Transient Pockets on Protein Surfaces Involved in Protein−Protein Interaction

Susanne Eyrisch and Volkhard Helms*

Center for Bioinformatics, Building C7 1, P.O. Box 151150, D-66041 Saarbruecken, Germany

Received January 24, 2007

A new pocket detection protocol successfully identified transient pockets on the protein surfaces of BCL-X_L, IL-2, and MDM2. Because the native inhibitor binding pocket was absent or only partly detectable in the unbound proteins, these crystal structures were used as starting points for 10 ns long molecular dynamics simulations. Trajectory snapshots were scanned for cavities on the protein surface using the program PASS. The detected cavities were clustered to determine several distinct transient pockets. They all opened within 2.5 ps, and most of them appeared multiple times. All three systems gave similar results overall. At the native binding site, pockets of similar size compared with a known inhibitor bound could be observed for all three systems. AutoDock could successfully place inhibitor molecules into these transient pockets with less than 2 Å rms deviation from their crystal structures, suggesting this protocol as a viable tool to identify transient ligand binding pockets on protein surfaces.

Introduction

Protein−protein interactions are of central importance for cellular processes such as signal transduction, immune response, bioenergetics, structural organization, and apoptosis. Hence, it has become of great current interest to study the potential of this class of interactions as general drug targets.[1] So far, a fair number of systems are known for which peptides as well as small molecules have been identified that either inhibit or enhance protein−protein interactions. Prominent examples are MDM2−p53, BCL-X_L−Bak, IL-2−IL-2Rα, LFA-1−ICAM-1, Ras−Raf, and TNF−TNFRc1 among others.[2,3] Of particular interest is the binding of p53 to MDM2 promoting its degradation and thus preventing the transcription of genes that control, for example, apoptosis and DNA repair.[4,5] P53 is the most frequently inactivated protein in cancer and its interaction with MDM2 has become an important drug target in anticancer therapy.[6,7] Several classes of molecules have been characterized that inhibit this interaction.[8−10]

In this study, we have investigated the structural basis of ligand design at protein−protein interfaces. Traditional structure-based drug design relies on the availability of an experimental high-resolution protein crystal structure. On the basis of the crystal structure, ligands may be docked by virtual screening packages into cavities of suitable size on the protein surface in order to identify potential drug candidates for in vitro screening. The crystal structures of protein−protein complexes, however, often lack deep clefts or clearly shaped binding pockets at the interface regions.[11] Therefore, the structure-based design of inhibitors of protein−protein interactions is generally considered to be quite difficult. Moreover, crystal structures represent time-averaged coordinates reflecting typically only one out of many possible conformations. In this respect it is now getting more and more appreciated that the conformational dynamics of protein molecules need to be properly considered during the ligand design process.

For virtual screening of ligand libraries, docking methods allowing for (partial) receptor flexibility have proven to be quite promising even if rigid docking to the free protein structure failed.[12−15] For example, Apostolakis et al. combined minimization with shifted nonbonded interactions and Monte Carlo minimization to model induced fit effects in the binding site and the ligand.[12] Zacharias used soft modes extracted as principal components from MD[a] simulations to allow for deformations of the binding pocket.[13] Meiler and Baker introduced a Monte Carlo minimization procedure in which the position of the ligand and the protein side chain conformations are optimized simultaneously.[14] Although most flexible docking methods are today computationally quite feasible, they depend on a definition of the known (or a potential) binding region of the protein. In the case of an unknown system, the usual strategy is either to sample the entire protein surface (what would be computationally more costly) or to consider only cavities identified in the protein's crystal structure. For this purpose several computer programs exist that can be applied to identify cavities on protein surfaces, such as PASS,[16] Pocket,[17] LigSite,[18] Surfnet,[19] and QSiteFinder.[20] Cavities that are only accessible in protein conformations different from the crystal conformation and would serve as more favorable binding sites will be missed in these approaches. So far, none of these cavity detection methods have been applied to design inhibitors for protein−protein interactions. In all small-molecule protein−protein interaction inhibition studies to date that employed structure-based design of ligands, the binding pocket was known.

Another way of accounting for appropriate conformational dynamics of protein molecules in structure-based drug design is to use a large number of snapshots taken from molecular dynamics (MD) simulations as basis for ligand docking.[21−23] For example, Frembgen−Kesner and Elcock successfully identified in MD simulations an alternative binding site of the p38 MAP kinase.[24] Because there is strong evidence for the high mobility of protein surfaces,[25] we assume here that transient pockets that are large and deep enough to bind small-molecule inhibitors may open from time to time.

The aim of this study is to provide a starting point for in silico drug design for cases in which standard screening methods would fail, for example, when no potential binding pocket could be identified. We therefore selected three model systems where

* To whom correspondence should be addressed. Phone: ++49 +681 302 64165. Fax: ++49 +681 302 64180. E-mail: volkhard.helms@bioinformatik.uni-saarland.de.

[a] Abbreviations: PDB, Protein Data Bank; MD, molecular dynamics; rmsd, root-mean-squared deviation; ASP, active site point; PID, pocket identifier; BP, binding pocket.

10.1021/jm070095g CCC: $37.00
Published on Web 06/30/2007

the crystal structure of the protein−protein complex, the crystal structure of the free form of the protein, and a structure with a small-molecule inhibitor bound are known. In one of the three cases investigated, application of the PASS program to the crystal structures of the apo protein did not reveal the native binding pocket, whereas it was clearly identified from the inhibitor−complex structures when the inhibitor was manually removed (see Results and Discussion). So far, it is not known whether these pockets only open in the presence of a nearby ligand or whether they also exist in the apo form of the protein. Even in the latter case, these openings could be rare events that do not occur on the typical nanosecond time scales of molecular dynamics simulations performed at room temperature. In this study, MD simulations of the solvated apo proteins were conducted over times of 10 ns, and cavities on the protein surfaces were identified with the PASS algorithm for 4000 equally spaced snapshots in each case. To our surprise, the analysis revealed that even at room temperature large pockets opened frequently on the protein surface. The lifetimes ranged from a few picoseconds to hundreds of picoseconds. The general impressions seen for all three systems are quite similar (see Results and Discussion). These identified transient pockets represent potential binding sites of new inhibitors. When attention is focused on the location of the native binding pocket, pockets of similar size compared with the known inhibitor bound could be observed in all three cases during 10 ns MD simulations.

Results and Discussion

Surface Pockets of the Crystal Structures. Applying the PASS algorithm to the crystal structures of the unliganded proteins revealed that the native binding pocket is partly present in apo BCL-X_L (36%) as well as in apo MDM2 (42%). In the structure of free IL-2 the binding pocket could not be detected at all. To show that PASS is able to identify the native binding pockets, it was then applied to all of the available small-molecule inhibitor-bound complexes, where the inhibitors were removed manually. These complexes were 1YSG, 1YSI, 1YSN, 1YSW for BCL-X_L, 1M48, 1M49, 1PW6, 1PY2, 1QVN for IL-2, and 1RV1, 1T4E, 1TTV for MDM2. For all tested structures except for chain B of 1M49 (chains A, B), chains B and C of 1PY2 (chains A−D), and chains A and B of 1QVN (chains A−D), the native binding pocket was identified by PASS. For 1PY2 (chains A and D) only one of the two subsites of the binding pocket was detected.

Surface Pockets for the MD Snapshots. Applying the pocket detection method to the MD snapshots revealed surprising results. For IL-2, 23 distinct transient pockets were detected in the first run and 31 in the second run. For BCL-X_L 23 (respectively, 20) were detected, and for MDM2 33 (respectively, 36) were detected. In comparison, the total numbers of pockets detected for the crystal structures of the apo forms were 2 for IL-2, 4 for BCL-X_L, and 5 for MDM2.

When the frequency of occurrences and the average pocket volumes were analyzed, quite similar results were found for all three systems (see Tables 1 and 2). Amounts of 35.0−52.2% of the transient pockets were rare events that were present in less than 1% of all MD snapshots. Their mean volume ranged between 335.5 and 365.3 Å^3, and thus, in general, they represent the smallest cavities for each system. In contrast, there are fewer frequently appearing pockets (present in more than 50% of all MD snapshots), but they tended to belong to the largest ones of the respective system.

All pockets were found to be open up to 440.0 ps, vanished, and reappeared again several times. Figure 1 illustrates the fast opening and closing behavior characteristic for the transient pockets shown for the example of pocket identifier (PID) 5 of MDM2 taken from run 1. Instead of opening slowly, the pockets suddenly opened to volumes up to 500 Å^3 within 2.5 ps.

Table 1. Average Volumes of the Pockets According to Their Frequency for the Two Independent Runs

	average volume (Å^3)							
	frequency, <1%		frequency, 1−10%		frequency, 10−50%		frequency, >50%	
system	run 1	run 2	run 1	run 2	run 1	run 2	run 1	run 2
BCL-X_L	361.4	340.2	405.1	384.4	451.5	469.9	527.7	423.8
IL-2	346.2	365.3	338.2	399.7	355.1	401.0	452.7	398.9
MDM2	335.5	354.3	400.7	365.3	422.3	405.9	468.7	639.1

Table 2. Relative Number of Pockets with These Frequencies per System for the Two Independent Runs

	relative number (%)							
	frequency, <1%		frequency, 1−10%		frequency, 10−50%		frequency, >50%	
system	run 1	run 2	run 1	run 2	run 1	run 2	run 1	run 2
BCL-X_L	52.2	35.0	13.0	25.0	21.8	30.0	13.0	10.0
IL-2	47.8	51.6	26.1	19.4	17.4	25.8	8.7	3.2
MDM2	45.5	47.2	24.2	19.4	21.2	27.8	9.1	5.6

The pocket polarity ratio (ratio of the sum of N, O, and S atoms to the sum of N, O, S, and C atoms) of all pockets ranged between 0.25 and 0.45. For comparison, the polarity ratio of the entire protein surface of the crystal structure of the apo form is 0.37 for BCL-X_L and MDM2 and is 0.38 for IL-2. Figure 2 shows the pocket volumes and the corresponding polarity ratios for selected PIDs. Since pockets of the same PID possess states of the same volume but distinct polarity ratio, these ratios were averaged to smooth the curves. This analysis reveals that in general the largest pockets (volumes of >800 Å^3) have a smaller polarity than the overall protein surface. Thus, the protein interior seems to partly open up, and these pockets may be "sticky" enough to bind ligands.

Figure 3 shows the pairwise similarities of the PIDs obtained in the two MD runs performed for each system. Considering PID_i as reproducible if the other run contains at least one PID_j with $sim(PID_i,PID_j) > 50\%$, then more than 77% of all PIDs are reproducible. Especially frequent pockets tended to occur in both simulations and were therefore reproduced in most cases.

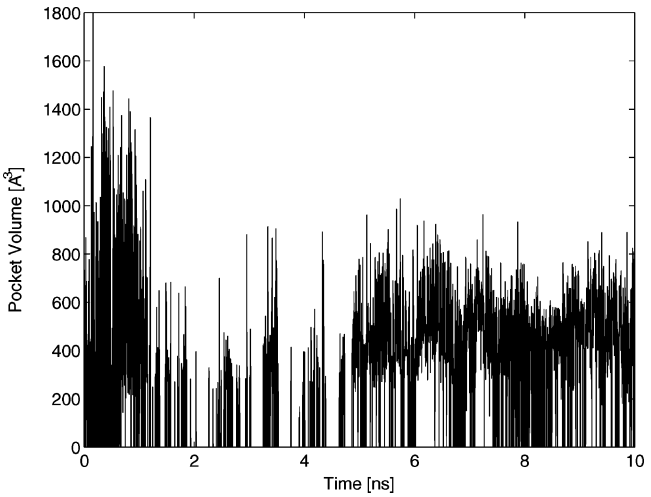

Figure 1. Pocket dynamics shown for the example of PID 5 (taken from run 1) of MDM2. The pocket volume is plotted against the simulation time.

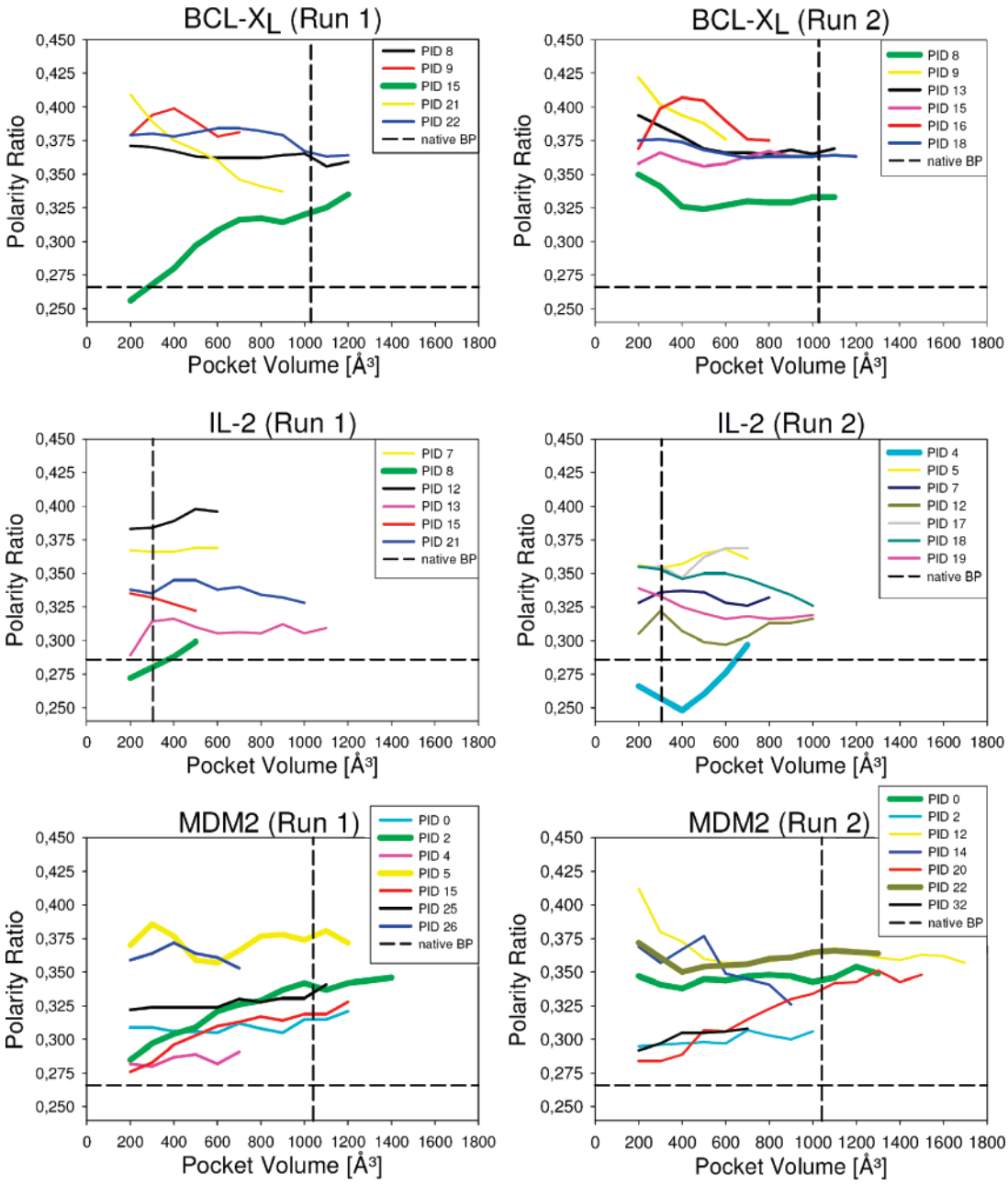

Figure 2. Changes in the average polarity ratio depending on the pocket volume. For each PID the polarity ratios of states having the same volume were averaged to smooth the curve. In order to obtain reliable values, only PIDs with frequency greater than 20% were used, resulting in a different number of PIDs for the different runs of the same system. The dashed reference lines indicate the polarity ratio and the volume of the native binding pocket (BP) as identified by the PASS algorithm. PIDs from different runs corresponding to each other are shown in the same color. The PIDs representing the native binding pocket are shown as thicker lines.

On the other hand, the nonreproducible pockets were most often rare event pockets (cf. Table 3). However, a PID may be reproduced by more than one PID (as Figure 3d shows, PIDs resulting from the same run may overlap, too) and their frequencies are often quite different.

To test whether the native binding pocket is among these transient pockets, we superimposed the inhibitor bound crystal structures onto the MD snapshots and onto the crystal structure of the apo protein and then determined the PASS probes overlapping the inhibitor atoms for the inhibitor bound structures, the apo structure, and the MD snapshots. New pocket volumes were calculated by considering only those overlapping PASS probes and comparing their volumes to those obtained for the inhibitor bound structures. The results (shown in Table 4) indicate that for all three systems the native binding pocket opened up during MD simulations. For BCL-X_L and MDM2, where the native binding pocket was already detectable in the crystal structure of the free protein, the mean overlap volume determined for the MD snapshots was more or less comparable to the overlap volume of the apo form of the protein (35.6% for BCL-X_L and 42.2% for MDM2). However, in some MD snapshots of BCL-X_L the native binding pocket was more than twice as large as in the apo form, although not quite as large as in the inhibitor bound crystal structure. In some MD snapshots of MDM2, the PASS volumes overlapping with the superimposed inhibitor were even of equal or larger size than in the inhibitor bound complex, indicating that the native binding pocket was sometimes large enough to fully accommodate the native inhibitor. Notably, for IL-2, where the native binding pocket was not detectable in the apo form, the native binding

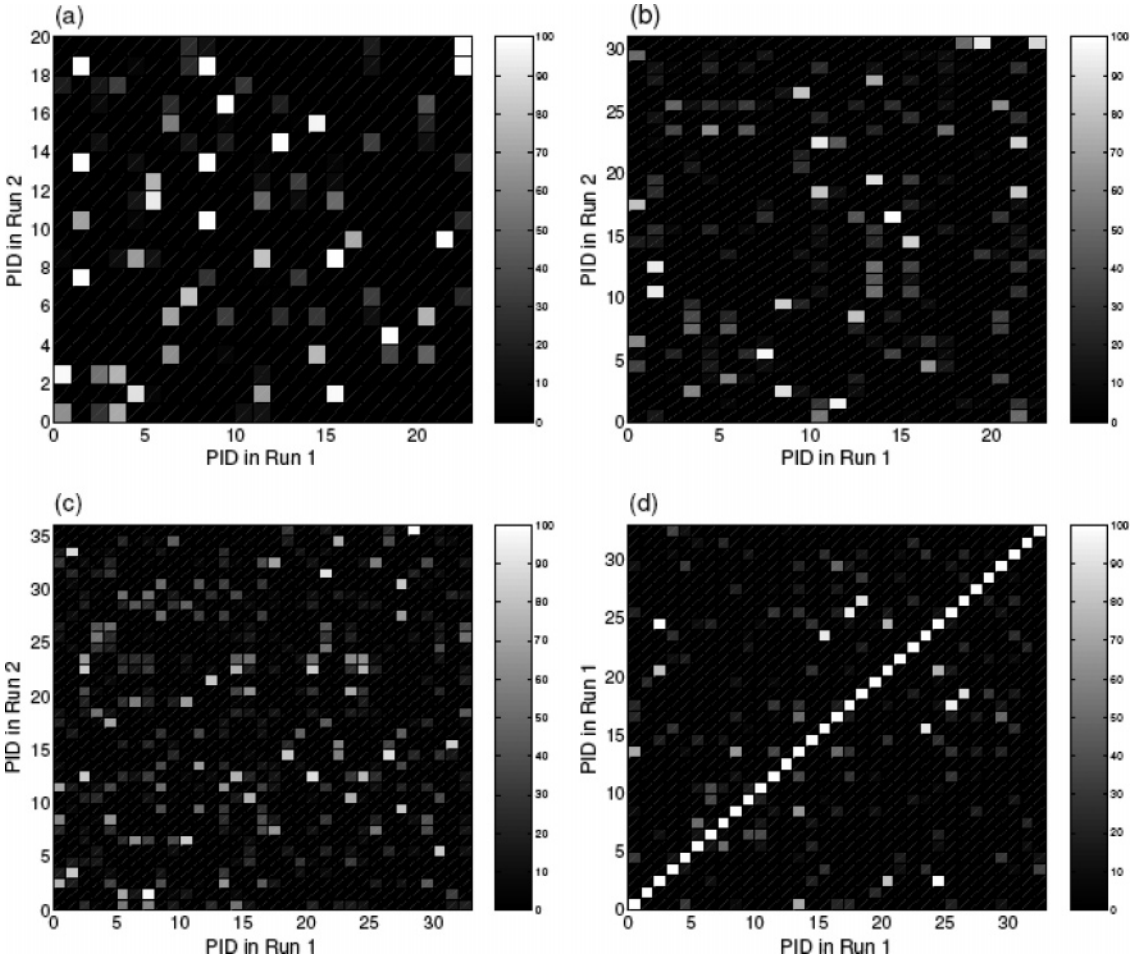

Figure 3. Comparison of the PIDs obtained in two different runs for BCL-X_L (a), IL-2 (b), and MDM2 (c) showing that most PIDs are reproducible and (d) showing the similarity of PIDs obtained within the same run (run 1) for MDM2. The shading scheme indicates the level of similarity.

Table 3. Reproducibility of the PIDs According to Their Frequency for the Two Independent Runs

	reproducibility (%)									
	frequency, <1%		frequency, 1−10%		frequency, 10−50%		frequency, >50%		total	
system	run 1	run 2	run 1	run 2	run 1	run 2	run 1	run 2	run 1	run 2
BCL-X_L	66.7	71.4	100.0	100.0	100.0	100.0	100.0	100.0	82.6	90.0
IL-2	81.8	62.5	100.0	83.3	100.0	100.0	100.0	100.0	91.3	77.4
MDM2	80.0	100.0	100.0	71.4	85.7	90.0	100.0	100.0	87.9	91.7

Table 4. Overlap between PASS Probes and Ligand Atoms Relative to the Overlap for Ligand Bound Structure

		MD mean overlap vol (%)		MD max overlap vol. (%)		no. MD snapshots with overlap		overlapping PIDs[a]	
system (PDB code)	apo overlap vol (%)	run 1	run 2	run 1	run 2	run 1	run 2	run 1	run 2
BCL-X_L−N3B (1YSI)	35.6	33.2	22.7	84.2	73.6	2,716	1,924	15	8, 11
IL-2−FRH (1PY2_A)	0	45.3	31.5	115.1[b]	130.7[b]	1,440	1,992	8	2, 4
MDM2−DIZ (1T4E_A)	42.2	53.7	39.4	136.2[b]	99.2	2,716	3,883	2, 5	0, 22

[a] Shown are the PIDs involved in the maximum overlap. [b] Overlapping volume is larger than in the complex crystal structure.

pocket was also found to fully open during the 10 ns MD simulations. Similar to MDM2, the overlapping volume was larger than in the inhibitor bound complex. Considering that the native binding pocket consists of two subsites, this result shows that the other subsite that was not detected in the inhibitor bound complex was detectable in some MD snapshots. This means that both subsites of the native binding pocket opened and that the binding pocket may be fully accessible in some of the MD snapshots.

The PIDs that possess the largest overlap with the native inhibitor were defined to correspond to the native binding pocket. Because the PIDs within the same run may overlap as well (resulting from the definition of the ASPs by the PASS algorithm), more than one PID may correspond to the native binding pocket. A comparison of the polarity ratios of these pockets to all others (Figure 2) indicates that for BCL-X_L and IL-2, the "native" PIDs represent the most nonpolar pockets. Note that the polarity plots for IL-2 show "native" PIDs corresponding to different subsites of the native binding pocket. In the first run of IL-2, the "native" PID corresponds to the subpocket identified in the inhibitor-bound crystal structure. Hence, it almost possesses the same polarity ratio as the native

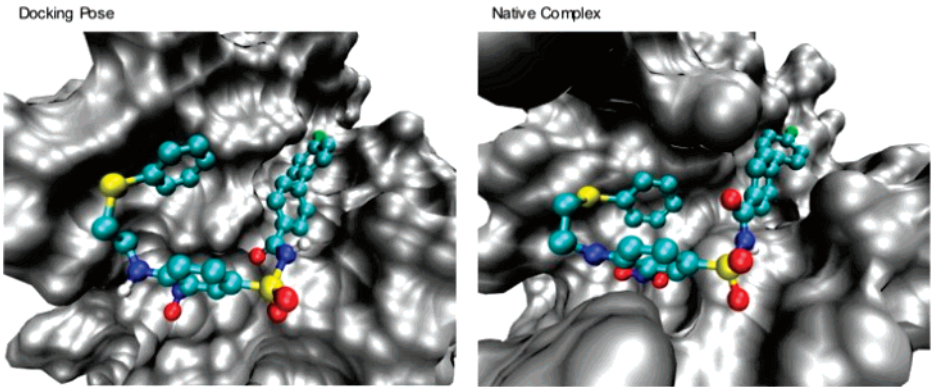

Figure 4. Comparison of a docking result of N3B to a BCL-X_L snapshot with 1.4 Å rmsd and the native BCL-X_L−N3B complex. Both structures are shown in the same orientation. N3B is shown in ball-and-stick representation, and BCL-X_L is shown in surface representation.

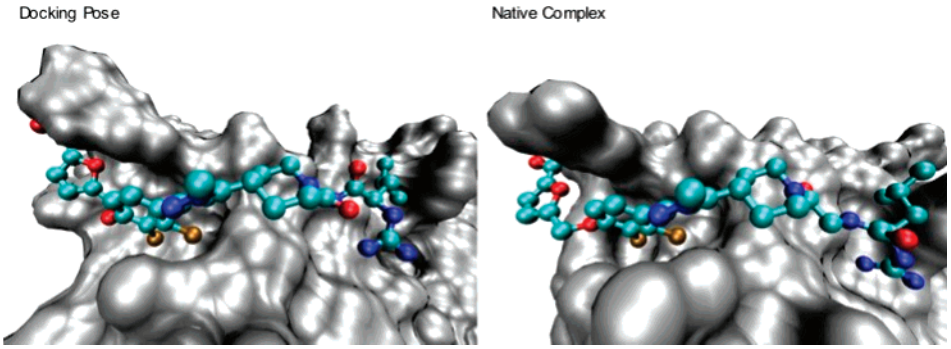

Figure 5. Comparison of a docking result of FRH to a IL-2 snapshot with 1.5 Å rmsd and the native IL-2−FRH complex. Both structures are shown in the same orientation. FRH is shown in ball-and-stick representation, and IL-2 is shown in surface representation.

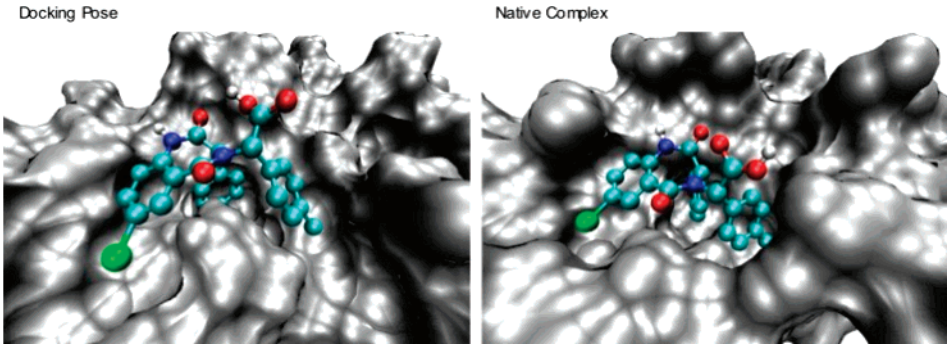

Figure 6. Comparison of a docking result of DIZ to a MDM2 snapshot with 1.9 Å rmsd and the native MDM2−DIZ complex. Both structures are shown in the same orientation. DIZ is shown in ball-and-stick representation, and MDM2 is shown in surface representation.

binding pocket at the reference volume. In the second run, the "native" PID corresponds to the less developed subpocket missed by the PASS algorithm in the inhibitor-bound crystal structure.

Except for IL-2, all detected transient pockets are more polar than the native binding pocket. This result suggests that only quite polar pockets open up during MD simulations in a polar solvent. Less polar pockets may require a less polar environment or the presence of a ligand to open, suggesting an induced fit mechanism.

Docking into MD Snapshots. Extensive docking studies were performed using the program AutoDock 3.0.5.[26] The results of the docking method into the MD snapshots are somehow biased toward the native bound ligand conformation because the bound conformations of the ligands were used to define the center of the search grid. The main purpose of these docking experiments was to validate whether these transient pockets are suitable to bind the native ligands and hence may be used for structure-based drug design. The results are found to be very promising. Figures 4−6 show the best docking poses in terms of root-mean-squared deviation (rmsd) along with the native complexes. The computed scoring energies are compared in Table 5. Note that this table only contains the docking results with the lowest rmsd. As indicated by the relative ranks of these results, it would not be possible to identify the correct docking solutions without prior knowledge. Let us, for example, consider the case of BCL-X_L−N3B. When taking the known center of mass of the ligand in the docked complex as the grid center of the docking run (see column termed "snapshot docking"), 4.7% of all docking poses have a better score than the docking pose with the smallest rmsd of 1.4 Å. In a de novo design project, this center of mass would of course not be known. It that case, one would use the center of mass of the transient pocket (see column termed "PID-docking"). Now, the best solution would only belong to the upper half of all docking poses. However, one should not exclusively focus on the single docking pose with the smallest rmsd. Instead, Table 6 lists the highest ranked docking solutions that can be classified as "correct" (rmsd ≤

Table 5. Best Docking Results for Redocking into Complex Structure, Docking into Apo Structure, and Docking into MD Snapshots

	redocking			apo-docking			snapshot-docking[b]			PID-docking[d]		
system	rmsd (Å)	score (kcal/mol)	rank[a]	rmsd (Å)	score (kcal/mol)	rank[a]	rmsd (Å)	score (kcal/mol)	rank[c] (%)	rmsd (Å)	score (kcal/mol)	rank[c] (%)
BCL-X_L−N3B	0.9	−10.5	2	3.3	−6.2	5	1.4	−8.7	4.7	1.5	−7.3	48.3
IL-2−FRH	1.1	−10.8	1	2.9	−6.2	1	1.5	−6.6	20.6	1.9	−6.5	14.1
MDM2−DIZ	1.1	−13.1	2	3.4	−6.7	5	1.9	−11.5	1.1	1.9	−11.5	0.7

[a] Rank of docking solution among 10 docking runs. [b] Docking into all MD snapshots (grid center coincident with center of mass of superimposed ligand). [c] Relative rank defined as the rank of this solution after sorting all results by increasing docking score in relation to the total number of docking results. [d] Docking into transient pockets (grid center coincident with center of mass of transient pocket).

Table 6. Highest Ranked Correct (rmsd ≤ 2.0 Å) Docking Results for Docking into Transient Pockets

	PID-docking[a]			
system	rmsd (Å)	score (kcal/mol)	score rank[b] (%)	final rank[c] (%)
BCL-X_L−N3B	1.8	−9.2	4.8	2.7
IL-2−FRH	2.0	−7.6	2.8	0.9
MDM2−DIZ	1.9	−11.5	0.7	0.7

[a] Docking into transient pockets (grid center coincident with center of mass of transient pocket). [b] Relative rank defined as the rank of this solution after sorting all results by increasing docking score in relation to the total number of docking results. [c] Number of docking results with better rank (i.e., better docking score) and higher fraction of buried nonpolar ligand atoms (i.e., relative number of nonpolar ligand atoms overlapping with PASS probes) than this one relative to the total number of docking results.

2.0 Å). At least one correct docking solution is always ranked among the best 5% of all docking results. Taking the fraction of buried nonpolar ligand atoms into account, the relative ranks can be reduced to less than 1% for IL-2 and MDM2 and to less than 3% for BCL-X_L.

Docking into transient pockets generally led to conformations much closer to the complex structure than docking into the crystal structures of the free proteins. Note that the shapes of the surfaces are somehow different so that some deviations are to be expected. The docking scores in Tables 5 and 6 indicate that pockets of appropriate shapes form spontaneously during molecular dynamics simulations of the free BCL-X_L and MDM2 proteins. For IL-2 only, we notice a significant improvement of the docking score when comparing the redocking result into the crystal structure of the complex to the other entries. Here, the docking scores for apo docking and MD snapshot docking suggest that formation of the native binding pocket requires the presence of the ligand with subsequent induced-fit effects. As an initial criterion for the suitability of this MD-based approach for detecting transient binding pockets, we suggest considering the lowering of docking scores for MD snapshots vs apo structures. This needs to be tested of course for a larger number of model cases.

Nevertheless, we suggest this protocol as a starting point for structure-based drug design especially in cases when no appropriate cavity can be identified on the surface of the crystal structure of the target protein. The regions in which transient pockets open may then be used as potential binding sites for virtual screening with flexible docking methods. Another possible application in virtual screening is to use these transient pockets and their properties as a prefiltering tool to reduce the number of ligands to be docked.

Conclusion

We have shown by applying standard molecular dynamics simulations to three protein systems that a surprisingly large number of transient pockets open up on protein surfaces on a 10 ns time scale. Reflecting the fluidlike properties of protein surfaces, these pockets open and close quickly. Their volumes may be as large as those of ligand binding pockets. More frequently open pockets tend to have larger volumes on average. As evidenced by a second set of control simulations, most transient pockets were reproducible. Open pockets are usually more hydrophilic than the actual native binding pocket, suggesting the importance of induced fit effects during ligand binding. Flexible ligand docking into the MD snapshots reproduced the native binding modes to about 2 Å rmsd from the crystal structure conformation. This encouraging result underlines the importance of properly accounting for protein flexibility in ligand design studies, particularly on the protein surface. Our pocket detection protocol may therefore be an interesting starting point for structure-based drug design especially for the design of protein−protein interaction inhibitors, when the crystal structure of the target protein lacks appropriate binding pockets.

Experimental Section

Preparation of Crystal Structures. The apo forms of the proteins were simulated using the following crystal structures from the PDB (Protein Data Bank):[27] 1R2D (apo BCL-X_L), 1Z1M (apo MDM2), and 1M47 (apo IL-2). All hetero atoms (waters and two sulfate anions in 1M47) were removed. Because residues 28−81 are missing in 1R2D, the two parts were modeled as two distinct chains. The missing residues in 1M47 were modeled as loops of the lowest AMBER/GBSA potential energy generated by the program RAPPER.[28] We note that these missing residues are far away from the native binding pocket.

Molecular Dynamics Simulations. Molecular dynamics (MD) simulations of the protein systems were performed with the GROMACS 3.3 package[29] using the OPLS-AA force field.[30] After placing the prepared proteins in cubic water boxes of 6.2−8.3 nm box dimensions, the system was minimized by 500 steps of steepest-descent energy minimization keeping the heavy proteins atoms fixed. Solvent molecules were modeled explicitly using the TIP4P water model.[31] If needed, counterions (Na^+ or Cl^-, respectively) were added to ensure a net neutral charge of the systems and the pre-equilibrating run was repeated. Equilibration continued during a 100 ps simulation in the *NPT* ensemble at a temperature of 300 K with periodic boundary conditions. Simulation snapshots were collected during a subsequent 10 ns simulation. Electrostatic interactions beyond the short-range cutoff of 0.9 nm were treated by the particle-mesh Ewald method.[32] van der Waals interactions were computed within a 0.9 nm cutoff. Temperature and pressure were kept constant at standard conditions (10^5 Pa, 300 K) by weak coupling to a temperature and pressure bath[33] with coupling constants of 0.1 ps for the temperature coupling and 1 ps for the pressure coupling. Protein, solvent, and counterions were coupled to separate baths. The LINCS procedure[34] was used to constrain all covalent bonds. Snapshots were taken every 2.5 ps, yielding a total of 4000 MD snapshots. Before these snapshots were further processed, they were superimposed based on the C_α atoms using the VMD program,[35] and all hydrogen atoms were removed. MD simulations were repeated once for every system to check the results for reproducibility.

Pocket Detection Using the PASS Algorithm. PASS (putative active sites with spheres) is a computational tool that uses geometric

criteria to characterize regions of buried volume in proteins and to identify positions likely to represent binding sites based on size, shape, and burial extent of these volumes.[16] After the respective atomic radii were assigned, the pocket screening process starts by coating the protein surface with an initial layer of spherical probes of radius 1.8 Å. Subsequently accretion layers of smaller probes of radius 0.7 Å are added onto the previously identified probes. After addition of each new layer all newly found probes are filtered before the next probe layer is added. These steps are repeated until a layer is encountered in which no additional probes pass the filters. For each probe in this final set of probes a weight is calculated that reflects their burial extent. A probe is defined as an active site point (ASP) if it has a weight above a given threshold and if it is separated by a minimum distance of 8.0 Å from other ASPs. These ASPs are meant to represent potential binding sites. We used PASS with the "-more" option to obtain an enhanced set of ASPs and probes.

Pocket Dynamics and Properties. The ASPs and probes positioned by the PASS algorithm were processed to identify contiguous pocket volumes. These volumes represent the negative image of a binding pocket as identified by the PASS algorithm. Because the pockets have irregular shapes and possess no well-defined boundary with the surrounding solvent, determination of their volume is not unambiguous. Here, we approximate this volume V from the PASS probes positioned in each pocket:

$$V \simeq N_{\text{layer}=1}\frac{4\pi}{3}(1.8\ \text{Å})^3 + N_{\text{layer}>1}\frac{4\pi}{3}(0.7\ \text{Å})^3 \qquad (1)$$

Here, $N_{\text{layer}=1}$ is the number of PASS probes belonging to the first layer (probes with radius 1.8 Å) and $N_{\text{layer}>1}$ is the number of PASS probes belonging to subsequent accretion layers (probes with radius 0.7 Å). Each pocket volume was represented by one ASP and the surrounding probes. Doing so, each probe was assigned to the volume involving the nearest ASP. Probes that were not adhered to any pocket volume (ASP) were ignored.

All protein atoms found within a distance of 5 Å from these pocket volumes were identified as "pocket lining atoms". A polarity ratio was assigned to each pocket depending on the chemical properties of these lining protein atoms, where N, O, and S atoms were considered as polar and all other atoms as nonpolar.

After application of the PASS algorithm to each MD snapshot, all pocket volumes occurring in the 4000 snapshots (about 11000−20000) were clustered using an agglomerative complete linkage approach. The similarity between two pocket volumes was defined as the relative agreement of their pocket lining atoms. Care was taken during the clustering step to ensure that the similarity was at least 85% and that no cluster contained more than one pocket volume taken from the same MD snapshot.

After the clustering step, all pocket volumes within the same cluster were labeled by the same unique pocket identifier (PID). The term PID therefore refers to a transient pocket. The dynamics of these pockets could be observed via these corresponding pocket volumes representing their different states taken from subsequent MD structures. To compare the PIDs to each other, subpockets sub(PID$_i$) were determined as the set of all protein atoms that lined the given PID$_i$ in at least 33% of all occurrences of this PID. The subpockets resulting from the two runs were compared to each other and the pairwise similarities of PIDs PID$_i$ and PID$_j$ (with PID$_i$ and PID$_j$ resulting from different runs) were determined as

$$\text{sim}(\text{PID}_i,\text{PID}_j) = \frac{|\text{sub}(\text{PID}_i) \cap \text{sub}(\text{PID}_j)|}{\max(|\text{sub}(\text{PID}_i)|,|\text{sub}(\textit{PID}_j)|)} \times 100 \qquad (2)$$

The complete analysis of the transient pockets including application of the PASS algorithm, clustering, and calculation of the properties took 16−20 h for each set of 4000 MD snapshots on one 2.8 GHz Xenon CPU.

Docking Method. Docking experiments were performed with AutoDock 3.0.5.[26] The ligands were extracted from the complex crystal structures, and hydrogens were added with the AutoDockTools (ADT 1.4.3) modules of the Python Molecular Viewer software.[36] The same software was used to compute the Gasteiger atomic charges.[37] Finally the rotatable bonds were assigned with AutoTors. The number of flexible torsions was 10 for N3B, 17 for FRH, and 5 for DIZ.

Four different docking experiments were performed: redocking into the inhibitor bound conformation, docking into the apo crystal structures, docking into all MD snapshots, and docking into all transient pockets identified by PASS. The apolar hydrogens of the MD snapshots were removed, and polar hydrogens were added for the crystal structures. Kollman united-atom partial charges and solvation parameter were assigned using the AutoDockTools utility. The grid maps were calculated with AutoGrid.

For all but the last docking experiment the grid center was chosen to coincide with the center of mass of the ligand in its bound conformation. For this purpose the complex structure was superimposed with the apo structure and the MD snapshots. The default grid dimensions of 21 Å × 21 Å × 21 Å and the default spacing of 0.375 Å between the grid points were used. For the docking into transient pockets, no prior information about the native bound ligand conformation was used. Therefore, the center of mass of the individual transient pocket was used as a grid center. Taking into account that only a terminal moiety of the ligands may be placed into the transient pocket, the grid dimensions were expanded to 30 Å × 30 Å × 30 Å for BCL-X_L and IL-2. For MDM2, the grid dimensions were reduced to 16.125 Å × 16.125 Å × 16.125 Å to confine the position of the smaller ligand to the transient pocket.

For the docking procedure the standard Lamarckian genetic algorithm protocol was used with an initial population of 50 randomly placed individuals, a maximum number of 250 000 energy evaluations, a mutation rate of 0.02, a crossover rate of 0.80, and an elitism value of 1. The probability of performing a local search on an individual was set to 0.06, and the maximum number of consecutive successes or failures before doubling or halving the local search step size was 4. The pseudo Solis and Wets algorithm was applied for these local searches with a maximum of 300 iterations.

Ten independent docking runs were carried out for each MD snapshot. Docking into a transient pocket took 1−3 min on one 2.8 GHz Xenon CPU depending on the flexibility of the ligand and the size of the grid box.

Acknowledgment. This work was funded by the Center for Bioinformatics Saar (DFG). We thank Dr. Michael Hutter and Dr. Iris Antes for helpful comments.

Supporting Information Available: Figure 1 showing the secondary structure of the simulated proteins over the simulation time and Figures 2−4 showing for each system the ligand binding mode, the binding pocket as detected by the PASS algorithm in the complex and apo crystal structure, and the transient pockets projected on the surface of the apo crystal structure. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Sharma, S. K.; Ramsey, T. M.; Bair, K. W. Protein−protein interactions: lessons learned. *Curr. Med. Chem.: Anti-Cancer Agents* **2002**, *2*, 311−330.
(2) Berg, T. Modulation of protein−protein interactions with small organic molecules. *Angew. Chem., Int. Ed. Engl.* **2003**, *42*, 2462−2481.
(3) Yin, H.; Hamilton, A. D. Strategies for targeting protein−protein interactions with synthetic agents. *Angew. Chem., Int. Ed. Engl.* **2005**, *44*, 4130−4163.
(4) Ryan, K. M.; Phillips, A. C.; Vousden, K. H. Regulation and function of the p53 tumor suppressor protein. *Curr. Opin. Cell Biol.* **2001**, *13*, 332−337.
(5) Michael, D.; Oren, M. The p53-Mdm2 module and the ubiquitin system. *Semin. Cancer Biol.* **2003**, *13*, 49−58.
(6) Hollstein, M.; Sidransky, D.; Vogelstein, B.; Harris, C. C. p53 mutations in human cancers. *Science* **1991**, *253*, 49−53.

(7) Zheleva, D. I.; Lane, D. P.; Fischer, P. M. The p53-Mdm2 pathway: targets for the development of new anticancer therapeutics. *Mini-Rev. Med. Chem.* **2003**, *3*, 257−270.
(8) Fry, D. C.; Emerson, S. D.; Palme, S.; Vu, B. T.; Liu, C. M.; Podlaski, F. J. NMR structure of a complex between MDM2 and a small molecule inhibitor. *J. Biomol. NMR* **2004**, *30*, 163−173.
(9) Vassilev, L. T.; Vu, B. T.; Graves, B.; Carvajal, D.; Podlaski, F.; Filipovic, Z.; Kong, N.; Kammlott, U.; Lukacs, C.; Klein, C.; Fotouhi, N.; Liu, E. A. In vivo activation of the p53 pathway by small-molecule antagonists of MDM2. *Science* **2004**, *303*, 844−848.
(10) Grasberger, B. L.; Lu, T.; Schubert, C.; Parks, D. J.; Carver, T. E.; Koblish, H. K.; Cummings, M. D.; LaFrance, L. V.; Milkiewicz, K. L.; Calvo, R. R.; Maguire, D.; Lattanze, J.; Franks, C. F.; Zhao, S.; Ramachandren, K.; Bylebyl, G. R.; Zhang, M.; Manthey, C. L.; Petrella, E. C.; Pantoliano, M. W.; Deckman, I. C.; Spurlino, J. C.; Maroney, A. C.; Tomczuk, B. E.; Molloy, C. J.; Bone, R. F. Discovery and cocrystal structure of benzodiazepinedione HDM2 antagonists that activate p53 in cells. *J. Med. Chem.* **2005**, *48*, 909−912.
(11) Arkin, M. R.; Wells, J. A. Small-molecule inhibitors of protein−protein interactions: progressing towards the dream. *Nat. Rev. Drug Discovery* **2004**, *3*, 301−317.
(12) Apostolakis, J.; Plueckthun, A.; Caflisch, A. Docking Small ligands in flexible binding sites. *J. Comput. Chem.* **1997**, *19*, 21−37.
(13) Zacharias, M. Rapid protein−ligand docking using soft modes from molecular dynamics simulations to account for protein deformability: binding of FK506 to FKBP. *Proteins* **2004**, *54*, 759−767.
(14) Meiler, J.; Baker, D. ROSETTALIGAND: protein−small molecule docking with full side-chain flexibility. *Proteins* **2006**, *65*, 538−548.
(15) Sousa, F. S.; Fernandes, P. A.; Ramos, M. J. Protein−ligand docking: current status and future challenges. *Proteins* **2006**, *65*, 15−26.
(16) Brady, G. P.; Stouten, P. F. W. Fast prediction and visualization of protein binding pockets with PASS. *J. Comput.-Aided Mol. Des.* **2000**, *14*, 383−401. http://www.ccl.net/cca/software/UNIX/pass/overview.shtml.
(17) Levitt, D. G.; Banaszak, L. J. POCKET: a computer graphics method for identifying and displaying protein cavities and their surrounding amino acids. *J. Mol. Graphics* **1992**, *10*, 229−234.
(18) Hendlich, M.; Rippmann, F.; Barnickel, G. LIGSITE: automatic and efficient detection of potential small molecule-binding sites in proteins. *J. Mol. Graphics Modell.* **1997**, *15*, 359−363.
(19) Laskowski, R. A. SURFNET: a program for visualizing molecular surfaces, cavities, and intermolecular interactions. *J. Mol. Graphics* **1995**, *13*, 323−330.
(20) Laurie, A. T. R.; Jackson, R. M. Q-SiteFinder: an energy-based method for the prediction of protein−ligand binding sites. *Bioinformatics* **2005**, *21*, 1908−1916.
(21) Meagher, K. L.; Carlson, H. A. Incorporating protein flexibility in structure-based drug discovery: using HIV-1 protease as a test case. *J. Am. Chem. Soc.* **2004**, *126*, 13276−13281.
(22) Sivanesan, D.; Rajnarayanan, R. V.; Doherty, J.; Pattabiraman, N. In-silico screening using flexible ligand binding pockets: a molecular dynamics-based approach. *J. Comput.-Aided Mol. Des.* **2005**, *19*, 213−228.
(23) Wong, C. F.; Kua, J.; Zhang, Y.; Straatsma, T. P.; McCammon, J. A. Molecular docking of balanol to dynamics snapshots of protein kinase A. *Proteins* **2005**, *61*, 850−858.
(24) Frembgen-Kesner, T.; Elcock, A. H. Computational sampling of a cryptic drug binding site in a protein receptor: explicit solvent molecular dynamics and inhibitor docking to p38 MAP kinase. *J. Mol. Biol.* **2006**, *359*, 202−214.
(25) Helms, V. Protein dynamics tightly connected to the dynamics of surrounding and internal water molecules. *ChemPhysChem* **2007**, *8*, 23−33.
(26) Morris, G. M.; Goodsell, D. S.; Halliday, R. S.; Huey, R.; Hart, W. E.; Belew, R. K.; Olson, A. J. Automated docking using a Lamarckian genetic algorithm and empirical binding free energy function. *J. Comput. Chem.* **1998**, *19*, 1639−1662. http://autodock.scripps.edu.
(27) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28*, 235−242. http://www.rcsb.org.
(28) de Bakker, P. I.; DePristo, M. A.; Burke, D. F.; Blundell, T. L. Ab initio construction of polypeptide fragments: accuracy of loop decoy discrimination by an all atom statistical potential and the AMBER force field with the generalized Born solvation model. *Proteins* **2003**, *51*, 21−40. http://mordred.bioc.cam.ac.uk/∼rapper.
(29) Lindahl, E.; Hess, B.; van der Spoel, D. GROMACS 3.0: A package for molecular simulation and trajectory analysis. *J. Mol. Model.* **2001**, *7*, 306−317. www.gromacs.org/.
(30) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *J. Am. Chem. Soc.* **1996**, *118*, 11225−11236.
(31) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926−935.
(32) Darden, T.; York, D.; Pedersen, L. Particle mesh ewald: An n log(n) method for Ewald sums in large systems. *J. Chem. Phys.* **1993**, *98*, 10089−10092.
(33) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **1984**, *81*, 3684−3690.
(34) Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M. LINCS: A linear constraint solver for molecular simulations. *J. Comput. Chem.* **1997**, *18*, 1463−1472.
(35) Humphrey, W.; Dalke, A.; Schulten, K. VMD. Visual molecular dynamics. *J. Mol. Graphics* **1996**, *14*, 33−38. http://www.ks.uiuc.edu/Research/vmd/.
(36) Sanner, M. F. Python: a programming language for software integration and development. *J. Mol. Graphics Modell.* **1999**, *17*, 57−61. http://mgltools.scripps.edu.
(37) Gasteiger, J.; Marsili, M. Iterative partial equilibration of orbital electronegativity−a rapid access to atomic charges. *Tetrahedron* **1980**, *36*, 3219−3228.

JM070095G