Energetic analysis of fragment docking and application to structure-based pharmacophore hypothesis generation

Kathryn Loving · Noeris K. Salam ·
Woody Sherman

Abstract We have developed a method that uses energetic analysis of structure-based fragment docking to elucidate key features for molecular recognition. This hybrid ligand- and structure-based methodology uses an atomic breakdown of the energy terms from the Glide XP scoring function to locate key pharmacophoric features from the docked fragments. First, we show that Glide accurately docks fragments, producing a root mean squared deviation (RMSD) of <1.0 Å for the top scoring pose to the native crystal structure. We then describe fragment-specific docking settings developed to generate poses that explore every pocket of a binding site while maintaining the docking accuracy of the top scoring pose. Next, we describe how the energy terms from the Glide XP scoring function are mapped onto pharmacophore sites from the docked fragments in order to rank their importance for binding. Using this energetic analysis we show that the most energetically favorable pharmacophore sites are consistent with features from known tight binding compounds. Finally, we describe a method to use the energetically selected sites from fragment docking to develop a pharmacophore hypothesis that can be used in virtual database screening to retrieve diverse compounds. We find that this method produces viable hypotheses that are consistent with known active compounds. In addition to retrieving diverse compounds that are not biased by the co-crystallized ligand, the method is able to recover known active compounds from a database screen, with an average enrichment of 8.1 in the top 1% of the database.

Introduction

Fragment-based drug design has seen increasing popularity in recent years with successful work in this area being reported by Astex Therapeutics, Abbott, Plexxicon, and many others [1–7]. Typically, low-molecular-weight compounds are identified as binders by either X-ray diffraction or nuclear magnetic resonance spectroscopy. These small compounds are considered "efficient" binders because the average energetic contribution per atom to binding is typically larger than for drug-like compounds. Because of this, fragments have greater scope for optimization than do larger hits obtained by high-throughput screening [8]. Structure-based drug design, including computational design, can then be used to create drug-like molecules based on these initial fragment hits.

Fragments have also been gaining in popularity because the size of the chemical space for fragments is significantly smaller than that for drug-like compounds. The entire chemical space for molecules with molecular weight below 160 Da is estimated at 14 million compounds [9], which is a tractable number to study with current computational tools and not unreasonable even with experimental approaches. This is in contrast to the 10^{60} estimated potential chemical entities of drug-like size, which is well beyond the realms of any computational or experimental screening approach [10].

K. Loving (✉) · N. K. Salam · W. Sherman
Schrödinger, Inc., 120 West 45th Street, 29th Floor, New York, NY 10036, USA
e-mail: kathryn.loving@schrodinger.com

Fragment-based drug design is challenging from a computational perspective for a number of reasons. First, modifications made to fragments often change their binding mode, making the lead-optimization process very difficult. For example, in estrogen receptor the fragment molecule in the 1GWQ co-crystal structure flips by 180° relative to the 1ERR structure where that same fragment is a part of the approved drug raloxifene. This also makes retrospective validation of computational methods difficult, since knowing the location of a drug-like compound from a co-crystal structure is not sufficient to be confident in the location of the constituent fragment parts if they are treated as separate molecules. Furthermore, advancements in computational fragment-based work have been challenged by the scarcity of fragment co-crystal structures in the public domain that can be used for validation and parameter optimization. Another challenge in fragment-based discovery is accounting for induced-fit effects, which are often substantial and could prevent finding the right pose if a rigid receptor model is used [11]. Despite the known challenges that induced-fit effects present to traditional docking algorithms, progress has been made in recent years to address this problem using more sophisticated methods that include receptor flexibility [12–14]. Finally, it is often observed in computational docking methods that one pocket of the binding site attracts the majority of fragment poses (i.e., a "hot spot"), thereby leaving much of the full binding site unexplored.

Docking programs have proven relatively successful in accurately reproducing known poses of drug-like molecules from co-crystal structures, with Glide consistently performing among the top of the programs [15, 16]; however, little work has been reported on docking accuracy for fragment molecules. One exception is a study by Marcou and Rognan where they used Glide 3.5, FlexX 1.13, Surflex 1.2, and Gold 3.1 for fragment docking [17]. Rather than using the standard RMSD metric, which has know deficiencies [18], this study measured docking accuracy using an interaction fingerprint score called Tc-IFP (similar to the structural interaction fingerprint methodology developed by Deng et al. [19]). It was found that Glide produced more poses with good Tc-IFP scores (>0.6) than other programs. Even with metrics that account for more than just RMSD, such as Tc-IFP described above or the RSR metric that uses electron density similarity [20], there are still substantial challenges in defining what is correct and incorrect when working with fragments as compared to drug-like molecules. This difficulty is primarily because it is much easier for fragments to adopt multiple binding modes [11] and in many cases the lowest energy pose for the fragment alone may change when substituents are added (as in the estrogen receptor example described above).

In the work presented here, we use the Glide XP 5.0 [21] docking program with default settings and show that accurate poses can be obtained for native fragment docking. We then describe modified docking settings that increase spatial sampling in order to more thoroughly explore the binding site when multiple poses per fragment are saved while maintaining the docking accuracy of the best-scoring pose. No changes were made to the Glide scoring function in this work. We then use a focused library of 548 diverse fragments [22] that has been curated from the literature [1, 5, 23–27] as a set of docking probe molecules to span a broad range of interaction types and chemically reasonable geometries while maintaining a relatively small number of molecules, thereby keeping the calculations computationally efficient. Multiple poses are saved for each fragment, resulting in a more thorough sampling of the binding site with a wide array of energetically feasible poses.

While the fragments have many pharmacophore sites, in most cases a number of the sites for a given docked pose do not make specific interactions with the receptor and therefore should not be considered important for binding. To address this, we map the energy terms computed by the Glide XP scoring function onto individual ligand atoms to determine the most favorable pharmacophoric features for a given pose. Using these energy-optimized pharmacophore sites from each of 15 spatially diverse fragment clusters we develop a pharmacophore hypothesis that is then used to screen a database seeded with active molecules. We show that this approach is effective for finding diverse compounds and, in many cases, known actives are recovered at the top of the database screen. This approach combines strengths of both ligand- and structure-based methods; namely, energetic contributions to binding can be accurately computed from the structure-based docking, while the pharmacophore-based screening is fast and allows a great deal of control. Furthermore, there is no need for knowledge of known active compounds, potentially reducing the bias toward finding compounds that are already covered by the existing patent space.

In the following sections we first describe validation work showing that Glide XP is able to accurately dock fragments. We then show that an energetic ranking of fragment features can reproduce pharmacophore features from known active compounds. Next, we describe the creation of pharmacophore hypotheses derived from energetic analysis of docked fragments. Finally, we apply this method to eight different crystal structures spanning a range of pharmaceutically relevant targets. We show that reasonable pharmacophore hypotheses are generated that can be used to retrieve diverse and active compounds from a database screen.

Methods

Docking

Each protein was prepared with the Protein Preparation Wizard in Maestro [28] using default options: bond orders were assigned, hydrogens were added, metals were treated, and water molecules 5 Å beyond hetero groups were deleted. Hydrogens were then optimized using the *Exhaustive sampling* option and the protein was minimized to an RMSD limit from the starting structure of 0.3 Å using the Impref module of Impact [29] with the OPLS_2001 force field. Fragments were prepared using LigPrep [30] with Epik [31] to expand protonation and tautomeric states at 7.0 ± 2.0 pH units. The fragments were then docked flexibly using Glide XP 5.0 [21, 32] with either default options or modified settings. For the modified settings, we increased the number of poses per ligand for the initial docking stage to 50,000, used a wider scoring window of 500.0 kcal/mol for keeping initial poses, and kept the best 1,000 poses per ligand for energy minimization. The keyword *roughmin* was added to the *maxkeep* line of the Glide input file, instructing Glide to bypass the sorting by the rough score and to minimize all *maxkeep* (50,000) poses on the Glide grid. This allows for a much larger number of poses to be scored with the more accurate scoring function in Glide. We chose to "Write XP descriptor information" during the docking run (an option in the Glide interface), which writes a file containing atom-level energy terms such as hydrogen-bond interactions, electrostatic interaction, hydrophobic enclosure, and pi–pi stacking interactions. We also subjected the top 100 poses per fragment to post-docking minimization and requested the top 100 poses to be returned. These non-default docking options are applicable for both Glide SP and Glide XP, although for this work we used Glide XP in order to obtain the atom-specific Glide XP descriptor terms. We did not modify the Glide-Score in this work, although rescoring by a ligand efficiency metric [33] would be prudent when the docking library compounds vary in size.

Volume clustering

Volume clustering of fragment poses was performed with a Python script [34]. First, the volume overlap of each pair of fragment poses is computed to generate a matrix of overlapping volume values between all pairs of structures. Volumes are computed by treating each molecule as a set of atomic spheres with sizes determined by their van der Waals radius. This matrix is then hierarchically clustered using complete linkage into 15 clusters. We chose a reasonable clustering level as described in the Results and discussion section: we looked at the seven diverse targets and found that a clustering level of 15 provided a good balance between coverage of the native features and the number of fragments to analyze.

Energetic feature scoring

Starting with the docked pose of the 15 fragments obtained from the above volume clustering, Phase 3.0 [35] was used to generate pharmacophore features. These included the standard set of six chemical features from Phase: hydrogen bond acceptor, hydrogen bond donor, hydrophobic, negative ionizable, positive ionizable, and aromatic ring. Hydrogen bond acceptor sites were represented as vectors along the hydrogen bond axis in accordance with the hybridization of the acceptor atom. Hydrogen bond donors were represented as projected points, located 1.4 Å away from the hydrogen, along the X–H bond. The Phase feature definitions were edited so that positive and negative groups are also labeled as donors or acceptors, respectively, allowing each feature type to be considered separately.

An energetic value is attributed to each pharmacophore feature based on the sum of individual Glide XP terms arising from atoms that comprise the feature. These terms are taken directly from the XP descriptors (viewable with the "XP Visualizer" program in Maestro), and a script is available on the Script Center [33] that assigns these XP descriptor terms to individual atoms. In addition to the XP descriptors, we also use the ChemScore [36] lipophilic term to score rings that make good interactions but do not receive a hydrophobic packing reward from Glide XP. To generate a hypothesis based on this rank-ordering of features, the four best-scoring features were chosen, along with rules that prevent overlapping features. The minimum distance between any two features is chosen to be 2.0 Å and the minimum distance between two features of the same type is 4.0 Å. Based on work where crystal structures of active compounds were analyzed, we found that most reasonable pharmacophoric hypotheses have between three and seven features (data not shown).

We created excluded volumes as part of each hypothesis that allow small overlaps with the receptor in order to approximately account for induced-fit effects. Excluded volume spheres were generated for receptor atom positions within a 20.0 Å shell around the 15 spatially diverse fragments, where each sphere's radius is half that of the van der Waals radius of that receptor atom. Additionally, receptor atoms whose surfaces are 1.5 Å or closer to a docked fragment surface were ignored. This creates a loose excluded volume surface representing the receptor, forcing potential matches to stay mostly within the binding site.

For hydrogen bonds or single hydrophobically-packed hydrogen bonds, the energy term is associated with one ligand atom and therefore the entire score was assigned to

that atom. For correlated hydrophobically enclosed hydrogen bonds the energy was split between the associated atoms. The hydrophobic enclosure reward term is generally associated with several ligand atoms, usually in a ring. We wanted to weight this term as being very important, so the full score was associated with each of the ligand atoms. For the electrostatic reward term a score of −1.0 kcal/mol was given to each atom associated with that term, and for the pi–pi stacking term a score of −0.1 kcal/mol was given to each associated atom. We also calculated the hydrophobic atom-atom pair interaction ChemScore [36] term for all rings. We assign this score (in kcal/mol) to the atoms in that ring. For each ring, the more favorable of either the ChemScore term or the hydrophobic enclosure reward term is assigned

Target selection

A diverse set of eight well-characterized complexes from the PDB (Table 2), which have previously been employed in docking accuracy and scoring evaluations [21, 37, 38] were selected for this study. One target (p38 MAP kinase) is represented by two co-crystallized receptor sites, including both the DFG-in and the DFG-out conformations [39]. Seven targets were used for the fragment feature analysis, and cyclooxygenase-2 was added as an additional target for the hypothesis generation and Phase database screening.

Database screening

Pharmacophore database screens were performed with the program Phase 3.0 [34], described by Dixon et al. [40]. Each hypothesis was comprised of the 4 most energetically favorable features based on the Glide XP methodology described above. Three of four sites had to match, with the best scoring feature being required. Excluded volumes were included in each hypothesis as described in the "Energetic feature scoring" section of the "Methods". The decoy screening database contained a set of 1,000 drug-like decoys, which have an average molecular weight of 400 Da. These are described in the original Glide docking paper [37] and are available as Supplementary Material S2. For each target, a set of ∼20 known binders that have been used in previous studies were added to the screening database [21, 37, 38]. Each of the active compounds exhibits affinities better than 10 μM, with the exception of neuraminidase ligands, which contain some compounds that bind more weakly. The databases were generated with Phase 3.0 using default options, with a maximum of 100 conformations per molecule generated using ConfGen with the OPLS_2005 force field and a distant-dependent dielectric solvation treatment. The enrichment in the top

X% of the database, EF(X%), is defined as the fraction of the actives recovered divided by the fraction X of decoys. For the EF(1%) primarily used in this work, it is the fraction of actives found before 10 decoys are found and this number is divided by 0.01.
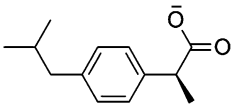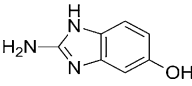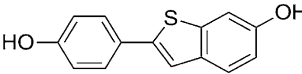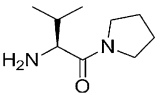
Results and discussion
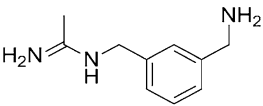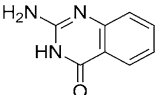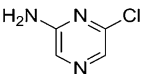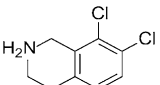
Fragment docking accuracy

A principle aim in structure-based fragment work is to find the energetically favorable binding mode(s), which can lead to valuable insights into the nature of the binding site and key interactions responsible for molecular recognition. To assess the ability of Glide XP to accurately dock fragment-sized molecules, we performed docking of fragments to their cognate crystal structures described in a recent paper by Congreve et al. [41] that contains a diverse set of targets. Twelve targets are represented in this set, shown in Table 1, with fragment molecular weights ranging from 95 to 242 Da and binding affinities from 0.003 to 350 mM. The diverse 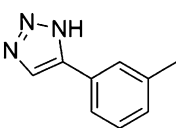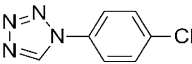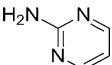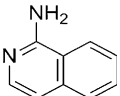binding sites contain water molecules, metals, and a heme. Water molecules greater than 5 Å from the fragment were removed prior to the docking calculations.

The RMSD of the best-scoring pose to the X-ray structure is reported in Table 1. The results in the second to last column are with the default settings in Glide XP whereas the results in the last column use the fragment-specific settings, described in the "Methods". In short, the fragment-specific settings were designed to increase the number of energetically-reasonable spatially diverse poses without degrading the docking accuracy results for the best scoring pose. We increased both the positional and orientational sampling throughout the hierarchical Glide docking procedure by passing more poses from each stage to the next; however, there were no changes to the Glide XP scoring function. It can be seen in Table 1 that the docking accuracy results are good (mostly below 1.0 Å RMSD) and that the fragment-specific settings have little effect on the quality of the results. However, as will be shown below, fragment-specific settings are useful for more thoroughly probing a binding site when multiple poses per fragment are saved. Also, we have found these fragment-specific settings to be useful in cases where certain subpockets of the binding site are not thoroughly sampled with the default Glide settings.

The successful docking results indicate that without any optimization for small fragment molecules the Glide docking algorithm and scoring function is able to consistently produce a top scoring pose that is in close agreement with the crystal structure pose. As shown in Fig. 1, the docked poses closely match the crystallographic ligand

Table 1 Native fragment docking results using Glide XP 5.0 on 12 targets

Fragment	PDB	Target	Chain	RMSD, default protocol	RMSD, fragment protocol
	1EQG	Cox-1	A	0.67	1.02
			B	1.09	1.31
	1FV9	Urokinase	A	0.54	0.45
	1GWQ	Estrogen receptor	A	0.36	0.45
			B	0.31	0.41
	1N1M	Dipeptidyl peptidase	A	0.60	0.39
			B	0.51	0.57
	1QWC	NO synthase oxygenase	A	0.47	0.61
	1S39	TGT	A	0.31	0.32
	1WCC	CDK2	A	0.37	0.37
	1YZ3	*N*-methyl-transferase	A	0.26	0.29
			B	0.43	0.39
	2ADU	Aminopeptidase	A	0.98	0.97
	2C90	Thrombin	A	0.50	0.57
	2JJC	HSP90	A	0.20	0.20
	2OHK	Beta-secretase	A	0.31	0.23

The RMSD results in the next to last column were generated using the default settings in Glide XP. The RMSD results in the last column were generated using the fragment-optimized settings for Glide, as described in the text

poses. This docking experiment represents a best-case scenario where the protein binding site has the optimal conformation for each fragment (native docking) and no induced-fit is required.

The single case that does not produce an RMSD of <1.0 Å is Cox-1 (1EQG). While most of the molecule is docked accurately, the isobutyl group does not make specific interactions with the receptor and in the docked structure is rotated compared to the native ligand, resulting in a slightly higher RMSD than the rest of the cases. The locations of the carboxylate and the benzene ring are in close agreement with the crystal structure. Flexible aliphatic atoms, such as the isobutyl group in 1EQG, are likely less important to position in exactly the same place as the crystal structure and therefore should not be weighted as heavily when looking at docking accuracy. While the standard RMSD metric used in docking studies treats all atoms equally, it would be better to use a method that weights the RMSD by the importance to get each atom in the right place. Developing such a method is beyond the scope of this work but has been discussed by others [20, 42].

Several cases highlight the importance of correctly predicting the ligand and protein protonation states. One example is TGT (1S39), where there is some ambiguity about whether Asp 102 should be protonated (as shown) or whether the ligand should be positively charged. The empirical pKa prediction program Epik [31] predicts the state penalty of the positively-charged ligand to be 6.2 kcal/mol at pH 7, which makes the protonated Asp (with a pKa of ∼4.0 and therefore an energy penalty of 4.2 kcal/mol) slightly more favorable at pH 7. The Protein Preparation Wizard in Maestro produces the structure shown in Fig. 1 with default settings. Considering that this structure was solved at pH 5.5 [43], the predicted structure with a protonated Asp is even more sensible.

Beta secretase (2OHK) is another case where accurate treatment of the ligand protonation state is important. The pKa of the isoquinoline nitrogen is computed to be 7.2 (favoring the positively-charged species by 0.3 kcal/mol at pH 7.0) by Epik, making it ambiguous from a ligand-only view whether it should be protonated or unprotonated in the bound state. The default pKa range of 7.0 ± 2.0 in the ligand preparation program LigPrep [30] produces both the unprotonated and protonated states; however, the Glide XP docked pose of the protonated ligand state is preferred by 2.4 kcal/mol. Given that both Epik and Glide XP prefer the positively charged state of the ligand, it is clear from the computational perspective that this is the species to choose. This state also produces the lowest RMSD.

N-methyl-transferase (1YZ3) is an additional example where there is ambiguity about whether the fragment should be neutral or positively charged, but in this case both Epik and Glide predictions are required to determine the appropriate state. The positively-charge fragment has an Epik state penalty of 3.3 kcal/mol at pH 7. This structure was solved at pH 5.5 [44], reducing the state penalty to approximately 1.2 kcal/mol. As in the previous example, we prepared the structure (as described in the "Methods") with both the protonated and deprotonated ligand forms, which alters the hydrogen-bond network near the binding site and the optimal orientation of water number 1,040. In both protein chains, the combination of the GlideScore plus the Epik penalty is more favorable for the positively-charged ligand (for the fragment-specific settings, −6.6 vs. −5.9 kcal/mol in chain A and −6.2 vs. −5.6 kcal/mol in chain B). We report the lowest-energy poses in Table 1 and Fig. 1, which also correspond to the poses with the lowest RMSD. This highlights the importance of combining multiple thermodynamic contributors to binding (in this case the docking score and the ligand state penalty) to get an accurate overall picture of the preferred species in the bound state.

Finally, aminopepsidase (2ADU) presents a difficult case because of the two cobalt ions in the binding site. The docking results are acceptable, with RMSDs slightly <1.0 Å, but the best-scoring ligand pose has an incorrect protonation state and the double chelating interaction with the cobalt is not made. Glide has not been parameterized for docking to cobalt-containing binding sites and no special reward was given for making an interaction with these metals. If the default reward of 2.5 kcal/mol for an optimal metal interaction had been added to the GlideScore, as one would expect in this case, the correct tautomer would have been chosen as the lowest energy structure by 1.6 kcal/mol when adding the Epik state penalty to the GlideScore.

Fragment pharmacophore feature identification

Next, in order to explore the capabilities of fragment docking to accurately place key features in a binding site, we used the fragment-specific settings for Glide XP to dock a set of 548 fragments [22] to seven targets. These fragments are available as Supplementary Material S1. The targets include estrogen receptor (1ERR), factor Xa (1FJS), HIV-1 protease (1HPX), MMP-3 (1G49), neuraminidase (1MWE), and two P38 structures (both DFG-in: 1A9U and DFG-out: 1KV2). A maximum of 100 poses for each fragment were requested with the objective that all sub-pockets of each binding site be occupied by at least one fragment pose. In Fig. 2 we show the 2,000 best-scoring fragment poses by GlideScore for two targets (neuraminidase and P38) to highlight that these fragment poses fill the binding site. In the work below we show that although most of the fragment features are not related to the features from

Fig. 1 Poses of compounds in Table 1 with the native ligand shown in gray carbons and the lowest-energy pose predicted by Glide XP 5.0 with fragment-specific settings shown in *green* carbons. Each structure is visually oriented similarly to the poses in Congreve et al. [41] with hydrogen bonds shown and metal interactions for the docked poses shown as *yellow dotted lines* and *purple dotted lines*, respectively. The blue spheres in the 2ADU image represent cobalt ions. All crystal waters within 5 Å of the fragment were included in the docking calculation, but only waters making interactions with the fragments are shown. In structures with two chains, only the chain A binding site is shown
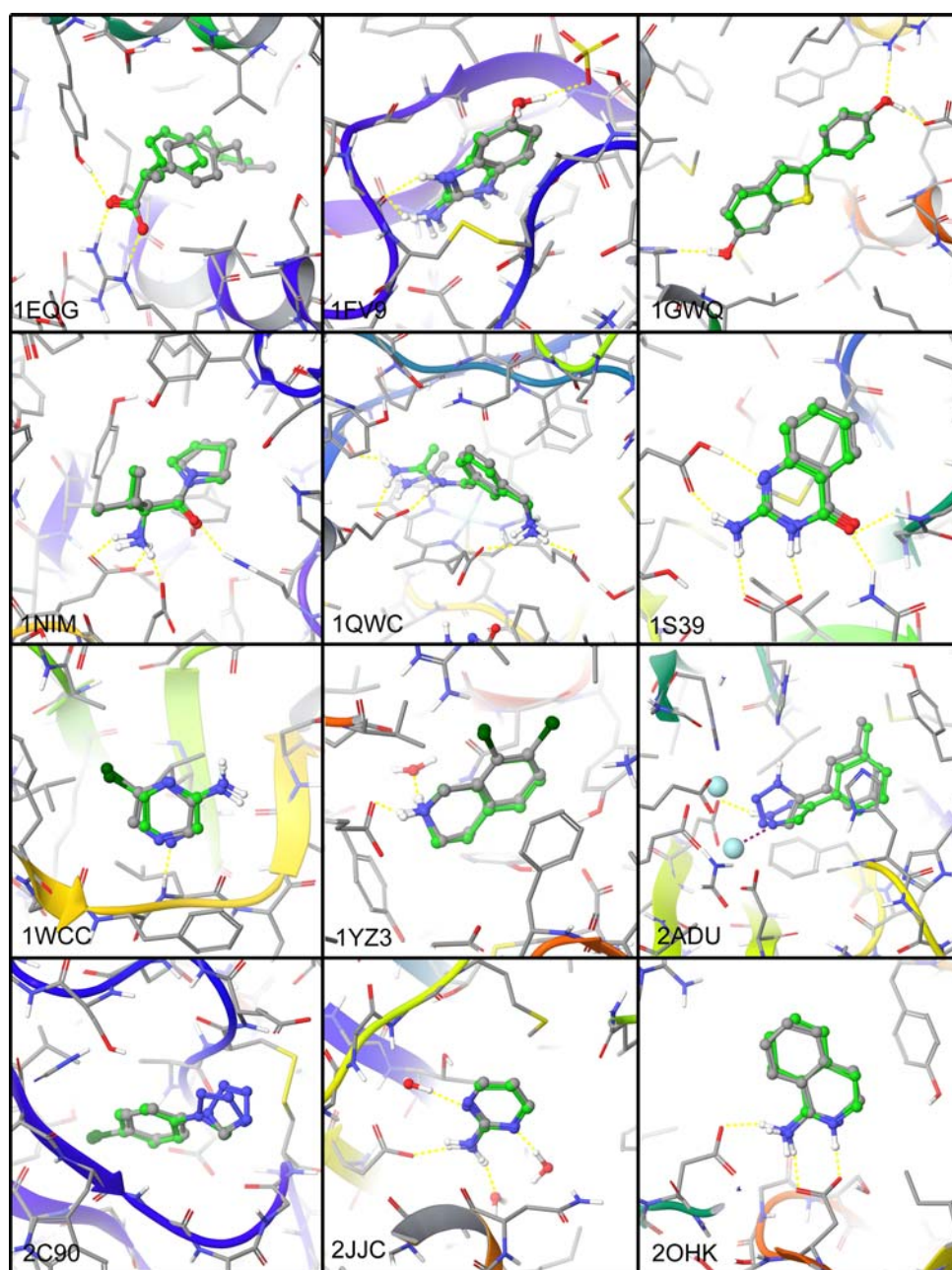
the crystallographic ligand, the fragment features that receive the best atom-level scores from Glide XP match the crystal features well.

In the first step we score each pharmacophore feature based on the individual energetic terms that comprise the GlideScore in Glide XP. Since each term is associated with one or more atoms, an energetic score in kcal/mol is created for each atom (see "Methods" for details). The individual energies for the atoms that comprise each pharmacophore feature are then summed to get a total score for each feature. In order to compare the results for pharmacophore feature placement from fragment docking with

active ligands, we performed this scoring on the co-crystallized molecules and for the best scoring fragment poses. Figure 3 shows the results of analyzing a variety of fragments sets extracted from the docking.

In order to match most of the energetically favorable native features, ∼50 of the top-scoring fragments were needed (see "50 best" in Fig. 3). Increasing this number beyond 50 does not substantially improve the results (data not shown). However, reducing the number does result in a considerable degradation in the matching of important native features (see "15 best" in Fig. 3). From a methodological perspective, 50 fragments are too many for the

pharmacophore hypothesis generation and we sought ways to reduce the number of fragments while maintaining the level of feature matching. Figure 3 "15 clustered" shows the results of generating 15 spatially diverse clusters based on volume overlap, taking the best-scoring fragment from each cluster. In this case, matching the native features is accomplished with many fewer fragments than when clustering is not used. Increasing the number of clusters does not substantially improve the matching (data not shown) whereas degradation is observed when reducing the number of clusters below 15 (Fig. 3 "10 clustered" and "5 clustered"). With the 15 top-scoring fragments from volume clustering we retrieve 100% of the most energetically important native feature, 91% of the top three, and 77% of the top six. Based on these results we will use this scoring and clustering protocol through the remainder of this work.

While the above analysis shows that fragments make many of the energetically important interactions observed in native compounds, they also make other interactions that the native compounds and known-active compounds do not make. For the seven targets listed above, the 15 spatially diverse fragments have an average of 3.0 features that are seen in the fragment but not seen in the native compound. A fragment feature falls into this category if it is of a different type from any native compound feature or if it is more than 2.5 Å from a native feature of the same type.

Fragment-based pharmacophore hypothesis determination

To develop a pharmacophore hypothesis from fragment docking we applied the feature scoring procedure described above to eight targets (cyclooxygenase-2 was added as an eighth target). A pharmacophore hypothesis derived from docked fragments could be used in a number of applications, including elucidation of key features for molecular recognition, fast database screening, and finding new compounds with a ligand-based screen that are not biased by the existing active ligands. For each target, the 548-fragment library was docked and the best GlideScore from each of the 15 spatially diverse clusters was retained, as described above. Phase pharmacophore hypotheses were created for each target using the four best-scoring pharmacophore features as well as receptor-based excluded volumes. In Fig. 4 we show the hypothesis for each of the eight targets along with all of the docked fragments that contribute at least one feature to the final hypothesis. The best-scoring feature in each hypothesis is highlighted with an arrow and the native ligand is shown for reference.

Phase screening was then performed on a set of ~20 active compounds and 1,000 drug-like decoys from the Glide enrichment validation set [37] for each target. Screening compounds had to match three of the four features in the hypothesis, requiring the best-scoring feature (highlighted with an arrow in Fig. 4) to be one of the three. The results are seen in Table 2, where we report the enrichment of active compounds in the top 1% of the database, called EF(1%), as well as the diversity of the top 1% of the hits. We use as our diversity metric the average Tanimoto similarity between the top 1% of the database screening hits and the native ligand. For comparison, we also did a fingerprint-based screen with the native ligand as the query. Canvas [45] linear fingerprints were used for both screens. While the fingerprint-based screen yields higher enrichments, the fragment-based pharmacophore hits have greater diversity while still retrieving actives in the top 1% of the database for most targets. A more detailed analysis of each case helps to illuminate the positive and negative aspects of the fragment-based pharmacophore approach.

In cyclooxygenase-2 (1CX2) the best-scoring pharmacophore feature is an aromatic ring feature in the upper

Fig. 2 Binding sites for two targets: **A** neuraminidase: 1MWE and **B** p38: 1A9U. The native ligand is shown in *green* carbons and the 2,000 best-scoring fragment poses are shown in wire with *gray* carbons
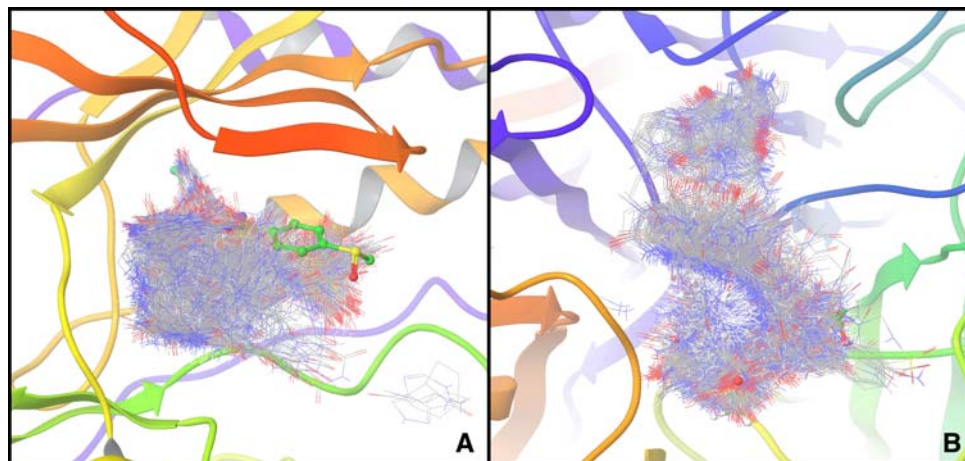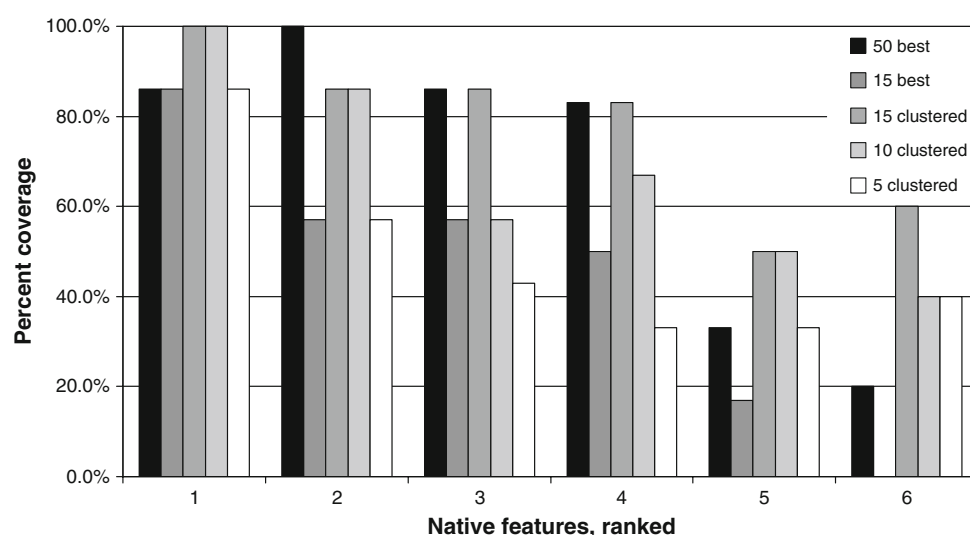
Fig. 3 Average coverage of native features by fragment features, for seven crystal structures: 1ERR, 1FJS, 1HPX, 1G49, 1MWE, 1A9U, and 1KV2. A fragment feature is said to match a native feature if the fragment feature is within 2.5 Å of a native feature of the same type. Native features were scored as described in the "Methods". For each of the six best-scoring features, we tested whether that native feature was covered by at least one fragment feature in the set of 50 best-scoring fragments by GlideScore (*black*), the set of 15 best-scoring fragments by GlideScore (*dark gray*), or sets of 15 (*medium gray*), 10 (*light gray*), or 5 (*white*) fragments chosen by volume-clustering of the 2,000 best-scoring fragments

portion of the active site channel where typical inhibitors place a ring to block the active site tyrosine (Tyr385). This feature is contributed by an indanone fragment. The second- and third-ranked features are a hydrophobic feature and another aromatic feature nearby. Although these are not exactly where typical Cox-2 inhibitors place these features, they are close enough to match some actives. The hydrophobic feature is contributed by a tetralone fragment and the second aromatic feature is contributed by a thiophenol. The fourth-ranked feature is an acceptor from a benzoxazole fragment near the top-ranked aromatic feature, making a hydrogen bond with the active site tyrosine. Several inhibitors place a halogen in this position. 15 of 33 active compounds match this hypothesis, with the top 2 being ranked ninth and eleventh by fitness, giving an EF(1%) of 6.1.

Several compounds known to be active against estrogen receptor (1ERR) have a series of carbon rings or aromatic groups, often with a hydroxyl group at both ends. The best-scoring feature in the fragment-based hypothesis is a hydrophobic group contributed by a spiro fragment. A nearby aromatic ring is the second ranked feature. A donor and acceptor representing one of the hydroxyl groups are the third and fourth ranked features, respectively. The donor is contributed by a purine-like fragment and the acceptor is contributed by a dihydroisoquinoline fragment. Seven of 20 active compounds match this hypothesis, but only one compound is ranked in the top 10 (ranked fifth) by fitness, giving an EF(1%) of 5.0.

The best-scoring feature for the factor Xa (1FJS) hypothesis is a donor in the S4 pocket, contributed by a

naphthyridinone fragment. A positive charge in the S1 pocket is contributed by an iminium group and scores as the second most favorable feature. This is a known feature in a large number of early factor Xa inhibitors. Finally, aromatic features from benzene rings are placed in each of the S1 and S4 binding pockets—these are the third and fourth ranked features, respectively. While not every active has the charge and donor features, most have at least one. Fifteen active molecules match the hypothesis, with four active compounds within the top 10 hits, resulting in an EF(1%) of 20.0. The initial hypothesis generated with the default protocol had an oxadiazole fragment in a small pocket where it made hydrogen bonds with the backbone of Trp225 and with Ser172. Although an acceptor feature on this fragment received a good Glide XP score because it is hydrophobically enclosed, this feature is more than 6 Å from any other fragment in the hypothesis, and for this fragment to match any active compound a linker would have to pass between antiparallel beta strands. We eliminated this feature from the hypothesis because its distance from the other features would preclude any known factor Xa compounds from matching. However, based on the favorable Glide XP energy it is possible that this represents a potentially desirable subpocket that has not yet been explored in the existing factor Xa literature and could be useful for future studies where it is desirable to find molecules that probe new interactions and potentially access new areas in the factor Xa IP space.

For the HIV-1 protease (1HPX) structure, fragment feature scoring selects two donors and two acceptors for

Fig. 4 Left column: Docked fragments superimposed on the native ligand for 1CX2, 1ERR, 1FJS, 1HPX, 1G49, 1MWE, 1A9U, and 1KV2. The top 2,000 fragments were clustered by volume-overlap into 15 clusters. The best scoring fragment was chosen from each cluster, and the fragments containing features that were chosen for the final hypothesis are shown in *shades of green*. The native ligand is shown in gray to provide a frame of reference. Right column: Fragment-derived pharmacophore hypotheses, superimposed on the native ligand for 1CX2, 1ERR, 1FJS, 1HPX, 1G49, 1MWE, 1A9U, and 1KV2. Features were chosen based on energetic analysis of the docked fragments shown in the left column. *Pink* features are acceptors, *light blue* features are donors, *red* features are negative, *blue* features are positive, *green* features are hydrophobic, and *orange rings* are aromatic ring features. The native ligand is shown in *gray* to provide a frame of reference. The best-scoring feature in each hypothesis is highlighted with an *arrow*
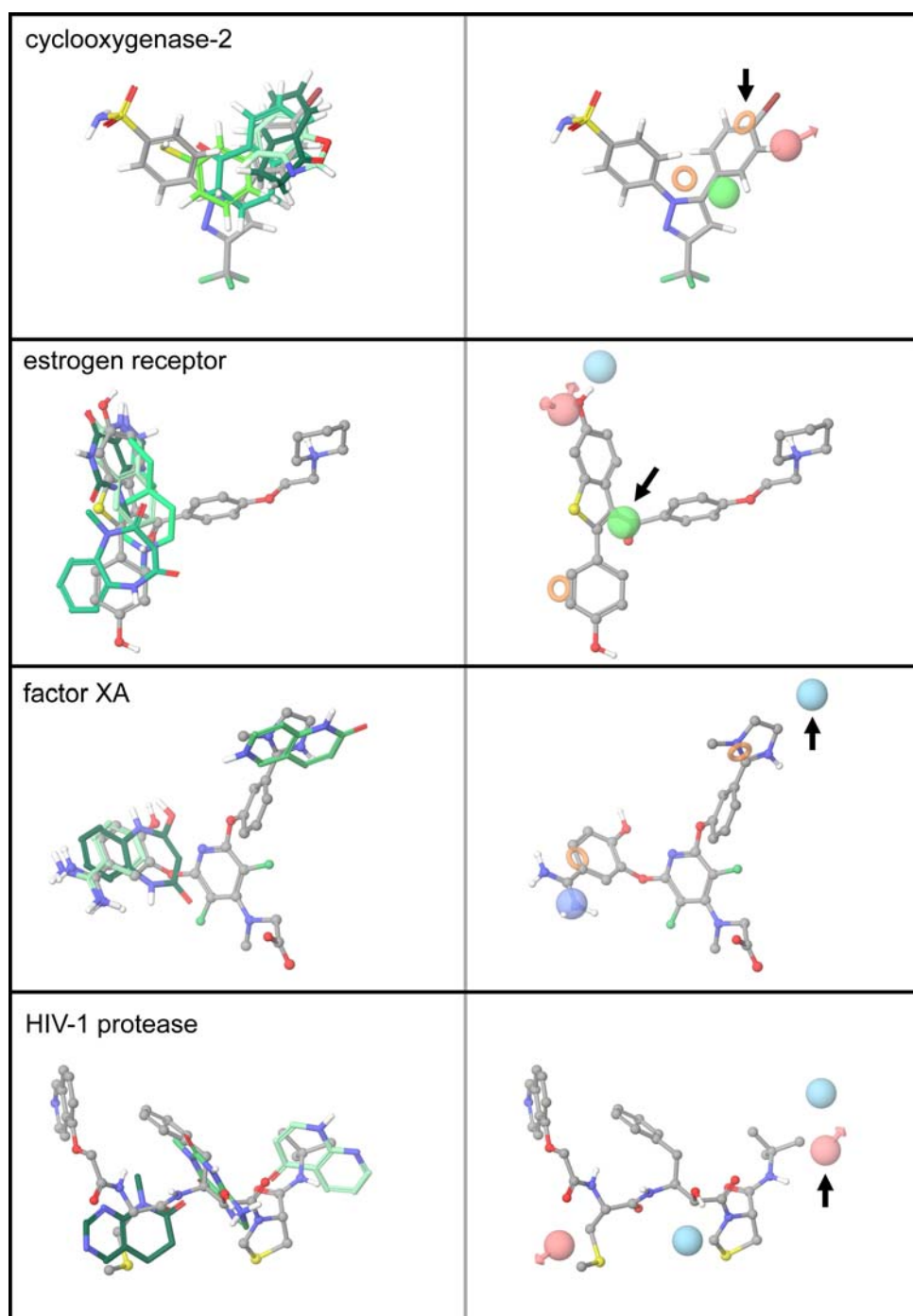
the hypothesis. While seven active compounds match this hypothesis, many decoys match this hypothesis as well, leading to no enrichment in the top 1% of the database. The best-scoring feature for HIV-1 protease is an acceptor from a naphthyridinone making a hydrogen bond to the back-bone nitrogen of Asp 30 in chain A and the second-best feature is a donor from a methylxanthine making a hydrogen bond with the backbone carbonyl of the same residue. The third-ranking feature is an acceptor from a pyrimidine ring interacting with the backbone of Asp30 of

chain B in the same way. The fourth is a donor that interacts with the catalytic Asp25 in chain B. This donor is contributed by the same naphthyridinone that contributes the top-ranking acceptor feature. While active compounds do match this hypothesis, the score is suboptimal because the features are poorly aligned. The lack of enrichment does not necessarily indicate that the features are irrelevant to active compounds; rather, it points to potential limitations in the general methodology where certain hypotheses may need further refinement. For example, HIV-1 protease

Fig. 4 continued
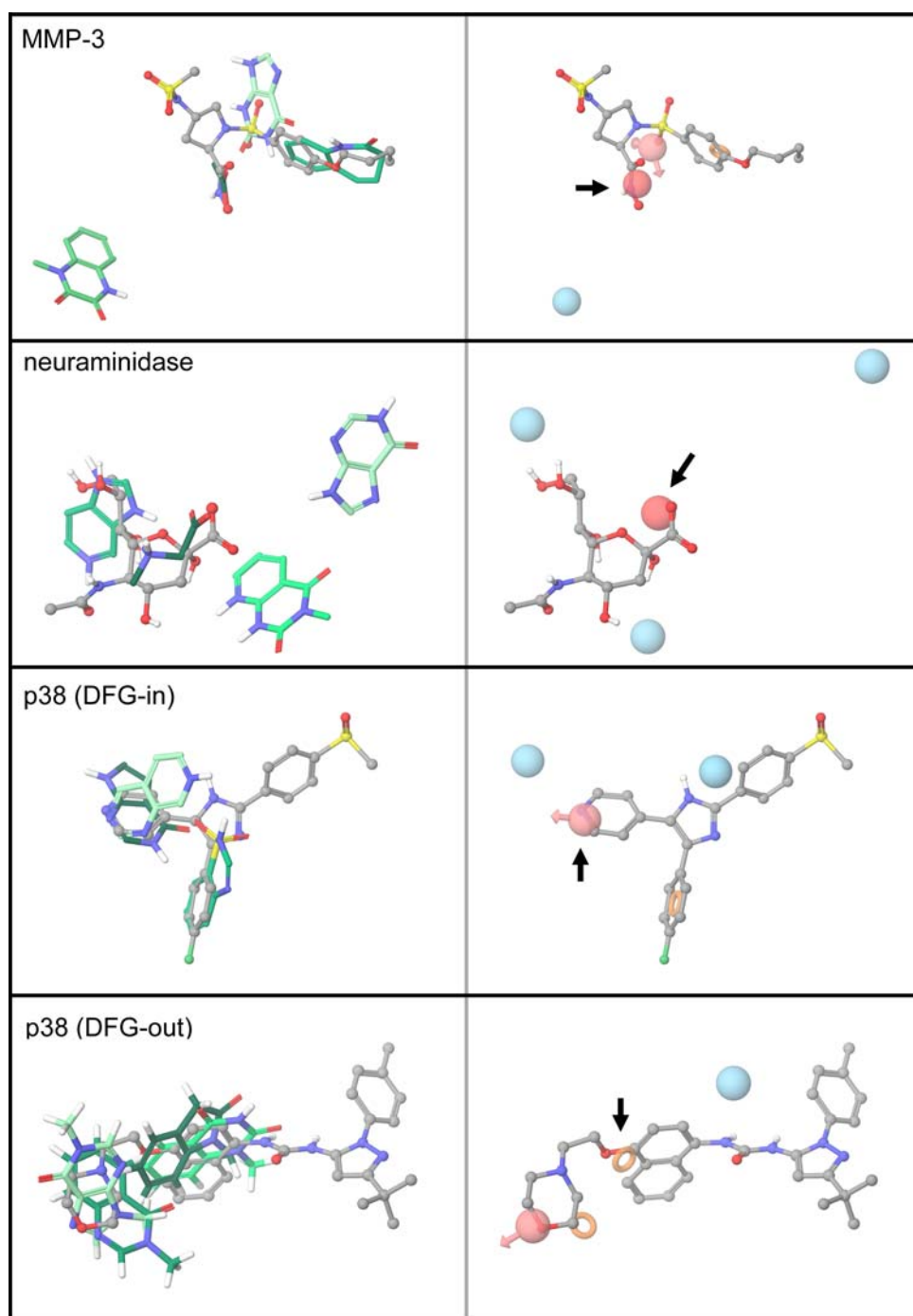
ligands are very large, with an average 14.8 rotatable bonds in our set of actives, which presents a significant challenge for a conformational sampling algorithm and it may be necessary to generate more than 100 conformations for the Phase database. Alternatively, since active compounds for HIV-1 protease have many functional groups, it may be necessary to add more features to distinguish actives from decoys. Indeed, when requiring all four features to be matched one active molecule is retrieved in the top 1% of

the database. Future work is likely needed to improve the methodology for retrieving very large active compounds.

In the MMP-3 structure (1G49), the methodology gives the best score to a negatively charged feature from an acetate fragment interacting with the zinc metal ion in the binding site. The second best-scoring feature is a donor feature from a pyrazine-like fragment that interacts with the backbone carbonyl of Pro656. We removed this donor feature from the hypothesis because it is too far away from

Table 2 Database enrichment results and diversity results for fragment-based pharmacophore hypotheses for eight targets

PDB	Target	EF(1%)	Similarity	fp EF(1%)	fp Similarity
1CX2	Cyclooxygenase-2	6.1	0.075	18.0	0.659
1ERR	Estrogen receptor	5.0	0.071	45.0	0.383
1FJS	Factor Xa	20.0	0.087	5.0	0.317
1HPX	HIV-1 protease	0.0	0.073	35.0	0.494
1G49	MMP-3	10.0	0.092	45.0	0.537
1MWE	Neuraminidase	14.0	0.139	66.0	0.521
1A9U	p38 MAP kinase, DFG-in	10.0	0.095	30.0	0.571
1KV2	p38 MAP kinase, DFG-out	0.0	0.059	30.0	0.185

Similarity is the average Tanimoto similarity between the top-scoring 1% of the hits from the database screen and the native ligand, with all compounds defined by Canvas [45] linear fingerprints. The third and fourth columns contain the results for the fragment-based pharmacophore screens. The fifth and sixth columns contain results for a fingerprint-based screen of the database, with the native ligand used as the query compound

the binding site region that is normally targeted for MMP-3—it is about 7 Å from the next closest fragment atom. However, as for factor Xa, it is possible that this represents a novel subpocket for gaining MMP-3 activity, which cannot be tested in a retrospective study as is being conducted in this work. The third-ranked feature is a donor from a quinoxalinedione fragment that interacts with the backbone carbonyl of Ala667. The final two features of this hypothesis are an aromatic ring that stacks against His701 and an acceptor feature from a xanthine fragment making a hydrogen bond with the backbone nitrogen of Ala665. Eight of 20 active compounds match our hypothesis with the best scoring active ranked sixth and an additional active in the top 10, resulting in an EF(1%) of 10.0. The negative feature, the acceptor, and the aromatic ring are the three features matched most often by active compounds.

In neuraminidase (1MWE), 14 of 29 active compounds match our hypothesis and 3 of them are within the top 10 hits in our database screen, resulting in an EF(1%) of 14.0. The best-scoring fragment feature is a negative charge that comes from the carboxylate of the n-methylglycine fragment that interacts with the side chain of Arg292. Every known active compound contains this negative feature. The second-ranked feature is a donor from a hypoxanthine fragment interacting with the backbone carbonyl of Ala369. The third and fourth ranked features are donors interacting with Glu276 and Asp151, respectively, and many of the active compounds make these interactions as well. An imidazopyridine fragment and a pyridine ring contribute these two donors.

In the p38 DFG-in binding site (1A9U), the best-scoring pharmacophore feature is an acceptor from a pyrrolopyrimidinone fragment making a hydrogen bond to the backbone hinge nitrogen of Met109; this is a common feature among all DFG-in p38 inhibitors. There is also an aromatic feature from a benzene ring in the pocket where the native ligand places a chlorobenzene. Our fragment-based

methodology scores this aromatic group as the third-ranked feature. Our hypothesis also contains a donor at the hinge interacting with the backbone carbonyl of Met109 and another donor interacting with the side chain of Asp168. The hinge donor is contributed by the same pyrrolopyrimidinone that contributes the top-ranking acceptor feature, and the other donor is from a pyridopyrimidine fragment. Most active compounds have at least one of these donor features. Fourteen of the 20 active compounds match this hypothesis and 2 hits are in the top 1% of the database (ranked fourth and fifth), giving an EF(1%) of 10.0. Several of the decoys map nicely onto the hypothesis, including the key acceptor and aromatic features, indicating that there may be other potential binders in our decoy set.

Finally, for the p38 DFG-out structure (1KV2), the best-scoring feature is an aromatic ring from a methylindolinone fragment in a hydrophobic pocket near the center of the native ligand. The second-ranked feature is an aromatic ring from a methylquinazolinone fragment placed at the hinge, near the morpholine group of the native ligand. The third-ranked feature is an acceptor from hypoxanthine that makes a hydrogen bond with the hinge backbone nitrogen of Gly110 and the fourth-ranked feature is a donor from quinazolinedione interacting with the backbone carbonyl of Asp168. This donor is near the amide/urea feature that many of the DFG-out active compounds share. The native ligand accepts a hydrogen bond from the backbone nitrogen of Asp168 whereas the quinazolinedione fragment donates to the carbonyl of the same residue. One of the 15 spatially diverse fragments does make a hydrogen bond to the backbone nitrogen of Asp168, as the native ligand does, but that feature does not get a very favorable energetic contribution (only −0.5 kcal/mol) and therefore is not included in our hypothesis. This is in contrast to the donor feature, which gets a reward for making a hydrophobically enclosed hydrogen bond, and so its score of −2.2 kcal/mol selects it for the hypothesis. While 15 of 20 active

compounds match the hypothesis, a large number of decoys also match equally well, resulting in no enrichment in the top 1% of the database. This indicates that the hypothesis is too general and needs more specificity to consistently score the active compounds over the decoys. One way to increase selectivity is to increase the number of sites in the hypothesis and to require more sites to match. We found that including another feature (for a total of five) and requiring compounds to match all five of them to be considered a match, one active compound was scored in the top 1% of the database, resulting in an EF(1%) of 5.0. The fifth feature is an acceptor in the amide-bond region where most active compounds place the amide or urea carbonyl.

Conclusion

We have presented results for a number of fragment-based applications. First, we showed that Glide XP is able to accurately dock fragment molecules. The accurate prediction of fragment poses is a key step in many structure-based fragment projects. For example, it can be used to gain insights into the binding mode when crystal or NMR structures are slow, expensive, or impractical to solve. It can also be used in computational methods, such as the fragment-based pharmacophore hypothesis generation work described here. However, there are many other applications that require the accurate prediction of one or more poses for a given fragment. For example, structure-based fragment optimization, fragment linking, and fragment joining all rely on accurate fragment poses.

We also described a way to enhance sampling in Glide XP to ensure a more thorough coverage of a binding site when multiple poses are retained. This is particularly important because saving only a single pose based on the best aggregate score for the entire molecule may cause individual features to make less than optimal interactions. Also, saving a large number of energetically viable poses can be useful in downstream applications, such as linking and joining of fragments, as it is often observed that fragments of drug-like molecules do not always go in the same place as when structures are solved with the fragment alone or as part of the full molecule.

Using a protocol for mapping atom-based energies from the Glide XP scoring function onto pharmacophore features, we were able to show that features from docked fragments are consistent with features from co-crystal ligands. The mapping of the Glide XP terms onto atoms could be useful for a number of other applications. Simply visualizing the energetic contributions of each atom to binding could aid in lead optimization. Using the atom-based energies could also help design more efficient molecules by focusing on improving or eliminating the parts of

the molecule that are not favorable for binding. While not explored in this work, it is possible to score fragments from crystal or NMR structures and assess the energetic contributions from each chemical feature. This could help the design of more efficient molecules, since features that are not energetically contributing to binding could be removed or changed.

Finally, we described how a pharmacophore hypothesis can be derived from an energetic analysis of fragment docking. The hypotheses for eight targets were screened against a database of drug-like molecules and active molecules were retrieved early in the database. Overall, the results for screening a database with the fragment-based pharmacophore hypotheses were good, with an average enrichment of 8.1 in the top 1% of the database and a diverse set of compounds was retrieved. While this is not an exceptionally high enrichment, especially compared to simpler methods like 2D fingerprint screening, we feel there are benefits to the methodology described here. Most importantly, the method is not biased by existing ligands, yet is able to take advantage of the speed of a ligand-based screening method. Also, compared to other ligand-based methods the results should be more diverse, which was shown when compared to 2D fingerprints.

Glide XP was shown to produce both accurate poses for fragment docking and produce a rank-ordered list of fragment features that is consistent with the features of known binding molecules. However, there may be improvements that could be made to the Glide XP scoring function specifically geared toward fragments. While there is an entropic term in the scoring function related to the number of rotatable bonds, it is possible that a more sophisticated entropy calculation is required for fragments. Fragments either have none or very few rotatable bonds, so the term computed in Glide XP will typically be small. However, because fragments bind more weakly than typical drug-like molecules and can more easily adopt multiple binding modes, it may help to include entropy terms related to vibrational modes, symmetry, and degenerate states. Another important consideration related to the scoring function is the treatment of ligand and protein desolvation. Glide XP does have terms to account for buried polar and charged groups on both the ligand and protein and it is likely that these terms will be valid for fragments because they are based on physical principles; however, this has not been explicitly tested and more fragment data would be needed to do so.

The retrospective virtual screening results presented in this work are highly dependent on the pre-existing actives in the database. This bias can, in fact, improve the results of some method, like 2D fingerprint screening where finding similar compounds is what the method does best; however, this can lead to poor results for a method that is aimed

toward producing diverse and novel hits. This is a perpetual problem in retrospective screening that is not possible to overcome without testing some of the novel compounds that are predicted to be active by the method but are not yet known to be active. Our method is particularly susceptible to this because small fragments are able to more thoroughly probe a binding site than drug-like molecules and therefore it is possible to place features in subpockets that have not been explored by existing compounds.

One potential limitation to the fragment-based pharmacophores described in this work is that the docked fragments alone contain no information about the chemical linkers that might be used to join fragment features together. This could potentially lead to geometrically inconsistent combinations of feature positions that cannot be matched by drug-like compounds. Another limitation is that there is no measure of potential drug-likeness accounted for in the combined fragments used to build the hypothesis. While filters for such properties could be applied either as a prefilter or postfilter of the database, it is not entirely clear how to build drug-likeness into the hypotheses.

References

1. Erlanson DA, McDowell RS, O'Brien T (2004) J Med Chem 47:3463. doi:10.1021/jm040031v
2. Boehm H-J, Boehringer M, Bur D, Gmuender H, Huber W, Klaus W, Kostrewa D, Kuehne H, Luebbers T, Meunier-Keller N, Mueller F (2000) J Med Chem 43:2664. doi:10.1021/jm000017s
3. Gill A (2004) Mini Rev Med Chem 4:301. doi:10.2174/1389557043487385
4. Card GL, Blasdel L, England BP, Zhang C, Suzuki Y, Gillette S, Fong D, Ibrahim PN, Artis DR, Bollag G, Milburn MV, Kim S-H, Schlessinger J, Zhang KYJ (2005) Nat Biotechnol 23:201. doi:10.1038/nbt1059
5. Hartshorn MJ, Murray CW, Cleasby A, Frederickson M, Tickle IJ, Jhoti H (2005) J Med Chem 48:403. doi:10.1021/jm0495778
6. Nienaber VL, Richardson PL, Klighofer V, Bouska JJ, Giranda VL, Greer J (2000) Nat Biotechnol 18:1105. doi:10.1038/80319
7. Tondi D, Morandi F, Bonnet R, Costi MP, Shoichet BK (2005) J Am Chem Soc 127:4632. doi:10.1021/ja042984o
8. Hajduk PJ, Greer J (2007) Nat Rev Drug Discov 6:211. doi:10.1038/nrd2220
9. Tobias Fink HB, Reymond J (2005) Angew Chem 117:1528. doi:10.1002/ange.200462457
10. Martin YC (1981) J Med Chem 24:229. doi:10.1021/jm00135a001
11. Babaoglu K, Shoichet BK (2006) Nat Chem Biol 2:720. doi:10.1038/nchembio831
12. Sherman W, Day T, Jacobson MP, Friesner RA, Farid R (2006) J Med Chem 49:534. doi:10.1021/jm050540c
13. Moitessier N, Therrien E, Hanessian S (2006) J Med Chem 49:5885
14. Nabuurs SB, Wagener M, de Vlieg J (2007) J Med Chem 50:6507
15. Zhou Z, Felts AK, Friesner RA, Levy RM (2007) J Chem Inf Model 47:1599
16. Perola E, Walters WP, Charifson PS (2004) Proteins 56:235
17. Marcou G, Rognan D (2007) J Chem Inf Model 47:195
18. Cole JC, Murray CW, Nissink JW, Taylor RD, Taylor R (2005) Proteins 60:325
19. Deng Z, Chuaqui C, Singh J (2004) J Med Chem 47:337
20. Yusuf D, Davis AM, Kleywegt GJ, Schmitt S (2008) J Chem Inf Model 48:1411
21. Friesner RA, Murphy RB, Repasky MP, Frye LL, Greenwood JR, Halgren TA, Sanschagrin PC, Mainz DT (2006) J Med Chem 49:6177
22. Schrödinger Fragment Library (2008) Schrödinger, Inc. http://www.schrodinger.com/ProductInfo.php?mID=6&sID=6&cID=53
23. Bemis GWM, Murcko MA (1996) J Med Chem 39:7
24. Fejzo J, Lepre CA, Peng JW, Bemis GW, Ajay, Murcko MA, Moore JM (1999) Chem Biol 6:755
25. Huth JR, Sun C (2002) Comb Chem High Throughput Screen 5:631
26. Hajduk PJ, Bures M, Praestgaard J, Fesik SW (2000) J Med Chem 43:3443
27. Jacoby E, Davies J, Blommers MJJ (2003) Curr Top Med Chem 3:11
28. Maestro v8.5, Schrödinger, Inc.: Portland, OR
29. Impact v5.0, Schrödinger, Inc.: Portland, OR
30. LigPrep v2.2, Schrödinger, Inc.: Portland, OR
31. Epik v1.6, Schrödinger, Inc.: Portland, OR
32. Glide v5.0, Schrödinger, Inc.: Portland, OR
33. Kuntz ID, Chen K, Sharp KA, Kollman PA (1999) Proc Natl Acad Sci USA 96:9997
34. Volume_cluster.py from the Schrödinger ScriptCenter (2008) Schrödinger, Inc. http://www.schrodinger.com/ScriptCenter.php
35. Phase v3.0, Schrödinger, Inc.: Portland, OR
36. Eldridge MD, Murray CW, Auton TR, Paolini GV, Mee RP (1997) J Comput-Aided Mol Des 11:425
37. Halgren TA, Murphy RB, Friesner RA, Beard HS, Frye LL, Pollard WT, Banks JL (2004) J Med Chem 47:1750
38. Friesner RA, Banks JL, Murphy RB, Halgren TA, Klicic JJ, Mainz DT, Repasky MP, Knoll EH, Shelley M, Perry JK, Shaw DE, Francis P, Shenkin PS (2004) J Med Chem 47:1739
39. Wang Z, Canagarajah BJ, Boehm JC, Kassisa S, Cobb MH, Young PR, Abdel-Meguid S, Adams JL, Goldsmith EJ (1998) Structure 39:12
40. Dixon S, Smondyrev A, Knoll E, Rao S, Shaw D, Friesner R (2006) J Comput-Aided Mol Des 20:647
41. Congreve M, Chessari G, Tisi D, Woodhead AJ (2008) J Med Chem 51:3661
42. Damm KL, Carlson HA (2006) Biophys J 90:4558
43. Meyer EA, Furler M, Diederich F, Brenk R, Klebe G (2004) Helv Chim Acta 87:1333
44. Wu Q, Gee CL, Lin F, Tyndall JD, Martin JL, Grunewald GL, McLeish MJ (2005) J Med Chem 48:7243
45. Canvas v1.1, Schrödinger, Inc.: Portland, OR