Ligand deconstruction: Why some fragment binding positions are conserved and others are not

Dima Kozakov[a,1], David R. Hall[b], Sefan Jehle[a,2], Lingqi Luo[c], Stefan O. Ochiana[d], Elizabeth V. Jones[d], Michael Pollastri[d], Karen N. Allen[e], Adrian Whitty[e,1], and Sandor Vajda[a,e,1]

[a]Department of Biomedical Engineering, Boston University, Boston, MA 02215; [b]Acpharis Inc., Holliston, MA 01746; [c]Program in Bioinformatics, Boston University, Boston, MA 02215; [d]Department of Chemistry and Chemical Biology, Northeastern University, Boston, MA 02215; and [e]Department of Chemistry, Boston University, Boston, MA 02215

Edited by Gregory A. Petsko, Weill Cornell Medical College, New York, NY, and approved April 3, 2015 (received for review January 23, 2015)

Fragment-based drug discovery (FBDD) relies on the premise that the fragment binding mode will be conserved on subsequent expansion to a larger ligand. However, no general condition has been established to explain when fragment binding modes will be conserved. We show that a remarkably simple condition can be developed in terms of how fragments coincide with binding energy hot spots—regions of the protein where interactions with a ligand contribute substantial binding free energy—the locations of which can easily be determined computationally. Because a substantial fraction of the free energy of ligand binding comes from interacting with the residues in the energetically most important hot spot, a ligand moiety that sufficiently overlaps with this region will retain its location even when other parts of the ligand are removed. This hypothesis is supported by eight case studies. The condition helps identify whether a protein is suitable for FBDD, predicts the size of fragments required for screening, and determines whether a fragment hit can be extended into a higher affinity ligand. Our results show that ligand binding sites can usefully be thought of in terms of an anchor site, which is the top-ranked hot spot and dominates the free energy of binding, surrounded by a number of weaker satellite sites that confer improved affinity and selectivity for a particular ligand and that it is the intrinsic binding potential of the protein surface that determines whether it can serve as a robust binding site for a suitably optimized ligand.

protein–ligand interaction | fragment-based drug discovery | druggability | binding hot spot | fragment library

Fragment-based drug discovery (FBDD) has become a major tool in modern medicinal chemistry (1–4). The method is based on screening relatively small libraries of low-molecular-weight (<300 Da) compounds, called fragments. Fragments bind with low affinities but frequently have better ligand efficiencies than traditional screening hits (3, 5, 6), and the hit rate is also higher (1, 3, 4). Due to their low affinity, fragment screening is commonly done using structural methods, primarily X-ray crystallography or NMR, and the structures of bound fragments provide starting points for drug discovery. Optimization of the fragment hits by growing them through addition of further atoms and groups or by linking two fragments that bind in adjacent pockets can provide potent lead compounds, as demonstrated by a number of successful discovery campaigns (4–6).

FBDD relies on the premise that the binding site and binding mode of the fragments are conserved as the fragment is grown into a full-sized lead. Thus, it is assumed that the orientation of fragments in their respective binding pockets will remain the same whether they bind on their own or as a moiety of a larger ligand. Examples of successfully expanding fragment hits into higher affinity compounds demonstrate that the assumption is valid in certain cases (4–6). The hypothesis has been rigorously tested in a number of studies by deconstructing larger ligands into their component fragments and characterizing how these fragments bind. However, results from different studies disagree, with the results of some supporting binding mode conservation, whereas others did not (7). On the supportive side, Hajduk (8) described 18 drug leads that were reduced in size step by step to the smallest compound retaining potency for the target. The affinities of the different fragments were measured, and the study highlighted a nearly linear relationship between potency and molecular weight, suggesting a high degree of binding mode conservation. Although Hajduk (8) did not show that the fragment positions are actually conserved, such data are available in a number of other studies. Andersen et al. (9) shortened the natural cyclopentapeptide argifin, a chitinase inhibitor, to a series of dimethylguanyl-urea containing peptides and finally to dimethylguanyl-urea. The binding of fragments was analyzed by X-ray crystallography, and their conformations were shown to be similar to that observed in argifin. Binding site conservation was also demonstrated by Lange et al. (10), who deconstructed inhibitors of the SH2 domain of ^{pp60}Src, and by Van Molle et al. (11), who described the deconstruction of inhibitors of the interaction between the von Hippel–Lindau (VHL) complex and hypoxia-inducible factor 1α (HIF-1α) (11).

In contrast, several studies reported that fragment positions are not conserved. Barelier et al. considered 22 fragments resulting from the dissection of nine inhibitors of the antiapoptotic protein Bcl-xL (12) and performed both ligand-observed and protein-observed NMR experiments to determine whether the

Significance

Fragment-based drug discovery (FBDD), in which initial screening is done with low-molecular-weight compounds called fragments, relies on the premise that the fragment binding mode will be conserved on subsequent expansion to a larger ligand. We describe a remarkably simple condition for fragment binding conservation that can be tested computationally. The condition can be used for detecting whether a protein is suitable for FBDD, for predicting the size of fragments required for screening, and for determining if a fragment hit can be extended into a higher affinity ligand. The findings also reveal general properties of binding sites, highlighting the role that critical interactions between anchor sites and anchor fragments play in protein–ligand interactions in general.

Author contributions: D.K., S.J., M.P., A.W., and S.V. designed research; D.K., D.R.H., S.J., L.L., S.O.O., E.V.J., and M.P. performed research; S.J., S.O.O., E.V.J., and M.P. contributed new reagents/analytic tools; D.K., D.R.H., S.J., L.L., M.P., K.N.A., A.W., and S.V. analyzed data; and D.K., S.J., M.P., K.N.A., A.W., and S.V. wrote the paper.

Conflict of interest statement: D.R.H. is a full-time employee of Acpharis, Inc. The company offer software similar to the FTMap program that was used in this paper. D.K. and S.V. own Acpharis stock. However, the FTMap software and server are free for use.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

[1]To whom correspondence may be addressed. Email: midas@bu.edu, whitty@bu.edu, or vajda@bu.edu.

[2]Present address: Bruker BioSpin Corporation, Billerica, MA 01821.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1501567112/-/DCSupplemental.

fragments retained affinity toward the protein and to identify their approximate binding locations. Although the full-sized inhibitors extended to four different binding pockets, all but one of the component fragments were seen to bind at a single site, and even the fragments that bound in the same overall region as they did in the original inhibitor did not retain their detailed positions. The lack of fragment conservation was also convincingly demonstrated by Babaoglu and Shoichet, who deconstructed a known β-lactamase inhibitor into three fragments and showed that all three fragments shifted in position and bound to previously unexplored binding sites (13). Similar conclusions were reached in a recent paper by Barelier et al. (14), who deconstructed substrates of six enzymes from three different superfamilies into 41 overlapping fragments that were tested for activity or binding. Even fragments containing the key reactive group had little activity, and most fragments did not bind measurably until they captured most of the substrate features. Two other studies have shown that small fragments may not bind to any detectable extent. Fry et al. analyzed the binding of fragments derived from the well-known protein–protein interaction inhibitor Nutlin, which inhibits the interaction between MDM2 and p53, and showed that binding to MDM2 and conservation of the position occurred only for a very large fragment with molecular weight >300 Da, above the usual size range for fragment libraries (15). Brandt et al. (16) similarly derived 21 fragments from three nonnucleoside inhibitors of HIV-1 reverse transcriptase and found no detectable fragment binding locations despite the strong binding seen for the full-sized inhibitors.

In view of the contradictory results in the literature, it is not clear whether specific fragments will display detectable affinity for a target, and if they do, under what circumstances they are likely to retain their initial binding mode on expansion to a larger ligand. Indeed, it has been suggested that a general condition for fragment conservation is unlikely to exist (7). However, in this paper, we show that a remarkably simple condition can be developed in terms of fragment occupancy of binding energy hot spots, i.e., regions of the protein that are major contributors to the binding free energy (17). Hot spots were originally introduced in the context of mutating protein-protein interface residues to alanine (18). On the basis of this method, a residue is considered a hot spot if its mutation to alanine gives rise to a substantial drop in binding affinity. An alternative experimental method, more directly related to the binding of small ligands, identifies hot spots based on the binding of fragment-sized organic molecules to the target protein (17). It is now well established that a fundamental property of hot spots is their capability to bind a variety of small organic probe molecules (17, 19–21). As mentioned, the binding of the small compounds is very weak, and hence the interactions are most frequently detected by X-ray crystallography (22–24) or NMR (20). In the multiple solvent crystal structures (MSCS) method, X-ray crystallography is used to determine the structure of the target protein separately soaked in aqueous solutions of each of six to eight organic solvents. By superimposing the structures, regions that bind multiple different solvent probes can be detected (22, 25). Similarly, in the structure-activity relationship (SAR) by NMR method, proteins are incubated in aqueous solutions containing small organic probe molecules, and perturbations in residue chemical shifts are used to identify the sites on the protein that can participate in small molecule binding (20). It has been shown that the small "probe" ligands cluster at binding energy hot spots, and the hit rate in this experimental fragment assay predicts the energetic importance of the site (20, 22). Although the experimental determination of hot spots can be technically demanding to carry out (24), it has been shown that hot spots can also be reliably determined computationally using the solvent mapping algorithm FTMap, a close computational analog of the X-ray– and NMR-based approaches (26–31). We have further shown that, for protein–protein interface sites, the small molecule hot spots identified by FTMap correspond closely to the protein–protein interaction hot spots identified by alanine scanning mutagenesis (32, 33).

Because hot spots are the energetically important regions of the binding site, the portion of the ligand that interacts with the strongest hot spot is expected to be the part that is most essential for binding (22, 25). Thus, we hypothesized that a fragment that comprises the portion of a ligand that fully overlaps with the strongest hot spot will retain its position when the rest of the molecule is removed, provided that the fragment retains the atoms needed to exploit the key binding interactions available at this site. To investigate this hypothesis we describe eight case studies. Despite substantial structural diversity among the systems, analysis of hot spots provided a unifying approach to accounting for conservation of fragment binding mode, enabling us to formulate an easily testable condition for determining if a specific fragment of a bound ligand will retain its position when isolated. We additionally show that the fragment that interacts with the strongest hot spot generally displays a ligand efficiency (LE) that is substantially higher than that for the larger ligand. Finally, we discuss how our findings can be used to determine whether a protein is suitable for FBDD to predict the size and shape of fragments that are most appropriate for screening a particular target and to determine the likelihood that a fragment hit can be extended to achieve a higher affinity ligand. These findings additionally reveal fundamental properties of protein binding sites, showing that ligand binding sites can usefully be thought of in terms of an anchor site, which is the top-ranked hot spot and dominates the free energy of binding, surrounded by a number of weaker satellite sites that confer improved affinity and selectivity for a particular ligand, and that the intrinsic binding potential of the protein surface is the dominant factor in determining whether it can serve as a robust binding site for a suitably optimized ligand.

Results

We first tried to identify all published ligand deconstruction studies that reported structures for a ligand and some of its component fragments cocrystallized with the target protein. Studies without such structural information were not considered. The results of this search are collected in Table 1, which shows the proteins, ligands, and fragments considered in our case studies. The ligands and their fragments are numbered **1** through **27** and are shown in Figs. S1–S4. Seven case studies are based on the results of deconstruction experiments reported in the literature, whereas the last uses data we generated for this study by synthesizing fragments derived from an inhibitor of the interaction between IL-2 and the α chain of its receptor (IL-2Rα) and characterizing their binding to IL-2. To develop the condition for the conservation of fragment positions we first identified the binding energy hot spots present on each protein using computational solvent mapping (26). In most cases we mapped an X-ray structure of the ligand-free protein, thereby eliminating any potential bias arising from ligand-induced conformational changes. The exceptions are chitinase and MDM2, where ligand-free structures were not available, for which we mapped structures extracted from complexes formed with argifin and a p53 peptide, respectively. The computational solvent mapping server FTMap determines the distribution of 16 small probe molecules by performing global sampling of probe positions on the surface of the target protein and evaluating the interaction energy using molecular mechanics energy functions with continuum electrostatics models. The method identifies the hot spots as consensus clusters formed by low energy clusters of several probe molecules, overlapping at the so-called consensus sites. In view of this definition, we will use the terms consensus cluster, consensus site, and hot spot interchangeably when discussing mapping re-

Table 1. Target proteins, characteristics of the main hot spots, and fractional overlap with fragments

Target protein name	PDB ID code of structure mapped	PDB ID code with ligand	PDB ID code with fragment	Clusters in main hot spot	Maximum size of main hot spot, Å	Fractional overlap with fragment
Chitinase	1W9V*	1W9V (**1**)	3CH9 (**2**)	22	7.7	0.82
			3CHC (**3**)			0.82
SH2 of ^{pp60}Src	1O4C	1O44 (**4**)	1O4P (**5**)	25	8.2	0.93
VHL:HIF-1α	3ZRF	3ZRC (**6**)	4AWJ (**7**)	16	7.0	0.99
DPP-4	1TK3	2HHA (**8**)	1N1M (**9**)	21	8.0	0.93
		2FJP (**10**)				
Thrombin	1HTS	2C8X (**11**)	2C8Z (**12**)	26	8.2	0.87
			2C93 (**13**)			
β-Lactamase	2BLS	1XGJ (**14**)		35	13.2	0.62
			2HDQ (**15**)			0.44
			2HDR (**16**)			0.23
			2HDS (**17**)			0.35
		CEPHALOTONIN (**18**)†				
			4OLG (**19**)			0.68
			4OKP (**20**)			0.67
			4OLD (**21**)			0.67
MDM2	1T4F‡	4IPF (**22**)	4J74 (**23**)	38	7.8	0.98
IL-2:IL-2Rα	1M47	1PY2 (**24**)	1M4B (**25**)	18	8.9	0.82
			F1 (**26**)			0.79
			F2 (**27**)			0.12

*From the complex of chitinase, cocrystallized with argifin.
†No X-ray structure of the complex.
‡From the complex of MDM2, cocrystallized with a p53 peptide.

sults. Similarly to the hit rate in X-ray– or NMR-based methods, the number of probe clusters in a consensus cluster predicts the energetic importance of the corresponding hot spot (26, 34). Therefore, the hot spots are ranked on the basis of the number of probe clusters they include; the one with the highest number of probe clusters is defined as the main hot spot, whereas those with fewer probe clusters define secondary hot spots. Table 1 shows the number of probe clusters in the main hot spot for each target protein. Although FTMap uses only 16 different probe types, the number of probe clusters can be higher than 16 because we use the relatively small clustering radius of 4-Å root mean square deviation (RMSD), and strong hot spots frequently include several clusters of the same probe type. It has been shown that a hot spot with at least 16 probe clusters is required for a site to be druggable, i.e., to be able to bind a ligand with micromolar or better affinity (34). For each of the proteins considered here, the ligand binds at the main hot spot, and we have shown previously that this is generally the case for high affinity ligands and their key fragments (31). Table 1 also shows the size of the main hot spot, measured by the maximum distance between any two probe atoms within the consensus cluster.

We hypothesized that, because a substantial fraction of the free energy of ligand binding comes from the moiety that interacts with the residues in the main hot spot, the portion of the ligand that engages in these interactions will retain its location even when other parts of the ligands are removed. To evaluate this hypothesis, we formulated a quantitative measure of the degree of spatial overlap between a fragment and the main hot spot in terms of fractional overlap, defined as $FO = N_F/N_T$, where N_T denotes the total number of nonhydrogen atoms of all probe molecules in the main hot spot, and N_F is the number of such atoms that are within 2 Å from any nonhydrogen atom of the fragment when it is part of the bound ligand. Thus, *FO* measures the fraction of the main hot spot occupied by the fragment, weighted according to the energetic importance of each region within the hot spot as measured by the local density of probe atoms. Our main hypothesis is that an *FO* value close to 1.0 assures that a fragment derived from a larger ligand will retain its binding mode even after the other atoms within the ligand are removed. One of the goals of the case studies considered here was to establish the minimum *FO* value required for such conservation. This analysis is demonstrated below by applications to the eight proteins listed in Table 1.

Chitinase Inhibitor Argifin. Andersen et al. (9) shortened the natural cyclopentapeptide argifin (**1** in Fig. S1), a chitinase inhibitor, in a stepwise manner to design shorter peptides each containing the dimethylguanylurea moiety **2**, and finally to dimethylguanylurea. The binding of the peptides and of dimethylguanylurea was analyzed by X-ray crystallography. Fig. 1*A* shows the bound argifin molecule as sticks, and the main hot spot of chitinase as a transparent surface that defines the volume encompassed by the probes in the consensus cluster. As shown in Fig. 1*B*, dimethylguanylurea (cyan) binds exactly at the same position on its own, in argifin (green), and in the monopeptide fragment **3** of argifin (yellow). Indeed, the overlap is so good that the three different colors are barely distinguishable. Thus, although dimethylguanylurea does not occupy the entire volume of the main hot spot, the 82% fractional overlap that this moiety achieves is evidently sufficient for full conservation of its binding mode. Considering the strength of binding, the interactions between dimethylguanylurea and chitinase are sufficient for weak but specific binding with an IC_{50} value of 500 μM, resulting in the high LE of 0.51 kcal/mol per heavy atom (Table S1), suggesting that this fragment acts as the "anchor fragment" in argifin and its truncated derivatives.

Inhibitors of the SH2 Domain of ^{pp60}Src. Lange et al. (10) used X-ray crystallography to screen a fragment library for inhibitors of human ^{pp60}Src SH2 (Src homology 2 domain) and compared the structures containing bound fragments to those of the corresponding full-length inhibitors. The fragments were selected because they showed at least millimolar affinity in a Biacore binding assay or were suggested by ab initio design (10). All identified fragments bound in the phosphotyrosine binding pocket, which is the main hot spot on the protein (Fig. 1*C*), indicating that this binding pocket acts as an anchor site similarly to the dimethylguanylurea binding pocket

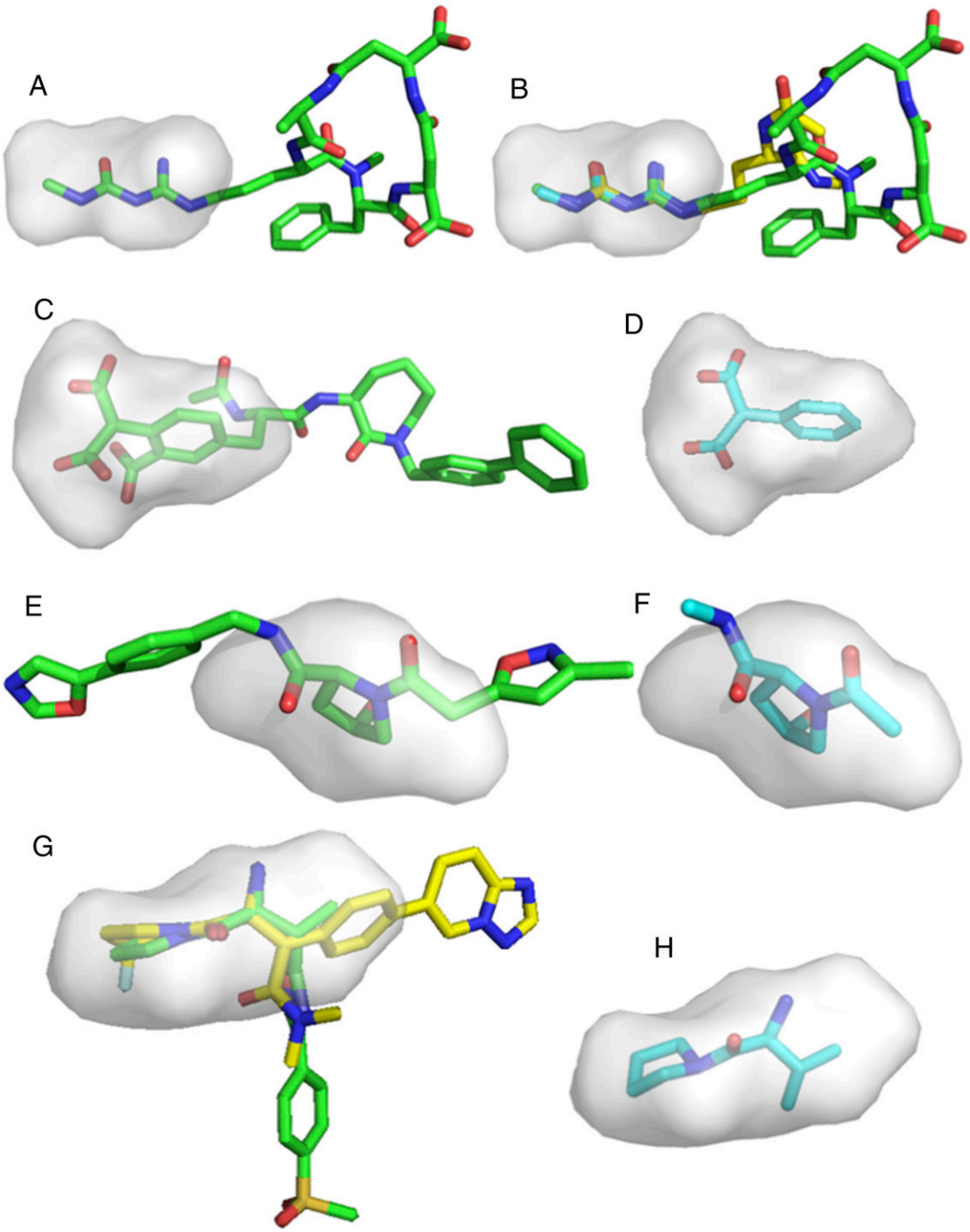

Fig. 1. Bound conformation of ligands and fragments. Compounds are shown in stick representation, superimposed on the main hot spots of their target proteins. Each main hot spot is shown as a transparent surface spanned by the representative probe molecules in the consensus cluster. (*A*) Chitinase inhibitor argifin (compound **1**). (*B*) Superposition of argifin (green, **1**), dimethylguanylurea (cyan, **2**), and the monopeptide fragment **3** of argifin (yellow). (*C*) Phenylmalonate-based inhibitor **4** of human ^{pp60}Src SH2. (*D*) Phenylmalonate **5,** binding at the main hot spot of ^{pp60}Src SH2. (*E*) Inhibitor **6** of the interaction between VHL protein and HIF-1α. (*F*) *N*-acetyl-Hyp-*N*-methyl **7**, binding at the main hot spot of the VHL protein. (*G*) Oxadiazol-based (**8**, green) and biarylphenylalanine amide-based (**10**, yellow) inhibitors of DPP-4. (*H*) Val-Pyr fragment **9**, common to both oxadiazol-based and biarylphenylalanine amide-based inhibitors, binding at the main hot spot of DPP-4.

for chitinase. In fact, the main hot spot was a good binding site in a general sense, with a number of fragments being identified that could successfully replace phosphotyrosine, resulting in non-peptidic high affinity inhibitors. For example, in inhibitor **4** (Fig. S1) this pocket accommodates a phenylmalonate group (Fig. 1*C*), which occupies 93% of the small main hot spot (*FO* = 0.93; Table 1). The additional *ortho*-carboxyl group in **4** actually protrudes outside the hot spot, and thus in truncating the inhibitor to the fragment 2-phenylmalonate, **5**, no functional group within the main hot spot is removed. In view of these properties, the fragment **5** is expected to retain its position when isolated. This expectation is confirmed by the X-ray structure of ^{pp60}Src SH2 cocrystallized with 2-phenylmalonate (Fig. 1*D*) (10). Thus, the large *FO* value and retention of all functional groups within the hot spot are sufficient

for conservation of fragment binding mode in this instance. 2-Phenylmalonate itself is a weak inhibitor, with IC_{50} between 0.2 and 2.5 mM, corresponding to LE = 0.28–0.39 kcal/mol per heavy atom (Table S1). In contrast, inhibitor **4** has IC_{50} = 3–5 nM, but its LE is only 0.25–0.26 kcal/mol per heavy atom (Table S1), consistent with the notion that the relatively atom-efficient interaction of the 2-phenylmalonyl moiety at the main hot spot can be considered to anchor the inhibitor.

Inhibitors of the Interaction Between VHL Protein and HIF-1α. E3 ubiquitin ligases, which bind protein targets and lead to their ubiquitination and subsequent degradation, are attractive drug targets due to their exquisite substrate specificity (35). The VHL complex is an E3 ubiquitin ligase with therapeutic potential. The primary substrate of VHL is HIF-1α, a transcription factor that up-regulates numerous genes (35). Van Molle et al. described the deconstruction of inhibitor **6** (Fig. S1), which binds to VHL and blocks its interaction with HIF-1α (11). The main hot spot on VHL, with 16 probe clusters, overlaps with the position of the *N*-acetyl-*L*-hydroxyproline *N*-methylamide (L-Hyp) central core of **6** (Fig. 1*E*). The fractional overlap of 0.99 (Table 1), while retaining all functional groups within the main hot spot (Fig. 1*E*), predicts that this moiety's binding mode will be conserved in an *N*-Hyp fragment. Indeed, an X-ray cocrystal structure shows that an *N*-Hyp fragment, **7** (Fig. S1), binds with a fully conserved position (Table 1 and Fig. 1*F*). The conservation occurs despite the fact that **7** is a very weak binder on its own (K_D = 4.9 mM) (36), resulting in the moderate LE of 0.24 kcal/mol per heavy atom, the same as for inhibitor **6** (Table S1). No X-ray structure could be obtained for VHL with any fragments of **6** without the L-Hyp moiety (11), consistent with the interaction of the L-Hyp moiety of **6** at the main hot spot serving an anchor role. With only 16 probe clusters, the main hot spot of VHL is relatively weak, placing the protein at the edge of druggability (34), and yet it anchors the inhibitor. For completeness we note that, more recently, Dias et al. decomposed an inhibitor also containing L-Hyp but slightly different from **6** (36) and, using higher sensitivity detection methods, observed weak binding of fragments on the two sides of the main hot spot, with K_D values of 4.3 and 2.7 mM, respectively. No X-ray structures of complexes with these fragments could be obtained, but the conservation of their binding modes was supported by ^{1}H-^{15}N heteronuclear single quantum coherence (HSQC) chemical shift perturbation experiments. In the bound inhibitor, the two fragments occupy very hydrophobic regions of the binding site, and because the mapping does not show any strong hot spot in those regions, our opinion is that fragment binding in those sites may not involve any very specific orientation. However, even the binding of these fragments with some level of specificity would not contradict to the results of this paper, as we propose only that our condition for fragment conservation is sufficient, and not that it is necessary. Thus, specific binding may occur without overlap with the main hot spot, although thus far we have seen very few examples of this. Furthermore, the weak hot spots on the two sides of the L-Hyp moiety also contribute to binding. In fact, optimization of **7** resulted in several inhibitors with high nanomolar affinities (37), all retaining L-Hyp as their anchor.

Dipeptidyl Peptidase IV (DPP-4). Dipeptidyl peptidase IV (DPP-4) is a multifunctional type II transmembrane serine peptidase that specifically cleaves a variety of peptides after proline or alanine (38). Inhibition of DPP-4 has been pursued as an important therapeutic approach for the treatment of type 2 diabetes (39). Here we consider the deconstruction of the oxadiazol-based inhibitor **8** (Fig. S2) shown as green sticks in Fig. 1*G* (IC_{50} = 120 nM). The main hot spot of ligand-free DPP-4 binds 21 probe clusters and overlaps very well with the Val-Pyr moiety of the inhibitor (Fig. 1*G*), resulting in a fractional overlap of 0.93. All functional groups within the hot spot are retained, and thus our condition predicts a conserved binding position for a Val-Pyr dipeptide fragment. Indeed, Val-Pyr **9** (Fig. S2) binds to DPP-4 at exactly the same position as in inhibitor **8** (Fig. 1*H*). Val-Pyr itself is an inhibitor, with IC_{50} = 3.0–5.4 μM, corresponding to the very high LE of 0.61–0.63 kcal/mol per heavy atom (Table S1). Notably, Val-Pyr also occupies the same position in biarylphenylalanine amide inhibitors such as **10** (40), shown as yellow sticks in Fig. 1*G*, although other parts of oxadiazol-based and biarylphenylalanine amide-based inhibitors extend into completely different regions of the active site. Thus, Val-Pyr serves as a strong anchor for both types of DPP-4 inhibitors.

Thrombin Inhibitors. Thrombin has been under intense investigation for decades as a potential target for anticoagulation agents. Recent efforts have included fragment-based approaches using X-ray crystallography that led to the discovery of several fragment hits and the development of novel inhibitors (41). In particular, inhibitor **11** (Fig. S2 and Fig. 2*A*) was based on the chlorophenyl fragment hit **12**. Fragment **12** binds in the S1 pocket of thrombin, yielding a fractional overlap of 0.87 with a main hot spot defined by 26 probe clusters (Table 1) (41) and is clearly seen in an X-ray structure to bind at exactly the same location as the corresponding part of the inhibitor. Although **12** is a very weak binder, with IC_{50} = 1–3 mM, its small size results in a respectable LE of 0.38–0.46 kcal/mol per heavy atom (Table S1). The second ranked hot spot on thrombin has 16 probe clusters (shown as magenta lines in Fig. 2 *A* and *B*) and is located in the S4 pocket of the protein, only 7.5 Å from the primary hot spot. This second hot spot supports the binding of fragment **13** (41), which extends toward the S1 pocket and also overlaps somewhat with the main hot spot. Due to the overlap with a secondary strong hot spot, the binding mode of fragment **13** (Fig. S2) is also conserved, as shown by an X-ray structure of the complex (41). In fact, inhibitor **11** was developed by linking fragment hits **12** and **13** (41). This result again demonstrates that, in some cases, fragment conservation can occur without substantial overlap with the main hot spot, particularly if the fragment overlaps with another strong hot spot. However, we note that the molecular weight of compound **13** is 392, and thus, based on the rule of three (42), is significantly larger than a true fragment. Our prediction is that a slightly smaller subfragment of **13** would likely shift toward the main hot spot. It is rare that two strong hot spots are adjacent to each other but bind different fragments in conserved positions, and hence linking of separately identified fragments can be seldom used to form higher affinity ligands from fragment hits (43).

β-Lactamase. β-Lactamases comprise the major resistance mechanism to β-lactam antibiotics and pose a growing threat to public health, and hence the discovery of β-lactamase inhibitors is a longstanding and important problem (44). The deconstruction of the AmpC β-lactamase inhibitor **14** was used by Babaoglu and Shoichet to generate a well-documented counterexample to fragment conservation (13). They studied fragments **15**, **16**, and **17** (Fig. S3) of inhibitor **14** and demonstrated that none of them binds in the position occupied by the corresponding structural moiety in the full-sized **14**. As shown in Table 1, β-lactamase has a more diffuse hot spot structure than the other proteins considered here, as its main hot spot is much longer and is separated only by 3.5 Å from the fourth ranked hot spot containing 11 probe clusters. The starting inhibitor **14**, shown in Fig. 2*C* as green sticks, itself has only 65% overlap with the large main hot spot, resulting in weak binding (K_i = 1 μM), although with the respectable LE of 0.37 kcal/mol per heavy atom (Table S1). Fragment **15** occupies only about 44% of the main hot spot. As shown in Fig. 2*D*, these interactions are not sufficient to stabilize the binding mode of the fragment, which by itself shifts to the

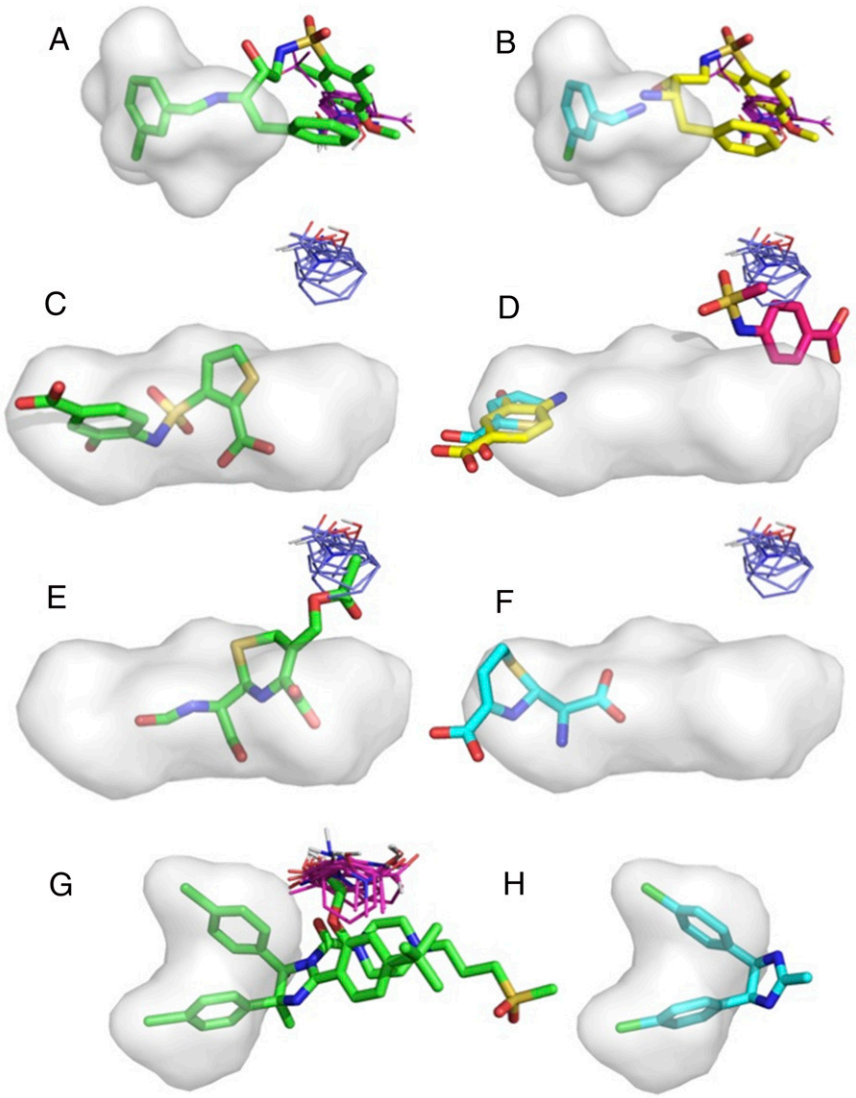

Fig. 2. Bound conformation of ligands and fragments. Secondary hot spots are shown as clusters of probe cluster representatives (thin sticks). (*A*) Thrombin inhibitor **11**, based on fragment screening and fragment linking. Representative probes of second strongest hot spot (16 probe clusters) are shown as magenta lines. (*B*) Fragment hits **12** and **13** used for the discovery of thrombin inhibitor **11**. The second hot spot supports the binding of fragment **13** (yellow sticks) that also protrudes into the main hot spot. (*C*) AmpC β-lactamase inhibitor **14**. Representative probes of the fourth ranked hot spot (11 probe clusters) are shown as blue lines. (*D*) AmpC β-lactamase inhibitor fragment **15** (cyan), **16** (yellow), and **17** (magenta). (*E*) AmpC β-lactamase inhibitor fragment 7-*N*-formyl-cephalosporanic acid **19**. (*F*) AmpC β-lactamase inhibitor fragment **20,** derived from fragment **19**. (*G*) Nutlin-3 (compound **22**), an inhibitor of the MDM2:p53 interaction. The second strongest hot spot of MDM2 (22 probe clusters) is shown as magenta lines. (*H*) Fragment **23**, the smallest Nutlin-3 fragment capable of binding to MDM2.

edge of the main hot spot (cyan). The fractional overlaps between the main hot spot and fragments **16** and **17** are only 23% and 35%, respectively. Fragment **16** binds close to the position occupied by the corresponding portion of the larger inhibitor, but changes its orientation to almost overlap with the bound pose of fragment **15** (Fig. 2*D*, yellow). Notably, fragment **17,** which is slightly larger than fragment **16** because it also includes the sulfoxy moiety (although it lacks the phenolic OH), shifts to hot spot 4 (Fig. 2*D*, shown as blue lines), as the latter seems to have preference for compounds with a single aromatic ring. Thus, none of the three fragments has a sufficient fractional overlap with the main hot spot, or interacts sufficiently with other strong hot spots, to stabilize its binding mode, and hence they all shift when isolated from the rest of the ligand. In fact, the ligand efficiencies of these fragments are lower than that of the starting inhibitor, the only such case in Table S1, suggesting that none of these fragments binds strongly enough to be considered as an anchor for a larger inhibitor.

In a more recent study, Barelier et al. fragmented cephalothin (**18**), a substrate of AmpC β-lactamase, and determined the X-ray structures of three fragments cocrystallized with the protein (14). The largest fragment, 7-*N*-formyl-cephalosporanic acid **19** (Fig. S3), almost entirely recapitulates cephalothin. Measurable activity was achieved for this fragment, and its X-ray structure showed a covalent bond between this fragment and the catalytic Ser64, suggesting that it retains the position it occupies in cephalothin. The reaction mechanism revealed that 7-*N*-formyl-cephalosporanic acid captures the stable acyl-enzyme intermediate step between the transition-state acylation and deacylation complexes (14). Although fragment **19** overlaps with only 68% of the main hot spot, it also extends into hot spot 4, and this may contribute to its specific binding (Fig. 2*E*). In fact, the moiety that

extends into hot spot 4 is removed in the two smaller fragments **20** and **21** (Fig. S3), and although both contain the core cephalosporin ring and have fractional overlaps of 57% with the main hot spot, neither of them reacted at a measurable rate with the β-lactamase. The hydrolyzed form of **20** shifts toward the left edge of the hot spot, close to the bound position of fragments **15** and **16** (Fig. 2*F*, cyan), whereas fragment **21** binds at the surface of the protein in a hydrophobic cavity where the mapping does not show any hot spot, and hence the binding is most likely not specific. We note that the Shoichet group also screened a fragment library for binding to AmpC β-lactamase, resulting in fairly high hit rate (45). However, cocrystallization of the protein with eight of the bound fragment-sized compounds demonstrated that different fragments can bind at different locations of the large hot spot region, and hence the results provided limited information for the design of higher affinity inhibitors. In fact, it appears that **14** remains the highest affinity noncovalent AmpC β-lactamase inhibitor currently available, despite substantial recent efforts (46, 47).

MDM2. The human version of the MDM2 (mouse double minute protein 2) influences transcription by binding to the tumor suppressor protein p53 (48). Vassilev et al. reported a series of *cis*-imidazoline analog inhibitors termed Nutlins (49). Fry et al. (15) recently performed systematic deconstruction of a Nutlin, compound **22** (Fig. S4), obtaining molecules that represented successively smaller fragments of the parent, providing an adequate coverage of the possible fragmentation pathways. Because no X-ray structure of ligand-free MDM2 is available, in this case, we mapped the structure of MDM2 cocrystallized with a p53 peptide (34). As shown in Fig. 2*G*, the main hot spot of MDM2 encompasses two pockets, each binding one chlorophenyl group of the inhibitor. In the MDM2:p53 complex, these pockets accommodate the side chains of Trp23 and Leu26 of the p53 peptide. Interestingly, this hot spot is not very large (Table 1), but due to its shape, it requires a fragment with a particular 3D structure that enables it to reach into both deep hydrophobic pockets and thus requires retention both chlorophenyl groups for good overlap with the hot spot. This expectation was fully confirmed by the deconstruction experiments (15), which showed **23** (Fig. S4) to be the smallest Nutlin fragment capable of achieving measurable binding to MDM2. In fact, the two halogenated phenyl groups stabilized in a particular conformation are so important that these are present in all high affinity MDM2 inhibitors currently available (50). Indeed, fragment **23** does not bind detectably if the two chlorines are removed (15). This result shows that, even when a fragment is fully encompassed within the main hot spot, removal of functional groups within the main hot spot may result in loss of specific binding.

IL-2. IL-2 (Interleukin-2) is an immunoregulatory cytokine that stimulates normal and pathogenic T cells and contributes to rejection of tissue grafts (51). A number of small molecules (e.g., compound **24**; Fig. S4) bind to IL-2 and inhibit its interactions with the IL-2 receptor α subunit IL-2Rα (52). The binding hot spots of ligand-free IL-2 were previously determined using the FTMap algorithm (34) The IL-2:IL-2Rα interface includes two hot spots. The main hot spot, containing 20 probe clusters and shown in Fig. 3*A*, is located in a largely polar and rigid pocket, and partially overlaps with the guanido end group of inhibitor **24**, which was discovered by Sunesis Pharmaceuticals using a site-specific fragment-discovery method called tethering (52, 53). The tethering method identifies fragments that are selected from a library of small compounds containing a common disulfide through equilibrium disulfide exchange with an engineered cysteine on the protein target. The adducts that predominate are those with fragments that additionally engage in stabilizing noncovalent interactions with the protein (53, 54). Compound **25** (Fig. S4) is one of the disulfide-bound fragments that was co-

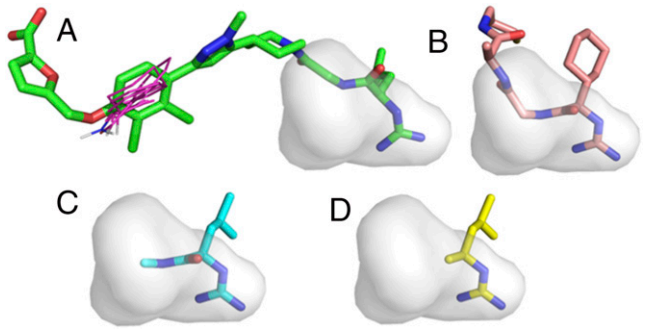

Fig. 3. Deconstructing an inhibitor of the interaction between IL-2 and its receptor IL-2Rα. (*A*) Inhibitor **24** of the IL-2:IL-2Rα interaction. Molecules representing the probe clusters in fourth ranked hot spot (10 probe clusters) of ligand-free IL-2 are shown as magenta lines. (*B*) Fragment **25** was tethered to and cocrystallized with IL-2. (*C*) Because the binding of fragment **26** was shown by chemical shift perturbations observed in $^{15}N/^{1}H$ HSQC NMR spectra, and no X-ray structure is available, the bound position shown is based on the binding mode of fragment **25**. (*D*) Hypothetical position of fragment **27**, which does not bind to IL-2, based on the binding mode of fragment **25**.

crystallized with IL-2 (53). Although **25** includes the guanido group, its structure differs from any portion of **24** and was covalently linked to the protein, although the linker does not seem to change the binding mode (52, 53, 55). To identify the smallest fragment derived from inhibitor **24** that can bind at the main hot spot, we synthesized fragments **26** and **27** (Fig. S4), which do not include any extra atoms and thus are true fragments of **24**. Analyzing the overlap of these fragments with the main hot spot, we expected that fragment **26** might bind in a conserved manner (Fig. 3*C*; $FO = 0.80$), but **27** should not (Fig. 3*D*; $FO = 0.12$). The two fragments were tested for their affinity for IL-2, and their approximate binding location determined, by chemical shift perturbation (CSP) mapping based on ^{15}N-^{1}H HSQC NMR experiments. Significant CSPs are observed for residues 43, 44, 46, 58–60, and 62–64 on titration of compound **26**, which indicates binding (Figs. S5 and S6). This result is in good agreement with our mapping results and shows that fragment **26** interacts with the same residues on its own and as part of fragment **24.** The NMR data gave a binding affinity of K_D ~170 μM, which is relatively high for such a small fragment and results in a LE of 0.4 kcal/mol per heavy atom (Table S1). In contrast, compound **27**, which is a truncated version of compound **26** obtained by removing an *N*-methyl amide group, does not cause CSPs on titration, which indicates that this compound does not bind detectably.

Discussion

In this paper, we extracted information from all deconstruction studies we were able to find in the literature that included X-ray structures for both a ligand and some of its component fragments bound to a protein receptor. In two cases, the binding of some fragments was shown only by NMR. These previous studies did not explain why fragment positions were conserved for some ligands but not for others. Here we show that determining the overlap of the fragments with the main binding hot spot offers a remarkably simple condition for conservation of binding mode. The basic idea is that, because a substantial fraction of the binding free energy is due to protein–ligand interactions within the main hot spot, a fragment that overlaps well with this hot spot and retains the interacting functional groups will retain its binding mode when the rest of the ligand is removed. To make the condition more rigorous, we introduced the concept of fractional overlap, *FO*, defined as the fraction of probes in the main hot spot that overlap with the fragment in the bound po-

sition of the ligand. Based on the eight case studies considered in this paper, fragments that have $FO > 0.8$ and retain all substantive functional groups within the main hot spot displayed conservation of binding mode compared with the same structural moiety contained within a larger ligand. This condition is sufficient but not necessary for retaining the binding mode, and thus the position of a fragment can be conserved even when the condition is not satisfied. However, in such cases the outcome is uncertain, and in most cases, the fragments shift within the main hot spot or fail to show detectable binding.

The eight case studies described here demonstrate different aspects of fragment binding mode conservation. The analysis of argifin and its substructures shows that a small moiety of the ligand, in this case dimethylguanylurea, can serve as a powerful anchor for the binding of much larger ligands, although its fractional overlap with the main hot spot is only 0.82. The analysis of the SH2 domain of ^{pp60}Src demonstrates that the phosphotyrosine binding pocket binds a variety of fragments, which is the fundamental property of hot spots, and that these fragments can be extended into higher affinity ligands while retaining their binding positions. In the VHL protein, the fragment *N*-acetyl-Hyp-*N*-methyl overlaps perfectly with the main hot spot in the middle of the binding site ($FO = 0.99$), and its position is well conserved as expected. In DPP-4 the Val-Pyr fragment ($FO = 0.93$) serves as the anchor for two very different classes of inhibitors that do not overlap beyond the main hot spot. The analysis of thrombin emphasizes that interactions with strong secondary hot spots can also lead to specific binding of fragments, and in such cases, higher affinity ligands can be obtained by fragment linking (41). β-lactamase differs from the other targets considered here because even the known inhibitors bind only to relatively small fractions of its very large and diffuse main hot spot ($FO < 0.68$). The fragmentation of these ligands further reduces the *FO* values, and all resulting fragments shift from their original positions. The large main hot spot can accommodate a variety of different fragments, leading to a high hit rate in fragment-based screening, but, in the absence of an additional stabilizing feature such as a reversible covalent bond with the protein (56), the positions of the bound fragments vary and hence provide limited information for the development of larger ligands. The main message from the analysis of MDM2 is that fragment based discovery of inhibitors that disrupt protein–protein interactions may require fragments with a well-defined 3D structure and with molecular weight exceeding 300 Da for the binding mode to be sufficiently robust for it to be conserved in larger ligands. All of the above studies demonstrate that the fragments in the main hot spot that have a conserved binding mode and serve to anchor the binding of the larger ligand do not have to be strong binders, but must have a high LE relative to the larger ligand.

Our proposed condition for conservation of fragment binding mode incorporates a number of assumptions. First, we assume that the larger ligand overlaps with the main hot spot and is a fairly good binder. We also assume that there is good steric, electrostatic, and chemical complementarity between the ligand and receptor functional groups in this region; that is, that the ligand is well optimized for binding to the protein target in the regions that interact with the primary hot spot. Finally, our analysis implicitly assumes that a fragment that is highly complementary to the main hot spot will not have alternative binding modes in that site that are as favorable as the position occupied in the context of the larger ligand. Our results show that these assumptions are reasonable for the structurally diverse set of targets and ligands we analyzed, supporting the notion that a level of overlap of $FO > 0.8$ implies a degree of complementarity between the fragment and the hot spot such that it is unlikely that any alternative fit can provide similarly high binding free energy. Importantly, even for fragments with FO > 0.8, the case studies analyzed above show that conservation of binding mode is not assured if any of the functional groups of the ligand that interact with residues in the main hot spot are removed when defining the fragment. However, combining the requirement $FO > 0.8$ with the requirement of retaining the functional groups that interact with the primary hot spot clearly yields a useful condition for predicting which portions of a ligand will retain their binding mode as isolated fragments, despite the simplifications incorporated in our analysis.

The FTMap algorithm, freely available as a server at ftmap.bu.edu/login.php, has been shown to reliably identify binding hot spots of proteins (26–31), in good agreement with experimental screening approaches, and we show here that this method provides an easy way to analyze potential fragment conservation for targets with known structure. We envision three important applications. First, initial mapping can detect whether a protein target is or is not suitable for application of a fragment based approach to drug discovery. It has already been shown that targets that lack a main hot spot of sufficient strength give a low hit rate in experimental fragment screens, and have poor prospects for inhibitor development (20, 22). We show here that targets containing a very large and diffuse main hot spot, substantially larger than the compounds that are usually considered as fragments, are also poor prospects for fragment-based drug discovery. Although fragment hits may be obtained for targets of this kind, they are unlikely to be useful starting points for the discovery of larger ligands because they cannot achieve the level of overlap required to ensure a unique and robust binding mode that will be conserved on expansion to a larger ligand. Among the proteins studied here, β-lactamase is clearly in this category. Second, the mapping results also predict whether standard fragment libraries are appropriate for screening a specific target. Such libraries generally contain flat heterocyclic compounds with molecular weights between 100 and 250 Da, and this may be too restrictive for some proteins such as MDM2 discussed in this paper, for which the shape of the main hot spot is complex, indicating that achieving high *FO* will require relatively large fragments with complex 3D shapes. Because we map ligand-free protein structures, this result potentially enables the designs of target-specific fragment libraries, containing fragments of a size and topological complexity chosen to match that of the main hot spot. The third and possibly the most important application is predicting the likelihood that a hit from an experimental fragment screen can be extended into a higher affinity ligand. Among the examples studied here, it is notable that, in almost every case in which the fragment binding mode was conserved, LE for the fragment was markedly higher than that for the larger ligand from which it was derived (Table S1). The exception is VHL, for which the conserved fragment and the larger ligand have identical LE values of 0.24. However, these LE values are low compared with the value of LE = 0.3 that is typically considered as the minimum LE for a drug or for a useful fragment hit (31), consistent with the fact that the primary hot spot on VHL is the weakest among all of the targets in Table 1. Our results therefore suggest that, to serve as a strong anchor that provides a good starting point for FBDD, a fragment must occupy a large fraction of the main hot spot (i.e., $FO > 0.8$), but additionally must have a relatively high LE, indicating strong complementarity with the main hot spot region and efficient exploitation of the binding potential it offers. The results in Table S1 suggest that LE = 0.4–0.6 is typical of a strong anchor fragment. As we have shown previously, the ability to extend a fragment beyond the main hot spot also requires the existence of at least one additional hot spot that can be reached by a druglike molecule (31, 34), and the mapping results can even be used to predict the changes in LE upon the extension (31).

In addition to the question of under what circumstances fragment binding modes will be conserved, our results reveal fundamental information about the nature of ligand binding sites

on proteins. Binding is conventionally thought of as a mutual property of both protein and ligand, in the sense that whether binding will occur, and with what affinity, depends on the amount of binding energy there is to be gained through interaction at a particular site, and also by how effectively a given ligand accesses this binding energy by achieving good overlap with and complementarity to the site. The results presented here indicate that the first of these—the intrinsic binding propensity of the site on the protein—is in some sense the dominant factor. Our analysis shows that fragment binding sites are conserved when they overlap with the strongest binding energy hot spot on the protein surface, a site that is identified from the structure of the protein alone without reference to any particular ligand. Moreover, the strongest hot spot tends to bind many different fragment structures, acting as a general "attractor." In fact, the binding modes of fragments that bind at secondary hot spots are often not conserved when they are not constrained as part of a larger ligand that interacts with the primary hot spot, leading these fragments to migrate to instead occupy the main hot spot. Thus, it is the intrinsic binding potential of the region on the protein surface that determines whether it can serve as a stable and robust binding site for a suitably optimized ligand, and strong hot spots are characterized by their ability to bind a variety of different fragment structures.

Additional evidence that the potential to form a stable complex is dominated by the binding potential of the protein is provided by a number of observations from the literature. First, experimental or computational mapping of protein binding sites using small fragment probes shows that individual probe structures are rather indiscriminate, such that binding subsites occupied by ligands are characterized by their ability to bind a variety of different probe structures. Indeed, the FTMap algorithm works by identifying such "consensus sites," and it is well validated that these consensus sites coincide with the experimentally observed binding sites for drugs and other ligands (31, 34), and with the locations of binding energy hot spots at protein–protein interfaces identified by alanine-scanning mutagenesis (32, 33). Moreover, the locations of FTMap consensus sites are robust to conformational changes in the protein, and the same sites can be reliably identified using experimental X-ray crystal structures that show the protein in any of a variety of bound or unbound conformational states (34). Second, it is well established that the overall prospects of identifying a high affinity drug-like ligand for a given protein target correlates with the hit rate observed in an experimental fragment screen; that is, on the variety of different fragment structures that can bind to that site, rather than the affinity with which any one particular fragment binds. Third, it has been found by multiple groups that advancing small fragment hits into larger drug-like ligands is typically more successful when done by growing the fragment by adding moieties that reach into adjacent hot spot regions rather than by linking together separate fragment hits that bind to distinct but nearby sites on the protein (57), which was how fragment-based lead discovery was originally proposed to work. The bias toward fragment growing has become embodied in the terminology of FBDD, in which the goal is often explicitly stated as being to identify an initial anchor fragment that provides the starting point for lead identification. Previously, the poor success that has typically been seen for fragment linking, compared with fragment growing, has been attributed to the difficulty in achieving linkers that connect the fragments in precisely the correct relative orientations required to maximize entropic benefit (58, 59). Our results provide an alternative explanation for the higher success of fragment growing. We show that an anchor fragment that binds at the main hot spot and efficiently exploits the binding energy to be gained there, thereby achieving high LE and a robust binding mode, is a characteristic feature of the better ligands in our test set, because this mirrors the binding potential of the protein target that typically contains one main hot spot accompanied by a number of weaker, secondary hot spots. Indeed, the current study can be viewed as a way to deconstruct a ligand to retrospectively identify the anchor fragment: the portion that possesses a binding geometry that is conserved on expansion of the fragment into a larger ligand.

Overall, our results show that ligand binding sites on proteins can usefully be thought of in terms of an anchor site, which is the top-ranked hot spot that dominates the contributions to the free energy of binding, surrounded by a number of weaker satellite sites that confer improved binding affinity and selectivity of the site for a particular ligand. Consequently, analysis of fragment overlap with the main hot spot identifies which regions of a larger ligand represent the primary anchor for ligand binding. This finding largely resolves the uncertainty in the literature regarding which fragments of a larger ligand will retain their position as isolated fragments and, when combined with consideration of LE, provide an objective and useful framework for identifying which experimental fragment hits represent the most promising starting points for ligand discovery.

Methods

Computational Solvent Mapping. Protein structures were downloaded from the Protein Data Bank (PDB) (60). All ligand and bound water molecules were removed. Mapping was performed using the FTMap algorithm (26, 61) through its online server (ftmap.bu.edu) (*SI Methods*).

Synthesis of Fragments. For the synthesis of (2*R*)-2- carbamimidamido-*N*,4-dimethylpentanamide (fragment **26**) and 1-(3-methylbutyl)guanidine (fragment **27**), see *SI Methods*.

CSP Mapping by NMR. We probed both compounds **26** and **27** for binding to IL-2 in an NMR titration experiment. CSPs were observed in $^{15}N/^{1}H$ HSQC NMR spectra to characterize the binding affinity and location of ligands that interact with the target protein IL-2 (*SI Methods*).

ACKNOWLEDGMENTS. This investigation was supported by National Institute of General Medical Sciences Grants GM064700 and GM094551 and National Institute of Allergy and Infectious Diseases Grant AI082577.

1. Hajduk PJ, Greer J (2007) A decade of fragment-based drug design: Strategic advances and lessons learned. *Nat Rev Drug Discov* 6(3):211–219.
2. Erlanson DA, McDowell RS, O'Brien T (2004) Fragment-based drug discovery. *J Med Chem* 47(14):3463–3482.
3. Erlanson DA (2012) Introduction to fragment-based drug discovery. *Top Curr Chem* 317:1–32.
4. Murray CW, Verdonk ML, Rees DC (2012) Experiences in fragment-based drug discovery. *Trends Pharmacol Sci* 33(5):224–232.
5. Hubbard RE, Murray JB (2011) Experiences in fragment-based lead discovery. *Methods Enzymol* 493:509–531.
6. Murray CW, et al. (2010) Fragment-based drug discovery applied to Hsp90. Discovery of two lead series with high ligand efficiency. *J Med Chem* 53(16):5942–5955.
7. Bisignano P, et al. (2012) In silico deconstruction of ATP-competitive inhibitors of glycogen synthase kinase-3β. *J Chem Inf Model* 52(12):3233–3244.
8. Hajduk PJ (2006) Fragment-based drug design: How big is too big? *J Med Chem* 49(24):6972–6976.
9. Andersen OA, Nathubhai A, Dixon MJ, Eggleston IM, van Aalten DMF (2008) Structure-based dissection of the natural product cyclopentapeptide chitinase inhibitor argifin. *Chem Biol* 15(3):295–301.
10. Lange G, et al. (2003) Requirements for specific binding of low affinity inhibitor fragments to the SH2 domain of (^{pp60})Src are identical to those for high affinity binding of full length inhibitors. *J Med Chem* 46(24):5184–5195.
11. Van Molle I, et al. (2012) Dissecting fragment-based lead discovery at the von Hippel-Lindau protein:hypoxia inducible factor 1α protein-protein interface. *Chem Biol* 19(10):1300–1312.
12. Barelier S, Pons J, Marcillat O, Lancelin JM, Krimm I (2010) Fragment-based deconstruction of Bcl-x_L inhibitors. *J Med Chem* 53(6):2577–2588.
13. Babaoglu K, Shoichet BK (2006) Deconstructing fragment-based inhibitor discovery. *Nat Chem Biol* 2(12):720–723.
14. Barelier S, et al. (2014) Substrate deconstruction and the nonadditivity of enzyme recognition. *J Am Chem Soc* 136(20):7374–7382.
15. Fry DC, et al. (2013) Deconstruction of a nutlin: Dissecting the binding determinants of a potent protein-protein interaction inhibitor. *ACS Med Chem Lett* 4(7):660–665.

16. Brandt P, Geitmann M, Danielson UH (2011) Deconstruction of non-nucleoside reverse transcriptase inhibitors of human immunodeficiency virus type 1 for exploration of the optimization landscape of fragments. *J Med Chem* 54(3):709–718.
17. DeLano WL (2002) Unraveling hot spots in binding interfaces: Progress and challenges. *Curr Opin Struct Biol* 12(1):14–20.
18. Clackson T, Wells JA (1995) A hot spot of binding energy in a hormone-receptor interface. *Science* 267(5196):383–386.
19. Vajda S, Guarnieri F (2006) Characterization of protein-ligand interaction sites using experimental and computational methods. *Curr Opin Drug Discov Devel* 9(3):354–362.
20. Hajduk PJ, Huth JR, Fesik SW (2005) Druggability indices for protein targets derived from NMR-based screening data. *J Med Chem* 48(7):2518–2525.
21. Seco J, Luque FJ, Barril X (2009) Binding site detection and druggability index from first principles. *J Med Chem* 52(8):2363–2371.
22. Mattos C, Ringe D (1996) Locating and characterizing binding sites on proteins. *Nat Biotechnol* 14(5):595–599.
23. Ciulli A, Williams G, Smith AG, Blundell TL, Abell C (2006) Probing hot spots at protein-ligand binding sites: A fragment-based approach using biophysical methods. *J Med Chem* 49(16):4992–5000.
24. Winter A, et al. (2012) Biophysical and computational fragment-based approaches to targeting protein-protein interactions: Applications in structure-guided drug discovery. *Q Rev Biophys* 45(4):383–426.
25. Allen KN, et al. (1996) An experimental approach to mapping the binding surfaces of crystalline proteins. *J Phys Chem-Us* 100(7):2605–2611.
26. Brenke R, et al. (2009) Fragment-based identification of druggable 'hot spots' of proteins using Fourier domain correlation techniques. *Bioinformatics* 25(5):621–627.
27. Chuang GY, et al. (2009) Binding hot spots and amantadine orientation in the influenza a virus M2 proton channel. *Biophys J* 97(10):2846–2853.
28. Landon MR, et al. (2009) Detection of ligand binding hot spots on protein surfaces via fragment-based methods: Application to DJ-1 and glucocerebrosidase. *J Comput Aided Mol Des* 23(8):491–500.
29. Hall DH, et al. (2011) Robust identification of binding hot spots using continuum electrostatics: Application to hen egg-white lysozyme. *J Am Chem Soc* 133(51):20668–20671.
30. Buhrman G, et al. (2011) Analysis of binding site hot spots on the surface of Ras GTPase. *J Mol Biol* 413(4):773–789.
31. Hall DR, Ngan CH, Zerbe BS, Kozakov D, Vajda S (2012) Hot spot analysis for driving the development of hits into leads in fragment-based drug discovery. *J Chem Inf Model* 52(1):199–209.
32. Zerbe BS, Hall DR, Vajda S, Whitty A, Kozakov D (2012) Relationship between hot spot residues and ligand binding hot spots in protein-protein interfaces. *J Chem Inf Model* 52(8):2236–2244.
33. Golden MS, et al. (2013) Comprehensive experimental and computational analysis of binding energy hot spots at the NF-κB essential modulator/IKKβ protein-protein interface. *J Am Chem Soc* 135(16):6242–6256.
34. Kozakov D, et al. (2011) Structural conservation of druggable hot spots in protein-protein interfaces. *Proc Natl Acad Sci USA* 108(33):13528–13533.
35. Buckley DL, et al. (2012) Targeting the von Hippel-Lindau E3 ubiquitin ligase using small molecules to disrupt the VHL/HIF-1α interaction. *J Am Chem Soc* 134(10):4465–4468.
36. Dias DM, et al. (2014) Is NMR fragment screening fine-tuned to assess druggability of protein-protein interactions? *ACS Med Chem Lett* 5(1):23–28.
37. Galdeano C, et al. (2014) Structure-guided design and optimization of small molecules targeting the protein-protein interaction between the von Hippel-Lindau (VHL) E3 ubiquitin ligase and the hypoxia inducible factor (HIF) alpha subunit with in vitro nanomolar affinities. *J Med Chem* 57(20):8657–8663.
38. Rasmussen HB, Branner S, Wiberg FC, Wagtmann N (2003) Crystal structure of human dipeptidyl peptidase IV/CD26 in complex with a substrate analog. *Nat Struct Biol* 10(1):19–25.
39. Xu J, et al. (2006) Discovery of potent, selective, and orally bioavailable oxadiazole-based dipeptidyl peptidase IV inhibitors. *Bioorg Med Chem Lett* 16(20):5373–5377.
40. Edmondson SD, et al. (2006) (2S,3S)-3-Amino-4-(3,3-difluoropyrrolidin-1-yl)-N,N-dimethyl-4-oxo-2-(4-[1,2,4]triazolo[1,5-a]-pyridin-6-ylphenyl)butanamide: A selective alpha-amino amide dipeptidyl peptidase IV inhibitor for the treatment of type 2 diabetes. *J Med Chem* 49(12):3614–3627.
41. Howard N, et al. (2006) Application of fragment screening and fragment linking to the discovery of novel thrombin inhibitors. *J Med Chem* 49(4):1346–1355.
42. Congreve M, Carr R, Murray C, Jhoti H (2003) A 'rule of three' for fragment-based lead discovery? *Drug Discov Today* 8(19):876–877.
43. Rees DC, Congreve M, Murray CW, Carr R (2004) Fragment-based lead discovery. *Nat Rev Drug Discov* 3(8):660–672.
44. Wang X, Minasov G, Shoichet BK (2002) Evolution of an antibiotic resistance enzyme constrained by stability and activity trade-offs. *J Mol Biol* 320(1):85–95.
45. Teotico DG, et al. (2009) Docking for fragment inhibitors of AmpC β-lactamase. *Proc Natl Acad Sci USA* 106(18):7455–7460.
46. Chan FY, et al. (2012) Validation of the AmpC β-lactamase binding site and identification of inhibitors with novel scaffolds. *J Chem Inf Model* 52(5):1367–1375.
47. Hendershot JM, Mishra UJ, Smart RP, Schroeder W, Powers RA (2014) Structure-based efforts to optimize a non-β-lactam inhibitor of AmpC β-lactamase. *Bioorg Med Chem* 22(13):3351–3359.
48. Kussie PH, et al. (1996) Structure of the MDM2 oncoprotein bound to the p53 tumor suppressor transactivation domain. *Science* 274(5289):948–953.
49. Vassilev LT, et al. (2004) In vivo activation of the p53 pathway by small-molecule antagonists of MDM2. *Science* 303(5659):844–848.
50. Zhao Y, Aguilar A, Bernard D, Wang S (2015) Small-molecule inhibitors of the MDM2-p53 protein-protein interaction (MDM2 Inhibitors) in clinical trials for cancer treatment. *J Med Chem* 58(3):1038–1052.
51. Hoyer KK, Dooms H, Barron L, Abbas AK (2008) Interleukin-2 in the development and control of inflammatory disease. *Immunol Rev* 226:19–28.
52. Thanos CD, Randal M, Wells JA (2003) Potent small-molecule binding to a dynamic hot spot on IL-2. *J Am Chem Soc* 125(50):15280–15281.
53. Arkin MR, et al. (2003) Binding of small molecules to an adaptive protein-protein interface. *Proc Natl Acad Sci USA* 100(4):1603–1608.
54. Braisted AC, et al. (2003) Discovery of a potent small molecule IL-2 inhibitor through fragment assembly. *J Am Chem Soc* 125(13):3714–3715.
55. Erlanson DA, Wells JA, Braisted AC (2004) Tethering: Fragment-based drug discovery. *Annu Rev Biophys Biomol Struct* 33:199–223.
56. Tondi D, et al. (2014) Targeting class A and C serine β-lactamases with a broad-spectrum boronic acid derivative. *J Med Chem* 57(12):5449–5458.
57. Orita M, Ohno K, Warizaya M, Amano Y, Niimi T (2011) Lead generation and examples opinion regarding how to follow up hits. *Methods Enzymol* 493:383–419.
58. Hung AW, et al. (2009) Application of fragment growing and fragment linking to the discovery of inhibitors of Mycobacterium tuberculosis pantothenate synthetase. *Angew Chem Int Ed Engl* 48(45):8452–8456.
59. Chung S, Parker JB, Bianchet M, Amzel LM, Stivers JT (2009) Impact of linker strain and flexibility in the design of a fragment-based inhibitor. *Nat Chem Biol* 5(6):407–413.
60. Berman HM, et al. (2000) The Protein Data Bank. *Nucleic Acids Res* 28(1):235–242.
61. Kozakov D, et al. (2015) The FTMap family of web servers for determining and characterizing ligand-binding hot spots of proteins. *Nat Protoc* 10(5):733–755.