



Published in final edited form as:

J Chem Inf Model. 2012 August 27; 52(8): 2236–2244. doi:10.1021/ci300175u.

Relationship between Hot Spot Residues and Ligand Binding Hot Spots in Protein-Protein Interfaces

Brandon S. Zerbe[†], David R. Hall[†], Sandor Vajda^{†,‡}, Adrian Whitty[‡], and Dima Kozakov[†]
Department of Biomedical Engineering, Boston University, Boston, MA, 02215, United States,
and Department of Chemistry, Boston University, Boston, MA, 02215, United States

Sandor Vajda: vajda@bu.edu; Adrian Whitty: whitty@bu.edu; Dima Kozakov: midas@bu.edu

Abstract

In the context of protein-protein interactions, the term “hot spot” refers to a residue or cluster of residues that makes a major contribution to the binding free energy, as determined by alanine scanning mutagenesis. In contrast, in pharmaceutical research a hot spot is a site on a target protein that has high propensity for ligand binding and hence is potentially important for drug discovery. Here we examine the relationship between these two hot spot concepts by comparing alanine scanning data for a set of 15 proteins with results from mapping the protein surfaces for sites that can bind fragment-sized small molecules. We find the two types of hot spots are largely complementary; the residues protruding into hot spot regions identified by computational mapping or experimental fragment screening are almost always themselves hot spot residues as defined by alanine scanning experiments. Conversely, a residue that is found by alanine scanning to contribute little to binding rarely interacts with hot spot regions on the partner protein identified by fragment mapping. In spite of the strong correlation between the two hot spot concepts, they fundamentally differ, however. In particular, while identification of a hot spot by alanine scanning establishes the potential to generate substantial interaction energy with a binding partner, there are additional topological requirements to be a hot spot for small molecule binding. Hence, only a minority of hot spots identified by alanine scanning represent sites that are potentially useful for small inhibitor binding, and it is this subset that is identified by experimental or computational fragment screening.

Introduction

Specific protein-protein interactions (PPIs) are critical events in most biological pathways, including disease pathways where therapeutic intervention could bring widespread benefit. Many PPI interfaces are biologically compelling targets for drug discovery, but the design of compounds that are capable of interfering with protein-protein interactions is notoriously difficult,^{1,2} due in part to our incomplete understanding of the sources of affinity and specificity at such interfaces.³ It is now generally recognized that PPI interfaces include smaller regions, termed “hot spots”, that comprise the subset of residues that contribute the bulk of the binding free energy. Because of the concentration of binding energy they represent, these hot spot regions have been proposed as prime targets for drug binding.^{1,4} The established approach to the identification of such hot spots is alanine scanning

Correspondence to: Sandor Vajda, vajda@bu.edu; Adrian Whitty, whitty@bu.edu; Dima Kozakov, midas@bu.edu.

[†]Department of Biomedical Engineering, Boston University

[‡]Department of Chemistry, Boston University

Supporting Information Available

The identity of the atoms taken as end groups of the residues and the density correlation for all scanned residues are provided in tables within a supplement. This material is available free of charge via the Internet at <http://pubs.acs.org>.

mutagenesis, which involves serially mutating each interface residue to alanine and then measuring the impact of each mutation on the affinity for binding to the partner protein.⁵⁻⁹ Based on this method, a residue is considered a hot spot residue if its mutation to alanine gives rise to a substantial drop in binding affinity (typically tenfold or higher⁴). As alanine scanning mutagenesis has become increasingly used for the analysis of PPIs, experimental data have been accumulated for a large number of complexes.^{10,11}

Concurrent with the development of the thermodynamic concept of hot spots through alanine scanning mutagenesis, two groups independently developed methods that explored protein binding sites using libraries of fragment-sized or even smaller organic “probe” molecules. Ringe and coworkers introduced the method called Multiple Solvent Crystal Structures (MSCS), which involves determining X-ray structures of a target protein in aqueous solutions containing high concentrations of organic co-solvents and then superimposing the structures to find consensus binding sites that accommodate a number of the organic probes binding in well-defined orientations.^{12,13} The MSCS method has been used by several groups, and it has been shown that the consensus sites identify the functionally most important regions of proteins, such as the subsites making up the active site of an enzyme.¹²⁻¹⁶ The same year, Fesik and coworkers published the first results using their Structure-Activity Relationship by Nuclear Magnetic Resonance (SAR by NMR) method, which screens large libraries of fragment-sized organic compounds for binding to target proteins using NMR.¹⁷ They showed that the fragments cluster at ligand binding sites, and that fragment binding rarely occurs anywhere else.¹⁸ Thus, both MSCS and SAR by NMR identify “consensus binding sites” that are capable of binding a variety of small molecules and that have been shown to frequently coincide with drug binding sites.^{13,16,19} As a result, these consensus sites are of prime interest for drug design. Multiple other approaches to performing experimental fragment screens have subsequently been developed^{20,21} but only methods based on X-ray crystallography¹³ or NMR^{17,22-25} give direct information about the location and orientation of the fragment binding and thereby identify the locations of the ligand-binding consensus sites. Fesik and colleagues described such regions as “hot spots on protein surfaces”,¹⁸ and this terminology has become well established in the fragment screening literature (see, e.g.,^{19,25}).

The idea of the importance of a consensus site has not been limited to protein-small molecule interactions. Wells and colleagues demonstrated that the hinge region on the Fc fragment of the human immunoglobulin G interacts with a consensus of at least four different protein scaffolds as well as peptides selected for high affinity.⁴ Based on this observation, DeLano determined that “the overlap of convergent binding sites and hot spots discovered through mutagenesis provides evidence that hot spots genuinely reflect innate properties of certain protein surfaces that greatly promote binding.”⁴ This echoes the earlier speculation by Mattos and Ringe concerning the relationship between consensus sites of small organic molecules and binding affinity.¹³

The term “hot spot” has thus become used within the drug design community to describe the results of two different measurement techniques. Alanine scanning mutagenesis examines contributions to the mutual interaction energy within a protein-protein complex, whereas a consensus site for fragment binding is the property of a single protein. The relationship between these two definitions of binding energy hot spots becomes of interest when considering the design of small molecules to disrupt or modulate PPIs. Substantial research efforts have been devoted to the identification of small molecules that bind to hot spots at protein-protein interfaces,^{1-3,26} and fragment based screening has been shown to be a relatively efficient approach to finding such binders.^{20,27-29} In these cases the hot spots were identified as consensus binding sites rather than based on alanine scanning although mutagenesis data were also available in some cases. Clearly, however, the complementary

hot spot residues on each side of a protein-protein interface would not necessarily define surface sites that are equally good for binding small molecules. How exactly the structural and physicochemical properties of hot spots identified by alanine scanning compare to those discovered by fragment screening remains unclear, which limits our ability to fully exploit data on hot spot locations for structure-based drug design.

The goal of this work is to perform a systematic comparison of hot spots identified by alanine scanning mutagenesis versus by small molecule fragment screening, to clarify the relationship between the protein surface sites identified in these different ways, and to gain a better understanding of their implications for the discovery of small molecule inhibitors of PPIs. For this purpose we analyze X-ray structures of protein-protein complexes from the Protein Data Bank (PDB)³⁰ for which alanine scanning data are available, and where one of the component proteins can be regarded as a receptor because it is known to also bind small molecules or short peptides. There are very few proteins for which there exist both alanine scanning data and public domain information on small molecule binding hot spots from NMR or X-ray fragment screening. We consider here one such system, ribonuclease A (RNase A) interacting with RNase inhibitor (RNI). We additionally analyze 14 other protein-protein complexes, determining the fragment binding sites on the receptor proteins by the method of computational solvent mapping, a virtual analog of the MSCS method that has been shown to accurately identify fragment binding hot spots at PPI surface sites.³¹ Computational mapping places molecular probes — small organic molecules that vary in size and shape — on a dense grid around the protein, finds favorable positions using empirical free energy functions, clusters the conformations, and ranks the clusters on the basis of their average empirical energy.³² The regions that bind a large number of probe clusters identify the consensus binding sites, termed “consensus sites” (CS), and the relative importance of these CSs can be ranked according to the number of bound probe clusters they contain.³² It has been extensively verified that such consensus sites reliably identify the experimentally determined small molecule binding hot spots,^{31–37} and hence we can use the computational solvent mapping results in place of direct data from X-ray or NMR screening experiments. The calculations are performed using the computational solvent mapping algorithm FTMap,³² implemented as a web-based server (see Methods).

We note that functionally important regions of proteins can also be identified based on the conservation of sequence^{38,39} or structure,^{40,41} and that such conserved region may also be called hot spots. Indeed, such methods have been used successfully for the detection of protein-protein, protein-small ligand, and protein-DNA binding sites.⁴⁰ While we fully recognize the importance of the approaches based on evolutionary conservation principles, in this paper we restrict consideration to comparing two hot spot concepts, defined in terms of alanine scanning and fragment screening, respectively, as these are the concepts most frequently mentioned in the context of targeting protein-protein interactions using small molecules.

Results and discussion

Comparison of experimental versus computational fragment mapping of ribonuclease A

To demonstrate the accuracy of the FTMap mapping server for identifying small molecule fragment binding sites, we first compared experimental and computational fragment mapping results for ribonuclease A (RNase A), one of the proteins that will be further studied in this paper. An experimental fragment screen of RNase A by the MSCS method^{12,13} has been reported,¹⁶ in which crystals of bovine pancreatic RNase A were soaked in aqueous solutions of the following organic compounds: 50% dioxane, 50% dimethylformamide, 70% dimethylsulfoxide, 70% 1,6-hexanediol, 70% isopropanol, 50% R,S,R-bisfuran alcohol, 70% t-butanol, 50% trifluoroethanol, or 1.0 M trimethylamine-N-

oxide. The resulting X-ray crystal structures were analyzed to identify positions in which an organic co-solvent molecule could be seen to bind at a specific location and with a well-defined orientation, and the different structures were superimposed to identify consensus sites where the binding locations for multiple different organic probe molecules overlapped. Our analysis focused on consensus sites lying within 4 Å of the binding interface with the partner protein ribonuclease inhibitor (RNI). The MSCS results revealed four consensus sites in this region: two in the B1 pocket, one in the P1 pocket, and one in the B2 pocket (Figure 1(a)).¹⁶ The two sites in the B1 pocket are very close together, and hence Figure 1(a) shows only three distinct sites, one in each of the three pockets.

We used the FTMap server³² (<http://ftmap.bu.edu>) to computationally map the X-ray structure of RNase A, taken from the unbound structure of RNase (PDB code 2e3w), using a standard set of 16 small probe molecules.³² We used a mask to favor the atoms of the RNase A surface that are within 5 Å of any atom of RNI in the complex (see Materials and Methods). Figure 1(b) shows that FTMap identified the same three consensus sites found experimentally by MSCS, as well as an additional, fourth consensus site. Ranked in order of the number of probe clusters in each consensus site, CS1 binds 26 probe clusters in the B1 pocket, CS2 includes 20 clusters in the P1 pocket, CS3 with 15 clusters is at the site that was not identified by MSCS, and CS4 binds 13 clusters in the B2 pocket. Note that Figure 1(b) shows only a single representative pose from each probe cluster rather than all bound probe positions. Although it is not completely clear why the site identified as CS3 was not seen in the MSCS study, we note that this site is surrounded by a long loop comprising residues 32–36 that participates in crystal contacts in the crystal form used in the MSCS experiments,¹⁶ and it is known that crystal contacts can prevent access to regions that are hot spots in a different crystal form.⁴² The notion that CS3 represents a real binding site is supported by the observation that this site on RNase A accommodates a hot spot residue on RNI identified by alanine scanning mutagenesis (Figure 1(d)), as will be further discussed below. The observation that FTMap accurately identifies consensus sites for fragment binding to RNase A is supported by extensive previous evaluations showing that FTMap reliably predicts the results of screening experiments,^{32,35,43} including identification of small molecule binding sites at protein-protein interfaces.³¹

Correlation between the two hot spot concepts

Using the FTMap server to identify consensus sites for fragment binding, we studied the 15 protein-protein complexes listed in Table 1, which lists their PDB codes, names, receptor/ligand role, and the source for alanine scanning data. The complexes are either enzymes with inhibiting proteins, or signal modulating protein complexes where it has been shown that the binding of the partner protein can be recapitulated by a small peptide. Mutated residues were classified as hot spot residues if mutation to alanine resulted in a change in binding free energy, $\Delta\Delta G$, exceeding $2.0 \frac{\text{kcal}}{\text{mol}}$ or were reported in their original source as having a strong effect on binding. The term “neutral residues” was adopted from Kortemme and Baker⁴⁴ to indicate residues that when mutated did not meet the hot spot residue criterion of $\Delta\Delta G \geq 2.0 \frac{\text{kcal}}{\text{mol}}$.

The receptor from each complex was mapped using the FTMap server,³² again using a mask to favor the region of the surface within 5 Å of the partner protein. To establish the extent to which the hot spot residues found by alanine scanning of the partner protein tend to bind to the most important consensus sites on the receptor, we developed an overlap measure to quantify the degree to which the atoms of the hot spot residue overlap with the region of space that comprises a given FTMap consensus site, illustrated in Figure 1(d) for the case of RNase A with RNI. We call this overlap measure the “density correlation” (DC). The DC for a given hot spot residue is defined as the total number of atoms from all probes in a

given FTMap consensus site (i.e. a quantitative measure of probe density) that lie within 2 Å of the end-group atom of the amino acid side chain of the hot spot residue (defined in Table S1 within the Supplement). As the above definition indicates, the DC is not normalized; its value increases with increasing importance of the consensus site around the selected side chain. The motivation for restricting consideration to neighborhoods of side chain atoms beyond the C β atom is that these are the atoms that are eliminated when the residue is mutated to an alanine, and hence it is primarily their contributions that determine the change in binding free energy that accompanies the mutation.

The density correlation values for all residues probed in the published alanine scanning data sets for the 15 complexes are shown in Tables S2 and S3 of the Supplement. and a histogram of the DC values for hot spots and neutral residues is shown in Figure 2. Figure 2 shows that the vast majority of hot spot residues have DC values of ≥ 1000 , indicating extensive spatial overlap with important consensus sites on the receptor. In sharp contrast, the majority of neutral residues have DC = 0, indicating that these residues do not interact with the sites on the receptor that bind small molecule fragments. Using a threshold of DC > 600 to define a residue that binds into a consensus site, comparison of the alanine scanning data with our mapping results shows that 92% (34 of 37) of hot spot residues protrude into a consensus site, and 92% (49 of 53) of neutral residues do not (see Table 2). A χ^2 test for independence was conducted on this data, and $\chi^2 = 59.5$ ($p = 0.001$), which implies that the association between residue type (i.e., hot spot or neutral) and the density correlation value is extremely statistically significant. Figure 2 also shows that this result does not heavily depend on the threshold DC > 600; any cut-off value between 30 and 3000 would provide very good separation between neutral and hot spot residues.

The overlap between hot spot residues and consensus sites for each of the proteins studied is illustrated in Figure 3, which for each structure shows the receptor protein, the probe densities obtained by FTMap, and the ligand side chain atoms beginning with C β from residues that have been mutated in alanine scanning experiments.

The above results reveal a strong correlation between the two hot spot concepts. Clearly, the biophysical forces that render a particular site capable of binding small molecule fragments also make it very likely that an amino acid residue of the partner protein, protruding into the same region, will make a major contribution to the binding free energy and thus will be identified as a hot spot residue by alanine scanning. Because a very large fraction of the hot spot residues protrude into consensus sites on the partner protein, one could consider FTMap as another predictor of hot spot residues, similar to the many other methods developed for predicting the results of alanine scanning experiments. An important difference is that these other methods are usually trained on alanine scanning data,^{44,46,51–53} whereas for FTMap this is not the case. In fact, FTMap has been developed for the identification and characterization of binding sites of proteins as a virtual analogue of X-ray or NMR based fragment screening experiments,^{12,13,17,54} and only recently has been extended for the analysis of ligand binding pockets in protein-protein interfaces.³¹

Differences between the two hot spot concepts

Despite the strong correlation between the two measurements of hot spots that was established in the previous section, there are important differences which must be understood if the relevance of hot spots for drug discovery is to be fully grasped. These differences arise primarily from the fact that alanine scanning mutagenesis probes reciprocal interactions within a specific protein-protein complex. In contrast, a consensus site found by small molecule fragment screening or its computational equivalent reflects an intrinsic property of an individual protein, and the site is expected to be important in any interaction that involves that region of the target independent of any partner protein. We can gain

additional insight into the relationship between the two measurement techniques for determining hot spots by examining the few instances in our data set in which the definitions were not in accord.

One obvious difference between the methods is that alanine scanning involves perturbing the structure of one of the proteins, whereas fragment screening does not. Mutating a surface residue at a protein-protein interface can affect binding affinity through a number of mechanisms. Some of these mechanisms reflect a strictly local change in one or several components of the interaction energy with the partner protein, such as loss of optimal van der Waals contacts, loss of electrostatic pairings — either because they provide increased binding energy compared to solvent or because they become destabilizing if left unsatisfied — or loss of buried nonpolar surface area affecting the hydrophobic contribution to the binding free energy.⁴ However, mutation of a single residue to alanine can also have longer-range effects by disrupting side chain packing at the interface. This disruption can cause conformational changes extending beyond the side chain in question, or can cause local unfolding or increased entropy of unbound states. Interpreting alanine scanning data as reflecting only the direct contributions of individual amino acid side chains to interaction energy with the partner protein is therefore an approximation. For the mutations included in our test set, the quality of the correlation between mutagenic hot spots and fragment binding consensus sites shown in Figure 3 implies that these particular mutations mostly reflect local, direct effects on binding interactions with the receptor. Nevertheless, when interpreting a particular result from alanine scanning, the possibility of longer range effects often cannot be ruled out. Clearly, there is no corresponding concern about mutational perturbation of the protein structure in the case of computational or experimental fragment screening, which uses the wild-type protein.

Furthermore, alanine scanning assumes that protein hot spots are established by the interaction of the side chain; however, there is no *a priori* reason that dictates main chain interactions are not important within PPIs. As seen in Figure 3 panels B1 and D1, regions of dense FTMap results extensively correlate with neutral residue main chain atoms from the ligand protein while the side chains extend out of these regions, and our use of end group atoms (see Table S1) to represent the character of these side chains is able to correctly identify these residues as neutral residues. In our recent work, though, we have seen that regions of dense FTMap results correspond to important interactions with small ligands,^{43,55} and this suggests that main chain interactions residing in such regions of dense prediction may be significantly contributing to the binding free energy of a protein-protein complex. The consensus sites identified by solvent mapping may provide insight into such hot spots that are otherwise missed by alanine scanning measurements; however, there is some evidence that our methodology developed within this work may be incorrectly identifying some neutral residues due to their proximity to hot spots in which the main chain may be involved. As shown in panel B1 of Figure 3, mapping ZipA and selecting the residues protruding into the resulting consensus sites correctly identifies all three FtsZ hot spot residues. However, the neutral residue Leu6 also overlaps with a consensus site, and thus appears to violate the expectation that residues that project into consensus sites will be hot spot residues. Closer examination of the structure in this region shows that the interaction likely involves the main chain atoms of Leu6 projecting into a polar pocket on ZipA. The conformation of the side chain of Leu6 is such that the end group atom used in the DC calculation, C γ , is less than 2 Å away from C β , causing this atom to overlap with the consensus site even though the side chain atoms are not driving the interaction. Mutation of Leu6 to alanine would not be expected to disrupt this interaction if it is driven by contacts with main chain atoms on Leu6, since alanine would be expected to recover these important contacts with ZipA. Thus, Leu6 appear to be an example of an alanine scanning false negative of the kind first suggested by Wells.⁵⁶ One might then expect that under-

identification by alanine scanning of residues are important for binding, but interact mostly through main chain atoms and $C\beta$, is more likely to occur for relatively small amino acids. Indeed, for the 15 proteins studied in this paper we found four neutral residues that protrude into consensus sites. Among these four, two are leucines, and one is cysteine. The potential for alanine scanning to underestimate the energetic role of small residues accords with the observation that hot spot residues are dominated by amino acids with large side chains, primarily tryptophans, tyrosines, and arginines.⁹

Another reason why a small number of residues might deviate from the correlation established in Figure 3 is if it is identified as a hot spot by alanine scanning mutagenesis not because it projects into a consensus site on the partner protein, but because it forms part of the structure of a consensus site on the protein to which it belongs. An example of this behavior is seen for the complex of RNase A with RNI. Alanine scanning identifies four residues on RNI as hot spots in this complex: Tyr430, Asp431, Tyr433 and Trp259. Computational solvent mapping identified Tyr430, Asp431, and Tyr433 as protruding into consensus sites on RNase A, but Trp259 contacts a region of RNase A that does not coincide with any consensus site on that protein (Figure 1(d) and Figure 3, panel B3). FTMap analysis of RNI itself, however, shows that Trp259 participates in the formation of a consensus site of RNI, which so far we had treated as a ligand rather than a receptor. In fact, the mapping of RNI placed the top ranking consensus site directly adjacent to Trp259, in a pocket which Gly88 and Ser89 of RNase A occupy in the complex (Figure 4). Thus, the relationship between the two hot spot concepts is not always as simple as described in the previous section, where a residue is a hot spot residue because it resides within a physicochemical environment, created by the partner protein, which is energetically favorable. A residue can also be identified as a hot spot by alanine scanning if it contributes to creating such a favorable binding environment by being among the residues forming a consensus site on the protein to which it belongs. These two different origins of “hotness” have been noted by Nussinov and co-workers, who observed that most hot spot residues either protrude into or are located in complemented pockets.^{57,58} They defined a pocket as “complemented” if it becomes filled with atoms of the partner upon binding, thus representing a favorable binding environment, in contrast to unfilled pockets that remain empty (or filled with solvent) after protein-protein complexation. Our analysis establishes that, in cases where one of the interacting proteins can clearly be identified as the “receptor” by virtue of its ability to bind to a small molecule ligand or substrate or to a ligand-derived peptide, the overwhelming majority of hot spot residues on the ligand protein derive from their role in interacting with consensus sites on the receptor that have an intrinsic tendency to interact strongly with other molecules. However, when both sides of a PPI interfaces are investigated, hot spot residues like Trp259 that contribute to forming a consensus site rather than projecting into one will be observed.

This duality of “hotness” that is apparent in the consensus identification of hot spots is the largest distinction between the two hot spot concepts. In alanine scanning experiments particularly those involving larger, flatter or more complex PPI interfaces, more or less complementary hot spots are observed on both proteins involved in the complex because an energetically important region of the interface can typically be disrupted by mutating the residues involved on either of the two opposing protein surfaces.^{6,8} But the fact that hot spots of similar size and energetic importance can be identified on the binding surfaces of both proteins by no means implies that both proteins will be equally amenable to binding small molecule ligands. A convex surface site on a protein typically will not bind small molecules strongly no matter how much binding energy the region generates in an interaction with a complementary convex site on its protein binding partner. Thus, observation of a hot spot by alanine scanning mutagenesis does not necessarily imply the existence of a small molecule fragment consensus site at that region.

Conclusions

Our analysis shows that there is a very strong relationship between consensus sites on the receptor (i.e. on the concave side of the interface) and hot spot residues on the binding partner identified by alanine scanning mutagenesis; however, these residues represent only a subset of the hot spot residues that are identified by alanine scanning mutagenesis when both binding partners are scanned. Consequently, only a subset of hot spots that can be identified by alanine scanning have the potential to bind small organic ligands. It is this subset of hot spots, which combine both a potential to generate substantial binding energy plus a substantially concave topology, that is identified as consensus sites by experimental fragment screening and by FTMap.

Alanine scanning data must therefore be used judiciously when searching for sites that might bind small molecules designed to disrupt a protein-protein interaction. The existence of a hotspot identified by alanine scanning mutagenesis does not imply the existence of a druggable site at that region. In contrast, the observation of strong consensus sites by fragment screening is a necessary and possibly also a sufficient condition for finding such druggable sites.^{18,59} We have shown previously that, based on mapping of the receptor using 16 different types of probe molecules, a druggable site can be defined as a region that comprises a main hot spot binding at least 16 probe clusters in the protein-protein interface together with one or two additional hot spots close enough to be reached from the first site by a drug-sized molecule.³¹ As shown in the current study, such sites can also be identified by considering regions that accommodate hot spot residues of the partner protein, provided that a local topology appropriate to binding a small ligand can reasonably be assumed. Nonetheless, such alanine scanning data will not provide information on main chain interactions that significantly contribute to the binding free energy, and the development of drug-sized molecules should benefit from the identification of such regions as sites into which the molecule may be expanded for optimal binding affinity. FTMap might provide such insight.

Finding small molecule inhibitors has a better chance if the hot spots are presented relatively close together within a contiguous stretch on the partner protein's surface.^{31,55} Accordingly, many of the currently known protein-protein inhibitors have been designed starting from a short peptide fragment of the ligand protein. In many cases the relevant portion of the ligand protein is largely unstructured when unbound, and hence the receptor was co-crystallized only with the peptide fragment. Such receptor/peptide pairs representing current PPI interaction targets include ZipA/FtsZ,⁴⁷ Bcl-xL/BAK,⁴⁵ MDM2/p53 peptide,⁴⁹ XIAP BIR3/SMAC,⁶⁰ PDZ domain of PSD95/peptide ligand,⁶¹ Pin1/phosphopeptide,⁶² eIF4E/4E-BP1 peptide,⁶³ and NEMO/IKK peptide.⁵⁰ In all of these complexes, important side chains of the peptide, in many cases shown to be hot spot residues by alanine scanning, protrude into consensus binding sites of the receptor protein, and the same sites on the receptor also bind the important functional groups of small molecular inhibitors.³¹ Computational mapping easily identifies such consensus sites, even when mapping the ligand-free receptor structure;³¹ nonetheless, it is better to map the peptide-bound structure, if such is available, since it yields exactly the same consensus sites as mapping the protein structure co-crystallized with small inhibitors.³¹

The idea of using hot spot residues as a starting point for inhibitor design has been extended in a recent paper.⁶⁴ Because alanine scanning data are not always available, Camacho and coworkers suggested the use of so-called "anchor" residues instead of hot spot residues.⁶⁴ An "anchor residue" is defined as a residue for which solvent accessible surface area (SASA) changes at least by 0.5 \AA^2 upon binding, and if based on an empirical energy calculation it contributes at least $0.5 \frac{\text{kcal}}{\text{mol}}$ to the binding free energy. The ANCHOR database

lists pre-computed anchor residues from more than 30,000 PDB entries with at least two protein chains.⁶⁴ The goal is clearly to identify potential hot spot residues, but due to the relatively low thresholds used, all protein interface include a substantial number of “anchor” residues. Since the receptor protein can be easily mapped with the FTMap server,³² and based on the results of this paper the residues overlapping with consensus sites are almost always hot spots, selecting a high probe density region which, in addition, accommodates one or more “anchor” residues presumably will improve the chances of locating a druggable site. While FTMap currently provides only limited details on chemical specificity within hot spots, “anchor” residues do provide specificity information concerning possible drug scaffolds.⁶⁴ These drug scaffolds can then be optimized into regions identified by FTMap for the development of a lead compound. Evidence for such an approach within fragment-based drug design has recently been published by our group,⁵⁵ and the combination of “anchor” residues as starting points and the identification of important surrounding sites by FTMap should hopefully provide a rational framework for the development of lead-like molecules for some PPIs.

Materials and Methods

Computational mapping

Receptors of the protein complexes listed in Table 1 were mapped using the FTMap server (available via <http://ftmap.bu.edu>) using a mask containing all receptor protein atoms further away than 5 Å of the corresponding ligand protein. This process was also applied to one of the ligand proteins, ribonuclease inhibitor (RNI), of the complex structure with PDB code 1dfj. Details on the FTMap algorithm have been previously reported.³² All atoms from the minimized probe positions for the 16 probes were retained to generate an atomic density. A density correlation measure was calculated for each residue of the ligand protein by superimposing the bound ligand structure on the probe distribution from the mapping, and counting all atoms within 2 Å of the residue’s end group as defined in Table S1. This data was then binned and visualized within the scientific software package SigmaPlot 11.0.

Density visualization

The atomic density was placed into a $0.2 \text{ \AA} \times 0.2 \text{ \AA} \times 0.2 \text{ \AA}$ grid. All probe atoms were added to the nearest grid point so that each grid point contained the total number of probe atoms within the $.008 \text{ \AA}^3$ volume. To account for uncertainty in the atomic position as well as to smooth the resulting grid, the grid point was convolved 3 times consecutively with a step function that is $\frac{1}{\sqrt{3}}$ for $i, j, k \in (-1, 0, 1)$ and 0 everywhere else. Resulting grid points with values greater than $\frac{1}{\sqrt{3}}$ (corresponding to 375 atoms/ \AA^3) and 10 (corresponding to 1,250 atoms/ \AA^3) were reported and visualized along with the receptor and ligand’s alanine scanned residues using PyMol (<http://www.pymol.org>).

Density correlation calculation

A calculation that determines the distance between every ligand and density atom is computationally intensive; however, most atoms will be farther away than 2 Å which is the distance used to mark if the ligand correlates with the density atom. Therefore, we implemented an algorithm to speed this calculation while exactly counting the total number of density atoms that correlate with a ligand atom. The idea is to associate each probe atom with a point on a $2 \text{ \AA} \times 2 \text{ \AA} \times 2 \text{ \AA}$ grid. When the calculation for the correlation of a specific ligand atom is made, only the distances between the ligand atom and the probe atoms at grid points no further than $2 \text{ \AA} + \sqrt{3}$ needs to be computed. This eliminates unnecessary distance calculations and return the exact number of probe atoms within 2 Å of the ligand atom. The density correlations for the hot spot residues and neutral residues calculated in this manner are displayed in Tables S2 and S3, respectively.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This investigation was supported by grants GM064700, GM094551, GM061867 and GM93147 from the National Institute of General Medicine.

References

1. Wells JA, McClendon CL. Reaching for high-hanging fruit in drug discovery at protein-protein interfaces. *Nature*. 2007; 450:1001–1009. [PubMed: 18075579]
2. Berg T. Small-molecule inhibitors of protein-protein interactions. *Curr Opin Drug Discov Devel*. 2008; 11:666–674.
3. Whitty A, Kumaravel G. Between a rock and a hard place? *Nat Chem Biol*. 2006; 2:112–118. [PubMed: 16484997]
4. DeLano WL. Unraveling hot spots in binding interfaces: progress and challenges. *Curr Opin Struct Biol*. 2002; 12:14–20. [PubMed: 11839484]
5. Cunningham BC, Wells JA. High-resolution epitope mapping of hGH-receptor interactions by alanine-scanning mutagenesis. *Science*. 1989; 244:1081–1085. [PubMed: 2471267]
6. Cunningham BC, Wells JA. Comparison of a structural and a functional epitope. *J Mol Biol*. 1993; 234:554–563. [PubMed: 7504735]
7. Jin L, Wells JA. Dissecting the energetics of an antibody-antigen interface by alanine shaving and molecular grafting. *Protein Sci*. 1994; 3:2351–2357. [PubMed: 7538848]
8. Clackson T, Wells JA. A hot spot of binding energy in a hormone-receptor interface. *Science*. 1995; 267:383–386. [PubMed: 7529940]
9. Bogan AA, Thorn KS. Anatomy of hot spots in protein interfaces. *J Mol Biol*. 1998; 280:1–9. [PubMed: 9653027]
10. Thorn KS, Bogan AA. ASEdb: a database of alanine mutations and their effects on the free energy of binding in protein interactions. *Bioinformatics*. 2001; 17:284–285. [PubMed: 11294795]
11. Fischer TB, Arunachalam KV, Bailey D, Mangual V, Bakhru S, Russo R, Huang D, Paczkowski M, Lalchandani V, Ramachandra C, Ellison B, Galer S, Shapley J, Fuentes E, Tsai J. The binding interface database (BID): a compilation of amino acid hot spots in protein interfaces. *Bioinformatics*. 2003; 19:1453–1454. [PubMed: 12874065]
12. Allen KN, Bellamacina CR, Ding X, Jeffrey CJ, Mattos C, Petsko GA, Ringe D. An experimental approach to mapping binding surfaces of crystalline proteins. *J Phys Chem*. 1996; 100:595–599.
13. Mattos C, Ringe D. Locating and characterizing binding sites on proteins. *Nat Biotechnol*. 1996; 14:595–599. [PubMed: 9630949]
14. English AC, Done SH, Caves LS, Groom CR, Hubbard RE. Locating interaction sites on proteins: the crystal structure of thermolysin soaked in 2% to 100% isopropanol. *Proteins: Struct, Funct Bioinf*. 1999; 37:628–640.
15. English AC, Groom CR, Hubbard RE. Experimental and computational mapping of the binding surface of a crystalline protein. *Protein Eng*. 2001; 14:47–59. [PubMed: 11287678]
16. 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 Bioinf*. 2009; 76:861–881.
17. Shuker SB, Hajduk PJ, Meadows RP, Fesik SW. Discovering high-affinity ligands for proteins: SAR by NMR. *Science*. 1996; 274:1531–1534. [PubMed: 8929414]
18. Hajduk PJ, Huth JR, Fesik SW. Druggability indices for protein targets derived from NMR-based screening data. *J Med Chem*. 2005; 48:2518–2525. [PubMed: 15801841]
19. Congreve M, Chessari G, Tisi D, Woodhead AJ. Recent developments in fragment-based drug discovery. *J Med Chem*. 2008; 51:3661–3680. [PubMed: 18457385]
20. Erlanson DA, McDowell RS, O'Brien T. Fragment-based drug discovery. *J Med Chem*. 2004; 47:3463–3482. [PubMed: 15214773]

21. Neumann T, Junker HD, Schmidt K, Sekul R. SPR-based fragment screening: advantages and applications. *Curr Top Med Chem.* 2007; 7:1630–1642. [PubMed: 17979772]
22. Hann MM, Leach AR, Harper G. Molecular complexity and its impact on the probability of finding leads for drug discovery. *J Chem Inf Comput Sci.* 2001; 41:856–864. [PubMed: 11410068]
23. Boehm HJ, Boehringer M, Bur D, Gmuender H, Huber W, Klaus W, Kostrewa D, Kuehne H, Luebbers T, Meunier-Keller N, Mueller F. Novel inhibitors of DNA gyrase: 3D structure based biased needle screening, hit validation by biophysical methods, and 3D guided optimization. A promising alternative to random screening. *J Med Chem.* 2000; 43:2664–2674. [PubMed: 10893304]
24. Hartshorn MJ, Murray CW, Cleasby A, Frederickson M, Tickle IJ, Jhoti H. Fragment-based lead discovery using X-ray crystallography. *J Med Chem.* 2005; 48:403–413. [PubMed: 15658854]
25. Ciulli A, Williams G, Smith AG, Blundell TL, Abell C. Probing hot spots at protein-ligand binding sites: a fragment-based approach using biophysical methods. *J Med Chem.* 2006; 49:4992–5000. [PubMed: 16884311]
26. Arkin MR, Wells JA. Small-molecule inhibitors of protein-protein interactions: progressing towards the dream. *Nat Rev Drug Discovery.* 2004; 3:301–317.
27. Lesuisse D, Lange G, Deprez P, Benard D, Schoot B, Delettre G, Marquette JP, Broto P, Jean-Baptiste V, Bichet P, Sarubbi E, Mandine E. SAR and X-ray. A new approach combining fragment-based screening and rational drug design: application to the discovery of nanomolar inhibitors of Src SH2. *J Med Chem.* 2002; 45:2379–2387. [PubMed: 12036348]
28. Schade M, Oschkinat H. NMR fragment screening: tackling protein-protein interaction targets. *Curr Opin Drug Discov Devel.* 2005; 8:365–373.
29. Petros AM, et al. Discovery of a potent inhibitor of the antiapoptotic protein Bcl-xL from NMR and parallel synthesis. *J Med Chem.* 2006; 49:656–663. [PubMed: 16420051]
30. Berman HM, Bhat TN, Bourne PE, Feng Z, Gilliland G, Weissig H, Westbrook J. The Protein Data Bank and the challenge of structural genomics. *Nat Struct Biol.* 2000; 7(Suppl):957–959. [PubMed: 11103999]
31. Kozakov D, Hall DR, Chuang GY, Cencic R, Brenke R, Grove LE, Beglov D, Pelletier J, Whitty A, Vajda S. Structural conservation of druggable hot spots in protein-protein interfaces. *Proc Natl Acad Sci USA.* 2011; 108:13528–13533. [PubMed: 21808046]
32. Brenke R, Kozakov D, Chuang GY, Beglov D, Hall D, Landon MR, Mattos C, Vajda S. Fragment-based identification of druggable ‘hot spots’ of proteins using Fourier domain correlation techniques. *Bioinformatics.* 2009; 25:621–627. [PubMed: 19176554]
33. Dennis S, Kortvelyesi T, Vajda S. Computational mapping identifies the binding sites of organic solvents on proteins. *Proc Natl Acad Sci USA.* 2002; 99:4290–4295. [PubMed: 11904374]
34. 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; 332:1095–1113. [PubMed: 14499612]
35. Landon MR, Lieberman RL, Hoang QQ, Ju S, Caaveiro JM, Orwig SD, Kozakov D, Brenke R, Chuang GY, Beglov D, Vajda S, Petsko GA, 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. [PubMed: 19521672]
36. Chuang GY, Kozakov D, Brenke R, Beglov D, Guarnieri F, Vajda S. Binding hot spots and amantadine orientation in the influenza A virus M2 proton channel. *Biophys J.* 2009; 97:2846–2853. [PubMed: 19917240]
37. Hall DH, Grove LE, Yueh C, Ngan CH, Kozakov D, Vajda S. Robust identification of binding hot spots using continuum electrostatics: application to hen egg-white lysozyme. *J Am Chem Soc.* 2011; 133:20668–20671. [PubMed: 22092261]
38. Lichtarge O, Bourne HR, Cohen FE. An evolutionary trace method defines binding surfaces common to protein families. *J Mol Biol.* 1996; 257:342–358. [PubMed: 8609628]
39. Casari G, Sander C, Valencia A. A method to predict functional residues in proteins. *Nat Struct Biol.* 1995; 2:171–178. [PubMed: 7749921]
40. Konc J, Janezic D. ProBiS algorithm for detection of structurally similar protein binding sites by local structural alignment. *Bioinformatics.* 2010; 26:1160–1168. [PubMed: 20305268]

41. Carl N, Konc J, Vehar B, Janezic D. Protein-protein binding site prediction by local structural alignment. *J Chem Inf Model*. 2010; 50:1906–1913. [PubMed: 20919700]
42. Buhrman G, O'Connor C, Zerbe B, Kearney BM, Napoleon R, Kovrigina EA, Vajda S, Kozakov D, Kovrigin EL, Mattos C. Analysis of binding site hot spots on the surface of Ras GTPase. *J Mol Biol*. 2011; 413:773–789. [PubMed: 21945529]
43. Ngan CH, Hall DR, Zerbe B, Grove LE, Kozakov D, Vajda S. FTSite: high accuracy detection of ligand binding sites on unbound protein structures. *Bioinformatics*. 2012; 28:286–287. [PubMed: 22113084]
44. Kortemme T, Baker D. A simple physical model for binding energy hot spots in protein-protein complexes. *Proc Natl Acad Sci USA*. 2002; 99:14116–14121. [PubMed: 12381794]
45. Sattler M, Liang H, Nettessheim D, Meadows RP, Harlan JE, Eberstadt M, Yoon HS, Shuker SB, Chang BS, Minn AJ, Thompson CB, Fesik SW. Structure of Bcl-xL-Bak peptide complex: recognition between regulators of apoptosis. *Science*. 1997; 275:983–986. [PubMed: 9020082]
46. Darnell SJ, Page D, Mitchell JC. An automated decision-tree approach to predicting protein interaction hot spots. *Proteins: Struct, Funct Bioinf*. 2007; 68:813–823.
47. Mosyak L, Zhang Y, Glasfeld E, Haney S, Stahl M, Seehra J, Somers WS. The bacterial cell-division protein ZipA and its interaction with an FtsZ fragment revealed by X-ray crystallography. *EMBO J*. 2000; 19:3179–3191. [PubMed: 10880432]
48. Gordon NC, Pan B, Hymowitz SG, Yin J, Kelley RF, Cochran AG, Yan M, Dixit VM, Fairbrother WJ, Starovasnik MA. BAFF/BLYS receptor 3 comprises a minimal TNF receptor-like module that encodes a highly focused ligand-binding site. *Biochemistry*. 2003; 42:5977–5983. [PubMed: 12755599]
49. Li C, Pazgier M, Li C, Yuan W, Liu M, Wei G, Lu WY, Lu W. Systematic mutational analysis of peptide inhibition of the p53-MDM2/MDMX interactions. *J Mol Biol*. 2010; 398:200–213. [PubMed: 20226197]
50. Rushe M, Silvian L, Bixler S, Chen LL, Cheung A, Bowes S, Cuervo H, Berkowitz S, Zheng T, Guckian K, Pellegrini M, Lugovskoy A. Structure of a NEMO/IKK-associating domain reveals architecture of the interaction site. *Structure*. 2008; 16:798–808. [PubMed: 18462684]
51. Cho KI, Kim D, Lee D. A feature-based approach to modeling protein-protein interaction hot spots. *Nucleic Acids Res*. 2009; 37:2672–2687. [PubMed: 19273533]
52. Tuncbag N, GURSOY A, Keskin O. Identification of computational hot spots in protein interfaces: combining solvent accessibility and inter-residue potentials improves the accuracy. *Bioinformatics*. 2009; 25:1513–1520. [PubMed: 19357097]
53. Xia JF, Zhao XM, Song J, Huang DS. APIS: accurate prediction of hot spots in protein interfaces by combining protrusion index with solvent accessibility. *BMC Bioinformatics*. 2010; 11:174. [PubMed: 20377884]
54. Liepinsh E, Otting G. Specificity of urea binding to proteins. *J Am Chem Soc*. 1994; 116:9670–9674.
55. Hall DR, Ngan CH, Zerbe BS, 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. [PubMed: 22145575]
56. Wells JA. Systematic mutational analyses of protein-protein interfaces. *Meth Enzymol*. 1991; 202:390–411. [PubMed: 1723781]
57. Li X, Keskin O, Ma B, Nussinov R, Liang J. Protein-protein interactions: hot spots and structurally conserved residues often locate in complemented pockets that pre-organized in the unbound states: implications for docking. *J Mol Biol*. 2004; 344:781–795. [PubMed: 15533445]
58. Shulman-Peleg A, Shatsky M, Nussinov R, Wolfson HJ. Spatial chemical conservation of hot spot interactions in protein-protein complexes. *BMC Biol*. 2007; 5:43. [PubMed: 17925020]
59. Landon MR, Lancia DR, Yu J, Thiel SC, Vajda S. Identification of hot spots within druggable binding regions by computational solvent mapping of proteins. *J Med Chem*. 2007; 50:1231–1240. [PubMed: 17305325]
60. Liu Z, Sun C, Olejniczak ET, Meadows RP, Betz SF, Oost T, Herrmann J, Wu JC, Fesik SW. Structural basis for binding of Smac/DIABLO to the XIAP BIR3 domain. *Nature*. 2000; 408:1004–1008. [PubMed: 11140637]

61. Doyle DA, Lee A, Lewis J, Kim E, Sheng M, MacKinnon R. Crystal structures of a complexed and peptide-free membrane protein-binding domain: molecular basis of peptide recognition by PDZ. *Cell*. 1996; 85:1067–1076. [PubMed: 8674113]
62. Verdecia MA, Bowman ME, Lu KP, Hunter T, Noel JP. Structural basis for phosphoserine-proline recognition by group IV WW domains. *Nat Struct Biol*. 2000; 7:639–643. [PubMed: 10932246]
63. Moerke NJ, Aktas H, Chen H, Cantel S, Reibarkh MY, Fahmy A, Gross JD, Degtrev A, Yuan J, Chorev M, Halperin JA, Wagner G. Small-molecule inhibition of the interaction between the translation initiation factors eIF4E and eIF4G. *Cell*. 2007; 128:257–267. [PubMed: 17254965]
64. Meireles LM, Domling AS, Camacho CJ. ANCHOR: a web server and database for analysis of protein-protein interaction binding pockets for drug discovery. *Nucleic Acids Res*. 2010; 38:W407–W411. [PubMed: 20525787]

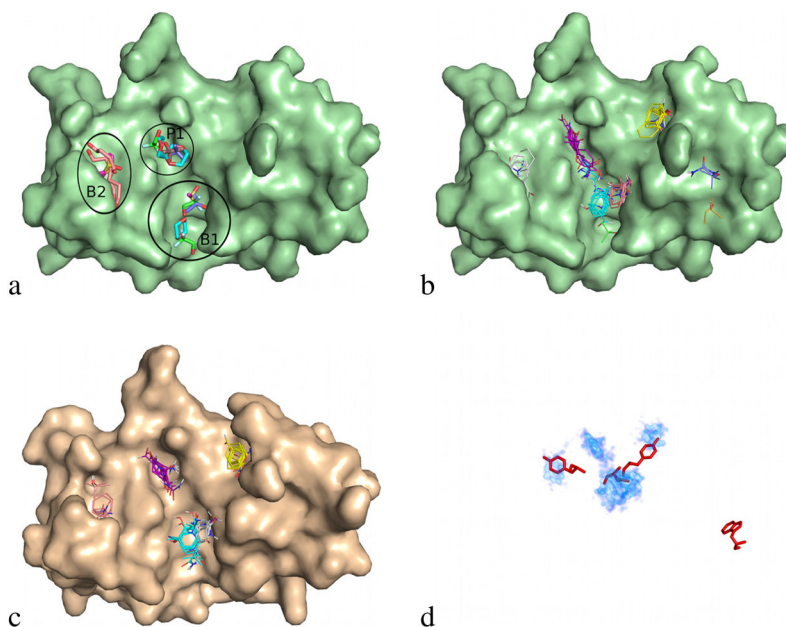


Figure 1.

Comparison of techniques that discover hot spots for ribonuclease A (RNase A). Unbound and ribonuclease inhibitor (RNI) bound RNase A are visualized as a pale green and wheat surfaces, respectively. (a) All ordered solvent molecules from the 9 structures solved in 9 different solvents for the MSCS method.¹⁶ Carbons are colored according to the solvent identity: dioxane colored green, isopropanol colored salmon, hexanediol colored magenta, and dimethylsulfoxide colored cyan. (b) Consensus clusters from applying the FTMap algorithm to unbound RNase A structure. Carbons of probes are colored by cluster rank: top cyan, second purple, third yellow, fourth salmon, fifth white, sixth green. Notice that the top-ranked (1–4) computationally identified consensus clusters identify similar regions to 1a except the third ranked consensus cluster. (c) Consensus clusters from applying the FTMap algorithm to the RNI-bound RNase A structure. Consensus clusters are colored by the same scheme as (b). (d) FTMap results visualized as four regions that have high probe atom density (cyan and blue cloud) with the residues (red sticks) of the aligned and superimposed RNI structure that were found to have $\geq 2.0 \frac{\text{kcal}}{\text{mol}}$ contribution to the binding free energy by alanine scanning mutagenesis.

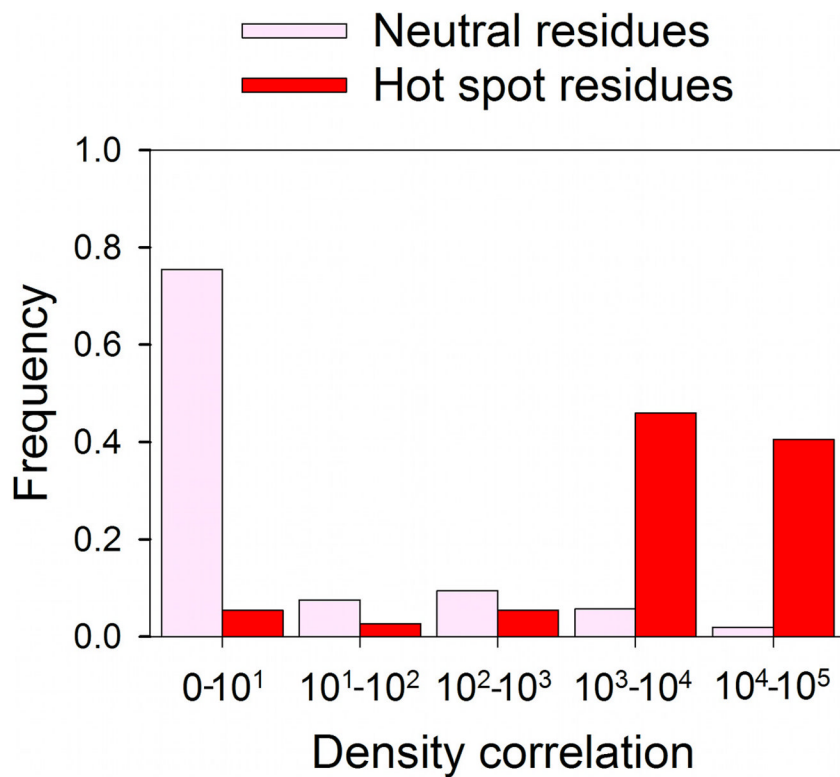


Figure 2. The distributions of density correlation values. Frequencies are shown in pink and red for neutral and hot spot residues, respectively. The two distributions only minimally overlap, indicating that almost all hot spot residues protrude into consensus sites and neutral residues rarely do.

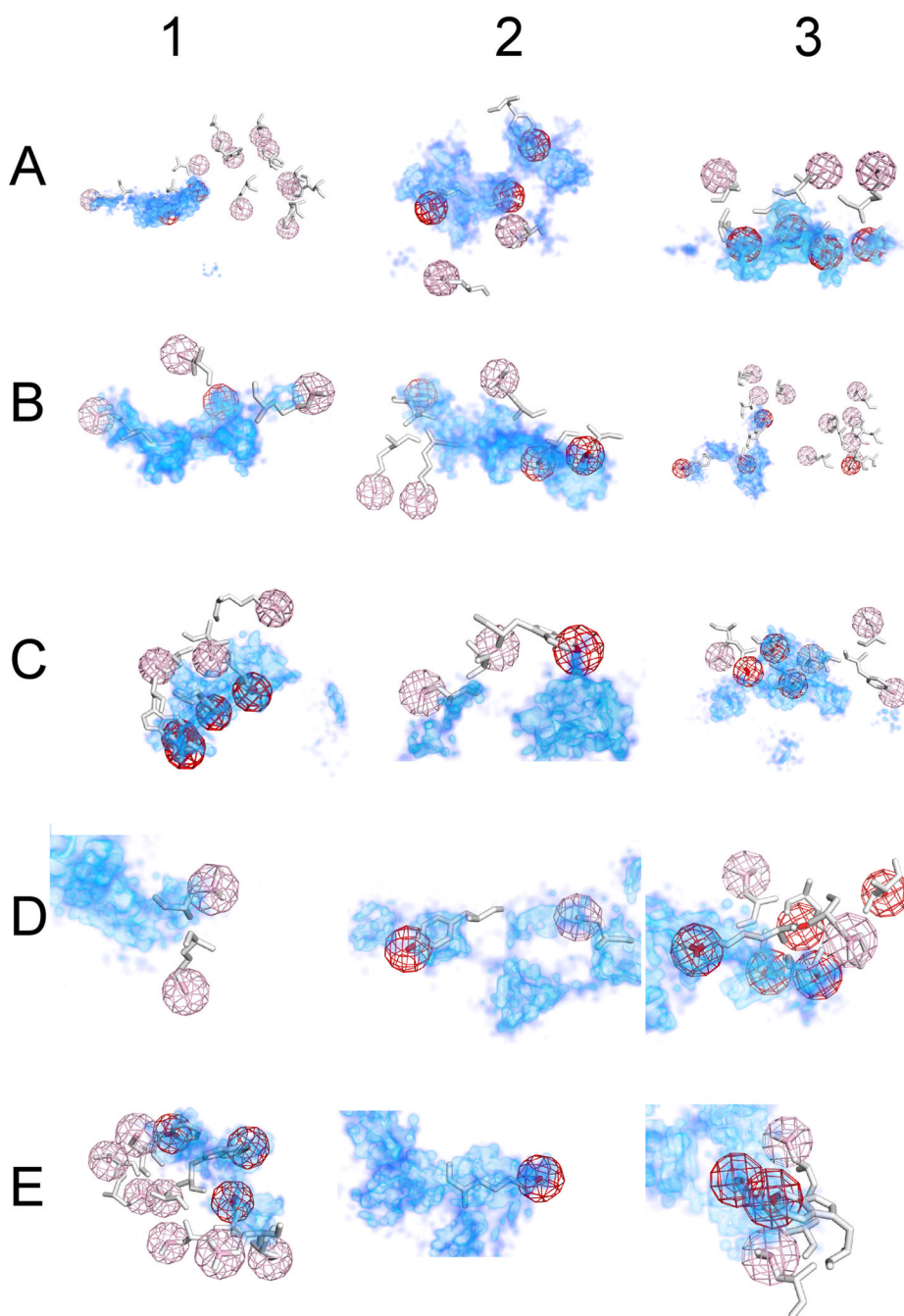


Figure 3.

FTMap probe atom density (cyan and blue cloud representation) for the 15 examined receptors showing that the side chains of residues identified as hot spots by alanine scanning mutagenesis largely overlap with the consensus sites identified by FTMap. The identity of the receptor can be found in Table 1. Alanine scanned residues of the ligand are displayed as white sticks, and the end group atom (defined in Table S1) of the residue is colored pink or red depending on whether the residue has been identified as a neutral or hot spot residue, respectively, by alanine scanning. A sphere around the end group atom having a radius of 2 Å is shown as a mesh with the same color as the end group atom, and all probe atoms within this sphere were counted to obtain the density correlation for the parent residue. The density

is shown contoured at two density levels: the red color is contoured about 3 sigma, and the cyan level is contoured at 1 sigma. The dark blue region indicates the region where any FTMap probe atoms were found.

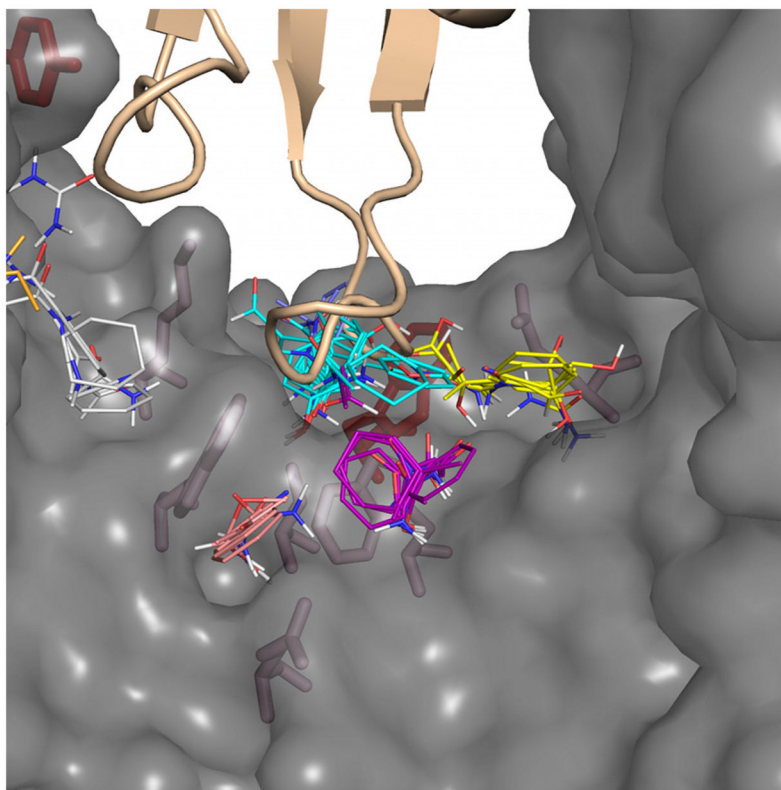


Figure 4. Trp259 of RNI lines the top ranked consensus cluster from the computational mapping results of RNI. RNI is shown as a translucent grey surface, and mapping results of RNI are shown as clusters colored by rank: top is cyan, second is purple, third is yellow, fourth is salmon, fifth is white, and sixth is orange. Alanine scanned residues from RNI are shown as sticks, and the hot spot residues are colored red while neutral residues are colored pink. RNI Trp259 is in the center colored red, and RNase A is shown as a wheat cartoon above it. Notice that a loop of RNase A passes through the top consensus cluster from the mapping results of RNI, and this site directly contacts Trp259.

Table 1

Identity and source of 15 alanine-scanned complexes used for analysis.

| PDB code ^a | Ligand Chain: Name ^b | Receptor Chain: Name ^b | Source ^c | Position in Fig. Figure 3 ^d |
|-----------------------|---------------------------------------|-----------------------------------|---------------------|--|
| 1a4y | A: Ribonuclease inhibitor | B: Angiogen | 10 | A1 |
| 1brs | D: Barstar | A: Barnase | 10 | A2 |
| 1bxl | B: Bak | A: Bcl-xL | 45 | A3 |
| 1cbw | D: BPTI | B,C: Chymotrypsin | 10 | B1 |
| 1cdl | E: Cdii fragment | A: CaM | 46 | B2 |
| 1dfj | I: Ribonuclease inhibitor | E: Ribonuclease A | 10 | B3 |
| 1dva | X: Hydrolase inhibitor | H: Hydrolase | 46 | C1 |
| 1ebp | C: Epo agonist mimetic peptide | A: Epo receptor | 46 | C2 |
| 1f47 | A: FtsZ fragment | B: ZipA | 47 | C3 |
| 1lqb | D: Hif-1 α peptide | B,C: ECVHL | 46 | D1 |
| 1nfi | F: Ikba | B: NF κ b p50 | 46 | D2 |
| 1osg | G: BR3 | A,B: Baff/Blys | 48 | D3 |
| 1ycr | B: p53 | A: MDM2 | 49 | E1 |
| 2ptc | I: Trypsin inhibitor | E: β -trypsin | 10 | E2 |
| 3brv | A,C: NF κ b- β inhibitor | B,D: NEMO | 50 | E3 |

^aPDB code is the four character accession code used within the protein data bank (www.rcsb.org).

^bThe ligand and receptor chains are the single character chain designation within the PDB file that refer to the given ligand/receptor.

^cThe column called source refers to the source of the alanine scanning energetics data used in our analysis.

^dThe final column refers to the placement of an image of the mapping results of the receptor with the superimposed alanine scanned residues from the ligand that is displayed in Figure 3.

Table 2

χ^2 table comparing alanine scanning hot spot/neutral residue categories to consensus-site/non-consensus-site categories determined by a density correlation (DC) threshold of 600

| | Neutral residues | Hot spot residues |
|----------|------------------|-------------------|
| DC < 600 | 49 | 3 |
| DC ≥ 600 | 4 | 34 |