

# Hot Spots and Transient Pockets: Predicting the Determinants of Small-Molecule Binding to a Protein–Protein Interface

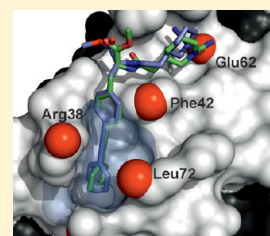
Alexander Metz,<sup>†,§</sup> Christopher Pfleger,<sup>†,§</sup> Hannes Kopitz,<sup>†</sup> Stefania Pfeiffer-Marek,<sup>‡</sup> Karl-Heinz Baringhaus,<sup>‡</sup> and Holger Gohlke<sup>\*,†</sup>

<sup>†</sup>Institute for Pharmaceutical and Medicinal Chemistry, Department of Mathematics and Natural Sciences, Heinrich-Heine-University, Düsseldorf, Germany

<sup>‡</sup>Sanofi-Aventis Deutschland GmbH, LGCR Drug Design, Frankfurt am Main, Germany

 Supporting Information

**ABSTRACT:** Protein–protein interfaces are considered difficult targets for small-molecule protein–protein interaction modulators (PPIMs). Here, we present for the first time a computational strategy that simultaneously considers aspects of energetics and plasticity in the context of PPIM binding to a protein interface. The strategy aims at identifying the determinants of small-molecule binding, hot spots, and transient pockets, in a protein–protein interface in order to make use of this knowledge for predicting binding modes of and ranking PPIMs with respect to their affinity. When applied to interleukin-2 (IL-2), the computationally inexpensive constrained geometric simulation method FRODA outperforms molecular dynamics simulations in sampling hydrophobic transient pockets. We introduce the PPIAnalyzer approach for identifying transient pockets on the basis of geometrical criteria only. A sequence of docking to identified transient pockets, starting structure selection based on hot spot information, RMSD clustering and intermolecular docking energies, and MM-PBSA calculations allows one to enrich IL-2 PPIMs from a set of decoys and to discriminate between subgroups of IL-2 PPIMs with low and high affinity. Our strategy will be applicable in a prospective manner where nothing else than a protein–protein complex structure is known; hence, it can well be the first step in a structure-based endeavor to identify PPIMs.



## INTRODUCTION

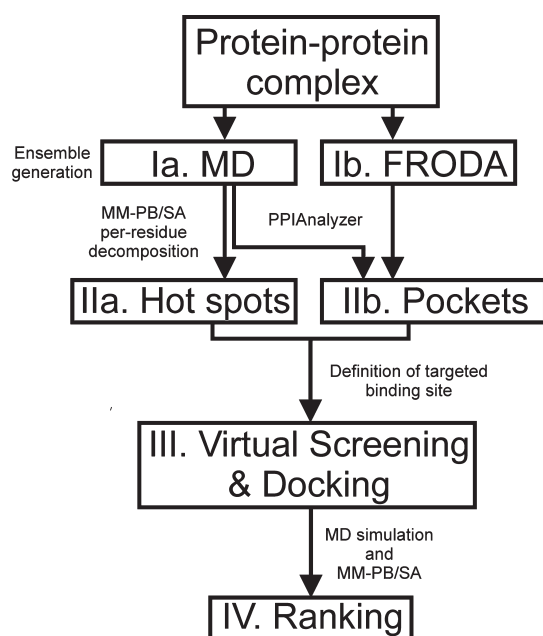
Protein–protein interactions (PPIs) are involved in nearly all biological processes. Due to their universal occurrence, protein–protein interfaces provide an important, yet neglected, new class of drug targets.<sup>1</sup> At present, the design of small-molecule protein–protein interaction modulators (PPIMs) encounters at least two challenges. First, in contrast to enzymes, protein–protein interfaces are rather flat and usually lack a distinct binding pocket.<sup>2</sup> Second, due to the often large size of protein–protein interfaces ( $\sim 1200$  to  $\sim 4660$  Å<sup>2</sup>)<sup>3,4</sup> interactions that are favorable for binding can be widely distributed over the interface.

Experimental evidence suggests that these challenges can be overcome.<sup>5–8</sup> Most strikingly, residues participating in important interactions have been shown to be spatially clustered in protein–protein interfaces, forming so-called “hot spot” regions.<sup>4,9–11</sup> Mimicking localized interactions at these hot spots provides a possibility for PPIM development.<sup>1,4,12–14</sup> Furthermore, an opening of so-called transient pockets was observed in protein–protein interfaces.<sup>5</sup> In fact, binding of several PPIMs to transient pockets in protein–protein interfaces has been reported.<sup>7</sup> The most prominent example is given by small-molecule inhibitors binding to interleukin-2 (IL-2). These PPIMs inhibit the interaction with the IL-2  $\alpha$ -receptor (IL-2R $\alpha$ ).<sup>5,15,16</sup> Notably, the small molecule-bound IL-2 exhibits pockets in the interface region that are present neither in the unbound nor in the IL-2R $\alpha$ -bound crystal structure.<sup>5,17</sup>

Computational methods can aid in finding and in the design of PPIMs if they are able to provide an accurate description of the energetics and dynamics of small-molecule binding to protein–protein interfaces.<sup>7,13,18–21</sup> While many computational studies have been reported that deal with the energetics<sup>13,22–33</sup> and dynamics<sup>7,13,19,34–39</sup> of protein–protein interfaces *per se*, only a few have focused on the aspect of small-molecule binding to protein–protein interface regions,<sup>7,14,20,40–46</sup> and none has considered aspects of energetics and interface plasticity simultaneously in this context so far. Thus, in the present study, we set out to evaluate the capability of state-of-the-art computational methods to predict small-molecule binding to protein–protein interfaces. In particular, we address five major questions that consider aspects of structure, dynamics, and energetics important for binding to protein–protein interfaces: (i) Can one identify hot spots in a protein–protein interface based on a protein–protein complex structure and make use of these hot spots for predicting the binding mode of small-molecule ligands? (ii) Can one sample the opening of transient pockets comparable to those observed in the protein–protein interface of a bound protein, as suggested by the conformational selection model,<sup>47</sup> starting from an unbound protein conformation? (iii) Is it possible to identify protein conformations with transient pockets by energetic or

Received: July 13, 2011

Published: November 17, 2011



**Figure 1.** Outline of the strategy for hot spot and transient pocket identification, docking, and ranking of PPIMs using only the structure of a protein–protein complex as a starting point. A similarity-based virtual screening was not performed in this study. See text for details.

geometrical criteria? (iv) By docking to transient pockets, can one reproduce binding modes of known PPIMs and enrich PPIMs by virtual screening? (v) Can one rank known PPIMs with respect to their affinity?

To answer these questions, we chose IL-2 as a model system because of the wealth of information available for this system in terms of crystal structures of unbound IL-2, IL-2 bound to IL-2R $\alpha$ , and IL-2 bound to five PPIMs as well as experimental binding and inhibition data for the wild type and mutant protein.<sup>5,5,17,48</sup> As to the methodological strategy (Figure 1), we apply and compare molecular dynamics (MD) and constrained geometric (FRODA) simulations for generating structural ensembles, introduce the PPIAnalyzer approach for investigating structural properties of protein–protein interfaces within these ensembles, and apply the MM-PB(GB)SA (molecular mechanics Poisson–Boltzmann (generalized Born) surface area) approach to identify hot spots and rank PPIMs. We note that when pursuing this strategy, we paid particular attention to mimicking a “real-life” scenario in structure-based ligand design. Thus, our strategy will be applicable also in a prospective study where nothing else than a protein–protein complex structure is known at the beginning.

## MATERIALS AND METHODS

**Overall Strategy.** As to the methodological strategy (Figure 1), we pursued the following steps:

- Ia/b. Starting from a given protein–protein complex structure, conformational ensembles are generated on the basis of MD and FRODA simulations.
- IIa. Hot spot residues of a protein–protein complex structure are identified by MM-PBSA free energy decomposition on the basis of the structural ensemble generated by MD simulation.

- IIb. Transient pockets in the protein–protein interface are identified by energetic or geometrical criteria in conformational ensembles generated either by MD or by FRODA.
- III. PPIM binding poses are predicted by molecular docking using the hot spot and transient pocket information for guidance. This docking setup is also used to enrich PPIMs from a large set of decoys, thus performing virtual screening.
- IV. PPIMs are ranked by their MM-PBSA binding effective energies ( $\Delta G_{\text{eff}}$  = gas phase energy + solvation free energy according to a continuum solvent model) calculated for conformational ensembles from MD simulations that were started from the docked binding poses.

These steps will now be described in more detail. Detailed information about structure preparation and protocols for molecular dynamic simulations and docking experiments is provided in the Supporting Information.

**FRODA Simulations.** FRODA is a geometrical simulation method that explores the internal mobility of biomolecular systems. Details of the algorithm can be found in ref 49. The FRODA simulation was performed with the FIRST 6.2 suite of programs. For the rigid cluster decomposition with FIRST, a hydrogen bond energy cutoff of  $-1.0 \text{ kcal mol}^{-1}$  was applied together with the “H 1” function for hydrophobic interactions. For each system, 10 000 000 conformations were sampled, of which every 10 000th conformation was stored. The random displacement distance of the mobile atoms was set to 0.1 Å, and the continuous motion (CM) method was applied.

**MM-PB(GB)SA Calculations.** MM-PBSA<sup>34,50</sup> calculations were carried out according to the “multiple trajectory method” and the “single trajectory method”. For the multiple trajectory method, snapshots of all of the IL-2 complexes, the unbound IL-2 structures, and the IL-2R $\alpha$  subunit as well as the small-molecule ligands were extracted from independent MD trajectories. Alternatively, for the single trajectory method, snapshots of the binding partners were extracted from MD trajectories of the complexes only. All counterions and water molecules were stripped from the snapshots. Snapshots were extracted every 10 ps. Autocorrelation analysis of the effective energy of snapshots revealed that this time interval is sufficient for generating statistically independent snapshots. The gas phase energy was calculated on the basis of the ff99SB force field<sup>51</sup> without applying any non-bonded cutoff. The polar part of the solvation free energy was determined by solving the linearized Poisson–Boltzmann (PB) equation<sup>52</sup> or by applying the “OBC” generalized Born (GB) method ( $igb = 5$ ) using mbondi2 radii.<sup>53</sup> A dielectric constant of 1 and 80 for the interior and exterior of the solute was applied, respectively. The polar contributions were computed at 100 mM ionic strength. Nonpolar solvation energies were calculated by a solvent-accessible surface area (SASA) dependent term, using a surface tension proportionality constant of  $\gamma = 0.0072 \text{ kcal mol}^{-1} \text{ Å}^{-2}$ . Contributions from vibrational entropy were neglected,<sup>54</sup> which can be justified with the small-molecule ligands being very similar. For calculating per residue contributions, the decomposition scheme<sup>22</sup> implemented in the SANDER and MM-PBSA code of AMBER 10 was extended to also consider the PB reaction field energy. This is done on the basis of the concept of induced surface charges on the dielectric boundary.<sup>55</sup> The contribution of a residue  $k$  to the reaction field energy  $E_{\text{rf}}$  is then calculated without additional computational costs as the sum of Coulomb interactions of all of

its atomic charges  $q_i$  with all induced surface charges  $q_j$  with  $r$  being the distance between the two charges (eq 1):

$$E_{\text{rf}}(k) = \sum_{i \in k} \sum_j \frac{q_i q_j}{4\pi\epsilon_0 r} \quad (1)$$

**Identification of Residues in the Interface of IL-2/IL-2R $\alpha$**  Residues that are  $\leq 5$  Å apart from IL-2R $\alpha$  in the IL-2/IL-2R $\alpha$  complex structure (PDB code: 1z92) were chosen as interface regions of IL-2. Residues pointing toward the interior of IL-2 and thus not contributing to the binding of IL-2R $\alpha$  or small ligands were excluded upon visual inspection. This resulted in a set of 31 IL-2 interface residues: Tyr31, Asn33, Pro34, Lys35, Thr37, Arg38, Met39, Thr41, Phe42, Lys43, Phe44, Tyr45, Glu60, Glu61, Glu62, Lys64, Pro65, Leu66, Glu67, Glu68, Val69, Asn71, Leu72, Met104, Cys105, Glu106, Tyr107, Ala108, Asp109, Glu110, and Thr111.

**Structural Analysis of Protein–Protein Interfaces.** For investigating structural properties of protein–protein interface regions from conformational ensembles, we developed the PPIAnalyzer method (Figure S1). The method comprises three steps:

- I. Structural changes of the interface are determined in terms of root mean-square deviations (RMSD) of backbone and side chain atoms.
- IIa. The steric quality of the generated conformations is assessed.
- IIb. Distinct interface conformations are selected on the basis of a clustering with respect to the RMSD of heavy atoms of interface residues.
- III. Transient pockets are identified in those conformations.

In more detail, heavy atom RMSD values of interface residues are calculated with respect to the small molecule-bound (PDB codes: 1m48, 1m49, 1pw6, 1py2, and 1qvn) and unbound crystal structures (PDB code: 1m47) after structurally aligning the interface regions. We note that this RMSD calculation was only done to retrospectively assess the sampling of bound interface conformations; at no time of the study was knowledge of bound conformations applied for the identification of protein conformations that were subsequently used in docking experiments. The stereochemical quality of each snapshot was assessed using PROCHECK.<sup>56</sup> Snapshots not satisfying all of the following stereochemical criteria were excluded from further investigation:

- I. At most, two bad contacts are present.
- II. Less than 5% of the amino acids are in disallowed regions of the Ramachandran plot.
- III. At most two unfavorable main chain or side chain parameters are present. Of the remaining snapshots, 100 representative structures were selected using a  $k$ -medoids clustering algorithm<sup>57</sup> with respect to the RMSD of the interface residues.

Finally, potential binding pockets were detected using the PocketAnalyzer program<sup>58</sup> that implements a pocket identification strategy similar to the one proposed by Hendlich et al.<sup>59</sup> Details of the algorithm can be found in ref 58. Here, the following parameters were used: minimal degree of buriedness: 9; minimal number of neighbors: 9; minimal cluster size: 50; grid spacing: 1.0 Å. Snapshots for the subsequent docking experiments were chosen with respect to their identified pocket volume.

**Data Set of Useful Decoys.** Following the procedure described to generate the directory of useful decoys (DUD),<sup>60</sup> we selected compounds from the “purchasable subset” of the ZINC

database<sup>61</sup> (as of May 19, 2010) that are similar with respect to physicochemical properties to the five IL-2 ligands in complex structures (Table 1). Descriptors for the ZINC compounds were downloaded from the ZINC Web site or were calculated for the IL-2 ligands using Molinspiration.<sup>62</sup> The number of functional groups was calculated using the OpenEye FILTER program.<sup>63</sup> The pairwise dissimilarity between compounds was calculated as the *weighted root mean-cubed difference* (RMCD; eq 2) of the differences of normalized descriptors  $X_i$  (weights  $w_i$  in parentheses) logP (8); molecular weight (4); number of hydrogen bond donors (4) and acceptors (4); number of rotatable bonds (4); number of amide (1), amino (1), and carboxylic acid (1) groups; and the sum of the numbers of amidino and guanidino groups (1).

$$\text{RMCD} = \sqrt[3]{\frac{\sum_i w_i |X_i|^3}{\sum_i w_i}} \quad (2)$$

The ZINC compounds were then sorted by their pairwise RMCDs to all five IL-2 ligands in complex structures. The 10 000 most similar ZINC compounds were clustered into 1000 clusters on the basis of the pairwise RMCD using hierarchical clustering according to Ward’s method as implemented in the hclust module of R.<sup>64</sup> Out of each cluster, the compound with the smallest RMSD of pairwise RMCD to the reference ligands was selected. Four compounds were not retrievable from the ZINC. For the remaining 996 unique decoy structures, a total of 1297 protonation and tautomerization states, as stored in the ZINC database, were considered.

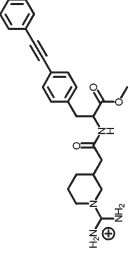
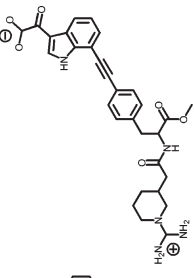
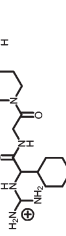
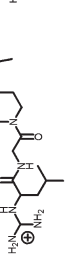
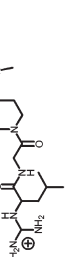
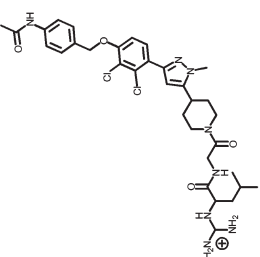
**Ranking of Docked Structures.** For ranking docked structures by MM-PBSA, appropriate starting structures for the MD-based snapshot generation must be chosen initially. For this, consider that 100 docking runs were performed for each of the 10 structures with the largest interface pocket volumes obtained by a FRODA simulation of the unbound IL-2 structure and for each ligand. The starting structure selection was based on hot spot information, RMSD clustering, and intermolecular docking energy. First, it was required that a ligand’s guanidinium group be within 5 Å of the side chain heavy atoms of the hot spot amino acid Glu62. Second, all of the remaining docking poses were clustered with respect to heavy atom RMSD of the PPIMs in the docked complexes after aligning only the proteins. Hierarchical complete linkage clustering was performed with R<sup>64</sup> with a cluster distance of 5 Å. Compared to clustering docking results above, a larger clustering distance of 5 Å was chosen to account for the fact that the clustering is performed over complex structures with different receptor conformations. Finally, the structure with the lowest intermolecular docking energy from the largest cluster was chosen as a starting structure for MD simulation and subsequent MM-PB-(GB)SA binding effective energy calculation. Equilibrations, production runs, and MM-PB(GB)SA calculations were performed as described in the Supporting Information for the crystal structures. To allow the complex structures to relax after geometrical FRODA simulations and subsequent docking, we performed 20 ns of unrestrained MD simulation, with only the last 10 ns being used for snapshot extraction.

## RESULTS AND DISCUSSION

**Structures from MD Simulations.** In order to investigate the opening of transient pockets in the IL-2 interface and to calculate effective energies by MM-PB(GB)SA, conformational ensembles of



Table 1. Structures of IL-2, IL-2 Complexes, and IL-2 Binding Partners with Respective Structural, Experimental, and Computed Data

PDB code	1m47 <sup>a</sup>	1m4c <sup>a</sup>	1m48 <sup>b</sup>	1m49 <sup>b</sup>	1pw6 <sup>b</sup>	1py2 <sup>b</sup>	1qvn <sup>b</sup>	1z92 <sup>c</sup>
Ligand	—	—						
Ligand abbreviation	—	—	FRG	CMM	FRB	FRH	FRI	IL-2Rα
Resolution <sup>d</sup>	1.99	2.40	1.95	2.00	2.60	2.80	2.70	2.80
Experimental affinity <sup>e</sup>	—	—	-6.90 <sup>j</sup>	-7.18 <sup>k</sup>	-7.02 <sup>l</sup>	-9.60 <sup>m</sup>	-8.03 <sup>n</sup>	-10.97 <sup>o</sup>
Calculated affinity	—	—	-44.50	-47.84	-49.72	-56.32	-49.13	-77.19
(SEM) <sup>f</sup>	—	—	(±0.14)	(±0.16)	(±0.13)	(±0.16)	(±0.17)	(±0.37)
MD length <sup>g</sup>	11.21	12.47	9.11	13.79	6.97	6.17	9.15	10.00
Drift <sup>h</sup>	—	—	0.20	0.19	-0.27	-0.63	0.74	0.49
IL-2 energy	-3181.50	-3182.45	-3203.77	-3211.09	-3212.32	-3210.79	-3188.06	-3190.96
(SEM) <sup>i</sup>	(± 1.06)	(± 1.02)	(± 1.25)	(± 1.19)	(± 1.36)	(± 1.47)	(± 1.36)	(± 1.33)

<sup>a</sup> Unbound IL-2. <sup>b</sup> IL-2 in complex with a small molecule. <sup>c</sup> IL-2 in complex with the extracellular IL-2 binding domain of IL-2Rα. <sup>d</sup> In Ångströms. <sup>e</sup> In kilocalories per mole. <sup>f</sup> Binding effective energy calculated by the MM-PBSA single trajectory method, in kilocalories per mole. Standard error of the mean (SEM) in parentheses. <sup>g</sup> In nanoseconds. <sup>h</sup> Average drift of the binding effective energy in the MD ensemble calculated by linear regression in kilocalories per mole per nanosecond. <sup>i</sup> Calculated average absolute effective energy of IL-2 in the unbound state or the bound state as extracted from the complex trajectories, in kilocalories per mole. Standard error of the mean (SEM) in parentheses. <sup>j</sup> Derived by van't Hoff analysis of surface plasmon resonance data. <sup>k</sup> Derived from the average of  $K_d = 7.5 \mu\text{M}$  and  $K_d = 4.2 \mu\text{M}$  as determined by equilibrium analytical ultracentrifugation. <sup>l</sup> Calculated from  $K_d = 7 \mu\text{M}$  as determined by surface plasmon resonance. <sup>m</sup> Calculated from  $K_d = 100 \text{ nM}$  as determined by surface plasmon resonance. <sup>n</sup> Calculated from  $K_d = 1.4 \mu\text{M}$ , which was calculated by  $K_{dA} = K_{dB}$  ( $\text{IC}_{50A}/\text{IC}_{50B}$ ) from the reported  $K_d = 7 \mu\text{M}$  of FRB and the reported  $\text{IC}_{50}$  values of FRB ( $\text{IC}_{50} = 10 \mu\text{M}$ )<sup>6</sup> and FRI ( $\text{IC}_{50} = 2 \mu\text{M}$ )<sup>15</sup> the latter two being measured by ELISA under identical conditions. <sup>o</sup> As denoted by Thanos et al.<sup>15</sup> and in agreement with the calculation from  $K_d = 10 \text{ nM}$  as denoted by Rickert et al.<sup>66</sup>

unbound IL-2 and IL-2 bound to either IL-2R $\alpha$  or five PPIMs were generated by MD simulations of at least 6 ns in length (Table 1).

For all systems, the RMSD of heavy atoms with respect to structures at the end of the equilibration procedure (see Supporting Information) was determined (Table S1, Supporting Information). Over all trajectories, IL-2 showed mean RMSD values of 2.55–3.46 Å. These values are in good agreement with those generally found during other MD simulations.<sup>65</sup> The interface regions of IL-2 showed generally lower RMSD values of 1.94–2.88 Å, with the interface region of the two unbound IL-2 structures showing the largest structural deviations (2.44 and 2.88 Å). This agrees well with observations from X-ray crystallography that show an opening of transient pockets in this region (see below). The small-molecule ligands bound to IL-2 showed RMSD values between 0.89 and 2.20 Å and, thus, stayed close to the initial binding region. Overall, after an initial rise during the first 2–4 ns, the RMSD values remain constant for the remainder of the MD simulations (Table S1).

To investigate the mobility of IL-2 and its interface region, we calculated root-mean-square fluctuations (RMSFs) of heavy atoms of protein residues. Starting structures of MD simulations with RMSF values mapped in a color-coded fashion are shown in Figure S2 (Supporting Information). Unsurprisingly, the largest fluctuations up to 9.58 Å were found for flexible loop regions and the termini. Many of these mobile regions have not been resolved in several of the crystallographic structures,<sup>5,66</sup> which already provides a hint as to their mobility. In contrast, all of the interface residues of IL-2 show RMSF values <2.50 Å. Interestingly, the mobility of Phe42 of IL-2, whose conformational transition is crucial for the opening of a transient pocket (see below), was found to be significantly higher in the unbound structure (RMSF = 1.86 Å (1.60 Å) for 1m47 (1m4c)) than in the bound structures (RMSF between 0.57 and 1.15 Å).

MM-PB(GB)SA calculations, which make use of a continuum electrostatic model for evaluating (de)solvation effects, may fail if structural waters are present in or close to the binding interface.<sup>67,68</sup> To identify such water molecules, we calculated the RMSFs of all water molecules and, subsequently, investigated the residence times of waters with low RMSFs by visual inspection. First, the analysis did not reveal any long-lasting (residence time >1 ns) water molecule on the outer surface of the IL-2 interface except in the case of the IL-2/IL-2R $\alpha$  complex. However, none of these water molecules formed strong interactions with the protein for the complete simulation time. Second, in the interior of IL-2 adjacent to the binding interface, long-lasting (residence time >1 ns) water molecules were found at three distinct sites (I, in the interior of IL-2 close to Glu62; II, inside of a loop region enclosed by Tyr45, Ala108, Asp109, and Glu110; III, at the N-terminal end of helix D enclosed by Met39, Phe42, and Leu114). However, none of these waters is in direct contact with any of the IL-2 ligands. Furthermore, these waters are conserved in almost all simulations of unbound and bound IL-2 so that potential effects on MM-PB(GB)SA results should cancel. Overall, these findings lead us to expect only minor influences due to structural waters on MM-PB(GB)SA results for hot spot prediction and ligand ranking.

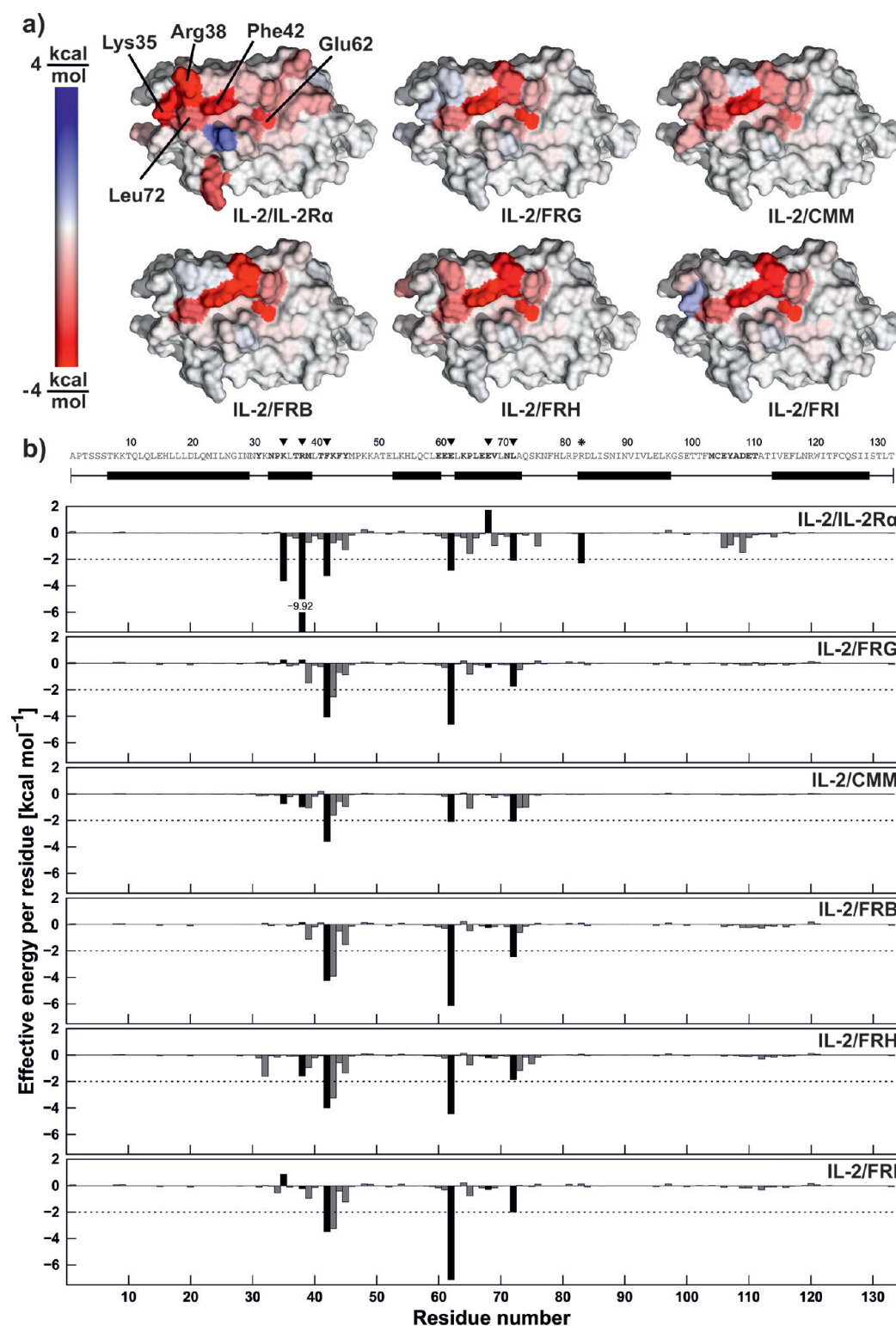
**Identification of Hot Spots by MM-PBSA Free Energy Decomposition.** Mimicking localized interactions in hot spot regions of protein–protein interfaces has proven valuable for PPIM development.<sup>4</sup> We thus set out to computationally identify

hot spots in the protein–protein interface of IL-2/IL-2R $\alpha$  and IL-2/small-molecule complexes (Figure 2). For this, we implemented and applied the MM-PBSA per residue effective energy decomposition,<sup>25,69</sup> which complements the MM-GBSA effective energy decomposition introduced by us.<sup>22</sup> Here, we applied the MM-PBSA single trajectory method. While this method neglects energetic contributions due to conformational changes of the binding partners, it leads to a drastic reduction in the statistical uncertainty of the free energy components.<sup>34,70,71</sup>

For validation, computed effective energy components were compared to experimentally determined changes in the binding free energy of IL-2/IL-2 R $\alpha$  and IL-2/FRH complexes upon mutations of IL-2 interface residues to alanine.<sup>6,15</sup> The experiments showed that for both IL-2R $\alpha$  and FRH binding was strongly disrupted ( $EC_{50,Ala}/EC_{50,WT} \geq 100$ , equivalent to  $\Delta G \geq 2.8$  kcal mol<sup>-1</sup> at 37 °C) when Phe42, Tyr45, or Glu62 were mutated to alanine. Encouragingly, the effective energy decomposition also identified Phe42 and Glu62 as hot spot residues ( $\Delta G_{eff} = -2.84$  to  $-4.45$  kcal mol<sup>-1</sup>) and only slightly underestimated the contribution of Tyr45 ( $\Delta G_{eff} \approx -1.3$  kcal mol<sup>-1</sup>). Thus, experimental and computational predictions of hot spots are in good agreement. Moderate changes in the binding affinity ( $EC_{50,Ala}/EC_{50,WT} \geq 10$ , equivalent to  $\Delta G \geq 1.4$  kcal mol<sup>-1</sup> at 37 °C) were observed for IL-2/IL-2R $\alpha$  when IL-2 residues Thr41, Lys43, or Phe44 were mutated to alanine.<sup>15</sup> Computed  $\Delta G_{eff}$  values are in the range of  $-0.4$  to  $-0.8$  kcal mol<sup>-1</sup> in these cases, demonstrating that smaller effects on the binding affinity could be well identified by the MM-PBSA effective energy decomposition, too. Finally, residues Lys35, Arg38, and Leu72 were also identified as hot spots by the MM-PBSA effective energy decomposition. However, except for Leu72, which moderately disrupted the IL-2/FRH complex ( $EC_{50,Ala}/EC_{50,WT} \geq 10$ ) when mutated, all others did only show a weak disrupting effect on IL-2R $\alpha$  and FRH binding ( $EC_{50,Ala}/EC_{50,WT} = 3-5$ ).

The seemingly prominent interactions of Arg83 ( $\Delta G_{eff} \approx -2.3$  kcal mol<sup>-1</sup>) are dominated by an intermittent salt bridge to Asp56 (IL-2R $\alpha$  residues are highlighted in italics, whereas IL-2 residues are depicted in “normal” font), for which the proteins needed to undergo structural changes during the MD simulation. As Arg83 is far apart from the localized cluster formed by the other hot spots, it was neglected for the guidance of the subsequent docking. Finally, our calculations predict Glu68 to contribute unfavorably ( $\Delta G_{eff} = +1.72$  kcal mol<sup>-1</sup>) to the binding of IL-2R $\alpha$ ; Glu68 thus is a “cold spot”. In summary, the identified hot spots cluster together and form a functional epitope localized on helices A' and B' with an approximate area of 500 Å<sup>2</sup> corresponding to 20% of the total protein–protein interface area.

**Mimicry of Localized Interactions in Hot Spot Regions by PPIMs.** Next, we investigated to what extent the PPIMs mimic IL-2R $\alpha$  as an interaction partner. If such mimicry existed, hot spots identified in a protein–protein complex could be used for guiding PPIM identification and development. Figure 2 reveals that the energetic fingerprint of IL-2/IL-2R $\alpha$  is indeed highly similar to the energetic fingerprints of the IL-2/small-molecule complexes. Three amino acids stand out in that respect: I. Phe42 is the center of a hydrophobic core forming contacts to other IL-2 residues (Met39, Val69, Leu72) as well as Met25, Asn27, Leu42, and His120, which explains its hot spot character in the IL-2/IL-2R $\alpha$  case. Moreover, Phe42 forms favorable interactions with the piperidine (FRI, FRH), pyrazole (FRB), or central phenyl



**Figure 2.** Per-residue contributions to the binding effective energy as calculated by MM-PBSA decomposition. The per-residue contributions were calculated by applying the single trajectory MM-PBSA method to the MD trajectories of IL-2 in complex with IL-2R $\alpha$  (PDB code: 1z92) and the five PPIMs FRG, CMM, FRB, FRH, and FRI (PDB codes: 1m48, 1m49, 1pw6, 1py2, and 1qvn). (a) The per-residue contribution is mapped onto the crystal structure of IL-2 bound to IL-2R $\alpha$  using a color-code with a linear scale for IL-2/IL-2R $\alpha$ , IL-2/FRG, IL-2/CMM, IL-2/FRB, IL-2/FRH, and IL-2/FRI. In (b) per-residue contributions of the six complexes are depicted as bar plots. At the top, the IL-2 sequence is depicted in single letter code with all interface residues highlighted (bold).  $\alpha$ -helices are marked as horizontal boxes in the line below. Hot spots (highlighted as black bars in the energy plot and marked by black triangles ( $\blacktriangledown$ ) above the sequence) were selected based on the per-residue decomposition of the IL-2/IL-2R $\alpha$  trajectory by applying an energy cutoff of 2 kcal mol $^{-1}$  (dashed line). See text for details on Arg83, which is marked by an asterisk (\*) above the sequence.

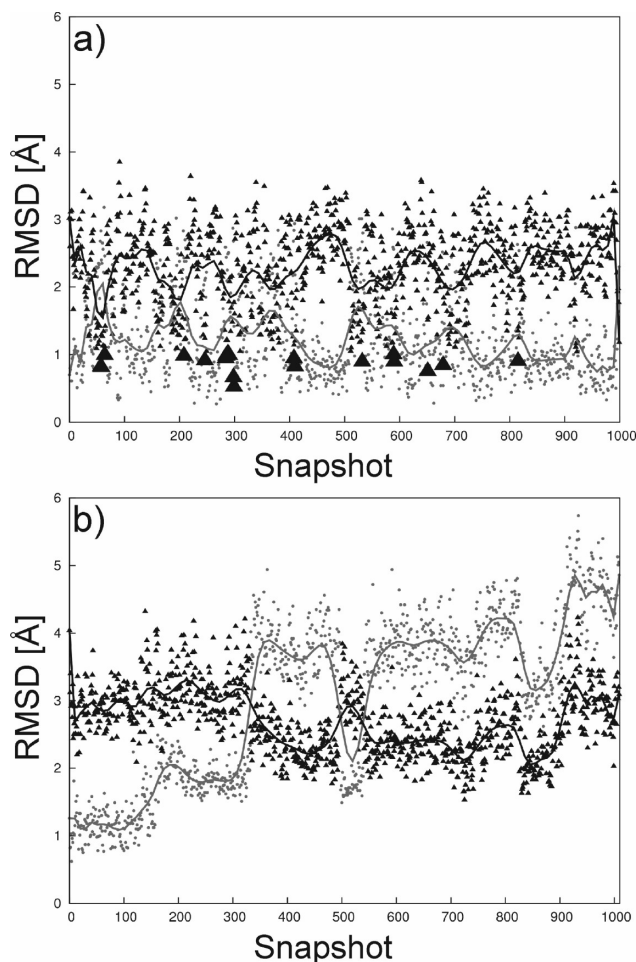


ring of the tolane moiety (FRG, CMM) of the PPIMs and their adjacent amide moieties, which makes Phe42 also a hot spot for the binding of these PPIMs. II. Glu62 forms a stable salt bridge with Arg36 in the protein–protein complex. Glu62 is also the strongest anchor for the binding of the five PPIMs by salt-bridge formation with the guanidinium groups present in all ligands. III. The hydrophobic interaction of Leu72 with Met25 and Leu2 is replaced by the phenyl and 1,2-dichlorophenyl moieties of FRI, the 1,2-dichlorophenyl moieties of FRH or FRB, the indolyl moiety of CMM, or the terminal phenyl ring of the tolane moiety of FRG. In contrast, interactions involving Lys35 and Arg38, which are important for IL-2/IL-2R $\alpha$  affinity, are not or only weakly mimicked by the PPIM.

In summary, these findings demonstrate that three out of five computationally identified hot spots of the IL-2/IL-2R $\alpha$  complex are equally important for small-molecule binding to IL-2. Furthermore, all five computed hot spots cluster in a subregion of the structural epitope of IL-2/IL-2R $\alpha$ . Together, this strongly suggests that hot spots computed from protein–protein complexes can be used for guiding the identification and optimization of PPIM.

**Opening of a Transient Pocket during Simulations Started from the Unbound State.** PPIMs have been found to be particularly effective when they bind to well-defined clefts or grooves in the protein–protein interface.<sup>5,9,72,73</sup> Here, we investigate whether an opening of transient pockets in the rather flat protein–protein interface of IL-2 can be observed when sampling the conformational space of unbound IL-2, following the “conformational selection” model.<sup>47</sup> Two conformational sampling techniques were used. First, as a state-of-the-art method, we applied all-atom MD simulations in explicit solvent of 10 ns in length. This setup is similar to a study by Helms and Eyrisch.<sup>7</sup> Second, as a computationally cheaper alternative, we applied the geometrical simulation method FRODA.<sup>49</sup> FRODA has already been successfully applied for identifying spontaneous and relevant *apo*-to-*holo* conformational transitions of HIV-1 TAR RNA.<sup>74</sup> FRODA relies upon a decomposition of a biomacromolecule into rigid and flexible regions.<sup>75</sup> In the unbound structure of IL-2, helices A, A', B, B', C, and D form discrete rigid clusters that are interconnected by flexible hinges (Figure S3, Supporting Information). With respect to the interface region, only 25% of the atoms belong to rigid clusters. This ensures that the majority of the interface atoms can move freely during the FRODA simulation.

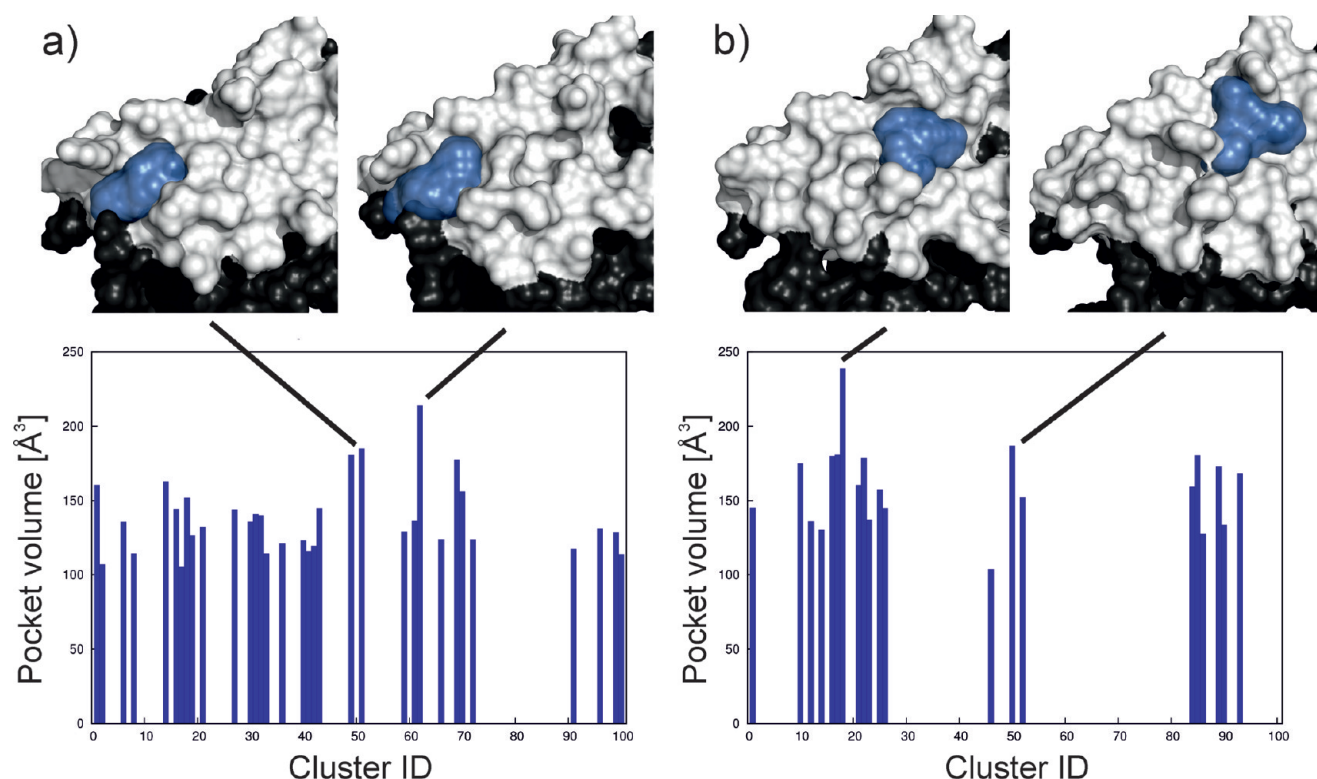
The interface heavy atom RMSD of all snapshots of MD and FRODA simulations were calculated with respect to the PPIM-bound and unbound IL-2 structures (Table S2, Supporting Information). The overall structural changes between bound and unbound IL-2 structures are small: in the best case, an interface conformation coming as close as 1.42 Å RMSD to a bound state was found, starting from a structural deviation between bound and unbound structures of 1.50 Å. This can be explained in that structural deviations between bound and unbound structures are uniformly distributed over the interface rather than caused by a large-scale collective movement. Interestingly, when comparing the performance of FRODA and MD simulations, FRODA snapshots were generally found to be more similar to four out of five bound IL-2 structures than were MD snapshots. We attribute this fact to an appropriate coarse-graining of the unbound IL-2 structure prior to the FRODA simulation. Apparently, residues were correctly identified to be part of a rigid cluster that is not involved in pocket opening, leading to a focusing of movements to that region where a pocket opens up (Figure S4,



**Figure 3.** Conformational analysis of the interface residue Phe42. Consecutive snapshots generated by (a) FRODA and (b) MD simulation of unbound IL-2 are compared to the bound (black, PDB code: 1m48) and unbound IL-2 structures (gray, PDB code: 1m47) regarding the RMSD of Phe42 based on a structural alignment of heavy atoms of the interface. For the FRODA simulation, the RMSD shows a clear antidromic character that indicates the flipping of the aromatic ring. In contrast, for the MD simulation, the antidromic character of the RMSD curve is much less pronounced. Phe42 conformations that come closer to the bound IL-2 state than 1 Å RMSD are marked by black triangles.

Supporting Information). In contrast, in the MD simulations, all residues are allowed to move freely, leading to larger overall structural deviations in the protein–protein interface that do not necessarily lead to a structure close to a bound conformation.

Of the residues that line the PPIM binding pocket, Phe42 has been described as functioning as a gate keeper by flipping its phenyl ring.<sup>5,15</sup> In Figure 3, RMSD time series calculated from all heavy atoms of Phe42 with respect to both unbound and bound IL-2 structures are shown for FRODA and MD simulations. The FRODA simulation shows an antidromic behavior of both time series, and the simulated conformations repeatedly approach the bound structure and depart from it. This leads to Phe42 approaching the small molecule bound conformation to <1.0 Å RMSD in 18 cases during the simulation. In contrast, the RMSD time series of the MD simulation showed a much less pronounced antidromic behavior, and a conformation of Phe42 with an RMSD <1.0 Å with respect to bound IL-2 was never detected during the simulation.



**Figure 4.** Detection of interface pockets in the cluster representatives of IL-2 structures generated by (a) FRODA and (b) MD simulation. The box plots depict pocket volumes computed by PocketAnalyzer. In addition, the two largest pockets found in IL-2 structures generated by either simulation method are shown. Notably, the locations of these pockets differ for both methods.

**Identification of Transient Pockets in Structural Ensembles.** Next, the question needs to be addressed as to how one can identify binding-competent conformations from the generated ensembles without already knowing the bound conformation from experimental results. An evaluation based on energetic criteria appears to be difficult<sup>34,76</sup> (see the Supporting Information on “Statistical significance of MM-PBSA results” and Figures S5 and S8) because the expected change in free energy accompanying the opening of a transient pocket is on the order of  $k_B T$ , which is several orders of magnitude smaller than the total conformational energy of the protein.<sup>77</sup> Instead, we resorted to a sequential scheme that involved checking the stereochemical quality of the simulated conformations, clustering of similar conformations, and identification of transient interface pockets based on volume and the degree of “buriedness”. Thus, we only considered geometrical parameters for the identification of transient pockets.

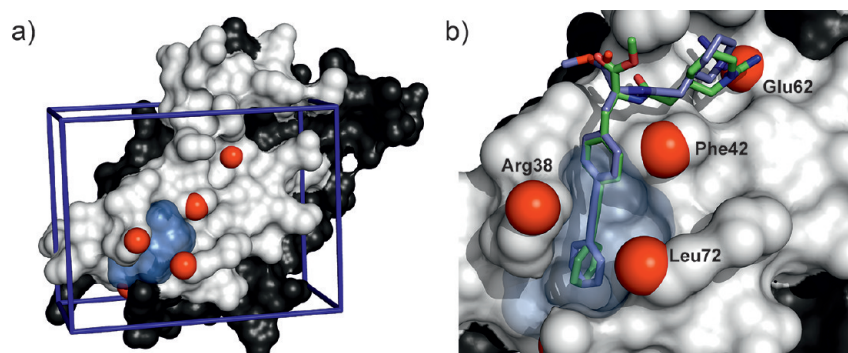
Conformations sampled by either MD or FRODA simulations generally showed a high degree of stereochemical quality, as determined by PROCHECK.<sup>56</sup> In fact, none of the FRODA conformations had to be excluded from further investigations, whereas only 45 MD conformations were discarded because they had more than two unfavorable main chain parameters. In the next step, 100 structurally varying interface conformations were selected as representatives from each simulation by *k*-medoids clustering. The interface RMSD of the selected representatives ranges from 0.85 to 3.43 Å for MD-generated conformations and from 0.78 to 2.14 Å for FRODA-generated conformations. Finally, the PocketAnalyzer program<sup>88</sup> was applied for pocket detection. Pockets embraced by at least 70% of the interface residues were identified as interface pockets. As for the crystal

structures, all small-molecule bound structures displayed interface pockets with volumes ranging from 107 to 234 Å<sup>3</sup>, with all of these pockets being located between residues Lys35, Arg38, and Phe42. No pocket was present in the unbound IL-2 structure as well as the IL-2/IL-2R $\alpha$  complex.

In 33% of the selected FRODA conformations, an interface pocket was detected (Figure 4). The average volume of these pockets is 138 Å<sup>3</sup>, with minimal (maximal) values of 104 Å<sup>3</sup> (215 Å<sup>3</sup>). Similar to the crystal structures, all interface pockets from FRODA-generated conformations were located between residues Lys35, Arg38, and Phe42. As for MD-generated conformations, an interface pocket was identified in 22% of the conformations, with an average volume of 159 Å<sup>3</sup> (min., 103 Å<sup>3</sup>; max., 240 Å<sup>3</sup>). Notably, 2/3 of these interface pockets are located between Lys43, Tyr45, and Phe42 and, thus, deviate in position from the pockets found in the bound crystal structures.

**Docking into Transient Pockets.** We then investigated whether simulated IL-2 conformations with transient pockets in the protein–protein interface can be used as receptor structures for docking. Therefore, we selected those 10 representative structures from each simulation that showed the largest interface pocket volume (Table S3, Supporting Information). Each of the five known IL-2 ligands (Table 1) was then docked into this set of conformations. To exclude any bias due to the knowledge of the experimentally determined complex structures, the placement of the potential grids for docking (Figure 5) was solely based on (I) all hot spots identified by MM-PBSA except Arg83 (see above, Figure 2) and (II) all amino acids lining the identified interface pocket (Table S4, Supporting Information). Docking was considered successful when the ligand pose with the lowest





**Figure 5.** Definition of the potential grid and exemplary docking result for IL-2. The interface region of IL-2/IL-2R $\alpha$  is colored white. (a) The hot spots identified by MM-PBSA decomposition (red spheres) and the transient pocket (blue surface) were used to define the location and size of the potential grid (see Materials and Methods). The grid dimensions are represented by a blue box. (b) Predicted binding pose of the ligand FRG (blue sticks) docked into a FRODA snapshot containing an identified transient pocket. The RMSD between the predicted and crystallographic binding pose (PDB code: 1m48, green sticks) is 1.28 Å.

**Table 2.** Results of Redocking and *apo*-Docking

redocking					<i>apo</i> -docking				
ligand	PDB code <sup>a</sup>	RMSD <sup>b</sup>	score <sup>c</sup>	cluster size	ligand	PDB code <sup>a</sup>	RMSD <sup>b</sup>	score <sup>c</sup>	cluster size
FRG	1m48	1.50	−13.84	84	FRG	1m47	2.58	−10.93	19
CMM	1m48	1.37	−15.69	82	CMM	1m47	3.08	−12.53	40
FRB	1pw6	0.59	−13.01	25	FRB	1m47	4.63	−10.82	51
FRH	1py2	0.87	−15.38	53	FRH	1m47	3.80	−14.10	9
FRI	1qvn	1.14	−14.65	27	FRI	1m47	8.10	−13.30	9

<sup>a</sup> PDB code of the IL-2 complex structure. <sup>b</sup> RMSD of the ligand pose with the lowest energy in the largest cluster with respect to the native structure, in Ångströms. <sup>c</sup> In kilocalories per mole.

intermolecular docking energy in the largest cluster had an RMSD < 2.0 Å to the native pose.

For comparison, we first redocked the five IL-2 ligands into the corresponding IL-2 complex structure. Likewise, we also performed docking of these ligands to an *apo* structure of IL-2 (PDB code: 1m47). Since no interface pocket could be detected in this *apo* structure, the same potential grid definition as for the redocking approach was used for the *apo*-docking. The redocking was successful in all cases, whereas *apo*-docking failed (Table 2). The latter is not unexpected due to the absence of any pronounced indentation in the protein–protein interface. Additionally, the success and convergence of the redocking is demonstrated by the occurrence of  $\geq 25$  poses in the largest cluster for IL-2/FRB, IL-2/FRI, and IL-2/FRH and >80 poses in the largest cluster for IL-2/FRG and IL-2/CMM.

The results of docking into transient pockets of simulated conformations are summarized in Table 3. Notably, docking into MD-generated conformations was not successful because in no case was a docked pose with RMSD < 2 Å obtained, and in only 2 out of 50 dockings was a pose with RMSD < 2.5 Å identified. This failure is a result of the interface pockets being located at a different position than the pockets found in the bound crystal structures. A more detailed inspection of the MD-generated conformations showed that correctly localized transient pockets do exist (data not shown). However, these pockets are much less pronounced than those occurring in FRODA-generated snapshots and, hence, are not among the 10 largest pockets chosen for the docking experiments. This is because a hydrophobic channel

**Table 3.** Number of Successful Attempts of Docking into Snapshots with Identified Transient Pockets

ligand	RMSD <sup>a</sup>							
	MD				FRODA			
	<2.0	<2.5	<3.5	$\geq 3.5$	<2.0	<2.5	<3.5	$\geq 3.5$
FRG	0	2	1	7	2	6	1	1
CMM	0	0	4	6	2	4	4	0
FRB	0	0	3	7	0	0	7	2
FRH	0	0	0	10	1	2	4	3
FRI	0	0	0	10	1	1	0	8

<sup>a</sup> RMSD of the ligand pose with the lowest energy in the largest cluster with respect to the native structure, in Ångströms.

embraced by Lys35, Arg38, and Phe42 must open for the transient pockets to be correctly localized.<sup>5</sup> Such an opening tends to be less pronounced in MD-generated ensembles, as demonstrated by the lower hydrophobicity of these pockets when compared to pockets found in crystal structures.<sup>7</sup>

In contrast, we were able to identify the hydrophobic channel in all of the selected FRODA-generated snapshots, possibly due to the absence of solvent in the simulation process. As a consequence, docking to at least one transient pocket was successful for all IL-2 ligands but FRB (Table 3). This is exemplarily shown for the ligand FRG in Figure 5b.

These results are encouraging in that a drop in docking accuracy compared to redocking is often found to be mirrored by the degree to which a protein moves upon ligand binding.<sup>78,79</sup> Thus, docking to an *apo* form usually shows the largest deterioration.<sup>80</sup> Being able to start from the *apo* IL-2 structure and identify transient pockets in trajectories of computationally inexpensive FRODA simulations that are adequate for ligand docking thus is a valuable achievement.

**Docking Enrichment in a Large Set of Decoys.** In order to demonstrate that the identified hot spots and transient pockets

**Table 4. Docking Enrichment of Known IL-2 Ligands<sup>a</sup>**

FRODA snapshot <sup>b</sup>	EF <sub>max</sub> <sup>c,d</sup>	EF <sub>20</sub> <sup>c,d</sup>	EF <sub>3</sub> <sup>c,d</sup>	EF <sub>1</sub> <sup>c,d</sup>	AUC <sup>d,e</sup>
6	23.7 (260.2)	4.6 (5.0)	17.4 (26.0)	22.0 (74.3)	0.95 (0.99)
117	23.8 (173.6)	4.4 (5.0)	18.0 (32.5)	22.1 (74.4)	0.95 (0.99)
169	23.7 (260.0)	4.8 (5.0)	22.0 (32.5)	23.7 (92.9)	0.98 (1.00)
301	23.8 (260.4)	4.9 (5.0)	19.7 (26.0)	18.7 (55.8)	0.99 (0.99)
418	23.8 (260.4)	4.8 (5.0)	10.4 (26.0)	11.9 (37.2)	0.94 (0.98)
514	19.0 (97.6)	5.0 (5.0)	16.2 (32.5)	17.0 (55.8)	0.98 (0.99)
534	23.8 (86.8)	4.4 (5.0)	11.0 (26.0)	13.6 (37.2)	0.92 (0.98)
657	16.2 (57.8)	4.0 (4.0)	10.4 (19.5)	15.3 (55.8)	0.89 (0.93)
698	23.8 (260.4)	4.1 (5.0)	16.2 (32.5)	23.8 (55.8)	0.94 (0.99)
729	23.8 (86.8)	4.6 (5.0)	12.2 (26.0)	13.6 (55.8)	0.95 (0.99)

<sup>a</sup> The set of known IL-2 ligands consists of five IL-2 ligands with available complex crystal structures as well as 52 structures with similar scaffolds and known IC<sub>50</sub>.<sup>6,17,48</sup> The set of decoys consists of 996 unique structures with a total of 1297 protonation and tautomerization states.

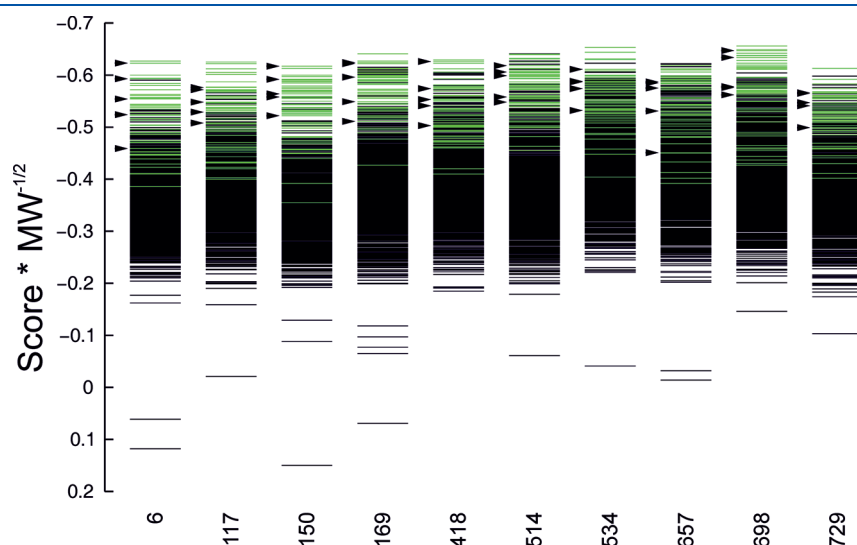
<sup>b</sup> Consecutive number of the snapshot from a total of 1000 snapshots uniformly extracted from the 10 000 000 FRODA-generated structures starting from the unbound IL-2 structure (PDB code: 1m47). The 10 snapshots with the largest pocket volumes were used. <sup>c</sup> EF<sub>1</sub>, EF<sub>3</sub>, EF<sub>20</sub>, and EF<sub>max</sub> correspond to the enrichment factors at 1%, 3%, and 20% of the ranked database and the maximal enrichment factors over the whole data set. <sup>d</sup> Values correspond to all 57 known IL-2 ligands, while values in parentheses correspond to the five IL-2 ligands with available complex crystal structures only. <sup>e</sup> Area under the receiver operator curve (ROC).

could also be used for structure-based virtual screening (VS), we performed a retrospective VS for IL-2 PPIMs. As known binders, the five IL-2 ligands in complex structures (Table 1) and 52 structures with similar scaffolds and known IC<sub>50</sub> values were used.<sup>6,17,48</sup> Decoys were selected from the “purchasable subset” of the ZINC database<sup>61</sup> following the procedure described for creating the directory of useful decoys (DUD).<sup>60</sup> The DUD procedure aims at selecting decoy structures that are physicochemically similar to known binders in order to avoid any bias in enrichment calculations. This led to 996 unique decoy structures with a total of 1297 protonation and tautomerization states. We note that during docking, the docking scores were normalized by the square root of the molecular weight of a ligand in order to correct for any size related bias, too.<sup>81</sup>

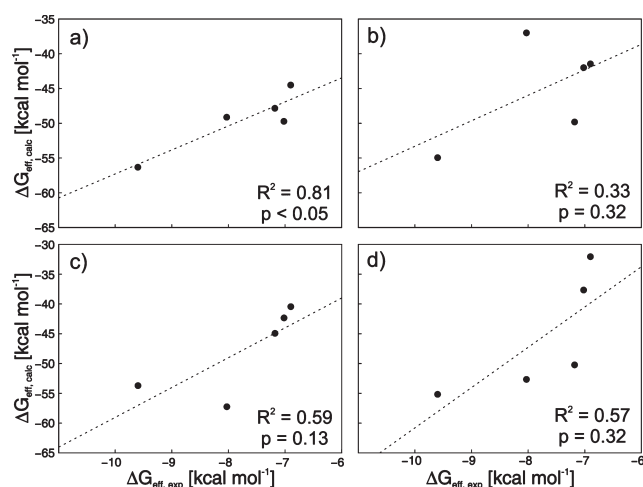
For the 57 (5) known IL-2 ligands (in complex structures) we found good enrichments for the individual transient binding pockets (Table 4, Figure 6, Figure S6, Supporting Information) with EF<sub>max</sub> = 16.2–23.8 (EF<sub>max</sub> = 57.8–260.4) and EF<sub>1</sub> = 13.6–23.8 (EF<sub>1</sub> = 37.2–92.2), and area under the curve values of receiver operator curves of AUC ≥ 0.89 (AUC ≥ 0.93).

We note that these enrichments may be too optimistic compared to a real-life scenario and, hence, should be interpreted cautiously because the VS has likely benefitted from the fact that the known IL-2 ligands were structurally optimized for binding to IL-2 and partially violate Lipinski’s rules.<sup>82</sup> Hence, even though the decoys were selected following the DUD procedure, in some cases, a perfect match of the property distribution curves between binders and decoys could not be achieved (Figure S7, Supporting Information). This is particularly true for the properties “molecular weight”, “no. of hydrogen bond donors”, and “no. of amidino and guanidino groups”. Still, with respect to the aim of this study, our results demonstrate that known IL-2 ligands could successfully be screened from a set of decoys using only information about hot spots and transient pockets on the protein side.

**Ranking of IL-2 Ligands.** Binding effective energies calculated by the MM-PBSA single trajectory method appear to be converged and remain stable throughout simulation lengths of 6–14 ns (Table S5 and Figure S8, Supporting Information).



**Figure 6.** Ranking of docked structures. The best poses (as defined in Materials and Methods) of docked binders and decoy structures were ranked by intermolecular energy divided by the square root of the molecular weight. The 10 FRODA structures with the largest transient pocket are indicated by their snapshot number at the bottom. The 57 known IL-2 ligands and the decoys are depicted with green and black lines, respectively. In addition, the five IL-2 ligands with available complex crystal structures are highlighted by arrows.



**Figure 7.** Correlation of computed binding effective energies ( $\Delta G_{\text{eff,calc}}$ ) with respect to experimental free enthalpies of binding ( $\Delta G_{\text{eff,exp}}$ ). Binding effective energies were calculated using the Poisson–Boltzmann continuum solvation model (a and c) and the generalized Born continuum solvation model (b and d) within the MM-PB(GB)SA single trajectory method. MM-PB(GB)SA calculations were based on either crystallographic (a and b) or docked (c and d) starting structures.

Absolute binding effective energies computed for the five IL-2 ligands starting from *crystal structures* of the complexes are about 45 kcal mol<sup>-1</sup> more negative than the experimentally determined free enthalpies of binding (Figure 7, Table 1). Two reasons account for this. First, unfavorable energetic contributions due to conformational strain of the binding partners are not taken into account in the single trajectory method. These contributions can be as high as 36 kcal mol<sup>-1</sup>.<sup>83</sup> Second, we neglect any changes in the configurational entropy of the binding partners, which accounted for contributions to the free energy of 20–30 kcal mol<sup>-1</sup> at 300 K in related studies.<sup>22,84</sup> However, these two points do not have a major impact on relative binding effective energies as demonstrated by a good ( $R^2 = 0.81$ ) and significant ( $p < 0.05$ ) linear correlation of computed and experimentally determined binding energies, allowing for a successful ranking of four out of the five ligands (Figure 7a). This result is all the more remarkable in that the range of experimentally determined binding affinities is only 3 kcal mol<sup>-1</sup>. To our knowledge, this is one of only a few reports so far of successfully applying MM-PBSA to rank PPIMs.<sup>69,85</sup> When using the GB model, no significant correlation is obtained. In particular, the binding effective energies of ligands FRI and CMM show strong deviations from the correlation line (Figure 7b). A per residue decomposition of the binding effective energies using the GB model did not allow one to assign the origin of these deviations to contributions by a particular set of residues (data not shown).

For ranking *docked* IL-2/small-molecule complex structures (Figure 7), first, reasonable poses obtained by docking into FRODA-generated structures with transient pockets were selected without making use of any knowledge of the bound crystal structures (Table S6, Supporting Information). Then, MD simulations and MM-PB(GB)SA calculations were applied as in the case of the crystal structures. The linear correlation of computed MM-PBSA binding effective energies with respect to experimentally determined binding free energies is fair ( $R^2 = 0.59$ ) and weakly significant ( $p = 0.13$ ; Figure 7c). Again, no significant correlation was obtained in the case of MM-GBSA. The largest deviations from the correlation line are observed for ligands

FRH and FRI in the case of MM-PBSA, which also show the largest structural deviations from the native pose in the starting structures (Table S6). These results demonstrate that for MM-PBSA calculation to be successful in ranking PPIMs, (at least) good starting structures (RMSD < 2.5 Å) are required. Nevertheless, it is encouraging to note that the quality of the generated docking poses was still sufficient to successfully discriminate between the subgroups of high and low affinity ligands.

## CONCLUSION

We have presented for the first time a computational strategy that simultaneously considers aspects of energetics and plasticity in the context of PPIM binding to a protein interface. In particular, our strategy aims at identifying the determinants of small-molecule binding, hot spots and transient pockets, in a protein–protein interface in order to make use of this knowledge for predicting binding modes of and ranking PPIMs with respect to their affinity. Although performed in a retrospective manner on the well-investigated system of IL-2, we note that at no point in the study did we utilize information from the experiment about the binding mode and affinity of PPIMs. Thus, our strategy will be applicable also in a prospective manner where nothing other than a protein–protein complex structure is known; hence, it can well be the first step in a structure-based endeavor to identify PPIMs.

Perhaps the most surprising result from a methodological point of view is that the computationally much cheaper constrained geometric simulation method FRODA outperforms state-of-the-art MD simulations in *sampling* transient pockets in the IL-2 interface. Apparently, the neglect of solvent in FRODA not only leads to a reduced computational burden but also facilitates the opening of a hydrophobic channel. Although applied to only one protein–protein interface in the present study, we note that the good performance of FRODA in sampling relevant conformational transitions is in line with results obtained by this<sup>74</sup> and a related method<sup>86</sup> on other systems.

It is encouraging that geometrical parameters summarized in the PPIAnalyzer method sufficed to successfully *identify* transient pockets. On the one hand, this finding alleviates the need to resort to conformational free/effective energies for identifying such pockets. Using energetic criteria is hampered by the demand for very precise computations due to the fact that small differences in conformational energies must be calculated from large absolute values. On the other hand, this finding reconfirms the FRODA results, as it demonstrates that the most pronounced pockets only opened up where expected. As FRODA strongly depends on a preceding flexibility analysis of the protein, it is thus tempting to speculate that regions that are prone to open transient pockets could be identified by such a flexibility analysis. This knowledge could then be used to focus other locally enhanced sampling schemes on that particular region.<sup>87</sup>

Finally, we consider it a valuable achievement that the sequence of, first, docking to identified transient pockets; second, starting structure selection based on hot spot information, RMSD clustering, and intermolecular docking energies; and third, MM-PBSA calculations allowed one to discriminate between subgroups of IL-2 PPIMs with low and high affinity. Also, we obtained good enrichments for the individual transient binding pockets in a docking-based, retrospective virtual screening for IL-2 PPIMs. Together with the fact that the known PPIMs of IL-2 were identified to mimic many of the interactions also found in the



IL-2/IL-2R $\alpha$  region, this suggests that current computational methods can assist the knowledge-driven process of PPIM identification when starting from a given protein–protein complex.

## ■ ASSOCIATED CONTENT

**S Supporting Information.** Tables with the heavy atom RMSD of IL-2 and its complexes during MD simulations (Table S1), heavy atom RMSD of the IL-2 interface region of experimentally determined bound IL-2 conformations (Table S2), the largest pocket volumes for selected FRODA and MD snapshots (Table S3), selected pocket residues of IL-2 for the definition of the potential energy grids (Table S4), the drift of the effective energies during MD simulations (Table S5), and the selection of the best docking poses in the FRODA snapshots with the largest pocket volume (Table S6) as well as graphical representations of the workflow of the PPIAnalyzer method (Figure S1); the RMSF values of IL-2 during MD simulations of the unbound and bound states (Figure S2); the rigid cluster decomposition of unbound IL-2 obtained by FIRST (Figure S3); the overlay of the protein–protein interface region of IL-2 in unbound, bound, and FRODA sampled conformation (Figure S4); the calculated mean absolute effective energies  $G_{\text{eff}}$  of IL-2 in its unbound and bound conformations (Figure S5); the enrichment of known IL-2 ligands (Figure S6); the property distributions of known IL-2 ligands and decoys (Figure S7); the time series of the effective energies (Figure S8); and the multiple sequence alignment of sequences of IL-2 crystal structures (Figure S9). This information is available free of charge via the Internet at <http://pubs.acs.org/>.

## ■ AUTHOR INFORMATION

### Corresponding Author

\*Phone: (+49) 211 81-13662. Fax: (+49) 211 81-13847. E-mail: [gohlke@uni-duesseldorf.de](mailto:gohlke@uni-duesseldorf.de).

### Author Contributions

<sup>S</sup>Both authors contributed equally to this work.

## ■ ACKNOWLEDGMENT

We are grateful to Teresa Jiménez Vaquero for providing an initial version of the grid-based pocket identification algorithm and to Dr. Simone Fulle for critically reading the manuscript. This work was supported by Sanofi-Aventis Deutschland GmbH (LGCR Drug Design). We are grateful to the “Zentrum fuer Informations- und Medientechnologie” (ZIM) at the Heinrich-Heine-University, Düsseldorf for computational support. We are grateful to OpenEye Scientific Software for granting a no-cost academic license to us. Figures were generated by gnuplot and PyMOL.

## ■ ABBREVIATIONS:

PPI, protein–protein interaction; PPIM, small-molecule protein–protein interaction modulator; RMSD, root mean-square deviation; RMCD, root mean-cubed difference; MD, molecular dynamics; IL-2, interleukin-2; IL-2R $\alpha$ ,  $\alpha$ -subunit of the interleukin-2 receptor; GB, generalized Born; PB, Poisson–Boltzmann; MM-PBSA, molecular mechanics Poisson–Boltzmann surface area; MM-GBSA, molecular mechanics generalized Born; DUD, directory of useful

decoys; ROC, receiver operator curve; AUC, area under the curve; EF<sub>X</sub>, enrichment factor at X% of the data set.

## ■ REFERENCES

- (1) Wells, J. A.; McClendon, C. L. Reaching for high-hanging fruit in drug discovery at protein-protein interfaces. *Nature* **2007**, *450* (7172), 1001–1009.
- (2) Jones, S.; Thornton, J. M. Principles of protein-protein interactions. *Proc. Natl. Acad. Sci. U. S. A.* **1996**, *93* (1), 13–20.
- (3) Lo Conte, L.; Chothia, C.; Janin, J. The atomic structure of protein-protein recognition sites. *J. Mol. Biol.* **1999**, *285* (5), 2177–2198.
- (4) Clackson, T.; Wells, J. A. A Hot-Spot of Binding-Energy in a Hormone-Receptor Interface. *Science* **1995**, *267* (5196), 383–386.
- (5) Arkin, M. R.; Randal, M.; DeLano, W. L.; Hyde, J.; Luong, T. N.; Oslob, J. D.; Raphael, D. R.; Taylor, L.; Wang, J.; McDowell, R. S.; Wells, J. A.; Braisted, A. C. Binding of small molecules to an adaptive protein-protein interface. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100* (4), 1603–1608.
- (6) Raimundo, B. C.; Oslob, J. D.; Braisted, A. C.; Hyde, J.; McDowell, R. S.; Randal, M.; Waal, N. D.; Wilkinson, J.; Yu, C. H.; Arkin, M. R. Integrating fragment assembly and biophysical methods in the chemical advancement of small-molecule antagonists of IL-2: An approach for inhibiting protein-protein interactions. *J. Med. Chem.* **2004**, *47* (12), 3111–3130.
- (7) Eyrisch, S.; Helms, V. What induces pocket openings on protein surface patches involved in protein-protein interactions? *J. Comput.-Aided Mol. Des.* **2009**, *23* (2), 73–86.
- (8) Oltersdorf, T.; Elmore, S. W.; Shoemaker, A. R.; Armstrong, R. C.; Augeri, D. J.; Belli, B. A.; Bruncko, M.; Deckwerth, T. L.; Dinges, J.; Hajduk, P. J.; Joseph, M. K.; Kitada, S.; Korsmeyer, S. J.; Kunzer, A. R.; Letai, A.; Li, C.; Mitten, M. J.; Nettesheim, D. G.; Ng, S.; Nimmer, P. M.; O'Connor, J. M.; Oleksijew, A.; Petros, A. M.; Reed, J. C.; Shen, W.; Tahir, S. K.; Thompson, C. B.; Tomaselli, K. J.; Wang, B.; Wendt, M. D.; Zhang, H.; Fesik, S. W.; Rosenberg, S. H. An inhibitor of Bcl-2 family proteins induces regression of solid tumours. *Nature* **2005**, *435* (7042), 677–681.
- (9) Keskin, Z.; Gursoy, A.; Ma, B.; Nussinov, R. Principles of protein-protein interactions: What are the preferred ways for proteins to interact? *Chem. Rev.* **2008**, *108* (4), 1225–1244.
- (10) Moreira, I. S.; Fernandes, P. A.; Ramos, M. J. Hot spots-A review of the protein-protein interface determinant amino-acid residues. *Proteins* **2007**, *68* (4), 803–812.
- (11) Ozbabacan, S. E.; Gursoy, A.; Keskin, O.; Nussinov, R. Conformational ensembles, signal transduction and residue hot spots: application to drug discovery. *Curr. Opin. Drug Discovery Dev.* **2010**, *13* (5), 527–537.
- (12) Berg, T. Small-molecule inhibitors of protein-protein interactions. *Curr. Opin. Drug Discovery Dev.* **2008**, *11* (5), 666–674.
- (13) Gonzalez-Ruiz, D.; Gohlke, H. Targeting protein-protein interactions with small molecules: Challenges and perspectives for computational binding epitope detection and ligand finding. *Curr. Med. Chem.* **2006**, *13* (22), 2607–2625.
- (14) Zhong, S.; Macias, A. T.; MacKerell, A. D., Jr. Computational identification of inhibitors of protein-protein interactions. *Curr. Top. Med. Chem.* **2007**, *7* (1), 63–82.
- (15) Thanos, C. D.; DeLano, W. L.; Wells, J. A. Hot-spot mimicry of a cytokine receptor by a small molecule. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103* (42), 15422–15427.
- (16) Wilson, C. G.; Arkin, M. R. Small-molecule inhibitors of IL-2/IL-2R: lessons learned and applied. *Curr. Top. Microbiol. Immunol.* **2011**, *348*, 25–59.
- (17) Braisted, A. C.; Oslob, J. D.; DeLano, W. L.; Hyde, J.; McDowell, R. S.; Waal, N.; Yu, C.; Arkin, M. R.; Raimundo, B. C. Discovery of a potent small molecule IL-2 inhibitor through fragment assembly. *J. Am. Chem. Soc.* **2003**, *125* (13), 3714–3715.
- (18) Fuller, J. C.; Burgoyne, N. J.; Jackson, R. M. Predicting druggable binding sites at the protein-protein interface. *Drug Discovery Today* **2009**, *14* (3–4), 155–161.

- (19) Brown, S. P.; Hajduk, P. J. Effects of conformational dynamics on predicted protein druggability. *ChemMedChem* **2006**, *1* (1), 70–72.
- (20) Czarna, A.; Beck, B.; Srivastava, S.; Popowicz, G. M.; Wolf, S.; Huang, Y.; Bista, M.; Holak, T. A.; Domling, A. Robust generation of lead compounds for protein-protein interactions by computational and MCR chemistry: p53/Hdm2 antagonists. *Angew. Chem., Int. Ed. Engl.* **2010**, *49* (31), 5352–5356.
- (21) Corradi, V.; Mancini, M.; Manetti, F.; Petta, S.; Santucci, M. A.; Botta, M. Identification of the first non-peptidic small molecule inhibitor of the c-Abl/14–3-3 protein-protein interactions able to drive sensitive and Imatinib-resistant leukemia cells to apoptosis. *Bioorg. Med. Chem. Lett.* **2010**, *20* (20), 6133–6137.
- (22) Gohlke, H.; Kiel, C.; Case, D. A. Insights into protein-protein binding by binding free energy calculation and free energy decomposition for the Ras-Raf and Ras-RalGDS complexes. *J. Mol. Biol.* **2003**, *330* (4), 891–913.
- (23) Gohlke, H.; Kuhn, L. A.; Case, D. A. Change in protein flexibility upon complex formation: Analysis of Ras-Raf using molecular dynamics and a molecular framework approach. *Proteins* **2004**, *56* (2), 322–337.
- (24) Ababou, A.; van der Vaart, A.; Gogonea, V.; Merz, K. M., Jr. Interaction energy decomposition in protein-protein association: a quantum mechanical study of barnase-barstar complex. *Biophys. Chem.* **2007**, *125* (1), 221–236.
- (25) Lafont, V.; Schaefer, M.; Stote, R. H.; Altschuh, D.; Dejaegere, A. Protein-protein recognition and interaction hot spots in an antigen-antibody complex: free energy decomposition identifies "efficient amino acids". *Proteins* **2007**, *67* (2), 418–434.
- (26) Wichmann, C.; Becker, Y.; Chen-Wichmann, L.; Vogel, V.; Vojtkova, A.; Herglotz, J.; Moore, S.; Koch, J.; Lausen, J.; Mantele, W.; Gohlke, H.; Grez, M. Dimer-tetramer transition controls RUNX1/ETO leukemogenic activity. *Blood* **2010**, *116* (4), 603–613.
- (27) Zoete, V.; Michielin, O. Comparison between computational alanine scanning and per-residue binding free energy decomposition for protein-protein association using MM-GBSA: application to the TCR-pMHC complex. *Proteins* **2007**, *67* (4), 1026–1047.
- (28) Krueger, D. M.; Gohlke, H. DrugScorePPI webserver: fast and accurate in silico alanine scanning for scoring protein-protein interactions. *Nucleic Acids Res.* **2010**, *38* (Suppl.), W480–W486.
- (29) Cole, D. J.; Skylaris, C. K.; Rajendra, E.; Venkitaraman, A. R.; Payne, M. C. Protein-protein interactions from linear-scaling first-principles quantum-mechanical calculations. *Europhys. Lett.* **2010**, *91* (3), 37004.
- (30) Dastidar, S. G.; Madhumalar, A.; Fuentes, G.; Lane, D. P.; Verma, C. S. Forces mediating protein-protein interactions: a computational study of p53 "approaching" MDM2. *Theor. Chem. Acc.* **2010**, *125* (3–6), 621–635.
- (31) Wong, S.; Amaro, R. E.; McCammon, J. A. MM-PBSA Captures Key Role of Intercalating Water Molecules at a Protein-Protein Interface. *J. Chem. Theory Comput.* **2009**, *5* (2), 422–429.
- (32) Moreira, I. S.; Fernandes, P. A.; Ramos, M. J. Protein-protein recognition: a computational mutagenesis study of the MDM2-PS3 complex. *Theor. Chem. Acc.* **2008**, *120* (4–6), 533–542.
- (33) Cui, Q. Z.; Sulea, T.; Schrag, J. D.; Munger, C.; Hung, M. N.; Naim, M.; Cygler, M.; Purisima, E. O. Molecular dynamics-solvated interaction energy studies of protein-protein interactions: The MP1-p14 scaffolding complex. *J. Mol. Biol.* **2008**, *379* (4), 787–802.
- (34) Gohlke, H.; Case, D. A. Converging free energy estimates: MM-PB(GB)SA studies on the protein-protein complex Ras-Raf. *J. Comput. Chem.* **2004**, *25* (2), 238–250.
- (35) Joce, C.; Stahl, J. A.; Shridhar, M.; Hutchinson, M. R.; Watkins, L. R.; Fedichev, P. O.; Yin, H. Application of a novel in silico high-throughput screen to identify selective inhibitors for protein-protein interactions. *Bioorg. Med. Chem. Lett.* **2010**, *20* (18), 5411–5413.
- (36) Neumann, J.; Gottschalk, K. E. The Effect of Different Force Applications on the Protein-Protein Complex Barnase-Barstar. *Biophys. J.* **2009**, *97* (6), 1687–1699.
- (37) Andrusier, N.; Mashlach, E.; Nussinov, R.; Wolfson, H. J. Principles of flexible protein-protein docking. *Proteins* **2008**, *73* (2), 271–289.
- (38) Chaudhury, S.; Gray, J. J. Conformer selection and induced fit in flexible backbone protein-protein docking using computational and NMR ensembles. *J. Mol. Biol.* **2008**, *381* (4), 1068–1087.
- (39) Chang, C. E. A.; McLaughlin, W. A.; Baron, R.; Wang, W.; McCammon, J. A. Entropic contributions and the influence of the hydrophobic environment in promiscuous protein-protein association. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105* (21), 7456–7461.
- (40) Enyedy, I. J.; Ling, Y.; Nacro, K.; Tomita, Y.; Wu, X. H.; Cao, Y. Y.; Guo, R. B.; Li, B. H.; Zhu, X. F.; Huang, Y.; Long, Y. Q.; Roller, P. P.; Yang, D. J.; Wang, S. M. Discovery of small-molecule inhibitors of bcl-2 through structure-based computer screening. *J. Med. Chem.* **2001**, *44* (25), 4313–4324.
- (41) Lugovskoy, A. A.; Degterev, A. I.; Fahmy, A. F.; Zhou, P.; Gross, J. D.; Yuan, J. Y.; Wagner, G. A novel approach for characterizing protein ligand complexes: Molecular basis for specificity of small-molecule Bcl-2 inhibitors. *J. Am. Chem. Soc.* **2002**, *124* (7), 1234–1240.
- (42) Koehler, N. K. U.; Yang, C. Y.; Varady, J.; Lu, Y. P.; Wu, X. W.; Liu, M.; Yin, D. X.; Bartels, M.; Xu, B. Y.; Roller, P. P.; Long, Y. Q.; Li, P.; Kattah, M.; Cohn, M. L.; Moran, K.; Tilley, E.; Richert, J. R.; Wang, S. M. Structure-based discovery of nonpeptidic small organic compounds to block the T cell response to myelin basic protein. *J. Med. Chem.* **2004**, *47* (21), 4989–4997.
- (43) Nikolovska-Coleska, Z.; Xu, L.; Hu, Z. J.; Tomita, Y.; Li, P.; Roller, P. P.; Wang, R. X.; Fang, X. L.; Guo, R. B.; Zhang, M. C.; Lippman, M. E.; Yang, D. J.; Wang, S. M. Discovery of embelin as a cell-permeable, small-molecular weight inhibitor of XLAP through structure-based computational screening of a traditional herbal medicine three-dimensional structure database. *J. Med. Chem.* **2004**, *47* (10), 2430–2440.
- (44) Gao, Y.; Dickerson, J. B.; Guo, F.; Zheng, J.; Zheng, Y. Rational design and characterization of a Rac GTPase-specific small molecule inhibitor. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101* (20), 7618–7623.
- (45) Fujii, N.; Haresco, J. J.; Novak, K. A. P.; Stokoe, D.; Kuntz, I. D.; Guy, R. K. A selective irreversible inhibitor targeting a PDZ protein interaction domain. *J. Am. Chem. Soc.* **2003**, *125* (40), 12074–12075.
- (46) Debnath, A. K.; Radigan, L.; Jiang, S. B. Structure-based identification of small molecule antiviral compounds targeted to the gp41 core structure of the human immunodeficiency virus type 1. *J. Med. Chem.* **1999**, *42* (17), 3203–3209.
- (47) Boehr, D. D.; Nussinov, R.; Wright, P. E. The role of dynamic conformational ensembles in biomolecular recognition. *Nat. Chem. Biol.* **2009**, *5* (11), 789–796.
- (48) Arkin, M.; Lear, J. D. A new data analysis method to determine binding constants of small molecules to proteins using equilibrium analytical ultracentrifugation with absorption optics. *Anal. Biochem.* **2001**, *299* (1), 98–107.
- (49) Wells, S.; Menor, S.; Hespeneide, B.; Thorpe, M. F. Constrained geometric simulation of diffusive motion in proteins. *Phys. Biol.* **2005**, *2* (4), S127–S136.
- (50) Kollman, P. A.; Massova, I.; Reyes, C.; Kuhn, B.; Huo, S. H.; Chong, L.; Lee, M.; Lee, T.; Duan, Y.; Wang, W.; Donini, O.; Cieplak, P.; Srinivasan, J.; Case, D. A.; Cheatham, T. E. Calculating structures and free energies of complex molecules: Combining molecular mechanics and continuum models. *Acc. Chem. Res.* **2000**, *33* (12), 889–897.
- (51) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. Comparison of multiple amber force fields and development of improved protein backbone parameters. *Proteins* **2006**, *65* (3), 712–725.
- (52) Lu, Q.; Luo, R. A Poisson-Boltzmann dynamics method with nonperiodic boundary condition. *J. Chem. Phys.* **2003**, *119* (21), 11035–11047.
- (53) Onufriev, A.; Bashford, D.; Case, D. A. Exploring protein native states and large-scale conformational changes with a modified generalized born model. *Proteins* **2004**, *55* (2), 383–394.
- (54) Zoete, V.; Irving, M. B.; Michielin, O. MM-GBSA binding free energy decomposition and T cell receptor engineering. *J. Mol. Recognit.* **2010**, *23* (2), 142–152.

- (55) Rocchia, W.; Sridharan, S.; Nicholls, A.; Alexov, E.; Chiabrera, A.; Honig, B. Rapid grid-based construction of the molecular surface and the use of induced surface charge to calculate reaction field energies: Applications to the molecular systems and geometric objects. *J. Comput. Chem.* **2002**, *23* (1), 128–137.
- (56) Laskowski, R. A.; Macarthur, M. W.; Moss, D. S.; Thornton, J. M. Procheck - a Program to Check the Stereochemical Quality of Protein Structures. *J. Appl. Crystallogr.* **1993**, *26*, 283–291.
- (57) de Hoon, M. J. L.; Imoto, S.; Nolan, J.; Miyano, S. Open source clustering software. *Bioinformatics* **2004**, *20* (9), 1453–1454.
- (58) Craig, I. R.; Pfleger, C.; Gohlke, H.; Essex, J. W.; Spiegel, K. Pocket-Space Maps To Identify Novel Binding-Site Conformations in Proteins. *J. Chem. Inf. Model.* **2011**, *51* (10), 2666–2679.
- (59) 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* (6), 359–363.
- (60) Huang, N.; Shoichet, B. K.; Irwin, J. J. Benchmarking sets for molecular docking. *J. Med. Chem.* **2006**, *49* (23), 6789–6801.
- (61) Irwin, J. J.; Shoichet, B. K. ZINC - A free database of commercially available compounds for virtual screening. *J. Chem. Inf. Model.* **2005**, *45* (1), 177–182.
- (62) Molinspiration, version 1; Molinspiration Cheminformatics: Slovensky Grob, Slovak Republic, 2008. www.molinspiration.com (accessed Nov. 2011).
- (63) Filter, version 2.0.2; OEChem, version 1.4.2; OpenEye Scientific Software, Inc.: Santa Fe, NM, 2010. www.eyesopen.com (accessed Nov. 2011).
- (64) R: A Language and Environment for Statistical Computing, version 2.6.2; R Development Core Team: Vienna, Austria, 2008.
- (65) Shaw, D. E.; Maragakis, P.; Lindorff-Larsen, K.; Piana, S.; Dror, R. O.; Eastwood, M. P.; Bank, J. A.; Jumper, J. M.; Salmon, J. K.; Shan, Y. B.; Wriggers, W. Atomic-Level Characterization of the Structural Dynamics of Proteins. *Science* **2010**, *330* (6002), 341–346.
- (66) Rickert, M.; Wang, X. Q.; Boulanger, M. J.; Goriatcheva, N.; Garcia, K. C. The structure of interleukin-2 complexed with its alpha receptor. *Science* **2005**, *308* (5727), 1477–1480.
- (67) Jiang, L.; Kuhlman, B.; Kortemme, T. A.; Baker, D. A "solvated rotamer" approach to modeling water-mediated hydrogen bonds at protein-protein interfaces. *Proteins* **2005**, *58* (4), 893–904.
- (68) Mobley, D. L.; Dill, K. A. Binding of Small-Molecule Ligands to Proteins: "What You See" Is Not Always "What You Get". *Structure* **2009**, *17* (4), 489–498.
- (69) Hu, G. D.; Wang, D. Y.; Liu, X. G.; Zhang, Q. G. A computational analysis of the binding model of MDM2 with inhibitors. *J. Comput.-Aided Mol. Des.* **2010**, *24* (8), 687–697.
- (70) Bradshaw, R. T.; Patel, B. H.; Tate, E. W.; Leatherbarrow, R. J.; Gould, I. R. Comparing experimental and computational alanine scanning techniques for probing a prototypical protein-protein interaction. *Protein Eng., Des. Sel.* **2011**, *24* (1–2), 197–207.
- (71) Hou, T. J.; Yu, R. Molecular dynamics and free energy studies on the wild-type and double mutant HIV-1 protease complexed with amprenavir and two amprenavir-related inhibitors: Mechanism for binding and drug resistance. *J. Med. Chem.* **2007**, *50* (6), 1177–1188.
- (72) Chene, P. Drugs targeting protein-protein interactions. *ChemMedChem* **2006**, *1* (4), 400–411.
- (73) Li, Y.; Huang, Y.; Swaminathan, C. P.; Smith-Gill, S. J.; Mariuzza, R. A. Magnitude of the hydrophobic effect at central versus peripheral sites in protein-protein interfaces. *Structure* **2005**, *13* (2), 297–307.
- (74) Fulle, S.; Christ, N. A.; Kestner, E.; Gohlke, H. HIV-1 TAR RNA spontaneously undergoes relevant apo-to-holo conformational transitions in molecular dynamics and constrained geometrical simulations. *J. Chem. Inf. Model.* **2010**, *50* (8), 1489–1501.
- (75) Gohlke, H.; Thorpe, M. F. A natural coarse graining for simulating large biomolecular motion. *Biophys. J.* **2006**, *91* (6), 2115–2120.
- (76) Zoete, V.; Meuwly, M.; Karplus, M. Study of the insulin dimerization: Binding free energy calculations and per-residue free energy decomposition. *Proteins* **2005**, *61* (1), 79–93.
- (77) Ballet, T.; Boulange, L.; Brechet, Y.; Bruckert, F.; Weidenhaupt, M. Protein conformational changes induced by adsorption onto material surfaces: an important issue for biomedical applications of material science. *Bull. Pol. Acad. Sci.: Tech. Sci.* **2010**, *58* (2), 303–315.
- (78) Ferrara, P.; Gohlke, H.; Price, D. J.; Klebe, G.; Brooks, C. L. Assessing scoring functions for protein-ligand interactions. *J. Med. Chem.* **2004**, *47* (12), 3032–3047.
- (79) Verdonk, M. L.; Mortenson, P. N.; Hall, R. J.; Hartshorn, M. J.; Murray, C. W. Protein-Ligand Docking against Non-Native Protein Conformers. *J. Chem. Inf. Model.* **2008**, *48* (11), 2214–2225.
- (80) Erickson, J. A.; Jalaie, M.; Robertson, D. H.; Lewis, R. A.; Vieth, M. Lessons in molecular recognition: The effects of ligand and protein flexibility on molecular docking accuracy. *J. Med. Chem.* **2004**, *47* (1), 45–55.
- (81) Jacobsson, M.; Karlen, A. Ligand bias of scoring functions in structure-based virtual screening. *J. Chem. Inf. Model.* **2006**, *46* (3), 1334–1343.
- (82) Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Delivery Rev.* **1997**, *23* (1–3), 3–25.
- (83) Ahmed, A.; Kazemi, S.; Gohlke, H. Protein flexibility and mobility in structure-based drug design. *Front. Drug Des. Discovery* **2007**, *3*, 455–476.
- (84) Kongsted, J.; Ryde, U. An improved method to predict the entropy term with the MM/PBSA approach. *J. Comput.-Aided Mol. Des.* **2009**, *23* (2), 63–71.
- (85) Ahmed, S.; Metpally, R. P. R.; Sangadala, S.; Reddy, B. V. B. Virtual screening and selection of drug-like compounds to block noggin interaction with bone morphogenetic proteins. *J. Mol. Graphics Modell.* **2010**, *28* (7), 670–682.
- (86) Gohlke, H.; Ahmed, A.; Rippmann, F.; Barnickel, G. A Normal Mode-Based Geometric Simulation Approach for Exploring Biologically Relevant Conformational Transitions in Proteins. *J. Chem. Inf. Model.* **2011**, *51* (7), 1604–1622.
- (87) Simmerling, C.; Elber, R. Hydrophobic Collapse in a Cyclic Hexapeptide - Computer-Simulations of Chdlfc and Caaaac in Water. *J. Am. Chem. Soc.* **1994**, *116* (6), 2534–2547.