*L}	CC1 and CC6 are close but too small (9.0 Å); CC1–CC4 = 11.5 Å, too far. Druggable only using a long ligand (peptide) or a macrocycle			
5 Protein phosphatase 1	3egg	3e7a	2(16), 6(7)	D ^{*L}	CC2–CC6 = 12.9 Å. Too far, druggable only using either a macrocycle or a very elongated ligand			
6 Type I signal peptidase	1kn9	1r7d	2(18)	D ^{*S}	CC2 is elongated, but MD = 10.0 Å. Small and weak hot spot			
7 Chitinase	1gav	1waw, 1wb0	1(23), 4(10), 5(6), 6(6)	D	Overlapping hot spots, total MD = 12.6 Å. Also druggable using druglike ligands			
8 Actin	1nwk	1qz5	3(14)	N	Inherently not druggable, additional surface area is needed			
		2q0u	2(16), 7(4)		CC2–CC7 = 8.0 Å, too weak hot spots, at most borderline			
		2q0r	10(2)		Inherently not druggable, additional surface area is needed			
9 AcetylCoA Carboxylase	3glk	3gid	3(15), 4(9), 5(8), 6(6)	B ^{*L}	CC3 and CC6 are close, but MD = 9.0 Å only. CC5 is close to CC4, but the distance between CC3 and CC5 is 9.0 Å, too large. Druggable only using a macrocycle or very long compound			
10 HSP90 (closed)	1yer	1yet, 3imw	1(23), 3(10), 5(8), 6(8)	D	CC1, CC3, and CC5 provide druggability, MD = 14.5 Å; CC1–CC6 is 8.9 Å, too far; macrocycle is useful, but not necessary			
HSP90 (open)	1yes	1yet, 3imw	1(22), 2(13), 3(12), 5(9)	D	CC1, CC2, and CC5 provide druggability, MD = 11.4 Å; CC2–CC3 = 7.67 Å, also part of the hot spot ensemble			

^aDruggability class based on FTMap analysis, as described in Table 1.

oxadiazol-2-yl)carbonyl]amino]-2-propenyl)-6-oxo-1,6-dihydro-4-pyrimidinecarboxamide (raltegravir),¹¹⁴ with low nanomolar IC₅₀ values. These newly druggable sites emphasize the importance of assessing druggability using biophysical rather than purely statistical methods. However, we agree with Cheng et al.⁸ that the other two targets are very difficult. FTMap results categorize Caspase 1 as borderline druggable at best due to the weak hot spots and the small size of the hot spot region. Indeed, although there are peptidomimetic inhibitors containing a carboxyl group with K_i values as low as 5 nM, their oral bioavailability is limited,¹¹⁵ and based on our mapping results, it is unlikely that more druglike inhibitors will be found. The druggability of PTP1B will be discussed in detail later in this perspective and hence is not considered here.

Validation of FTMap Druggability Assessment against DrugPred. Krasowski et al.¹² compiled a nonredundant data set containing crystal structures of 71 druggable and 44 less druggable proteins. This data set was used to train a structure-based druggability predictor called DrugPred. We did not apply FTMap to this entire set because, in contrast to the set by Cheng et al.,⁸ it includes many proteins that have not been targets of substantive screening campaigns. However, we mapped the targets for which DrugPred clearly failed. These included six targets (DPP-IV, urokinase plasminogen activator, factor Xa, human thymidine phosphorylase, thiamine pyrophosphokinase, and thymidine kinase) that were included as druggable in either training or validation sets by Krasowski et al.¹² but were predicted to be undruggable by DrugPred and two proteins (peptide deformylase and CDP-D-glucose synthase) that were included as undruggable but were predicted to be druggable (see Table S4). These analyses were performed to obtain more information on the likely druggability of these targets and to explore whether the FTMap results differ from those of DrugPred. Of the six targets that were considered to have false negative predictions,¹² FTMap correctly predicts two (DPP-IV and urokinase plasminogen activator) to be druggable. As already shown, factor Xa is categorized as D*_L, but it still binds druglike ligands, albeit at the upper limit of molecular weight.¹⁰⁹ Human thymidine phosphorylase is classified as B*_S, and it is known to require charged ligands or prodrugs for nanomolar affinity.¹¹⁶ The two remaining false negatives (thiamine pyrophosphokinase and thymidine kinase) are predicted not to be druggable also by FTMap. While we did not find any evidence that thiamine pyrophosphokinase has potent inhibitors, the thymidine kinase of the herpes simplex virus has inhibitors with K_i values as low as 5 nM.¹¹⁷ According to FTMap, the thymidine kinase binding site has only weak hot spots both in apo and holo structures (Table S4), making this target the sole real failure of our method. The apparent reason is that the thymidine kinase inhibitors are planar molecules that bind in an extremely narrow and fully buried pocket that can accommodate only the few probes that are also planar. This problem is very rare, and at this point, we have no solution to avoid it. The last two targets in Table S4 were considered to be false positives by Krasowski et al.¹² and were also predicted to be druggable by FTMap. Of these, peptide deformylase now has low nanomolar inhibitors.¹¹⁸ CDP-D-glucose synthase binds cytidine-5'-triphosphate with $K_i = 35 \mu\text{M}$,¹¹⁹ but it has not been considered to be a drug target and hence its druggability has not been established experimentally.

■ TARGETTING CONVENTIONALLY UNDRUGGABLE TARGETS WITH NON-CANONICAL DRUG CHEMOTYPES

Because the FTMap algorithm evaluates the binding potential of protein surface sites based on structural and physicochemical principles, rather than by empirical parametrization using a particular test set of target proteins, the method has the potential to provide information on the druggability of sites that differ in structure or properties from those that previously have been successfully targeted in drug discovery efforts. We therefore evaluated whether the FTMap method can shed light on precisely why sites that were not assessed to be druggable failed to qualify. This information is potentially useful in deciding whether the site might be targeted by a non-canonical drug chemotype, such as a peptide, a macrocycle, or a covalent inhibitor, or whether the target is likely to be intractable to all approaches currently available for small molecule drug discovery. Of the targets in Table 2 that were rated as not druggable, four of them, E2-31, the noncatalytic site of PTP1B, the SARS N-terminal domain, and the bir3 site of survivin, simply lack a hot spot of sufficient strength to serve as a strong anchor site for ligand binding. Such targets are unlikely to bind strongly to any small ligand, so we consider them likely to be undruggable by any such means. However, PDZ-PSD95 and the WW domain of Pin1 have strong hot spots that are positioned too far apart to be easily spanned by a ligand of typical drug size, and in the “other” site of survivin, the separation of hot spots also slightly exceeds 8 Å. In these cases, one might expect that these sites will not be conventionally druggable but might bind strongly to a somewhat larger ligand such as a peptide-derived foldamer or a large macrocycle; hence, they are classified as non-canonically druggable—large, denoted D*_L. Other examples include the PP⁶⁰Src SH2 domain, elongation factor TU, the catalytic domain of protein phosphatase-1, and the site on interleukin-2 (IL-2) that binds the IL-2 receptor α -chain. In each of these cases, FTMap analysis indicates that the binding site contains a sufficient number of strong hot spots to support high-affinity binding of a ligand, but these hot spots are spaced too far apart to be encompassed by a ligand the size of a typical drug; thus, the site requires a larger chemotype to achieve strong binding (see Tables S1–S6 for details). Furthermore, we have recently reported and further discuss below that the binding sites on proteins binding natural product macrocycles of MW > 600 g/mol differ from conventional drug binding sites in having hot spots that are spaced farther apart, supporting the notion that large ligands such as large macrocycles can bind to sites that have sufficiently strong hot spots but fail the standard druggability criterion due to the spacing between the hot spots being too large.⁶¹

Druggability of Proteins Binding Macrocyclic Compounds. It is generally recognized that macrocyclic compounds can be useful to expand the class of druggable targets and provide additional options against DNA, RNA, and carbohydrate binding proteins and protein–protein interactions.^{18,20,21} In Table 3, we show druggability results for a number of proteins that bind macrocyclic natural products with high affinity (except for actin, where the binding is at best micromolar). Additional details on individual targets are given in Table S5. These targets include the already discussed FK506 binding protein (Figure 1C), which FTMap indicates is druggable even using druglike compounds, as confirmed by

published experimental data,^{68,69} albeit not with the very high affinity of the macrocyclic FK506 (Figure 1A).⁶⁶ We find a similar result for a number of other targets in Table 3. Cyclophilin A, pancreatic elastase, Chitinase, and HSP90 are all predicted to be conventionally druggable, and, indeed, in addition to their potent nanomolar macrocyclic inhibitors, these proteins are known to also have somewhat less potent nonmacrocyclic inhibitors (see Table S5). However, in the case of three proteins, elongation factor TU, protein phosphatase 1, and the biotin carboxylase (BC) domain of human acetyl-CoA carboxylase 2, the hot spots are too far from each other to render the sites conventionally druggable; hence, these sites can be targeted only using larger compounds such as macrocycles or peptide foldamers.¹²⁰ According to FTMap, actin is not druggable by any means. In fact, although the protein has been cocrystallized with several large macrocycles and the compounds have biological activity, the binding is very weak in all cases (Table S5). These results support our previous suggestion⁶¹ that sites that fall outside the FTMap criteria for conventional druggability for the sole reason that the hot spots are spaced too far apart might be good targets for macrocyclic compounds or other larger ligands. For this reason, we classify such sites as non-canonically druggable, D*_L (Table 1), to indicate that they are not druggable using conventional druglike compounds but that they could be druggable using other, larger but still pharmaceutically relevant, drug chemotypes.

Druggability of Protein–Protein Interaction Targets. A substantial number of PPI interfaces are biologically compelling targets for drug discovery, and in some cases, small molecules are known to inhibit the interactions of two proteins with moderate to high potency.^{121,122} The druggability of such targets is still poorly understood, however, and PPI target sites are generally considered to be highly challenging if not undruggable. The problem in assessing the druggability of PPI targets is that protein–protein interfaces generally have much smaller pockets than traditional drug targets, and it is difficult to determine which one, if any, will be able to bind a small compound with sufficient affinity to modulate the interaction at pharmacologically relevant doses.¹²³ We recently showed that FTMap can identify the most important ligandable interface binding sites based on analysis of the unbound protein structure.⁴⁸ Here, we review these results in the light of our current and more extensive druggability analysis (Table S6). PPI targets T1–T8 of Table S6 (Bcl-xL, XIAP Bir3, MDM2, Pin1, PDZ–PSD95, the survivin Bir3 domain, the “other” site on survivin, and urokinase) are already in the Hajduk et al.⁷ data set of Table 1, and we have previously studied the druggability of targets T9–T12 (IL-2, HPV-11 E2, TNF- α , and stromelysin 1 catalytic domain).⁴⁸ Targets T13–T15 (ZipA, K-Ras, and HIV integrase) will be described in more detail in the next section. On the basis of the mapping results collected in Table S6, five of the 15 PPI targets (Bcl-xL, XIAP Bir3, MDM2, urokinase, and HPV-11 E2) appear to be druggable using druglike compounds. However, the binding of small ligands in protein–protein interfaces may introduce conformational changes that affect the location and relative strength of the hot spots. While these changes do not affect the predicted druggability, they may impact the class of compounds required for strong binding. For example, in the ligand-free structure of Bcl-xL, the three hot spots are close to each other, indicating druggability. Indeed, it has been reported that fragments of known inhibitors all bind in the pocket that binds LS78 of the BAK peptide.¹²⁴ Upon binding of the inhibitor ABT737, shown

in Figure 1B,¹²⁵ the hot spots move farther apart (Table S6). Nevertheless, Bcl-xL is assessed to be druggable also on the basis of mapping the ligand-bound structure, since the distance from CC1 to CC5 is 6.3 Å, from CC5 to CC4 is 7.7 Å, and from CC4 to CC2 is 7.0 Å. However, to achieve high potency, ABT737 must extend from CC1 all the way to CC2, a distance of 11.3 Å, and this requires long molecules with MW > 650 g/mol.^{48,126,127} In XIAP Bir3, the two hot spots are close to each other, and several potent monomeric inhibitors have been developed.¹²⁸ However, the hot spot region is quite small (MD = 10.6 Å); thus, XIAP Bir3 is close to the D*_S category. Accordingly, the inhibitors are druglike but relatively large. For example, the inhibitor (2S)-1-[(2S)-2-cyclohexyl-2-[(2S)-2-(methylamino)propanoyl]amino]acetyl]-N-(4-phenylthiadiazol-5-yl)pyrrolidine-2-carboxamide (GDC-0152)¹²⁸ has a MW of 498.2, and recently macrocyclic inhibitors have also been developed.¹²⁹ Mapping results for MDM2 are similar: the protein has strong hot spots close to each other, but the region is small (MD = 10.6 Å), and the inhibitors in the Nutlin family¹³⁰ and related compounds spread beyond the hot spots, resulting in MW of over 600 g/mol.¹³¹ However, MD exceeds the threshold of 10.0 Å, and more recent optimizations succeeded to yield inhibitors with MW below 600 g/mol and even slightly below 500 g/mol.^{132,133} Urokinase is druggable, and druglike nanomolar inhibitors are known.¹³⁴ HPV-11 E2 is also druggable, but the hot spot region is small (MD = 10.8 Å), and known inhibitors with high nanomolar affinity have MW > 500 g/mol.¹³⁵ On the basis of FTMap, three targets (the Smac peptide binding site of survivin, the FtsZ interface of ZipA, and the receptor binding interface of TNF- α) are not druggable with noncovalent inhibitors under any circumstances.⁴⁸ In fact, all three interactions have been targets of extensive screening campaigns that did not produce any potent compounds. The SOS-binding interface of KRAS, which is discussed in greater detail below, is borderline druggable, and the best noncovalent inhibitor found so far has K_d = 190 μ M.¹³⁶ Three targets in Table S6 (the WW domain of Pin1, PDZ-PSD95, the interface of IL-2 with IL-2R α) contain strong hot spots but with spacing that exceeds 8 Å; thus, they fall into the non-canonically druggable-large (D*_L) category and are druggable only using large (peptide-like or macrocyclic) ligands. Since the 8.3 Å distance between the two strong hot spots in the “other” site of survivin barely puts this target into the D*_L category, relatively small (although not <500 g/mol) potent ligands have been developed.⁸⁴ In the unbound structure of the stromelysin 1 catalytic domain, the ligand binding pocket is partially closed, and since the hot spot ensemble is small, FTMap predicts druggability in the D*_S category. We also mapped the structure complexed with the inhibitor 5-methyl-5-(4-phenoxyphenyl)-pyrimidine-2,4,6(1H,3H,5H)-trione (PDB ID 1g4k),¹³⁷ which showed a more extended hot spot region and hence suggested a higher level of druggability. However, it appears that the most potent inhibitor still has only K_i = 192 nM.¹³⁷ Although we have only a limited number of examples, the ones considered here suggest that the even the druggable PPI targets frequently require the use of compounds lying outside of traditional druglike chemical space.

■ DRUGGABILITY ANALYSIS OF SELECTED HIGH-VALUE TARGETS

In this final section, we apply our analysis to a number of targets that have been of high pharmaceutical interest and for which the question of druggability is not fully settled. Most of

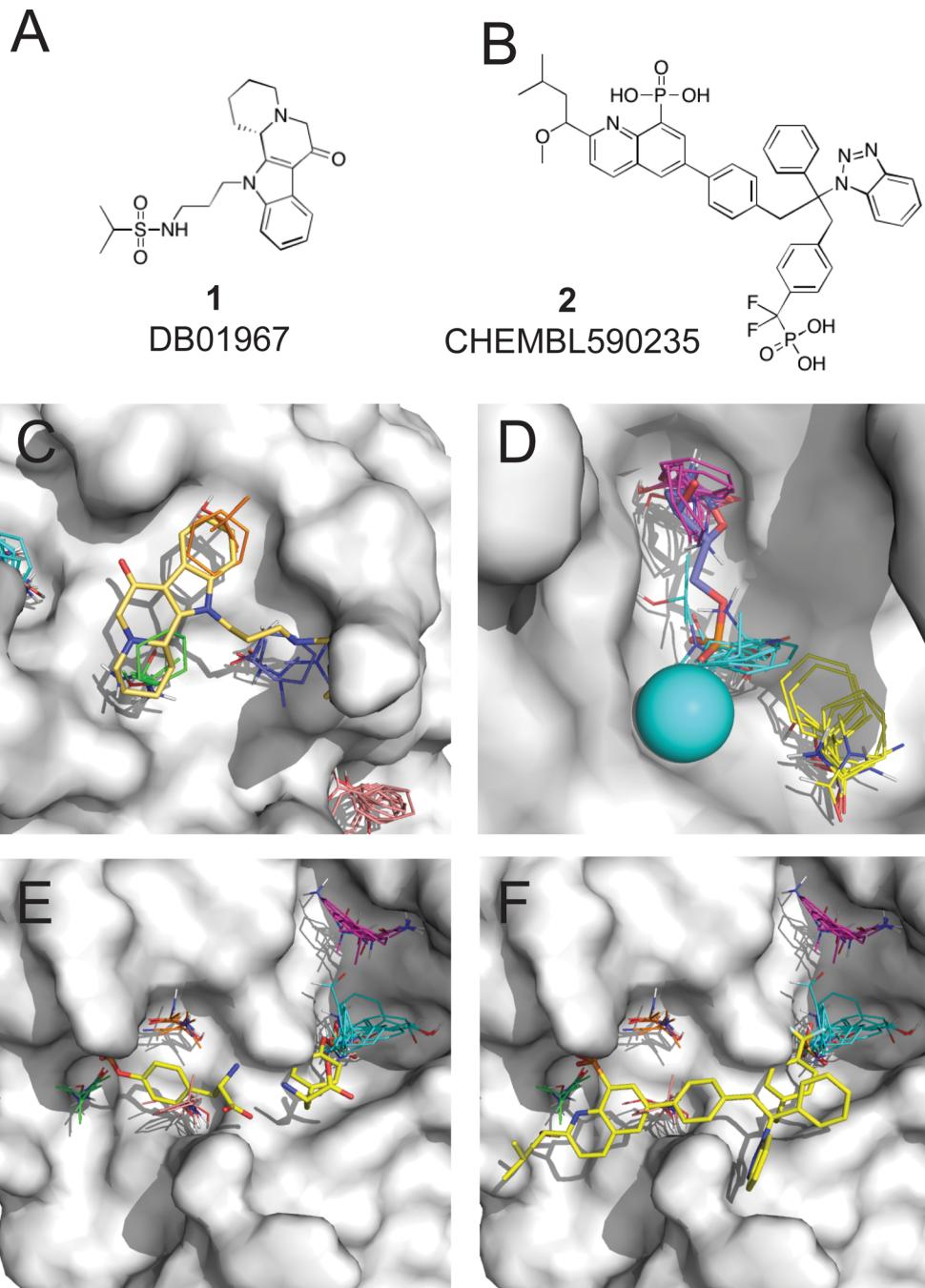


Figure 3. Druggability assessment for important targets. (A) Structure of the weak inhibitor **1** of the ZipA/FtsZ protein–protein interaction. (B) Structure of the PTP1B inhibitor **2**. (C) Mapping of the ligand free ZipA protein (PDB ID 1f46), with inhibitor **1** (from PDB ID 1s1s) superimposed for reference. The consensus clusters shown are CC8 (green, 6 probe clusters), CC9 (blue, 5 clusters), and CC10 (orange, 3 clusters). We also show the strongest hot spot CC1 (cyan, 22 clusters), which is located outside the ZipA/FtsZ interface. CC4 (salmon, 9 clusters) is also too far from the interface. (D) Mapping of the apurinic/apyrimidinic endonuclease 1 (APE1) structure from PDB ID 1de9, cocrystallized with abasic DNA (fragment of the DNA is shown as blue sticks for reference), and an Mn²⁺ ion (cyan sphere). As always, the DNA and the ion were removed prior to the mapping. The hot spots shown are CC1 (cyan, 16 clusters), CC2 (magenta, 16 clusters), and CC3 (yellow, 13 clusters). (E) Mapping of the ligand-free protein tyrosine phosphatase 1B (PTP1B, PDB ID 3a5j), with two bound phosphotyrosine (pTyr) molecules superimposed for reference (from PDB ID 1pty). The hot spots shown are CC1 (cyan, 20 clusters) at the catalytic PTyr site, CC4 (salmon, 9 clusters), CC7 (orange, 7 clusters), and CC8 (green, 7 clusters) close to the site that binds the noncatalytic pTyr. The hot spots CC2 (magenta, 16 clusters) and CC7 (orange, 7 clusters) are closed in any ligand-bound structure. (F) Same as panel E but with the bound structure of the potent inhibitor **2** shown for reference from the structure with PDB ID 1q6t.

the targets considered in this section are complex and undergo significant conformational change upon ligand binding, which may affect the strength and location of hot spots. Therefore, it is important to recall that, based on our experience drawn from

the analysis of more than 150 proteins, as described above, for druggability it is necessary to identify appropriate hot spots in ligand-free protein structures, and it is the hot spot properties of the unliganded proteins that provide the most valuable

Table 4. Detailed Analysis of Important Targets

target	site(s)	PDB	consensus sites		D ^a	comment
			U	B		
1 ZipA	FtsZ peptide binding site	1f46	8(6), 9(5), 11(3) 5(13), 6(6), 7(3)	N	All hot spots in the ZipA:FtsZ interface are very weak. There is a strong hot spot, 1(22) and 2(16), but are more than 10 Å from any hot spot in the interface	
2 APE-1	A: PO3 of abasic nucleotide; B: sugar of abasic nucleotide; U: PO3 of two nucl. upstream; Y: at Y128	1s1s	1f47 4(10), 8(4), 10(2) A: 1(18), B: 2(16); U: 5(10)	D	For the ensemble formed by CC1 and CC2, MD = 11.6 Å, druggable but weak; CC5 from CC2 = 8 Å, CCS5 is closed in DNA bound structures	
3 PTP1B	A site: catalytic pTyr; B site: noncatalytic pTyr; C site: Tyr46, Arg47, and Asp48	3asj	1e9n	A: 4(15), B: 2(17), U: 6(4)		
		1de9	A: 1(16), B: 2(16), Y: 3(13)	A: 10(6), B: 2(16); Y: 7(5)	With bound DNA and Mn ²⁺ ; CCS3 is close to CC1, not present in the DNA-free structures, MD = 14.5 Å	
		1dew	A: 3(14), B: 1(22), Y: 7(5)	A: 10(4), B: 2(16); Y: 10(4)	With bound DNA, no Mn ²⁺ , MD = 14.4 Å	
		1py	A: 1(20), 2(16); B: 4(9), 7(7), 8(7)	A: 1(21), 2(16); B: 8(4), 10(4)	A and B sites are too far (9.0 Å), but A is druggable on its own	
		2veu	A: 1(16), 10(3); B: 3(12), 6(7), 9(4)	A: 1(13), 7(7); B: 8(7), 9(7)	Co-crystallized with 2 pTyr amino acids; A site contracts upon pTyr binding due to the move of loop 180–184	
		4i8n	A: 1(26); B: 3(10), 8(6)	A: 1(16), 10(3), 9(1)	IC ₅₀ = 100 nM, not selective; inhibitor binds only at A and C, A–B = 9.1 Å is too far, MW = 604.6	
		2view	A: 1(25); B: 4(9), 6(7)	A: 1(26); B: 3(10), 8(6)	Binds at A and the weaker hot spot of B, not selective; A–B = 6.4 Å, MW = 483.5	
		2fjn	A: 1(15), 6(9); B: 5(11), 7(7)	A: 1(25); B: 4(9), 6(7)	IC ₅₀ = 64 nM; binds A and the weaker hot spot of B and part of C, A–B = 6.6 Å, MW = 584.6	
		2qfp	A: 1(18), 5(8); B: 4(10), 9(5)	A: 1(15), 6(9); B: 5(11), 7(7)	IC ₅₀ = 39 nM; A and one hot spot of B and part of C, A–B = 7.0 Å, MW = 603.6	
		1q6t	A: 1(20), 10(3); B: 5(11), 9(5)	A: 1(20), 9(5); B: 5(7), 7(7)	K _i = 4 nM, both A and B; both hot spots in B are used, A–B = 7.7 Å, MW = 609.5	
		1pxh	A: 1(20), 9(5); B: 5(7), 7(7)	S1: 1(21); S2: 4(12); S3: 7(5); S4: 2(16), 5(11)	IC ₅₀ = 7 nM, pTyr analogue, both A and B and some C, A–B = 7.7 Å, MW = 826.8	
		4	14-3-3 protein	S1: pS/T binding; S2: pS/T – 1; S3: pS/T + 1; S4: fusicoccin; SS: pS/T-X-COOH	S1: 2(14); S3: 1(29); S4: 3(13); SS: 5(10)	K _i = 1.4 nM, binds at A and weaker hot spot of B and C, A–B = 6.6 Å, MW = 873.2
		3pIn	4dhu	S2: 4(13), 8(3); S3: 5(11); S4: 1(18), 3(160), 7(5)	D ^{*L} CC1–CC4 = 8.2 Å, all other hot spots are even farther: CC1–CC2 = 10.7 Å, CC1–CC5 = 9.7 Å	
					S2: 2(14); S3: 1(29); S4: 3(13); SS: 5(10)	Bound phosphate group moves the side chain of Arg129 and closes down the pS/T pocket; S3 is strong
						The small inhibitor places its phosphate group into the pS/T pocket (S1 site), which becomes closed
5 KRAS	G: GTP/GDP; S1: Site 1; S2: Site 2 (SOS interface); S3: Site 3; S4: Site 4; SS: Site 5; S6: not in Raf or SOS interface	3gft	G: 1(24), 7(6), 9(5); S2: 2(13); S3: 4(9), 7(6); SS: 6(7), 8(6)	S2: 4(13); S3: 5(11); S4: 2(12), S5: 8(5); SS: 6(9)	Inhibitor binding site S2 is a weak and isolated hot spot, MD = 7.0 Å	
		4epr		G: 1(26), 4(10), 5(9); S3: 7(6); SS: 3(11)	S2 does not include any hot spot	
		4dst		G: 1(19), 5(6), 8(3); S2: 3(17), S3: 6(5), 7(4); SS: 2(19)	S2 stronger, but still isolated, MD = 6.7 Å	

Table 4. continued

target	site(s)	PDB	U	B	consensus sites	D ^a	comment
			4epw	G: 2(17), 4(12); S1: 7(5), S2: 1(21), 8(3); S5: 3(14), 5(9)	G: 2(17), 4(12); S1: 7(5), S2: 1(21), 8(3); S5: 3(14), 5(9)	D ^a	
		4epy	G: 2(16), 3(12), 4(10), S1: 1(20), 8(5); S3: 10(3); S5: 6(7), 9(4)	G: 2(16), 3(12), 4(10), S1: 1(20), 8(5); S3: 10(3); S5: 6(7), 9(4)			There are no hot spots for further extension. Adding different groups to the end of the inhibitor, the best K_d value achieved is 190 μM
Mapping with masked GTP site, i.e., F site is not available to probes	3gft		S2: 4(14), S6: 1(19), S3: 7(5); S5: 2(18)	S2: 4(14), S6: 1(19), S3: 7(5); S5: 2(18)			Strongest site is the one identified by MSCS; S2 site is essentially invariant, remains weak
		4dst	S2: 2(17), S3: 5(11), 6(5); S5: 1(23)	S2: 2(17), S3: 5(11), 6(5); S5: 1(23)			Almost no change relative to free GTP site
		4epw	S2: 1(17), 4(13); S3: 3(14); S5: 2(15)	S2: 1(17), 4(13); S3: 3(14); S5: 2(15)			The secondary site in S2 appears to be stronger, but S2 is still only borderline druggable
		4epy	S2: 1(16), 3(14); S3: 2(16), 7(9); S5: 4(11)	S2: 1(16), 3(14); S3: 2(16), 7(9); S5: 4(11)			The same as 4epw, MD = 11.4 Å
6	HIV Integrase	1b14	I(18), 5(9)	I(18), 5(9)		D* _S	Ligand binding site is a deep pocket, hot spot at the bottom; rather small hot spot ensemble, MD = 9.6 Å, predicted to require charged ligand for potency
		3avm	I(20)	I(20)			Bound to LEDGF peptide; elongated hot spot, but still only MD = 9.2 Å
		2b4j	I(19)	I(19)			Bound to LEDGF, MD = 8.9 Å
		3lpu	3(8), 9(3)	3(8), 9(3)			Inhibitor bound, MD = 9.5 Å, related inhibitor with EC ₅₀ = 69 nM, all ligands are carboxylic acid derivatives

^aDruggability class based on FTMap analysis, as described in Table 1.

information for establishing druggability. As discussed, the analysis of ligand-bound structures can be useful, however, for understanding the changes in the relative importance of the binding energy hot spots upon ligand binding. In addition, as shown for PTP1B, the analysis of hot spots can also indicate that a target that can bind small ligands with high affinity can nevertheless be undruggable because any such ligands will not be able to satisfy additional requirements for specificity and/or for membrane permeability. We additionally note that, even for targets with good druggability, lead and drug discovery can be challenging when pursued by particular approaches. For example, although high druggability generally correlates with the observation of a high hit rate in fragment-based screening,⁷ whether the resulting fragment hits will have sufficiently robust binding modes to serve as good starting points for lead discovery depends additionally on how compact or diffuse the main hot spot is and how fully the fragment occupies it.¹³⁸ This topic is beyond the scope of the current perspective, however, where we are concerned with whether high-affinity binding can occur rather than with differences between druggable sites in the technical ease or difficulty in identifying a good ligand.

Cell Division Protein ZipA. The ZipA/FtsZ protein–protein interaction has been considered to be a potential target for antibacterial agents.¹²¹ We have mapped three different structures of ZipA: unliganded, cocrystallized with a peptide fragment of FtsZ,¹³⁹ and cocrystallized with the indoloquinolizin inhibitor DB01967 (**1**), shown in Figure 3A.¹⁴⁰ In all three structures, FTMap finds only three weak hot spots at the protein–peptide interface (Table 4 and Figure 3C). Although ZipA has strong hot spots elsewhere, these are at least 11 Å distant from the interface. Thus, we conclude that the interface between ZipA and FtsZ is not a druggable target. In fact, the K_d of the interaction between ZipA and a 17-mer peptide of FtsZ has been measured to be $\sim 7 \mu\text{M}$,¹⁴¹ and efforts to identify and optimize small molecule PPI inhibitors for this system, including HTS of 250 000 compounds, resulted only in inhibitors with $K_d \geq 12 \mu\text{M}$.^{140,142–144}

Apurinic/Apyrimidinic (AP) Endonuclease 1 (APE1). APE1 is an essential protein that operates in the base excision repair (BER) pathway and is responsible for $\geq 95\%$ of the total AP endonuclease activity in human cells. BER is a major pathway that copes with DNA damage induced by several cancer causing agents, including ionizing radiation and Temozolomide. Overexpression of APE1 and enhanced AP endonuclease activity have been linked to increased resistance of tumor cells to treatment with monofunctional alkylators, implicating inhibition of APE1 as a valid strategy for cancer therapy.¹⁴⁵ We analyzed two different unbound structures of APE1 by FTMap (Table 4). The results show two main hot spots at the locations that bind the phosphate group (site A) of the abasic nucleotide, close to the divalent metal ion in the active site, and the sugar ring (site B) of the same nucleotide (Table 4). In ligand-free structures, there is also a weak hot spot (site U) at the phosphate group of the nucleotide two positions upstream from the abasic group, but the site is far (8 Å) from site B and closes down upon DNA binding. The mapping of DNA bound structures reveals another hot spot (site Y), formed by turning the side chain of Tyr128 out of the active site. This site is somewhat stronger when a bound Mn²⁺ ion is present (Figure 3D). However, the two main hot spots remain weak, and although utilizing the latter hot spot may provide additional potency, there is nothing else near the small active site. Thus, we consider the active site of APE-1 to be

druggable, but very weak, close to borderline druggability. In spite of recent hopes,¹⁴⁶ information available in the literature fully confirms this prediction. As described in an excellent review,¹⁴⁷ the reproducibility and biological value of a few potent compounds reported in early studies remain questionable and therefore APE1 was targeted in a number of more recent screens, which collectively returned over 50 compounds. However, even the best inhibitors had only low micromolar IC₅₀ values. Almost all compounds share a relatively small size, an aromatic hydrophobic central core that most likely binds at site B, and a negative ionizable group, frequently a carboxylate, which is thus predicted to bind at site A and to interact with the metal. On the basis of the weak druggability of the active site, we believe it will be difficult to get much higher potency, and APE1 selectivity may also become an issue.¹⁴⁷ However, it is possible that larger compounds, primarily macrocycles, will be able to take advantage of the additional U and Y sites, improving both potency and selectivity.

Protein Tyrosine Phosphatase 1B (PTP1B). PTP1B is a highly validated drug target with a very long history of failures to identify potent and selective inhibitors. The enzyme acts as a negative regulator in insulin signaling pathways. Inhibition of PTP1B activity exhibits potential for enhancing insulin action by prolonging the phosphorylated state of the insulin receptor.¹⁴⁸ A large variety of potent PTP1B inhibitors have been developed, such as **2** shown in Figure 3B with IC₅₀ = 7 nM,¹⁴⁹ and it is known that the primary challenges are cell permeability and selectivity.¹⁵⁰ The challenge of permeability is a consequence of the highly polar nature of the active site. Achieving selectivity is difficult because PTP1B and the most highly related tyrosine phosphatase, T-cell PTP (TCPTP), share 72% sequence identity overall and 94% identity for the active site residues. We analyzed two different unbound PTP1B structures by FTMap (Table 4). The results show that top two hot spots in the ligand-free PTP1B structures are located at the catalytic site (A site), with additional, weaker hot spots at a second, noncatalytic site for phosphotyrosine binding (B site) (Table 4 and Figure 3E). We agree with Hajduk et al.⁷ that the A site is druggable on its own, i.e., it binds potent (albeit charged) ligands. As seen for other ligands with a phosphate group, the structure of PTP1B in complex with phosphotyrosines (PDB entry 1pty) shows that the binding changes the conformation of the loop comprising residues 180–184, which tightens the pocket, but the site can open up to accommodate larger inhibitors. All inhibitors must interact with the charged A site because it is the main source of binding affinity. Attempts have been made to replace the phosphate group with 1,1-dioxido-3-oxothiazolidin in site A and expand the inhibitor toward residues Tyr46, Arg47, and Asp 48, defined as the C site, without any strong hot spots. Such inhibitors (e.g., PDB entry 2veu) have IC₅₀ values between 100 and 2900 nM, but MW > 600 g/mol.¹⁵¹ Alternatively, a number of inhibitors have been developed that have the phosphate group in site A but that expand into the less polar B site, thereby improving both affinity and cell permeability. To reach even the weaker hot spot at the B site, around 7 Å away from the A site, requires a molecule with MW > 500 g/mol. In addition, the residues that differ between PTP1B and TCPTP (Phe52Tyr and Ala27Ser) lie at the distal end of the B site. Thus, in order to improve selectivity, inhibitors first have to bridge from the A site to the first hot spot of the B site (7.7 Å) and then continue to the second, stronger hot spot of the latter site (another 6 Å). Adding interactions with the C site further improves binding,

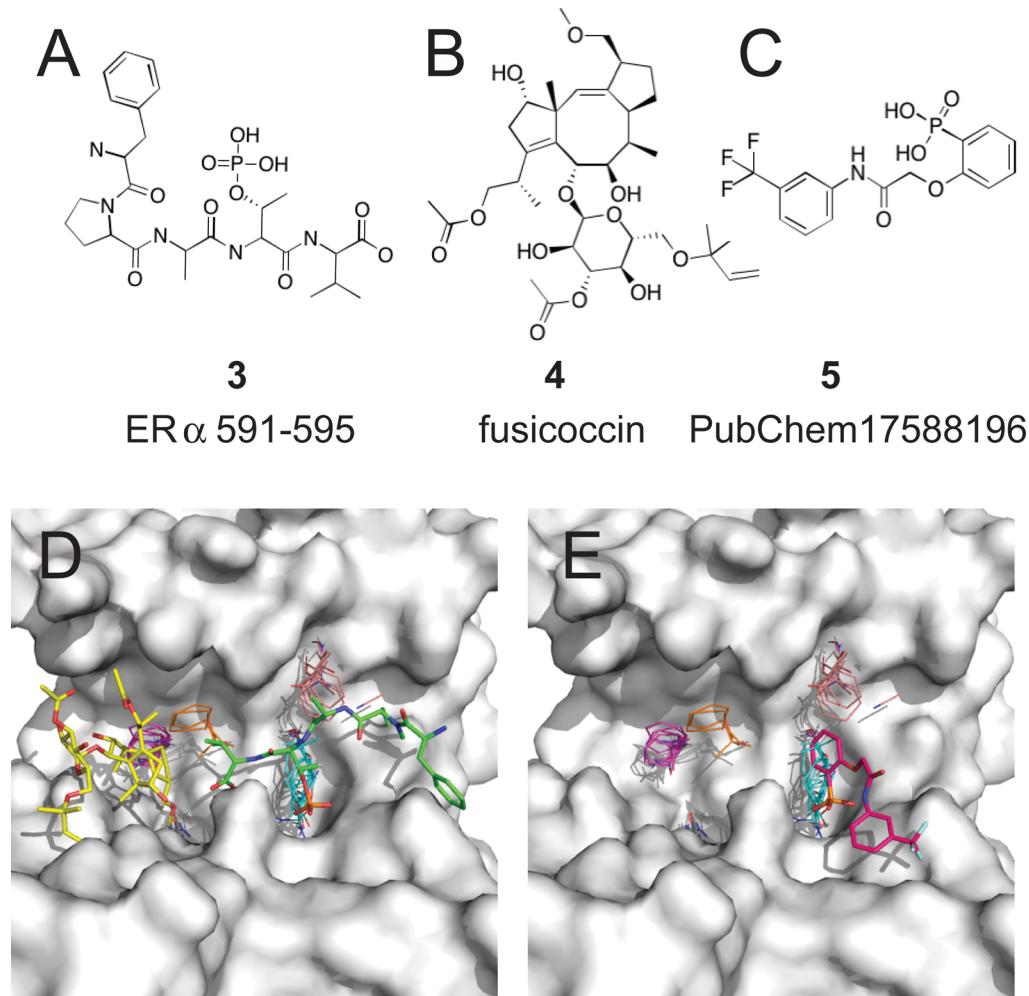


Figure 4. Druggability assessment for 14-3-3 protein. (A) Structure of the phosphothreonine peptide **3** that was cocrystallized with 14-3-3. (B) Structure of fusicoccin (**4**), a compound that binds to 14-3-3 and stabilizes the binding of **3**. (C) Structure of inhibitor **5**. (D) Mapping of the ligand free structure of 14-3-3 (PDB ID 1yz5). For reference we superimpose the bound compounds **3** and **4**, shown as green and yellow sticks, respectively, from PDB ID 4jdd. The hot spots shown are CC1 (cyan, 21 clusters) at the deep pocket that binds **3**, CC2 (magenta, 16 clusters), and CC5 (white, 11 clusters) at the region that binds **4**. Two more hot spots are close to **3**: CC4 (salmon, 12 clusters) and CC7 (orange, 5 clusters). (E) Same as panel D but with inhibitor **5** (magenta).

and such compounds reach an IC_{50} of 5 nM and 7-fold selectivity over TCPTP.¹⁴⁹ However, inhibitor **2** of this type from PDB entry 1q6t,¹⁴⁹ shown in Figure 3B, still has the charged phosphate group and the high MW of 826.76 g/mol. Thus, we conclude that PTP1B is formally druggable, in the sense that very potent inhibitors can be and have been developed. However, the analysis of the hot spot strengths and locations, and the residue differences between PTP1B and its close homologue TCPTP, show that the requirements of high affinity, selectivity, and cell permeability cannot be simultaneously satisfied by an ordinary druglike compound.

14-3-3 Protein. The 14-3-3 family of phosphoserine/threonine-recognition proteins engage multiple nodes in signaling networks that control diverse physiological and pathophysiological functions and have emerged as promising therapeutic targets for such diseases as cancer and neurodegenerative disorders.¹⁵² 14-3-3 proteins have been cocrystallized with several phosphoserine or phosphothreonine (pS/T) containing peptides.¹⁵³ Figure 4A shows such a phosphopeptide, ER α 591–595 (**3**), with phosphothreonine at position 594, and Figure 4B shows the compound named fusicoccin (**4**), originally identified as a fungal toxin, that stabilizes the binding

of **3**.¹⁵⁴ The 14-3-3 protein forms a weak homodimer, but each monomer can bind **3** independently in a well-defined elongated site with a deep pS/T binding cavity, together with **4** (Figure 4D). The protein has been recently cocrystallized with the inhibitor PubChem17588196 (**5**), shown in Figure 4C.¹⁵⁵ Figure 4E shows the resulting complex.¹⁵⁵ Although the IC_{50} value of **5** is only 5 μ M, it was shown to disrupt the interaction of 14-3-3 with aminopeptidase N in a cellular assay.¹⁵⁵

As shown in Table 4 and Figure 4D, the strongest hot spot of the ligand-free 14-3-3 protein is in the pS/T binding pocket, with additional hot spots, both at the side chains adjacent to the bound **3**, and in other regions in the binding site. The binding of the phosphate group is clearly an important part of the interactions with peptides and small inhibitors. However, the binding of the phosphate group, both in the peptides and the inhibitors, attracts the side chain of Arg129 into the pS/T binding pocket, eliminating the hot spot at that location (Table 4). This example emphasizes that the analysis of ligand-free protein structure, which here predicts the site to be druggable, provides useful druggability information, whereas restricting consideration to pocket properties in ligand-bound structures can be misleading when the protein binds charged ligands. Indeed, the

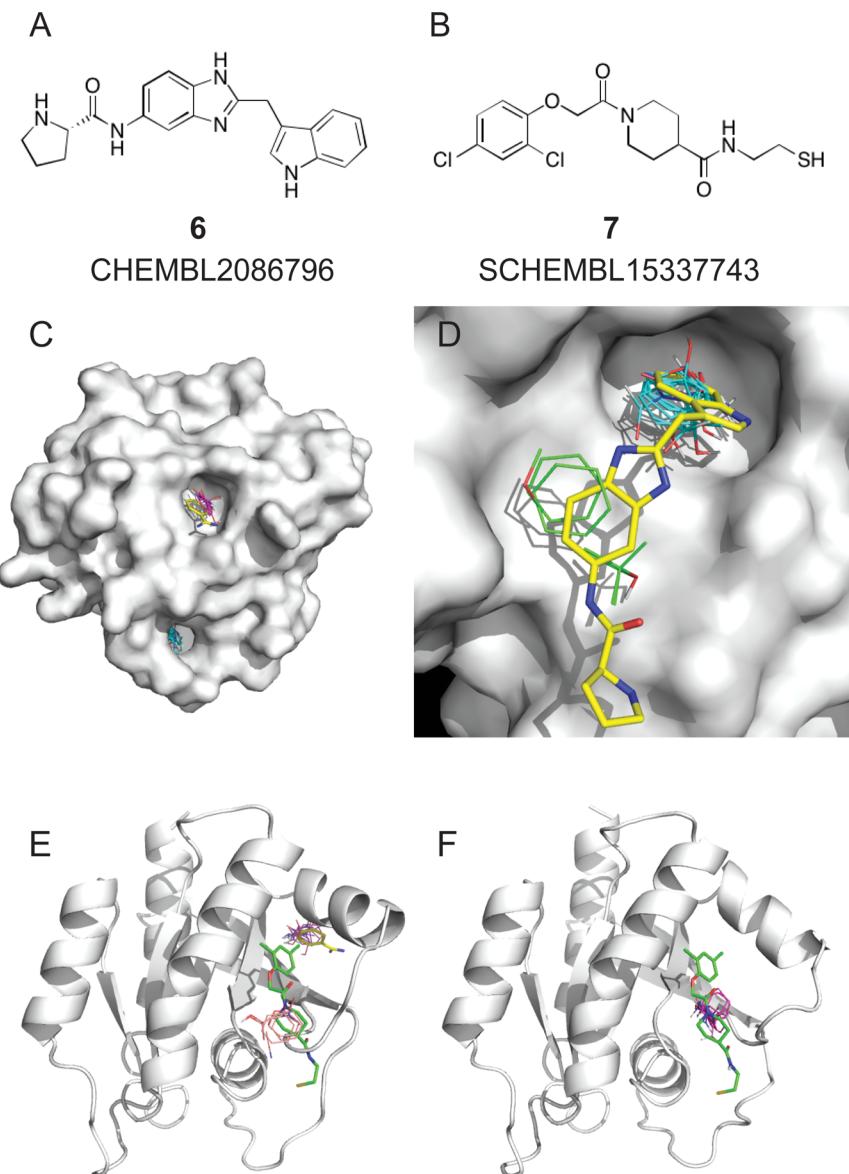


Figure 5. Druggability assessment for KRAS. (A) Structure of **6**, one of the highest affinity inhibitors binding to KRAS at the interface with SOS. (B) Structure of inhibitor **7** binding covalently to KRAS(G12C). (C) Mapping of the ligand-free KRAS (PDB ID 3gft). The cavities with probes bound are the GTP/GDP binding site, with part of CC1 (cyan, 24 clusters) visible, and the targeted hot spot CC2 (magenta, 13 clusters) in Site 2 located in the KRAS/SOS interface. For reference, we superimpose a small ligand, benzamidine (shown as yellow sticks, from the structure with PDB ID 4dso), binding to Site 2. (D) Mapping of the KRAS structure from its complex with inhibitor **6** from PDB ID 4epy ($K_d = 340 \mu\text{M}$, superimposed as magenta sticks). Due to the binding of **6**, the strongest hot spot CC1 (cyan, 20 clusters) is now in the KRAS/SOS interface, and the interface includes a second weak hot spot CC8 (green, 7 clusters) located close to CC1. (E) The mapping of the ligand-free KRAS (PDB ID 3gft) places CC4 (salmon, 9 clusters) at the S3 site, in the pocket binding the KRAS(G12C) inhibitor **7** that allosterically controls GTP affinity and effector interactions. Inhibitor **7** from PDB ID 4lv6 is superimposed for reference and is shown as green sticks. Notice that in this closed conformation of KRAS the loop 58–65 would clash with **7** and thus prevents its binding. For reference, the figure also shows CC2 (magenta) and the benzamidine ligand (yellow) on the other side of the protein. (F) Homology model of KRAS based on a HRAS structure (PDB ID 4q21) to model the open state. In this structure, the opening of the 58–65 loop creates a stronger hot spot, CC2 (magenta, 16 clusters). Inhibitor **7** from PDB ID 4lv6 is shown again for reference. Although the site in this open structure binds 16 probe clusters, the maximum dimension of CC2 is only 5.1 Å; thus, the site is not druggable without the covalent interaction with G12C.

inhibitors are located at the strongest hot spot, CC1, of the unliganded protein, rather than those present in the ligand-bound structure (Figure 4E). The hot spot closest to CC1 is CC5, but even that distance is 8.2 Å; all other hot spots are even farther and are not used by the small inhibitor. However, the S2 and S3 sites at the pS/T – 1 and pS/T + 1 positions, respectively, remain relatively strong in all structures, suggesting

that higher potency but larger inhibitors can most likely be found.

KRAS. The Ras (rat sarcoma) protein family members are all low-molecular-weight GTP-binding proteins that play a role in regulating cell differentiation, proliferation, and survival. KRAS, a key member of the Ras family, is an attractive cancer target, as frequent point mutations in the KRAS gene render the protein constitutively active.¹⁵⁶ However, even though the structure

and function of RAS have been studied intensively for nearly 30 years, it has remained elusive as a drug target.⁴⁰ The most obvious target site would be the deep pocket where GTP and GDP bind, but these nucleotides bind with nanomolar affinity¹⁵⁷ and are present at very high concentrations in cells. Therefore, we mapped the protein both with the GTP/GDP site free and also with the site masked to prevent probe binding (Table 4). There have been many attempts to find alternative sites to target, with recent successes by two groups, who reported small molecules binding to a pocket in the surface of KRAS that interacts with the SOS protein,^{136,158} referred to as Site 2.¹⁵⁹ Studying KRAS is complicated by the fact that the protein acts as a molecular switch with complex dynamics between multiple conformational states.⁴⁰ We analyzed three different unbound structures by FTMap. As shown in Table 4, in mapping KRAS the GTP/GDP site dominates, if the site is not masked, with several strong hot spots at this location. In two of the three unbound KRAS structures, Site 2 in the KRAS/SOS interface had the second strongest hot spots. However, the hot spot has at most 13 probe clusters and has a small size (MD = 7.0 Å); hence, on the basis of the FTMap criteria, this site is classified as B*_S. Nevertheless, the KRAS/SOS interface has been targeted in recent campaigns, and two groups independently identified series of druglike inhibitors binding at Site 2,^{136,158} such as CHEMBL2086796 (6)¹³⁶ shown in Figure 5A. To show the hot spot in the unliganded KRAS structure, we superimpose a small ligand, benzamidine (Figure 5C). As the size of ligands increases, the number of probes in the main hot spot increases to 20, and second hot spot also opens in the structure cocrystallized with 6, but it is very weak and includes at most 5 probe clusters (Figure 5D). In addition, the maximum dimension of the hot spot region increases to 10.9 Å, and on the basis of these values, the site would be considered druggable. However, as our analyses of other targets shows, the mapping of the ligand-free structure usually provides more valid assessment of druggability than the mapping of the ligand-bound state, and on the basis of the former, the interface is only borderline druggable (B*_S). This prediction is supported by the fact that, in spite of major efforts, the best-known noncovalent inhibitor binding at this site has a K_d value of 190 μM and is positively charged at neutral pH.¹³⁶ The hot spots are even weaker if the GTP/GDP site is masked (Table 4), so the weakness of the hot spots at Site 2 is not due to the binding of probes elsewhere.

We note that the mapping of the ligand-free KRAS structure places CC4 (salmon, 9 clusters) at the S3 site, in the pocket that binds the recently developed covalent KRAS(G12C) inhibitors such as SCHEMBL15337743 (7), shown in Figure 5B.¹⁶⁰ In this closed conformation of KRAS, the loop 58–65 would clash with the allosteric inhibitor 7 and thus prevent binding (Figure 5E). In spite of this unfavorable geometry, FTMap places a hot spot at the allosteric site. To show that the opening of the site would allow for the binding of 7, we constructed a homology model of the KRAS open state based on an open structure of HRAS. Figure 5F shows the mapping of the resulting structure with 7 superimposed. As expected, the opening of the 58–65 loop creates a stronger hot spot, CC2 (magenta, 16 clusters). However, the maximum dimension of this CC2 is only 5.1 Å; thus, the site is not druggable without covalent protein–ligand interaction.¹⁶⁰

HIV Integrase, LEDGF/p75 Interaction Site. Integration of HIV-1 DNA is required to maintain the viral DNA in the infected cell; hence, the integrase (IN) is an attractive target for

HIV treatment. Two different mechanisms have been explored for IN inhibition. The design and discovery of IN inhibitors were first focused at targeting the catalytic site of IN, with the goal of achieving an effect on strand transfer.¹⁶¹ Cheng et al.⁸ considered this site to be undruggable, whereas, based on FTMap, we predicted it to be druggable (Table S3). In fact, since the publication of the work by Cheng et al.,⁸ several compounds targeting the active site, called integrase strand transfer inhibitors (INSTIs), have been approved.¹⁶² Since INSTI resistance evolves readily, a proposed alternative approach is to disrupt the interaction between IN and the transcriptional coactivator protein LEDGF/p75, which results in inhibition through an allosteric mechanism.^{161,162} According to FTMap, the interface region of IN has a strong hot spot, but the hot spot ensemble is small (MD = 9.6 Å) and thus we classify the site as non-canonically druggable—small (D*_S), requiring charged inhibitors. In agreement with this prediction, the recently discovered potent inhibitors called LEDGINs include a carboxyl group.¹⁶³ An example of this class is (2S)-2-(6-chloro-2-methyl-4-phenylquinolin-3-yl)pentanoic acid (PubChem45281242),¹⁶¹ which binds to IN with EC₅₀ = 69 nM. Note that the binding of the charged ligands reduces the hot spot, similarly to the 14-3-3 protein, again emphasizing that the ligand-free structure is likely to provide more accurate druggability prediction.

CONCLUSIONS

We show here that the druggability of binding sites on proteins can be evaluated with good reliability based on fragment hit rates, whether obtained experimentally (e.g., by NMR) or computationally using the FTMap algorithm, and establish clear FTMap-based criteria for druggability. Importantly, because the FTMap method is based on the biophysics of binding rather than on empirical parametrization to a particular training set of protein–drug complexes, meaningful information can be gained about classes of proteins and classes of compounds beyond those resembling validated targets and conventionally druglike ligands. In particular, the information provided by the mapping identifies targets that, while not druggable by druglike ligands, may become druggable using pharmaceutically attractive compound classes such as macrocycles or other large molecules that fall outside conventional definitions of druglikeness such as the rule-of-five. Application of the method to ZipA and PTP-1B show why these targets are not druggable. In contrast, results for other targets of high current biomedical interest, including apurinic/aprimidinic endonuclease 1 (APE1), the SOS interaction site of KRAS, the phosphoserine/threonine-recognition protein 14-3-3, and the LEDGF/p75 interaction site of HIV integrase, show that, in these cases, it may be necessary to consider ligands lying outside conventional druglike chemical space and provide guidance as to what compound classes are most likely to succeed. Overall, due to the progress that has been made in recent years across multiple research groups and institutions, we believe that evaluation of protein druggability has entered an era in which, for the majority of cases, the prospects for identifying a pharmaceutically useful small molecule ligand can now be ascertained with good confidence based only on the X-ray crystal structure of the unliganded protein. These advances substantially lower the technical risk associated with launching a drug discovery project involving a novel protein target and provide valuable information concerning which approaches and chemotypes are most likely to succeed against a given site.

■ ASSOCIATED CONTENT

● Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: [10.1021/acs.jmedchem.5b00586](https://doi.org/10.1021/acs.jmedchem.5b00586).

Additional druggability results and histograms ([PDF](#)).

■ AUTHOR INFORMATION

Corresponding Authors

*(D.K.) Phone: 617-353-4842; E-mail: midas@bu.edu.

*(A.W.) Phone: 617-353-2488; E-mail: whitty@bu.edu.

*(S.V.) Phone: 617-353-4757; E-mail: vajda@bu.edu.

Notes

The authors declare the following competing financial interest(s): D.R.H. is a full-time employee of Acphearis, Inc. The company offers software similar to the FTMap program that was used in this perspective. D.K. and S.V. own Acphearis stock. However, the FTMap software and server are free for use.

Biographies

Dima Kozakov received an M.S. in Applied Mathematics and Physics from the Moscow Institute of Physics and Technology and a Ph.D. in Biomedical Engineering from Boston University. Currently, he is Assistant Professor of Applied Mathematics & Statistics at Stony Brook University. Dr. Kozakov has been working on the development of algorithms and software for molecular modeling and drug design. He is the primary architect of ClusPro, Version 2.0, protein docking server, currently the best automatic server according to latest round of evaluations of the blind protein docking assessment CAPRI. ClusPro has more than 10 000 users and has run more than 100 000 jobs in the last 3 years. Dr. Kozakov is also the primary developer of the FTMap approach, which was used in this work. The FTMap server currently has more than 2000 users.

David R. Hall received a Ph.D. in Biomedical Engineering from Boston University in 2012. Since 2012, he has led Acpharis, a computer-aided drug design company specializing in computational solvent mapping and protein–protein docking. In this role, he has led a long-term collaboration with Biogen on evaluating the binding sites of proteins using computational solvent mapping.

Raeanne L. Napoleon received a Ph.D. in Theoretical Physical Chemistry from Boston University in 2012. She received her B.S. (*cum laude*) from the University of the Sciences in Philadelphia, PA, in 2005. She is currently Assistant Professor of Chemistry at Santa Barbara City College, where she focuses on curriculum development in STEM education, with an emphasis on both chemistry and nanoscience disciplines, and promoting women in STEM. She also continues to work on computational simulations of protein systems and interfacial condensed phase simulations.

Christine Yueh received a B.S. in Chemical Engineering from UC Santa Barbara in 2010. She is currently a Ph.D. candidate in the Biomedical Engineering program at Boston University, where her research involves using computational modeling to study protein–protein and protein–small ligand interactions.

Adrian Whitty received a B.Sc. in Chemistry from King's College London in 1985, followed by M.S. and Ph.D. degrees in Chemistry from the University of Illinois at Chicago. After completing postdoctoral work with William P. Jencks FRS at Brandeis University, he joined the biopharmaceutical company Biogen, where he worked in the Department of Drug Discovery, rising to the position of Director of Physical Biochemistry. In 2008, he moved to Boston University,

where he is an Associate Professor in the Department of Chemistry. His research interests include protein–protein and protein–ligand binding, especially as they are related to the discovery of small molecule inhibitors of protein–protein interactions, and also the quantitative, mechanistic analysis of growth factor receptor activation and signaling.

Sandor Vajda received an M.S. in Applied Mathematics and Ph.D. in chemistry in Budapest, Hungary. He held research positions in the Department of Engineering, University of Warwick, England, and the Department of Chemistry, Princeton University. He was a faculty member in Hungary and at the Mount Sinai School of Medicine, New York, NY. Currently, he is Professor of Biomedical Engineering and Chemistry at Boston University. He also directs the Biomolecular Engineering Research Center at BU. Dr. Vajda has been active in method development for modeling biological macromolecules, with emphasis on molecular interactions and drug design. In 2004, he and his graduate students founded SolMap Pharmaceuticals, a start-up company focusing on fragment-based drug design. SolMap was acquired by Forma Therapeutics, Cambridge, MA, in 2008.

■ ACKNOWLEDGMENTS

This investigation was supported by grant nos. GM064700 to S.V. and GM094551 to A.W. and S.V. from the National Institute of General Medical Sciences.

■ ABBREVIATIONS USED

PTP-1B, protein tyrosine phosphatase-1B; APE1, apurinic/apyrimidinic endonuclease 1; SOS, son of sevenless; LEDGF/p75, lens epithelium-derived growth factor; KRAS, Kirsten rat sarcoma viral oncogene homologue; MSCS, multiple solvent crystal structures; CC, consensus cluster; PDB, Protein Data Bank; RMSD, root-mean-square deviation; CD, center-to-center distance; MD, maximum distance; DS, druggability score

■ REFERENCES

- (1) Hopkins, A. L.; Groom, C. R. The druggable genome. *Nat. Rev. Drug Discovery* **2002**, *1*, 727–30.
- (2) Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Delivery Rev.* **2001**, *46*, 3–26.
- (3) Russ, A. P.; Lampel, S. The druggable genome: An update. *Drug Discovery Today* **2005**, *10*, 1607–10.
- (4) Johns, M. A.; Russ, A.; Fu, H. A. Current drug targets and the druggable genome. *Chemical Genomics* **2012**, 320–32.
- (5) Sheridan, R. P.; Maiorov, V. N.; Holloway, M. K.; Cornell, W. D.; Gao, Y. D. Drug-like density: A method of quantifying the "bindability" of a protein target based on a very large set of pockets and drug-like ligands from the Protein Data Bank. *J. Chem. Inf. Model.* **2010**, *50*, 2029–40.
- (6) Hajduk, P. J.; Huth, J. R.; Tse, C. Predicting protein druggability. *Drug Discovery Today* **2005**, *10*, 1675–82.
- (7) 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*, 2518–25.
- (8) 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*, 71–5.
- (9) Seco, J.; Luque, F. J.; Barril, X. Binding site detection and druggability index from first principles. *J. Med. Chem.* **2009**, *52*, 2363–71.
- (10) Halgren, T. A. Identifying and characterizing binding sites and assessing druggability. *J. Chem. Inf. Model.* **2009**, *49*, 377–89.

- (11) Schmidtke, P.; Barril, X. Understanding and predicting druggability. A high-throughput method for detection of drug binding sites. *J. Med. Chem.* **2010**, *53*, 5858–67.
- (12) Krasowski, A.; Muthas, D.; Sarkar, A.; Schmitt, S.; Brenk, R. Drugpred: A structure-based approach to predict protein druggability developed using an extensive nonredundant data set. *J. Chem. Inf. Model.* **2011**, *51*, 2829–42.
- (13) Fauman, E. B.; Rai, B. K.; Huang, E. S. Structure-based druggability assessment - identifying suitable targets for small molecule therapeutics. *Curr. Opin. Chem. Biol.* **2011**, *15*, 463–8.
- (14) Volkamer, A.; Kuhn, D.; Rippmann, F.; Rarey, M. Dogsitescorer: A web server for automatic binding site prediction, analysis and druggability assessment. *Bioinformatics* **2012**, *28*, 2074–5.
- (15) Nisius, B.; Sha, F.; Gohlke, H. Structure-based computational analysis of protein binding sites for function and druggability prediction. *J. Biotechnol.* **2012**, *159*, 123–34.
- (16) Perola, E.; Herman, L.; Weiss, J. Development of a rule-based method for the assessment of protein druggability. *J. Chem. Inf. Model.* **2012**, *52*, 1027–38.
- (17) Volkamer, A.; Kuhn, D.; Grömbacher, T.; Rippmann, F.; Rarey, M. Combining global and local measures for structure-based druggability predictions. *J. Chem. Inf. Model.* **2012**, *52*, 360–72.
- (18) Driggers, E. M.; Hale, S. P.; Lee, J.; Terrett, N. K. The exploration of macrocycles for drug discovery—an underexploited structural class. *Nat. Rev. Drug Discovery* **2008**, *7*, 608–24.
- (19) Wessjohann, L. A.; Ruijter, E.; García-Rivera, D.; Brandt, W. What can a chemist learn from nature's macrocycles?—A brief, conceptual view. *Mol. Diversity* **2005**, *9*, 171–86.
- (20) Marsault, E.; Peterson, M. L. Macrocycles are great cycles: Applications, opportunities, and challenges of synthetic macrocycles in drug discovery. *J. Med. Chem.* **2011**, *54*, 1961–2004.
- (21) Mallinson, J.; Collins, I. Macrocycles in new drug discovery. *Future Med. Chem.* **2012**, *4*, 1409–38.
- (22) Kalgutkar, A. S.; Dalvie, D. K. Drug discovery for a new generation of covalent drugs. *Expert Opin. Drug Discovery* **2012**, *7*, 561–81.
- (23) Singh, J.; Petter, R. C.; Baillie, T. A.; Whitty, A. The resurgence of covalent drugs. *Nat. Rev. Drug Discovery* **2011**, *10*, 307–17.
- (24) Way, J. C. Covalent modification as a strategy to block protein-protein interactions with small-molecule drugs. *Curr. Opin. Chem. Biol.* **2000**, *4*, 40–6.
- (25) Edwards, T. A.; Wilson, A. J. Helix-mediated protein–protein interactions as targets for intervention using foldamers. *Amino Acids* **2011**, *41*, 743–54.
- (26) Murray, J. K.; Gellman, S. H. Targeting protein-protein interactions: Lessons from p53/MDM2. *Biopolymers* **2007**, *88*, 657–86.
- (27) Raj, M.; Bullock, B. N.; Arora, P. S. Plucking the high hanging fruit: A systematic approach for targeting protein-protein interactions. *Bioorg. Med. Chem.* **2013**, *21*, 4051–7.
- (28) Walensky, L. D.; Kung, A. L.; Escher, I.; Malia, T. J.; Barbuto, S.; Wright, R. D.; Wagner, G.; Verdine, G. L.; Korsmeyer, S. J. Activation of apoptosis in vivo by a hydrocarbon-stapled BH3 helix. *Science* **2004**, *305*, 1466–70.
- (29) Brown, S. P.; Hajduk, P. J. Effects of conformational dynamics on predicted protein druggability. *ChemMedChem* **2006**, *1*, 70–2.
- (30) Clackson, T.; Wells, J. A. A hot spot of binding energy in a hormone-receptor interface. *Science* **1995**, *267*, 383–6.
- (31) DeLano, W. L. Unraveling hot spots in binding interfaces: Progress and challenges. *Curr. Opin. Struct. Biol.* **2002**, *12*, 14–20.
- (32) Keskin, O.; Ma, B. Y.; Nussinov, R. Hot regions in protein-protein interactions: The organization and contribution of structurally conserved hot spot residues. *J. Mol. Biol.* **2005**, *345*, 1281–94.
- (33) Bogan, A. A.; Thorn, K. S. Anatomy of hot spots in protein interfaces. *J. Mol. Biol.* **1998**, *280*, 1–9.
- (34) Kortemme, T.; Baker, D. A simple physical model for binding energy hot spots in protein-protein complexes. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 14116–21.
- (35) Mattos, C.; Ringe, D. Locating and characterizing binding sites on proteins. *Nat. Biotechnol.* **1996**, *14*, 595–9.
- (36) Allen, K. N.; Bellamacina, C. R.; Ding, X. C.; Jeffery, C. J.; Mattos, C.; Petsko, G. A.; Ringe, D. An experimental approach to mapping the binding surfaces of crystalline proteins. *J. Phys. Chem.* **1996**, *100*, 2605–11.
- (37) English, A. C.; Done, S. H.; Caves, L. S.; Groom, C. R.; Hubbard, R. E. Locating interaction sites on proteins: The crystal structure of thermolysin soaked in 2% to 100% isopropanol. *Proteins: Struct., Funct., Genet.* **1999**, *37*, 628–40.
- (38) English, A. C.; Groom, C. R.; Hubbard, R. E. Experimental and computational mapping of the binding surface of a crystalline protein. *Protein Eng., Des. Sel.* **2001**, *14*, 47–59.
- (39) Dechene, M.; Wink, G.; Smith, M.; Swartz, P.; Mattos, C. Multiple solvent crystal structures of ribonuclease A: An assessment of the method. *Proteins: Struct., Funct., Genet.* **2009**, *76*, 861–81.
- (40) Buhrman, G.; O'Connor, C.; Zerbe, B.; Kearney, B. M.; Napoleon, R.; Kovrigina, E. A.; Vajda, S.; Kozakov, D.; Kovrigin, E. L.; Mattos, C. Analysis of binding site hot spots on the surface of Ras GTPase. *J. Mol. Biol.* **2011**, *413*, 773–89.
- (41) Edfeldt, F. N. B.; Folmer, R. H. A.; Breeze, A. L. Fragment screening to predict druggability (ligandability) and lead discovery success. *Drug Discovery Today* **2011**, *16*, 284–287.
- (42) Barelier, S.; Krimm, I. Ligand specificity, privileged substructures and protein druggability from fragment-based screening. *Curr. Opin. Chem. Biol.* **2011**, *15*, 469–74.
- (43) Dennis, S.; Kortvelyesi, T.; Vajda, S. Computational mapping identifies the binding sites of organic solvents on proteins. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 4290–5.
- (44) Landon, M. R.; Lancia, D. R., Jr.; Yu, J.; Thiel, S. C.; Vajda, S. Identification of hot spots within druggable binding regions by computational solvent mapping of proteins. *J. Med. Chem.* **2007**, *50*, 1231–40.
- (45) Brenke, R.; Kozakov, D.; Chuang, G. Y.; Beglov, D.; Hall, D.; Landon, M. R.; Mattos, C.; Vajda, S. Fragment-based identification of druggable 'hot spots' of proteins using Fourier domain correlation techniques. *Bioinformatics* **2009**, *25*, 621–7.
- (46) Landon, M. R.; Lieberman, R. L.; Hoang, Q. Q.; Ju, S.; Caaveiro, J. M.; Orwig, S. D.; Kozakov, D.; Brenke, R.; Chuang, G. Y.; Beglov, D.; Vajda, S.; Petsko, G. A.; Ringe, D. Detection of ligand binding hot spots on protein surfaces via fragment-based methods: Application to DJ-1 and glucocerebrosidase. *J. Comput.-Aided Mol. Des.* **2009**, *23*, 491–500.
- (47) Kozakov, D.; Chuang, G. Y.; Beglov, D.; Vajda, S. Where does amantadine bind to the influenza virus M2 proton channel? *Trends Biochem. Sci.* **2010**, *35*, 471–5.
- (48) Kozakov, D.; Hall, D. R.; Chuang, G. Y.; Cencic, R.; Brenke, R.; Grove, L. E.; Beglov, D.; Pelletier, J.; Whitty, A.; Vajda, S. Structural conservation of druggable hot spots in protein-protein interfaces. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 13528–33.
- (49) Hall, D. H.; Grove, L. E.; Yueh, C.; Ngan, C. H.; Kozakov, D.; Vajda, S. Robust identification of binding hot spots using continuum electrostatics: Application to hen egg-white lysozyme. *J. Am. Chem. Soc.* **2011**, *133*, 20668–71.
- (50) Hall, D. R.; Ngan, C. H.; Zerbe, B. S.; Kozakov, D.; Vajda, S. Hot spot analysis for driving the development of hits into leads in fragment-based drug discovery. *J. Chem. Inf. Model.* **2012**, *52*, 199–209.
- (51) Golden, M. S.; Cote, S. M.; Sayeg, M.; Zerbe, B. S.; Villar, E. A.; Beglov, D.; Sazinsky, S. L.; Georgiadis, R. M.; Vajda, S.; Kozakov, D.; Whitty, A. Comprehensive experimental and computational analysis of binding energy hot spots at the NF-κB essential modulator/IKKβ protein-protein interface. *J. Am. Chem. Soc.* **2013**, *135*, 6242–56.
- (52) Kozakov, D.; Grove, L. E.; Hall, D. R.; Bohnuud, T.; Mottarella, S. E.; Luo, L.; Xia, B.; Beglov, D.; Vajda, S. The FTMap family of web servers for determining and characterizing ligand-binding hot spots of proteins. *Nat. Protoc.* **2015**, *10*, 733–55.
- (53) Liepinsh, E.; Otting, G. Organic solvents identify specific ligand binding sites on protein surfaces. *Nat. Biotechnol.* **1997**, *15*, 264–8.

- (54) Mattos, C.; Bellamacina, C. R.; Peisach, E.; Pereira, A.; Vitkup, D.; Petsko, G. A.; Ringe, D. Multiple solvent crystal structures: Probing binding sites, plasticity and hydration. *J. Mol. Biol.* **2006**, *357*, 1471–82.
- (55) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28*, 235–42.
- (56) Murray, C. W.; Rees, D. C. The rise of fragment-based drug discovery. *Nat. Chem.* **2009**, *1*, 187–192.
- (57) Chen, Y.; Shoichet, B. K. Molecular docking and ligand specificity in fragment-based inhibitor discovery. *Nat. Chem. Biol.* **2009**, *5*, 358–64.
- (58) Warr, W. A. Fragment-based drug discovery. *J. Comput.-Aided Mol. Des.* **2009**, *23*, 453–8.
- (59) Congreve, M.; Chessari, G.; Tisi, D.; Woodhead, A. J. Recent developments in fragment-based drug discovery. *J. Med. Chem.* **2008**, *51*, 3661–80.
- (60) Carr, R. A.; Congreve, M.; Murray, C. W.; Rees, D. C. Fragment-based lead discovery: Leads by design. *Drug Discovery Today* **2005**, *10*, 987–92.
- (61) Villar, E. A.; Beglov, D.; Chennamadhavuni, S.; Porco, J. A.; Kozakov, D.; Vajda, S.; Whitty, A. How proteins bind macrocycles. *Nat. Chem. Biol.* **2014**, *10*, 723–31.
- (62) Ganeshan, A. The impact of natural products upon modern drug discovery. *Curr. Opin. Chem. Biol.* **2008**, *12*, 306–17.
- (63) Goodman, C. M.; Choi, S.; Shandler, S.; DeGrado, W. F. Foldamers as versatile frameworks for the design and evolution of function. *Nat. Chem. Biol.* **2007**, *3*, 252–262.
- (64) Zuckermann, R. N.; Kodadek, T. Peptoids as potential therapeutics. *Curr. Opin. Mol. Ther.* **2009**, *11*, 299–307.
- (65) Dietrich, U.; Durr, R.; Koch, J. Peptides as drugs: From screening to application. *Curr. Pharm. Biotechnol.* **2013**, *14*, 501–12.
- (66) Wilson, K. P.; Yamashita, M. M.; Sintchak, M. D.; Rotstein, S. H.; Murcko, M. A.; Boger, J.; Thomson, J. A.; Fitzgibbon, M. J.; Black, J. R.; Navia, M. A. Comparative x-ray structures of the major binding-protein for the immunosuppressant FK506 (tacrolimus) in unliganded form and in complex with FK506 and rapamycin. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **1995**, *51*, 511–21.
- (67) Sattler, M.; Liang, H.; Nettesheim, D.; Meadows, R. P.; Harlan, J. E.; Eberstadt, M.; Yoon, H. S.; Shuker, S. B.; Chang, B. S.; Minn, A. J.; Thompson, C. B.; Fesik, S. W. Structure of Bcl-x(L)-BAK peptide complex: Recognition between regulators of apoptosis. *Science* **1997**, *275*, 983–6.
- (68) Holt, D. A.; Luengo, J. I.; Yamashita, D. S.; Oh, H. J.; Konalians, A. L.; Yen, H. K.; Rozamus, L. W.; Brandt, M.; Bossard, M. J.; Levy, M. A.; Eggleston, D. S.; Liang, J.; Schultz, L. W.; Stout, T. J.; Clardy, J. Design, synthesis, and kinetic evaluation of high-affinity FKBP ligands and the X-ray crystal-structures of their complexes with FKBP12. *J. Am. Chem. Soc.* **1993**, *115*, 9925–38.
- (69) Sich, C.; Impronta, S.; Cowley, D. J.; Guenet, C.; Merly, J. P.; Teufel, M.; Saudek, V. Solution structure of a neurotrophic ligand bound to FKBP12 and its effects on protein dynamics. *Eur. J. Biochem.* **2000**, *267*, 5342–54.
- (70) Silberstein, M.; Dennis, S.; Brown, L.; Kortvelyesi, T.; Clodfelter, K.; Vajda, S. Identification of substrate binding sites in enzymes by computational solvent mapping. *J. Mol. Biol.* **2003**, *332*, 1095–113.
- (71) Bohnuud, T.; Kozakov, D.; Vajda, S. Evidence of conformational selection driving the formation of ligand binding sites in protein-protein interfaces. *PLoS Comput. Biol.* **2014**, *10*, e1003872.
- (72) Wang, R.; Fang, X.; Lu, Y.; Wang, S. The PDBbind database: Collection of binding affinities for protein-ligand complexes with known three-dimensional structures. *J. Med. Chem.* **2004**, *47*, 2977–80.
- (73) Liu, T.; Lin, Y.; Wen, X.; Jorissen, R. N.; Gilson, M. K. BindingDB: A web-accessible database of experimentally determined protein-ligand binding affinities. *Nucleic Acids Res.* **2007**, *35*, D198–201.
- (74) Benson, M. L.; Smith, R. D.; Khazanov, N. A.; Dimcheff, B.; Beaver, J.; Dresslar, P.; Nerothin, J.; Carlson, H. A. Binding MOAD, a high-quality protein-ligand database. *Nucleic Acids Res.* **2007**, *36*, D674–8.
- (75) Topalis, D.; Kumamoto, H.; Velasco, M. F. A.; Dugue, L.; Haouz, A.; Alexandre, J. A. C.; Gallois-Montbrun, S.; Alzari, P. M.; Pochet, S.; Agrofoglio, L. A.; Deville-Bonne, D. Nucleotide binding to human UMP-CMP kinase using fluorescent derivatives - a screening based on affinity for the UMP-CMP binding site. *FEBS J.* **2007**, *274*, 3704–14.
- (76) Mashhoon, N.; Pruss, C.; Carroll, M.; Johnson, P. H.; Reich, N. O. Selective inhibitors of bacterial DNA adenine methyltransferases. *J. Biomol. Screening* **2006**, *11*, 497–510.
- (77) Miao, B. C.; Skidan, I.; Yang, J. S.; Lugovskoy, A.; Reibarkh, M.; Long, K.; Brazell, T.; Durugkar, K. A.; Maki, J.; Ramana, C. V.; Schaffhausen, B.; Wagner, G.; Torchilin, V.; Yuan, J. Y.; Degterev, A. Small molecule inhibition of phosphatidylinositol-3,4,5-triphosphate (PIP₃) binding to pleckstrin homology domains. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 20126–31.
- (78) Lerner, C. G.; Hajduk, P. J.; Wagner, R.; Wagenaar, F. L.; Woodall, C.; Gu, Y. G.; Searle, X. B.; Florjancic, A. S.; Zhang, T.; Clark, R. F.; Cooper, C. S.; Mack, J. C.; Yu, L.; Cai, M.; Betz, S. F.; Chovan, L. E.; McCall, J. O.; Black-Schaefer, C. L.; Kakavas, S. J.; Schurdak, M. E.; Comess, K. M.; Walter, K. A.; Edalji, R.; Dorwin, S. A.; Smith, R. A.; Hebert, E. J.; Harlan, J. E.; Metzger, R. E.; Merta, P. J.; Baranowski, J. L.; Coen, M. L.; Thornewell, S. J.; Shivakumar, A. G.; Saiki, A. Y.; Soni, N.; Bui, M.; Balli, D. J.; Sanders, W. J.; Nilius, A. M.; Holzman, T. F.; Fesik, S. W.; Beutel, B. A. From bacterial genomes to novel antibacterial agents: Discovery, characterization, and antibacterial activity of compounds that bind to HI0065 (YjeE) from *Haemophilus influenzae*. *Chem. Biol. Drug Des.* **2007**, *69*, 395–404.
- (79) Piserchio, A.; Salinas, G. D.; Li, T.; Marshall, J.; Spaller, M. R.; Mierke, D. F. Targeting specific PDZ domains of PSD-95: Structural basis for enhanced affinity and enzymatic stability of a cyclic peptide. *Chem. Biol.* **2004**, *11*, 469–73.
- (80) Wintjens, R.; Wieruszkeski, J. M.; Drobecq, H.; Rousselot-Pailley, P.; Buee, L.; Lippens, G.; Landrieu, I. ¹H NMR study on the binding of Pin1 Trp-Trp domain with phosphothreonine peptides. *J. Biol. Chem.* **2001**, *276*, 25150–6.
- (81) Guo, C. X.; Hou, X. J.; Dong, L. M.; Dagostino, E.; Greasley, S.; Ferre, R.; Marakovits, J.; Johnson, M. C.; Matthews, D.; Mroczkowski, B.; Parge, H.; VanArsdale, T.; Popoff, I.; Piraino, J.; Margosiak, S.; Thomson, J.; Los, G.; Murray, B. W. Structure-based design of novel human Pin1 inhibitors (I). *Bioorg. Med. Chem. Lett.* **2009**, *19*, 5613–6.
- (82) Dong, L. M.; Marakovits, J.; Hou, X. J.; Guo, C. X.; Greasley, S.; Dagostino, E.; Ferre, R.; Johnson, M. C.; Kraynov, E.; Thomson, J.; Pathak, V.; Murray, B. W. Structure-based design of novel human Pin1 inhibitors (II). *Bioorg. Med. Chem. Lett.* **2010**, *20*, 2210–4.
- (83) Guo, C. X.; Hou, X. J.; Dong, L. M.; Marakovits, J.; Greasley, S.; Dagostino, E.; Ferre, R.; Johnson, M. C.; Humphries, P. S.; Li, H. T.; Paderes, G. D.; Piraino, J.; Kraynov, E.; Murray, B. W. Structure-based design of novel human Pin1 inhibitors (III): Optimizing affinity beyond the phosphate recognition pocket. *Bioorg. Med. Chem. Lett.* **2014**, *24*, 4187–91.
- (84) Wendt, M. D.; Sun, C. H.; Kunzer, A.; Sauer, D.; Sarris, K.; Hoff, E.; Yu, L. P.; Nettesheim, D. G.; Chen, J.; Jin, S.; Comess, K. M.; Fan, Y. H.; Anderson, S. N.; Isaac, B.; Olejniczak, E. T.; Hajduk, P. J.; Rosenberg, S. H.; Elmore, S. W. Discovery of a novel small molecule binding site of human survivin. *Bioorg. Med. Chem. Lett.* **2007**, *17*, 3122–9.
- (85) Huth, J. R.; Yu, L.; Collins, I.; Mack, J.; Mendoza, R.; Isaac, B.; Braddock, D. T.; Muchmore, S. W.; Comess, K. M.; Fesik, S. W.; Clore, G. M.; Levens, D.; Hajduk, P. J. NMR-driven discovery of benzoylanthranilic acid inhibitors of far upstream element binding protein binding to the human oncogene c-myc promoter. *J. Med. Chem.* **2004**, *47*, 4851–7.
- (86) Hajduk, P. J.; Dinges, J.; Miknis, G. F.; Merlock, M.; Middleton, T.; Kempf, D. J.; Egan, D. A.; Walter, K. A.; Robins, T. S.; Shuker, S. B.; Holzman, T. F.; Fesik, S. W. NMR-based discovery of lead

- inhibitors that block DNA binding of the human papillomavirus E2 protein. *J. Med. Chem.* **1997**, *40*, 3144–50.
- (87) Schaal, T. D.; Mallet, W. G.; McMinn, D. L.; Nguyen, N. V.; Sopko, M. M.; John, S.; Parekh, B. S. Inhibition of human papilloma virus E2 DNA binding protein by covalently linked polyamides. *Nucleic Acids Res.* **2003**, *31*, 1282–91.
- (88) Kurg, R.; Langel, U.; Ustav, M. Inhibition of the bovine papillomavirus E2 protein activity by peptide nucleic acid. *Virus Res.* **2000**, *66*, 39–50.
- (89) de Dios, A.; Prieto, L.; Martin, J. A.; Rubio, A.; Ezquerra, J.; Tebbe, M.; Lopez de Uralde, B.; Martin, J.; Sanchez, A.; LeTourneau, D. L.; McGee, J. E.; Boylan, C.; Parr, T. R., Jr.; Smith, M. C. 4-substituted D-glutamic acid analogues: The first potent inhibitors of glutamate racemase (MurI) enzyme with antibacterial activity. *J. Med. Chem.* **2002**, *45*, 4559–70.
- (90) Geng, B.; Breault, G.; Comita-Prevoir, J.; Petrichko, R.; Eyermann, C.; Lundqvist, T.; Doig, P.; Gorseth, E.; Noonan, B. Exploring 9-benzyl purines as inhibitors of glutamate racemase (MurI) in gram-positive bacteria. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 4368–4372.
- (91) Breault, G. A.; Comita-Prevoir, J.; Eyermann, C. J.; Geng, B.; Petrichko, R.; Doig, P.; Gorseth, E.; Noonan, B. Exploring 8-benzyl pteridine-6,7-diones as inhibitors of glutamate racemase (MurI) in gram-positive bacteria. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 6100–6103.
- (92) McRee, D. E.; Jensen, G. M.; Fitzgerald, M. M.; Siegel, H. A.; Goodin, D. B. Construction of a bisquo heme enzyme and binding by exogenous ligands. *Proc. Natl. Acad. Sci. U. S. A.* **1994**, *91*, 12847–51.
- (93) Appleby, C. A.; Wittenberg, B. A.; Wittenberg, J. B. Nicotinic acid as a ligand affecting leghemoglobin structure and oxygen reactivity. *Proc. Natl. Acad. Sci. U. S. A.* **1973**, *70*, 564–8.
- (94) Maldonado, S.; Lostao, A.; Irun, M. P.; Fernandez-Recio, J.; Genzor, C. G.; Gonzalez, E. B.; Rubio, J. A.; Luquita, A.; Daoudi, F.; Sancho, J. Apoflavodoxin: Structure, stability, and FMN binding. *Biochimie* **1998**, *80*, 813–20.
- (95) Irie, K.; Oie, K.; Nakahara, A.; Yanai, Y.; Ohigashi, H.; Wender, P. A.; Fukuda, H.; Konishi, H.; Kikkawa, U. Molecular basis for protein kinase c isozyme-selective binding: The synthesis, folding, and phorbol ester binding of the cysteine-rich domains of all protein kinase c isozymes. *J. Am. Chem. Soc.* **1998**, *120*, 9159–67.
- (96) Murray, I. A.; Cann, P. A.; Day, P. J.; Derrick, J. P.; Sutcliffe, M. J.; Shaw, W. V.; Leslie, A. G. W. Steroid recognition by chloramphenicol acetyltransferase - engineering and structural-analysis of a high-affinity fusidic acid-binding site. *J. Mol. Biol.* **1995**, *254*, 993–1005.
- (97) Vyas, M. N.; Vyas, N. K.; Quiocho, F. A. Crystallographic analysis of the epimeric and anomeric specificity of the periplasmic transport chemosensory protein-receptor for D-glucose and D-galactose. *Biochemistry* **1994**, *33*, 4762–8.
- (98) Chung, S. J.; Fromme, J. C.; Verdine, G. L. Structure of human cytidine deaminase bound to a potent inhibitor. *J. Med. Chem.* **2005**, *48*, 658–60.
- (99) Masaki, T.; Tanaka, T.; Tsunashima, S.; Sakiyama, F.; Soejima, M. Inhibition of achromobacter protease-I by lysinal derivatives. *Biosci., Biotechnol., Biochem.* **1992**, *56*, 1604–7.
- (100) Lunney, E. A.; Para, K. S.; Rubin, J. R.; Humblet, C.; Fergus, J. H.; Marks, J. S.; Sawyer, T. K. Structure-based design of a novel series of nonpeptide ligands that bind to the pp60src SH2 domain. *J. Am. Chem. Soc.* **1997**, *119*, 12471–6.
- (101) Fauman, E. B.; Rai, B. K.; Huang, E. S. Structure-based druggability assessment—identifying suitable targets for small molecule therapeutics. *Curr. Opin. Chem. Biol.* **2011**, *15*, 463–8.
- (102) Schonthal, A. H. Direct non-cyclooxygenase-2 targets of celecoxib and their potential relevance for cancer therapy. *Br. J. Cancer* **2007**, *97*, 1465–8.
- (103) Stangier, J.; Clemens, A. Pharmacology, pharmacokinetics, and pharmacodynamics of dabigatran etexilate, an oral direct thrombin inhibitor. *Clin. Appl. Thromb./Hemostasis* **2009**, *15*, 9S–16S.
- (104) Kong, Y.; Chen, H.; Wang, Y. Q.; Meng, L.; Wei, J. F. Direct thrombin inhibitors: Patents 2002–2012 (review). *Mol. Med. Rep.* **2014**, *9*, 1506–14.
- (105) Fu, K. P.; Neu, H. C. Piperacillin, a new penicillin active against many bacteria resistant to other penicillins. *Antimicrob. Agents Chemother.* **1978**, *13*, 358–67.
- (106) van Berkel, S. S.; Nettleship, J. E.; Leung, I. K.; Brem, J.; Choi, H.; Stuart, D. I.; Claridge, T. D.; McDonough, M. A.; Owens, R. J.; Ren, J.; Schofield, C. J. Binding of (S)-penicilloic acid to penicillin binding protein 3. *ACS Chem. Biol.* **2013**, *8*, 2112–6.
- (107) Patchett, A. A.; Harris, E.; Tristram, E. W.; Wyvratt, M. J.; Wu, M. T.; Taub, D.; Peterson, E. R.; Ikeler, T. J.; ten Broeke, J.; Payne, L. G.; Ondeyka, D. L.; Thorsett, E. D.; Greenlee, W. J.; Lohr, N. S.; Hoffsommer, R. D.; Joshua, H.; Ruyle, W. V.; Rothrock, J. W.; Aster, S. D.; Maycock, A. L.; Robinson, F. M.; Hirschmann, R.; Sweet, C. S.; Ulm, E. H.; Gross, D. M.; Vassil, T. C.; Stone, C. A. A new class of angiotensin-converting enzyme inhibitors. *Nature* **1980**, *288*, 280–3.
- (108) Gubareva, L. V.; Kaiser, L.; Hayden, F. G. Influenza virus neuraminidase inhibitors. *Lancet* **2000**, *355*, 827–35.
- (109) Maignan, S.; Guilloteau, J. P.; Pouzieux, S.; Choi-Sledeski, Y. M.; Becker, M. R.; Klein, S. I.; Ewing, W. R.; Pauls, H. W.; Spada, A. P.; Mikol, V. Crystal structures of human factor Xa complexed with potent inhibitors. *J. Med. Chem.* **2000**, *43*, 3226–32.
- (110) Istvan, E. S.; Deisenhofer, J. Structural mechanism for statin inhibition of HMG-CoA reductase. *Science* **2001**, *292*, 1160–4.
- (111) Barrett, D. G.; Deaton, D. N.; Hassell, A. M.; McFadyen, R. B.; Miller, A. B.; Miller, L. R.; Payne, J. A.; Shewchuk, L. M.; Willard, D. H., Jr.; Wright, L. L. Acyclic cyanamide-based inhibitors of cathepsin K. *Bioorg. Med. Chem. Lett.* **2005**, *15*, 3039–43.
- (112) Gauthier, J. Y.; Chauret, N.; Cromlish, W.; Desmarais, S.; Duong, L. T.; Falgueyret, J. P.; Kimmel, D. B.; Lamontagne, S.; Leger, S.; LeRiche, T.; Li, C. S.; Masse, F.; McKay, D. J.; Nicoll-Griffith, D. A.; Oballa, R. M.; Palmer, J. T.; Percival, M. D.; Riendeau, D.; Robichaud, J.; Rodan, G. A.; Rodan, S. B.; Seto, C.; Therien, M.; Truong, V. L.; Venuti, M. C.; Wesolowski, G.; Young, R. N.; Zamboni, R.; Black, W. C. The discovery of odanacatib (MK-0822), a selective inhibitor of cathepsin K. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 923–8.
- (113) Brixen, K.; Chapurlat, R.; Cheung, A. M.; Keaveny, T. M.; Fuerst, T.; Engelke, K.; Recker, R.; Dardzinski, B.; Verbruggen, N.; Ather, S.; Rosenberg, E.; de Papp, A. E. Bone density, turnover, and estimated strength in postmenopausal women treated with odanacatib: A randomized trial. *J. Clin. Endocrinol. Metab.* **2013**, *98*, 571–80.
- (114) Summa, V.; Petrocchi, A.; Bonelli, F.; Crescenzi, B.; Donghi, M.; Ferrara, M.; Fiore, F.; Gardelli, C.; Gonzalez Paz, O.; Hazuda, D. J.; Jones, P.; Kinzel, O.; Laufer, R.; Monteagudo, E.; Muraglia, E.; Nizi, E.; Orvieto, F.; Pace, P.; Pescatore, G.; Scarpelli, R.; Stillmack, K.; Witmer, M. V.; Rowley, M. Discovery of raltegravir, a potent, selective orally bioavailable HIV-integrase inhibitor for the treatment of HIV-AIDS infection. *J. Med. Chem.* **2008**, *51*, 5843–55.
- (115) Leung, D.; Abbenante, G.; Fairlie, D. P. Protease inhibitors: Current status and future prospects. *J. Med. Chem.* **2000**, *43*, 305–41.
- (116) Reigan, P.; Edwards, P. N.; Gbjaj, A.; Cole, C.; Barry, S. T.; Page, K. M.; Ashton, S. E.; Luke, R. W. A.; Douglas, K. T.; Stratford, I. J.; Jaffar, M.; Bryce, R. A.; Freeman, S. Aminoimidazolylmethyluracil analogues as potent inhibitors of thymidine phosphorylase and their bioreductive nitroimidazolyl prodrugs. *J. Med. Chem.* **2005**, *48*, 392–402.
- (117) Manikowski, A.; Verri, A.; Lossani, A.; Gebhardt, B. M.; Gambino, J.; Focher, F.; Spadari, S.; Wright, G. E. Inhibition of herpes simplex virus thymidine kinases by 2-phenylamino-6-oxopurines and related compounds: Structure-activity relationships and antiherpetic activity in vivo. *J. Med. Chem.* **2005**, *48*, 3919–29.
- (118) Clements, J. M.; Beckett, R. P.; Brown, A.; Catlin, G.; Lobell, M.; Palan, S.; Thomas, W.; Whittaker, M.; Wood, S.; Salama, S.; Baker, P. J.; Rodgers, H. F.; Barynin, V.; Rice, D. W.; Hunter, M. G. Antibiotic activity and characterization of BB-3497, a novel peptide deformylase inhibitor. *Antimicrob. Agents Chemother.* **2001**, *45*, 563–70.
- (119) Koropatkin, N. M.; Cleland, W. W.; Holden, H. M. Kinetic and structural analysis of alpha-D-glucose-1-phosphate cytidyltransferase from salmonella typhi. *J. Biol. Chem.* **2005**, *280*, 10774–80.

- (120) Heffron, S. E.; Jurnak, F. Structure of an EF-Tu complex with a thiazolyl peptide antibiotic determined at 2.35 Å resolution: Atomic basis for GE2270A inhibition of EF-Tu. *Biochemistry* **2000**, *39*, 37–45.
- (121) Wells, J. A.; McClelland, C. L. Reaching for high-hanging fruit in drug discovery at protein-protein interfaces. *Nature* **2007**, *450*, 1001–9.
- (122) Whitty, A.; Kumaravel, G. Between a rock and a hard place? *Nat. Chem. Biol.* **2006**, *2*, 112–8.
- (123) Fuller, J. C.; Burgoyne, N. J.; Jackson, R. M. Predicting druggable binding sites at the protein-protein interface. *Drug Discovery Today* **2009**, *14*, 155–61.
- (124) Barelier, S.; Pons, J.; Marcillat, O.; Lancelin, J. M.; Krimm, I. Fragment-based deconstruction of Bcl-x(L) inhibitors. *J. Med. Chem.* **2010**, *53*, 2577–88.
- (125) Oltersdorf, T.; Elmore, S. W.; Shoemaker, A. R.; Armstrong, R. C.; Augeri, D. J.; Belli, B. A.; Bruncko, M.; Deckwerth, T. L.; Dinges, J.; Hajduk, P. J.; Joseph, M. K.; Kitada, S.; Korsmeyer, S. J.; Kunzer, A. R.; Letai, A.; Li, C.; Mitten, M. J.; Nettesheim, D. G.; Ng, S.; Nimmer, P. M.; O'Connor, J. M.; Oleksijew, A.; Petros, A. M.; Reed, J. C.; Shen, W.; Tahir, S. K.; Thompson, C. B.; Tomaselli, K. J.; Wang, B.; Wendt, M. D.; Zhang, H.; Fesik, S. W.; Rosenberg, S. H. An inhibitor of Bcl-2 family proteins induces regression of solid tumours. *Nature* **2005**, *435*, 677–81.
- (126) Tse, C.; Shoemaker, A. R.; Adickes, J.; Anderson, M. G.; Chen, J.; Jin, S.; Johnson, E. F.; Marsh, K. C.; Mitten, M. J.; Nimmer, P.; Roberts, L.; Tahir, S. K.; Xiao, Y.; Yang, X.; Zhang, H.; Fesik, S.; Rosenberg, S. H.; Elmore, S. W. Abt-263: A potent and orally bioavailable Bcl-2 family inhibitor. *Cancer Res.* **2008**, *68*, 3421–8.
- (127) Tao, Z. F.; Hasvold, L.; Wang, L.; Wang, X.; Petros, A. M.; Park, C. H.; Boghaert, E. R.; Catron, N. D.; Chen, J.; Colman, P. M.; Czabotar, P. E.; Deshayes, K.; Fairbrother, W. J.; Flygare, J. A.; Hymowitz, S. G.; Jin, S.; Judge, R. A.; Koehler, M. F.; Kovar, P. J.; Lessene, G.; Mitten, M. J.; Ndubaku, C. O.; Nimmer, P.; Purkey, H. E.; Oleksijew, A.; Phillips, D. C.; Sleebs, B. E.; Smith, B. J.; Smith, M. L.; Tahir, S. K.; Watson, K. G.; Xiao, Y.; Xue, J.; Zhang, H.; Zobel, K.; Rosenberg, S. H.; Tse, C.; Leverson, J. D.; Elmore, S. W.; Souers, A. J. Discovery of a potent and selective bcl-xL inhibitor with in vivo activity. *ACS Med. Chem. Lett.* **2014**, *5*, 1088–93.
- (128) Flygare, J. A.; Beresini, M.; Budha, N.; Chan, H.; Chan, I. T.; Cheeti, S.; Cohen, F.; Deshayes, K.; Doerner, K.; Eckhardt, S. G.; Elliott, L. O.; Feng, B. N.; Franklin, M. C.; Reisner, S. F.; Gazzard, L.; Halladay, J.; Hymowitz, S. G.; La, H.; LoRusso, P.; Maurer, B.; Murray, L.; Plise, E.; Quan, C.; Stephan, J. P.; Young, S. G.; Tom, J.; Tsui, V.; Um, J.; Varfolomeev, E.; Vucic, D.; Wagner, A. J.; Wallweber, H. J. A.; Wang, L.; Ware, J.; Wen, Z. Y.; Wong, H.; Wong, J. M.; Wong, M.; Wong, S.; Yu, R.; Zobel, K.; Fairbrother, W. J. Discovery of a potent small-molecule antagonist of inhibitor of apoptosis (IAP) proteins and clinical candidate for the treatment of cancer (GDC-0152). *J. Med. Chem.* **2012**, *55*, 4101–13.
- (129) Seigal, B. A.; Connors, W. H.; Fraley, A.; Borzilleri, R. M.; Carter, P. H.; Emanuel, S. L.; Farnoli, J.; Kim, K.; Lei, M.; Naglich, J. G.; Pokross, M. E.; Posy, S. L.; Shen, H.; Surti, N.; Talbot, R.; Zhang, Y.; Terrett, N. K. The discovery of macrocyclic XIAP antagonists from a DNA-programmed chemistry library, and their optimization to give lead compounds with in vivo antitumor activity. *J. Med. Chem.* **2015**, *58*, 2855–61.
- (130) Vassilev, L. T. MDM2 inhibitors for cancer therapy. *Trends Mol. Med.* **2007**, *13*, 23–31.
- (131) Zhao, Y.; Aguilar, A.; Bernard, D.; Wang, S. Small-molecule inhibitors of the MDM2-p53 protein-protein interaction (MDM2 inhibitors) in clinical trials for cancer treatment. *J. Med. Chem.* **2015**, *58*, 1038–52.
- (132) Sun, D.; Li, Z.; Rew, Y.; Gribble, M.; Bartberger, M. D.; Beck, H. P.; Canon, J.; Chen, A.; Chen, X.; Chow, D.; Deignan, J.; Duquette, J.; Eksterowicz, J.; Fisher, B.; Fox, B. M.; Fu, J.; Gonzalez, A. Z.; Gonzalez-Lopez De Turiso, F.; Houze, J. B.; Huang, X.; Jiang, M.; Jin, L.; Kayser, F.; Liu, J. J.; Lo, M. C.; Long, A. M.; Lucas, B.; McGee, L. R.; McIntosh, J.; Mihalic, J.; Oliner, J. D.; Osgood, T.; Peterson, M. L.; Roveto, P.; Saiki, A. Y.; Shaffer, P.; Toteva, M.; Wang, Y.; Wang, Y. C.; Wortman, S.; Yakowec, P.; Yan, X.; Ye, Q.; Yu, D.; Yu, M.; Zhao, X.; Zhou, J.; Zhu, J.; Olson, S. H.; Medina, J. C. Discovery of AMG 232, a potent, selective, and orally bioavailable MDM2-p53 inhibitor in clinical development. *J. Med. Chem.* **2014**, *57*, 1454–72.
- (133) Yu, M.; Wang, Y.; Zhu, J.; Bartberger, M. D.; Canon, J.; Chen, A.; Chow, D.; Eksterowicz, J.; Fox, B.; Fu, J.; Gribble, M.; Huang, X.; Li, Z.; Liu, J. J.; Lo, M. C.; McMinn, D.; Oliner, J. D.; Osgood, T.; Rew, Y.; Saiki, A. Y.; Shaffer, P.; Yan, X.; Ye, Q.; Yu, D.; Zhao, X.; Zhou, J.; Olson, S. H.; Medina, J. C.; Sun, D. Discovery of potent and simplified piperidinone-based inhibitors of the MDM2-p53 interaction. *ACS Med. Chem. Lett.* **2014**, *5*, 894–9.
- (134) Wendt, M. D.; Rockway, T. W.; Geyer, A.; McClellan, W.; Weitzberg, M.; Zhao, X.; Mantei, R.; Nienaber, V. L.; Stewart, K.; Klinghofer, V.; Giranda, V. L. Identification of novel binding interactions in the development of potent, selective 2-naphthamide inhibitors of urokinase. Synthesis, structural analysis, and SAR of N-phenyl amide 6-substitution. *J. Med. Chem.* **2004**, *47*, 303–24.
- (135) D'Abromo, C. M.; Archambault, J. Small molecule inhibitors of human papillomavirus protein - protein interactions. *Open Virol. J.* **2011**, *5*, 80–95.
- (136) Sun, Q.; Burke, J. P.; Phan, J.; Burns, M. C.; Olejniczak, E. T.; Waterson, A. G.; Lee, T.; Rossanese, O. W.; Fesik, S. W. Discovery of small molecules that bind to K-Ras and inhibit SOS-mediated activation. *Angew. Chem., Int. Ed.* **2012**, *51*, 6140–3.
- (137) Pikul, S.; Dunham, K. M.; Almstead, N. G.; De, B.; Natchus, M. G.; Taiwo, Y. O.; Williams, L. E.; Hynd, B. A.; Hsieh, L. C.; Janusz, M. J.; Gu, F.; Mieling, G. E. Heterocycle-based MMP inhibitors with P2' substituents. *Bioorg. Med. Chem. Lett.* **2001**, *11*, 1009–13.
- (138) Kozakov, D.; Hall, D. R.; Jehle, S.; Luo, L.; Ochiana, S. O.; Jones, E. V.; Pollastri, M.; Allen, K. N.; Whitty, A.; Vajda, S. Ligand deconstruction: Why some fragment binding positions are conserved and others are not. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112*, E2585–94.
- (139) Mosyak, L.; Zhang, Y.; Glasfeld, E.; Haney, S.; Stahl, M.; Sehra, J.; Somers, W. S. The bacterial cell-division protein ZipA and its interaction with an FtsZ fragment revealed by X-ray crystallography. *EMBO J.* **2000**, *19*, 3179–91.
- (140) Jennings, L. D.; Foreman, K. W.; Rush, T. S., 3rd; Tsao, D. H.; Mosyak, L.; Li, Y.; Sukhdeo, M. N.; Ding, W.; Dushin, E. G.; Kenny, C. H.; Moghazeh, S. L.; Petersen, P. J.; Ruzin, A. V.; Tuckman, M.; Sutherland, A. G. Design and synthesis of indolo[2,3-a]quinolizin-7-one inhibitors of the ZipA-FtsZ interaction. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 1427–31.
- (141) Kenny, C. H.; Ding, W. D.; Kelleher, K.; Benard, S.; Dushin, E. G.; Sutherland, A. G.; Mosyak, L.; Kriz, R.; Ellestad, G. Development of a fluorescence polarization assay to screen for inhibitors of the FtsZ/ZipA interaction. *Anal. Biochem.* **2003**, *323*, 224–33.
- (142) Tsao, D. H.; Sutherland, A. G.; Jennings, L. D.; Li, Y.; Rush, T. S., 3rd; Alvarez, J. C.; Ding, W.; Dushin, E. G.; Dushin, R. G.; Haney, S. A.; Kenny, C. H.; Malakian, A. K.; Nilakantan, R.; Mosyak, L. Discovery of novel inhibitors of the ZipA/FtsZ complex by NMR fragment screening coupled with structure-based design. *Bioorg. Med. Chem. Lett.* **2006**, *14*, 7953–61.
- (143) Jennings, L. D.; Foreman, K. W.; Rush, T. S.; Tsao, D. H. H.; Mosyak, L.; Kincaid, S. L.; Sukhdeo, M. N.; Sutherland, A. G.; Ding, W. D.; Kenny, C. H.; Sabus, C. L.; Liu, H. L.; Dushin, E. G.; Moghazeh, S. L.; Labthavikul, P.; Petersen, P. J.; Tuckman, M.; Haney, S. A.; Ruzin, A. V. Combinatorial synthesis of substituted 3-(2-indolyl)piperidines and 2-phenyl indoles as inhibitors of ZipA-FtsZ interaction. *Bioorg. Med. Chem.* **2004**, *12*, 5115–31.
- (144) Rush, T. S.; Grant, J. A.; Mosyak, L.; Nicholls, A. A shape-based 3-D scaffold hopping method and its application to a bacterial protein-protein interaction. *J. Med. Chem.* **2005**, *48*, 1489–95.
- (145) Rai, G.; Vyjayanti, V. N.; Dorjsuren, D.; Simeonov, A.; Jadhav, A.; Wilson, D. M., 3rd; Maloney, D. J. Synthesis, biological evaluation, and structure-activity relationships of a novel class of apurinic/apyrimidinic endonuclease 1 inhibitors. *J. Med. Chem.* **2012**, *55*, 3101–12.

- (146) Abbotts, R.; Madhusudan, S. Human AP endonuclease 1 (APE1): From mechanistic insights to druggable target in cancer. *Cancer Treat. Rev.* **2010**, *36*, 425–35.
- (147) Wilson, D. M.; Simeonov, A. Small molecule inhibitors of DNA repair nuclease activities of APE1. *Cell. Mol. Life Sci.* **2010**, *67*, 3621–31.
- (148) Patel, D.; Jain, M.; Shah, S. R.; Bahekar, R.; Jadav, P.; Joharapurkar, A.; Dhanesha, N.; Shaikh, M.; Sairam, K. V.; Kapadnis, P. Discovery of potent, selective and orally bioavailable triaryl-sulfonamide based PTP1B inhibitors. *Bioorg. Med. Chem. Lett.* **2012**, *22*, 1111–7.
- (149) Scapin, G.; Patel, S. B.; Becker, J. W.; Wang, Q. P.; Desponts, C.; Waddleton, D.; Skorey, K.; Cromlish, W.; Bayly, C.; Therien, M.; Gauthier, J. Y.; Li, C. S.; Lau, C. K.; Ramachandran, C.; Kennedy, B. P.; Asante-Appiah, E. The structural basis for the selectivity of benzotriazole inhibitors of PTP1B. *Biochemistry* **2003**, *42*, 11451–9.
- (150) Barr, A. J. Protein tyrosine phosphatases as drug targets: Strategies and challenges of inhibitor development. *Future Med. Chem.* **2010**, *2*, 1563–76.
- (151) Douty, B.; Wayland, B.; Ala, P. J.; Bower, M. J.; Pruitt, J.; Bostrom, L.; Wei, M.; Klabe, R.; Gonville, L.; Wynn, R.; Burn, T. C.; Liu, P. C.; Combs, A. P.; Yue, E. W. Isothiazolidinone inhibitors of PTP1B containing imidazoles and imidazolines. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 66–71.
- (152) Zhao, J.; Du, Y. H.; Horton, J. R.; Upadhyay, A. K.; Lou, B.; Bai, Y.; Zhang, X.; Du, L. P.; Li, M. Y.; Wang, B. H.; Zhang, L. X.; Barbieri, J. T.; Khuri, F. R.; Cheng, X. D.; Fu, H. A. Discovery and structural characterization of a small molecule 14–3–3 protein-protein interaction inhibitor. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 16212–6.
- (153) Gardino, A. K.; Smerdon, S. J.; Yaffe, M. B. Structural determinants of 14–3–3 binding specificities and regulation of subcellular localization of 14–3–3-ligand complexes: A comparison of the x-ray crystal structures of all human 14–3–3 isoforms. *Semin. Cancer Biol.* **2006**, *16*, 173–82.
- (154) De Vries-van Leeuwen, I. J.; da Costa Pereira, D.; Flach, K. D.; Piersma, S. R.; Haase, C.; Bier, D.; Yalcin, Z.; Michalides, R.; Feenstra, K. A.; Jimenez, C. R.; de Greef, T. F.; Brunsved, L.; Ottmann, C.; Zwart, W.; de Boer, A. H. Interaction of 14–3–3 proteins with the estrogen receptor alpha f domain provides a drug target interface. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, 8894–9.
- (155) Thiel, P.; Roglin, L.; Meissner, N.; Hennig, S.; Kohlbacher, O.; Ottmann, C. Virtual screening and experimental validation reveal novel small-molecule inhibitors of 14–3–3 protein-protein interactions. *Chem. Commun. (Cambridge, U. K.)* **2013**, *49*, 8468–70.
- (156) Wang, Y.; Kaiser, C. E.; Frett, B.; Li, H. Y. Targeting mutant kras for anticancer therapeutics: A review of novel small molecule modulators. *J. Med. Chem.* **2013**, *56*, 5219–30.
- (157) Feig, L. A.; Cooper, G. M. Inhibition of NIH-3T3 cell-proliferation by a mutant Ras protein with preferential affinity for GDP. *Mol. Cell. Biol.* **1988**, *8*, 3235–43.
- (158) Maurer, T.; Garrenton, L. S.; Oh, A.; Pitts, K.; Anderson, D. J.; Skelton, N. J.; Fauber, B. P.; Pan, B.; Malek, S.; Stokoe, D.; Ludlam, M. J.; Bowman, K. K.; Wu, J.; Giannetti, A. M.; Starovasnik, M. A.; Mellman, I.; Jackson, P. K.; Rudolph, J.; Wang, W.; Fang, G. Small-molecule ligands bind to a distinct pocket in Ras and inhibit SOS-mediated nucleotide exchange activity. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 5299–304.
- (159) Wang, W. R.; Fang, G. W.; Rudolph, J. Ras inhibition via direct Ras binding—is there a path forward? *Bioorg. Med. Chem. Lett.* **2012**, *22*, 5766–76.
- (160) Ostrem, J. M.; Peters, U.; Sos, M. L.; Wells, J. A.; Shokat, K. M. K-Ras(G12C) inhibitors allosterically control GTP affinity and effector interactions. *Nature* **2013**, *503*, 548–51.
- (161) Christ, F.; Debyser, Z. The LEDGF/p75 integrase interaction, a novel target for anti-HIV therapy. *Virology* **2013**, *435*, 102–9.
- (162) Malet, I.; Calvez, V.; Marcellin, A. G. The future of integrase inhibitors of HIV-1. *Curr. Opin. Virol.* **2012**, *2*, 580–7.
- (163) Demeulemeester, J.; Chaltin, P.; Marchand, A.; De Maeyer, M.; Debyser, Z.; Christ, F. LEDGINs, non-catalytic site inhibitors of HIV-1 integrase: A patent review (2006 - 2014). *Expert Opin. Ther. Pat.* **2014**, *24*, 609–32.