Supporting Information

Hot spots and transient pockets: Predicting the determinants of small-molecule binding to a protein-protein interface

Alexander Metz,[1§] Christopher Pfleger,[1§] Hannes Kopitz,[1] Stefania Pfeiffer-Marek,[2] Karl-Heinz Baringhaus,[2] Holger Gohlke[1*]

[1]Institute for Pharmaceutical and Medicinal Chemistry, Department of Mathematics and Natural Sciences, Heinrich-Heine-University, Düsseldorf, Germany

[2]Sanofi-Aventis Deutschland GmbH, CAS Drug Design, Frankfurt am Main, Germany

[§]Both authors contributed equally to this work.

[*]Universitätsstr. 1, 40225 Düsseldorf, Germany. Phone: (+49) 211 81-13662. Fax: (+49) 211 81-13847. E-mail: gohlke@uni-duesseldorf.de

Structure preparation

Starting structures for the simulations of human IL-2 and its complexes were taken from the Protein Data Bank[1] (PDB codes: 1m47, 1m4c, 1m48, 1m49, 1pw6, 1py2, 1qvn, and 1z92). These structures were modified to achieve consistency with respect to the sequence and number of amino acids. Solvent and buffer molecules were removed except for crystal waters bound to protein chains, which were considered in the MD simulations. Histidine protonation and rotation states were assigned manually such that all IL-2 chains have the same constitution and that histidines can form optimal local interactions. In the case of multiple identical chains, the one with the lowest number of unresolved residues was chosen. Missing residues (Figure S9) were modeled with MODELLER 7v7[2] using other IL-2 structures as templates, as was done for the Ala69Val mutation in the structure with PDB code 1QVN. The flexible loop between *Ser64* and *Leu100* (all IL-2Rα residues are highlighted in italics, whereas all IL-2 residues are depicted in "normal" font) of IL-2Rα was not resolved in the crystal structure (PDB code: 1z92).[3] As the loop does not contact the binding interface,[3] it was not considered any further. This should not influence the structural integrity of IL-2Rα during MD simulations because either end of the loop is bound to the residual IL-2Rα structure by a disulfide bond. Ligand structures were extracted from the complexes. For docking, the ligands were converted to MOL2 files using the PRODRG2[4] server. Atom types were corrected manually if necessary. Flexible torsions were determined by AutoTors from the AutoDock suite of programs.[5]

Molecular dynamics simulations

MD simulations were performed with the AMBER 9 package of molecular simulation programs[6] using the Cornell *et al.* force field[7] with modifications introduced by Hornak *et al.* (ff99SB)[8] and the general amber force field (GAFF)[9] for proteins and small molecules, respectively. Partial charges of small molecules were generated according to the RESP procedure.[9-10] The structures were solvated in a truncated octahedron of TIP3P water[11] such that the distance between the edges of the box and the closest solute atom was at least 11 Å. Periodic boundary conditions were applied using the particle mesh Ewald (PME) method[12] to treat long-range electrostatic interactions. Bond lengths involving bonds to hydrogen atoms were constrained by SHAKE.[13-14] The time step for all MD simulations was 2 fs, and a direct-space non-bonded cutoff of 8 Å was applied. After minimization the system was heated from 100 K to 300 K using canonical ensemble (NVT) MD. Then, the solvent density was adjusted

using isothermal-isobaric ensemble (NPT) MD. Positional restraints applied during equilibration were reduced in a stepwise manner over 50 ps followed by 50 ps of unrestrained canonical ensemble (NVT) MD at 300 K with a time constant of 2 ps for heat bath coupling. Snapshots were extracted every 10 ps from production runs for further analysis (Table 1).

Docking

All docking runs were performed with AutoDock 3.05[5] using DrugScore pair potentials[15] as a scoring function.[16-17] The docking protocol for flexible ligand docking comprised 100 independent runs per ligand using an initial population size of 100 individuals, 5.0×10^3 generations, a maximum number of 10.0×10^6 energy evaluations, a mutation rate of 0.02, a crossover rate of 0.8, and an elitism value of 1. For the enrichment evaluation the maximum number of energy evaluations and the population size were reduced to 3.0×10^6 and 50, respectively. Before calculating the DrugScore potential grids, all structures were aligned to the *x*/*y*-plane of the Cartesian coordinate system such that the rms distance between the interface amino acids and the plane is minimal. By doing so, the potential grids are optimally positioned for the mainly flat interface region of IL-2. The dimensions of the grids were chosen such that the grids extend beyond all hot spots as well as amino acids lining the identified interface pockets by at least 2.5 Å. In the case of *apo*-docking where no transient pocket is available, the same potential grid definition was chosen as for the re-docking approach. We note that this way no information about the known binding modes of the PPIM was considered for setting up the docking. The grid spacing was set to 0.375 Å. Similar docking poses (RMSD < 1 Å) were clustered, and the intermolecular docking energy was calculated. As the final docking result, the ligand pose with the lowest intermolecular docking energy from the largest cluster was chosen. A docking experiment was considered successful when this ligand pose had an RMSD < 2.0 Å to the native pose.

Statistical significance of MM-PB/SA results

To investigate the energetics of IL-2/IL-2Rα and IL-2/small-molecule complex formation, the MM-PB/SA method was applied to compute effective energies as the sum of gas-phase energies and solvation free energies. Entropic terms resulting from translational, rotational, and vibrational contributions of the solutes were omitted. The gas-phase and solvation free energy values were averaged over 617 – 1379 snapshots (Table 1) taken at 10 ps intervals from the trajectories of the MD simulations. The correlation time for relaxation of effective energy fluctuations was computed to < 10 ps (data not shown), in agreement with related studies.[18] Hence, the extracted snapshots should be uncorrelated, and mean values of binding effective energies computed by the single trajectory MM-PB/SA method can be estimated to within a standard error of the mean (SEM) between 0.13 – 0.37 kcal mol^{-1} (Table 1).

Time-series of effective energies computed using the MM-PB/SA method are displayed in Figure S8 for snapshots of the unbound solutes and the IL-2/IL-2Rα and IL-2/small-molecule complexes. In all cases, significant drifts and fluctuations in the absolute effective energies were found, which demonstrates the sensitivity of these values to conformational details and reflects structural variations throughout the MD trajectories. The observed energy drift (Table S5) depends on the size and conformational complexity of the solutes (Table 1) with IL-2/IL-2Rα showing the largest drift (-11.02 kcal mol^{-1} ns^{-1}), unbound IL-2 and the IL-2/small-molecule complexes showing drifts of -0.56 – -6.03 kcal mol^{-1} ns^{-1}, and the small molecules showing negligible drifts of -0.30 – 0.20 kcal mol^{-1} ns^{-1}.

These analyses indicate as to why MM-PB/SA binding effective energies computed by the multiple trajectory method for IL-2/small-molecule complexes do not correlate with experimental results ($R^2 < 0.1$, data not shown). In contrast, in the case of the single trajectory method, binding effective energies show a much smaller drift (-0.63 – 0.74 kcal mol^{-1} ns^{-1}, Table S5) due to a cancellation of internal energies.[18] These results also provide an explanation as to why differentiating between conformational states of IL-2 based on absolute effective energies is not successful (Figure S8), in addition to the error introduced by neglecting changes in the solute's configurational entropy. As the energy drifts are mainly caused by conformational transitions of the solute that occur, in particular, in modeled regions, loops, and termini, much longer simulation times would be required to obtain mean absolute effective energies that are stable over time. However, even when simulating for up to 10 ns in related studies,[18-23] this problem could not be alleviated, and comparable drifts were observed.

Tables

Table S1: Heavy atom RMSD during MD simulation

	RMSD[a]				
PDB code	Complex	IL-2[b]	IL-2 interface[c]	Bound ligand[d]	Unbound ligand[d]
1m47	–	3.17 (3.63)	2.88 (3.70)	–	–
1m4c	–	3.16 (3.83)	2.44 (3.13)	–	–
1m48	2.57 (3.08)	2.55 (3.03)	2.42 (3.00)	1.46 (2.02)	2.66 (4.21)
1m49	2.66 (3.06)	2.68 (3.09)	2.13 (2.60)	1.06 (1.89)	2.64 (4.29)
1pw6	2.70 (3.25)	2.72 (3.27)	2.20 (2.86)	0.89 (1.46)	2.01 (4.76)
1py2	2.59 (3.09)	2.61 (3.12)	1.94 (2.36)	1.55 (2.03)	2.92 (4.94)
1qvn	2.88 (3.21)	2.81 (3.17)	2.27 (3.03)	2.20 (3.94)	2.81 (5.52)
1z92	3.50 (4.52)	3.46 (4.32)	2.49 (3.18)	2.98 (4.07)	3.77 (5.73)

[a] Mean heavy atom RMSD with respect to the equilibrated structure; in Å. Five N-terminal amino acids of IL-2 were omitted. Maximum RMSD in parentheses.

[b] Unbound IL-2 or IL-2 extracted from the MD trajectory of the complex.

[c] IL-2 residues Tyr31, Asn33-Lys35, Thr37-Met39, Thr41-Tyr45, Glu60-Glu62, Lys64-Val69, Asn71, Leu72, and Met104-Thr111.

[d] Aligned with respect to the ligand.

Table S2: RMSD of the IL-2 interface region of experimentally determined bound IL-2 conformations[a]

IL-2 structure	IL-2/FRG	IL-2/CMM	IL-2/FRB	IL-2/FRH	IL-2/FRI
Unbound structure[b]	1.77	1.69	3.08	1.50	1.69
MD[c]	1.83	1.74	2.99	1.74	1.78
FRODA[c]	1.66	1.58	2.98	1.42	1.51

[a] All heavy atoms of the interface region (IL-2 residues Tyr31, Asn33-Lys35, Thr37-Met39, Thr41-Tyr45, Glu60-Glu62, Lys64-Val69, Asn71, Leu72, Met104-Thr111) are considered; in Å.

[b] RMSD from the unbound IL-2 conformation (PDB code: 1m47).

[c] Minimal RMSD obtained from snapshots generated by either MD or FRODA simulation starting from the unbound IL-2 conformation (PDB code: 1m47).

Table S3: Ten largest pocket volumes of selected FRODA and MD snapshots

FRODA snapshot[a]	Volume[b]	MD snapshot[c]	Volume[b]
6	159	123	175
117	157	199	182
169	148	215	180
301	145	234	240
418	148	269	181
514	185	498	184
534	188	782	163
657	215	794	180
698	176	843	173
729	156	940	167

[a] Consecutive number of the snapshot from a total of 1,000 snapshots uniformly extracted from the 10,000,000 FRODA-generated snapshots starting from the unbound IL-2 structure (PDB code: 1m47).

[b] In $Å^3$.

[c] Consecutive number of the snapshot from a total of 1,021 snapshots 10 ps apart that were generated by MD starting from the unbound IL-2 structure (PDB code: 1m47).

Table S4: Pocket residues of IL-2 selected for the definition of the potential energy grids

Protein structure	**Residues**
Crystal structures[a]	Ile28, Tyr31-Tyr45, Cys58, Glu61-Pro65, Glu68-Lys76, Tyr107, Ile114
MD	Tyr31, Pro34-Lys35, Arg38-Met39, Thr41-Tyr45, Glu60-Glu62, Lys64-Val69, Asn71-Leu72, Cys105, Tyr107, Thr111
FRODA	Tyr31, Pro34-Lys35, Thr37-Met39, Thr41-Tyr45, Glu61-Glu62, Lys64-Val69, Asn71-Leu72

[a] Pocket residues identified in PDB codes 1m48 (chain A,B), 1m49 (chain A,B), 1pw6 (chain A), 1py2 (chain A,B,C,D), and 1qvn (chain B,C,D).

Table S5: Drift of the effective energy

PDB code	**Drift of effective energy**[a]					
	Multiple trajectory method[b]			Single trajectory method[c]		Binding effective energy from single trajectory method[c]
	IL-2	ligand	complex	IL-2	ligand	–
1m47	-1.45	–	–	–	–	–
1m4c	-0.56	–	–	–	–	–
1m48	–	0.03	-2.38	-2.33	0.13	0.20
1m49	–	-0.09	-3.08	-3.22	-0.05	0.19
1pw6	–	0.20	-4.55	-4.05	-0.23	-0.27
1py2	–	-0.30	-5.33	-4.96	0.16	-0.63
1qvn	–	-0.04	-6.03	-6.78	0.00	0.74
1z92	–	0.16	-11.02	-8.24	-3.27	0.49

[a] In kcal mol^{-1} ns^{-1}.

[b] Structures of IL-2, ligand, and complex were generated by separate MD simulations.

[c] Structures of IL-2 and ligand were extracted from the MD trajectory of the respective complexes.

Table S6: Selection of poses from docking into FRODA snapshots

Ligand	PDB code[a]	RMSD[b]	Score[c]	Clustered poses[d]	Cluster size[e]
FRG	1m48	2.08	-13.20	213	204
CMM	1m48	2.59	-14.81	335	327
FRB	1pw6	2.48	-12.78	105	94
FRH	1py2	4.30	-15.77	272	78
FRI	1qvn	3.22	-14.92	133	46

[a] PDB code of the corresponding IL-2 complex structure.

[b] RMSD of the ligand pose with the lowest intermolecular docking energy in the largest cluster with respect to the native pose; in Å.

[c] In kcal mol^{-1}.

[d] The number of poses out of 1,000 docked poses (10 FRODA simulated structures with largest pocket volume times 100 docking runs) where the ligand's guanidinium group is within 5 Å of the side chain heavy atoms of Glu62 that were subjected to hierarchical complete linkage clustering with R^{24} with a cluster distance of 5 Å.

[e] Number of ligand poses in the largest cluster.

Figures

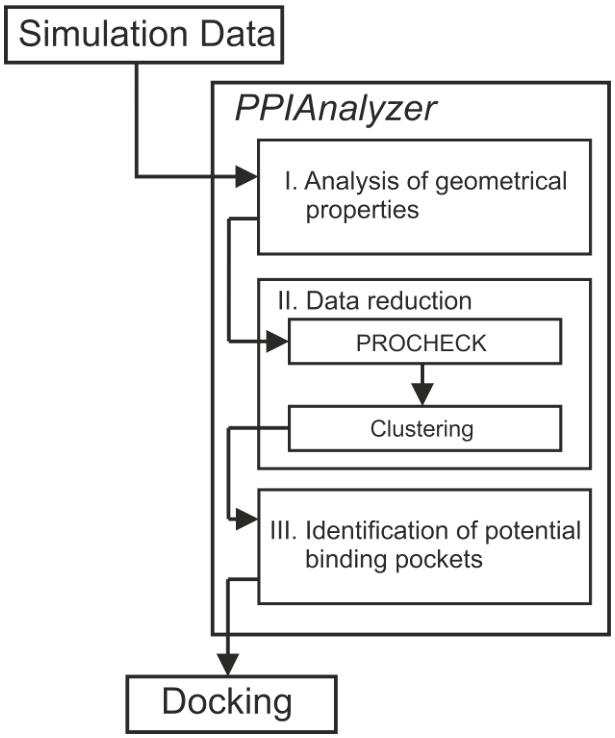

Figure S1: Workflow of the PPIAnalyzer method. The method contains three main steps: I. Analysis of geometrical properties in terms of root mean-square deviations (RMSD) and rotamer analysis. II. Reduction of the dataset by assessing the steric quality of the generated conformations and clustering with respect to the RMSD of heavy atoms of interface residues. III. Identification of transient pockets in the remaining conformations. Representative structures that show the largest interface pocket volume are then selected for the subsequent docking experiments.

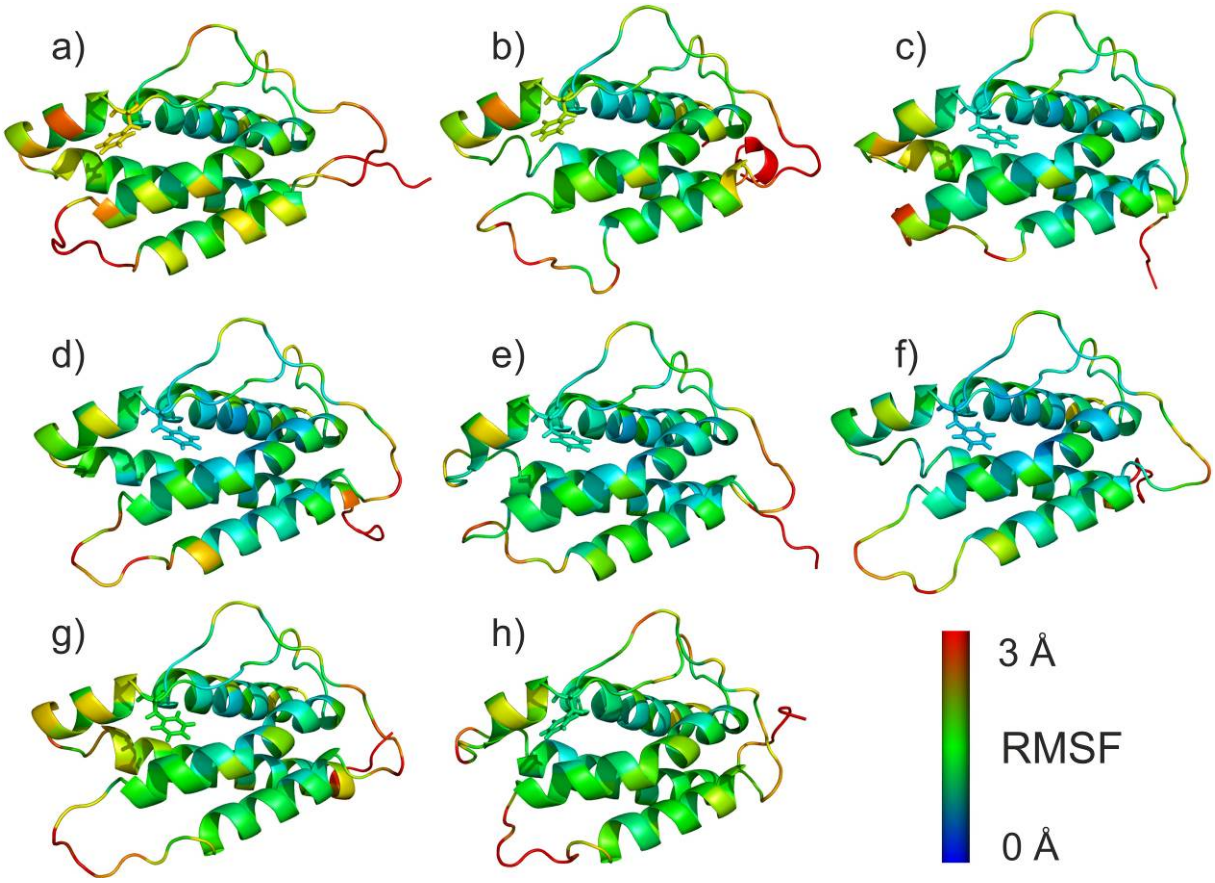

Figure S2: RMSF values of IL-2 residues obtained by MD simulations of the unbound and bound states. The RMSF value of each residue is calculated as the average over all atoms. RMSF values are color-coded onto the respective starting structure of the MD simulations: a) 1m47 and b) 1m4c for unbound IL-2; c) 1m48, d) 1m49, e) 1pw6, f) 1py2, and (g) 1qvn for IL-2 bound to PPIM; h) 1z92 for IL-2 bound to IL-2Rα. The RMSF values were calculated for snapshots 10 ps apart. Prior to the RMSF calculations, all snapshots were structurally aligned to the starting structure of the MD simulation considering all heavy protein atoms. The highly mobile five N-terminal residues of IL-2 were neglected in the structural alignment. The protein is depicted in cartoon representation. Phe42 is depicted in stick representation to indicate the location of the small molecule binding pocket. Figures were generated by PyMOL.[25]

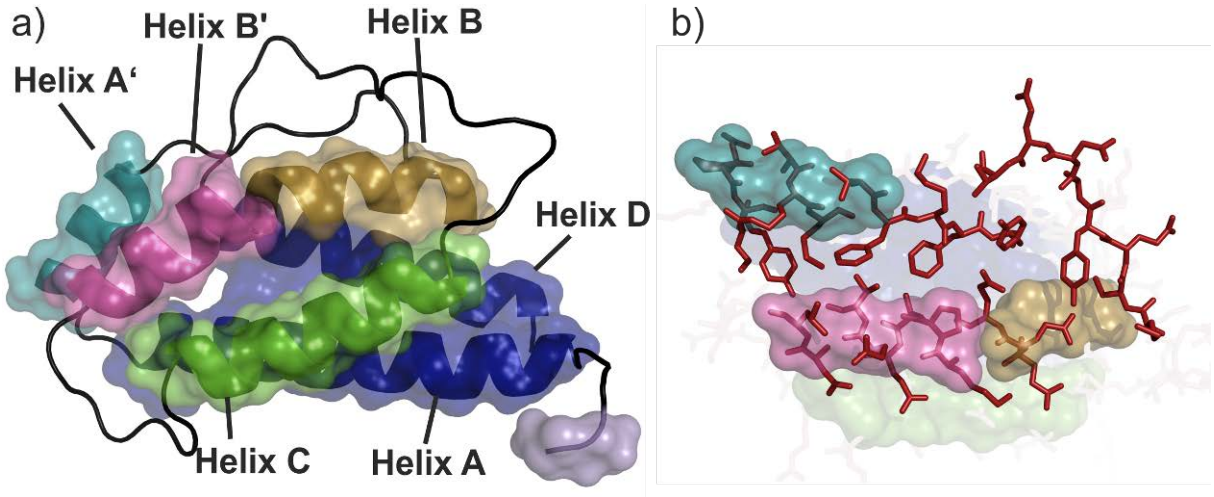

Figure S3: Rigid cluster decomposition obtained by FIRST. (a) The rigid clusters (transparent surfaces) are denominated RC1-6 in the order of decreasing size. RC1 (blue) covers helices A and D, RC2 (green) covers helix C, RC3 (magenta) covers parts of helix B', RC4 (turquoise) covers helix A', RC5 (gold) covers parts of helix B, and RC6 (light blue) is located at the N-terminus of IL-2. (b) 25.5% of all interface atoms are part of the rigid clusters RC3, RC4, and RC5. All flexible atoms (red) can move freely in FRODA simulations. Figures were generated by PyMOL.[25]

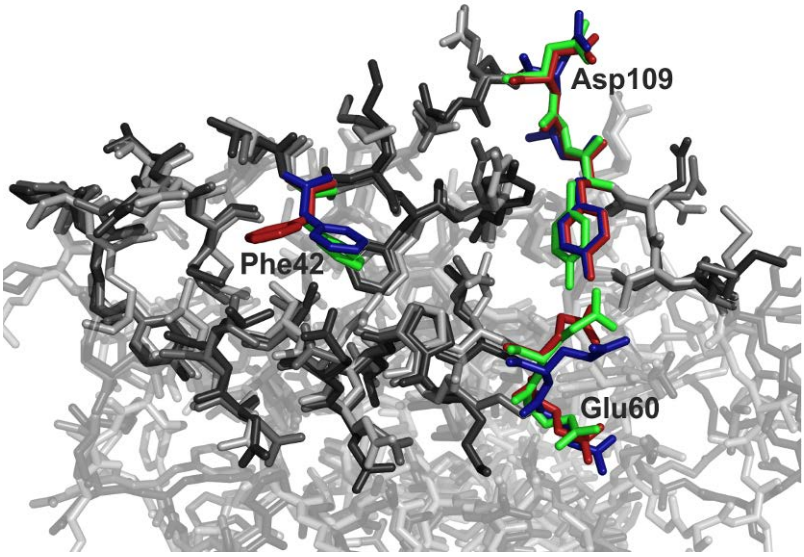

Figure S4: Overlay of the protein-protein interface region of IL-2 in unbound (red) and bound (green) conformation. Exemplarily, one snapshot from a FRODA simulation started from the unbound state is shown (blue), demonstrating that the movement of Phe42 can even be observed in the absence of the ligand, leading to a transient pocket opening. Regions for which no movements were observed by experiment (around Glu60 and Asp109) also remain immobile during the simulation. Figure was generated by PyMOL.[25]

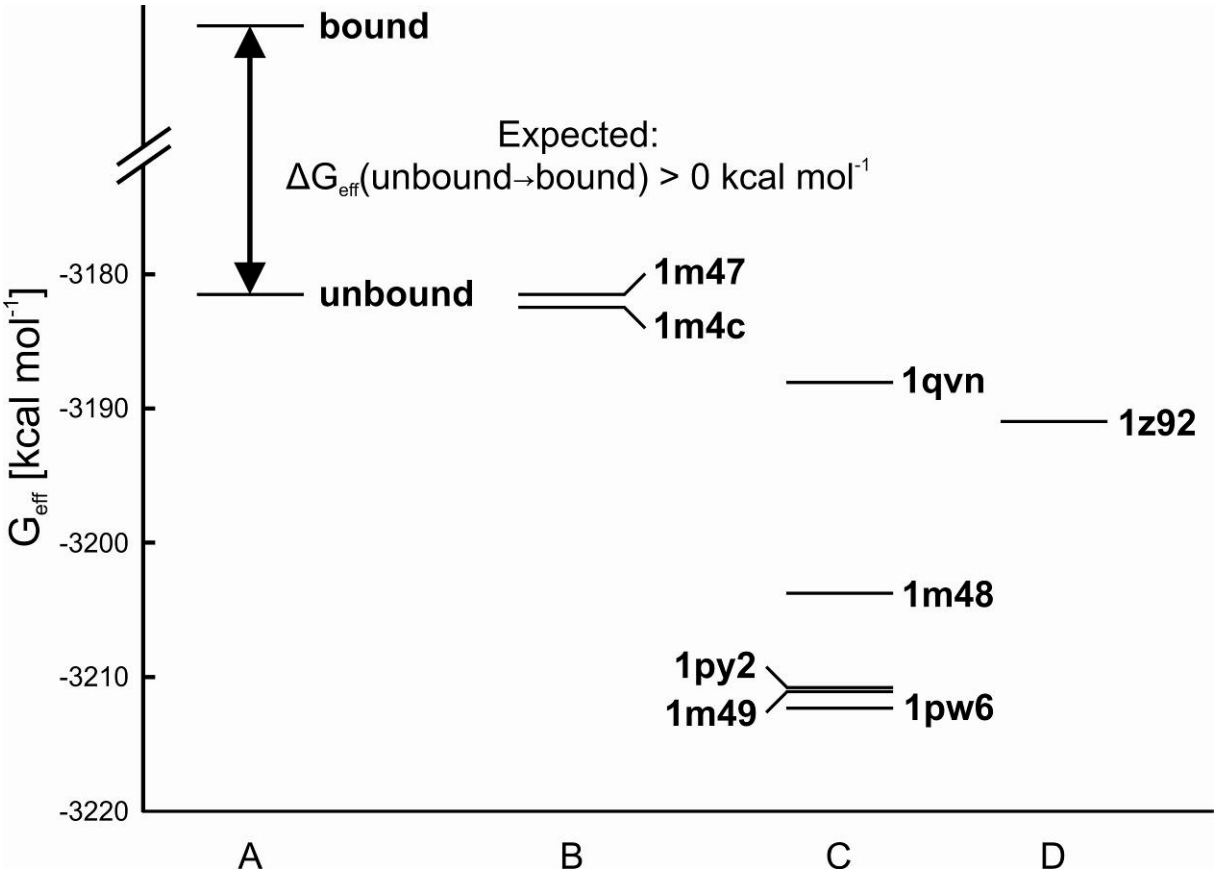

Figure S5: Mean absolute effective energies G_{eff} of IL-2 in its unbound and bound conformations. Conformational stress and changes in solvation and configurational entropy are expected to increase the free energy of a bound conformation over an unbound one (lane A). In contrast, computed G_{eff} of IL-2 extracted from MD trajectories of IL-2/small molecule complexes (lane C) or from the IL-2/IL-2Rα complex (lane D) are lower than G_{eff} of unbound IL-2 (lane B). We attribute this observation to neglecting changes in configurational entropy upon the conformational transitions and the occurrence of significant drifts of G_{eff} over time (see Table S5 and Figure S8). Figure was generated by gnuplot.[26]

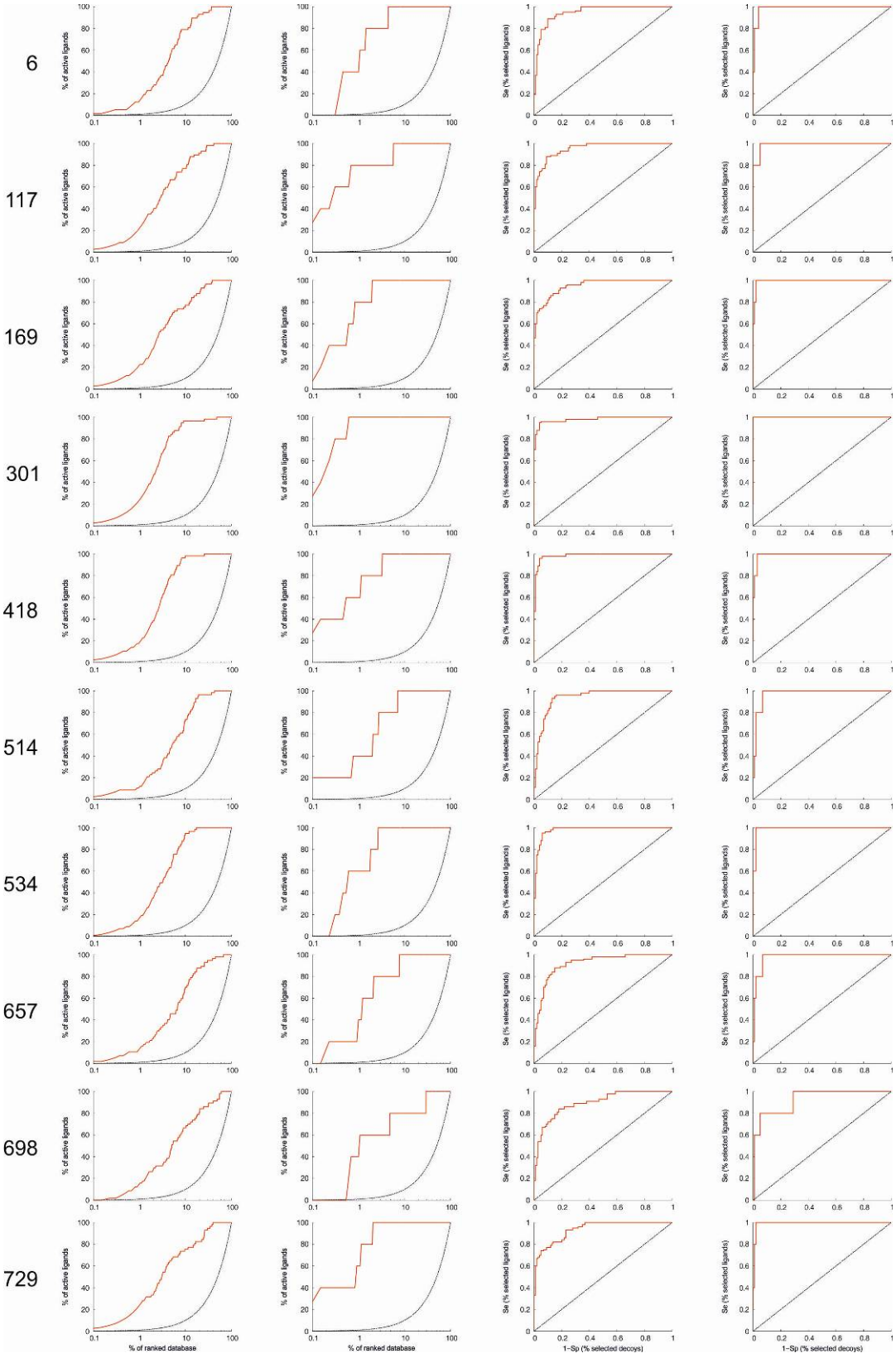

Figure S6: Docking enrichment of known IL-2 ligands. The number of the FRODA snapshot with a transient pocket used for docking is indicated in the left row. Enrichment plots for all 57 IL-2 ligands (1st vertical lane) and the five IL-2 ligands with available complex crystal structure (2nd lane) as well as ROC curves for all 57 IL-2 ligands (3rd lane) and the five IL-2 ligands with available complex crystal structure (4th lane) are given.

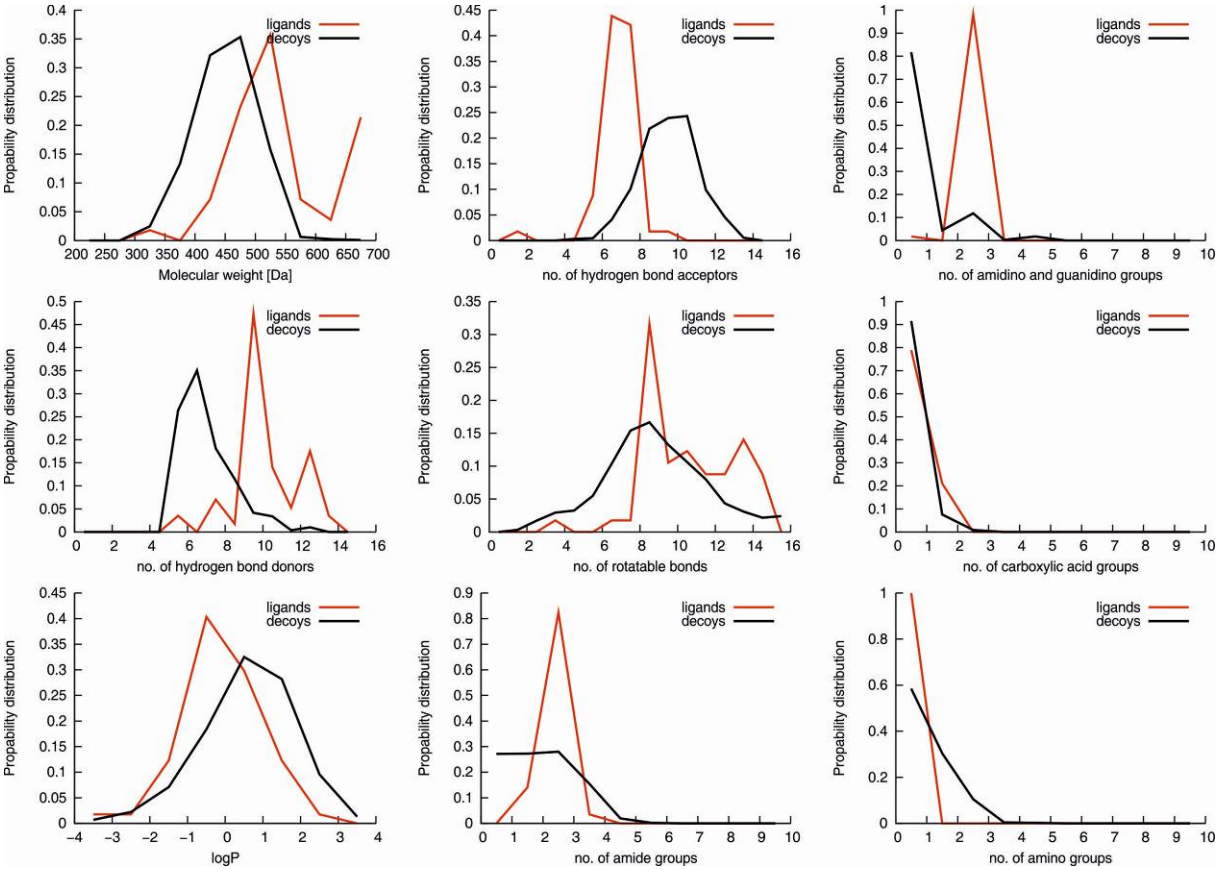

Figure S7: Property distribution of the known IL-2 ligands and decoys. The red line represents all 57 IL-2 ligands. The black line represents the decoy set generated with the aim of similar physicochemical properties to the five IL-2 ligands with available complex crystal structures following the DUD procedure. Figures were generated by gnuplot.[26]

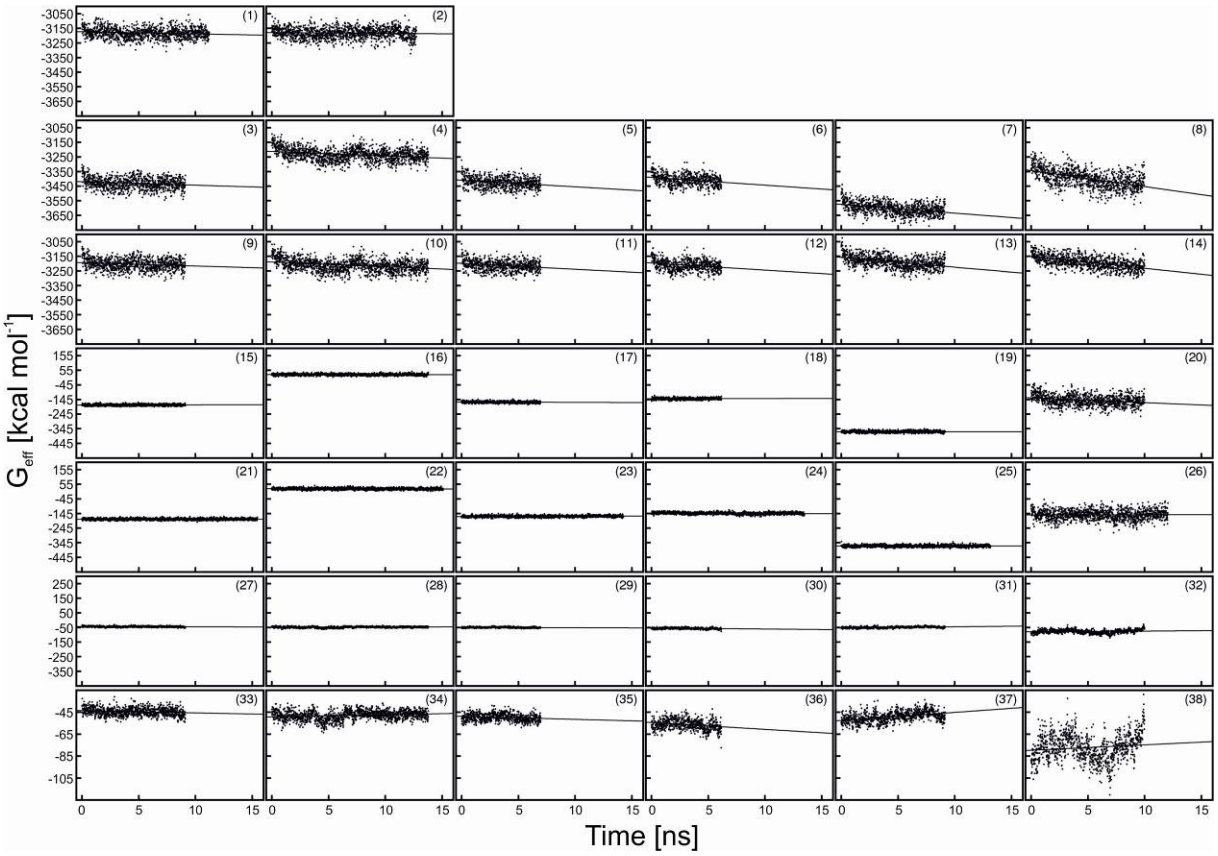

Figure S8: Time series of effective energies. The effective energies were calculated by applying the MM-PB/SA method to snapshots extracted every 10 ps from MD trajectories for: (1) unbound IL-2 [PDB-code: 1m47]; (2) unbound IL-2 [1m4c]; IL-2 in complex with (3) FRG [1m48], (4) CMM [1m49], (5) FRB [1pw6], (6) FRH [1py2], (7) FRI [1qvn], and (8) IL-2Rα [1z92]; IL-2 extracted from the trajectories of the complexes of IL-2 with (9) FRG [1m48], (10) CMM [1m49], (11) FRB [1pw6], (12) FRH [1py2], (13) FRI [1qvn], and (14) IL-2Rα [1z92]; IL-2 ligands extracted from the trajectories of the complexes of IL-2 with (15) FRG [1m48], (16) CMM [1m49], (17) FRB [1pw6], (18) FRH [1py2], (19) FRI [1qvn], and (20) IL-2Rα [1z92]; unbound ligands of IL-2 (21) FRG [1m48], (22) CMM [1m49], (23) FRB [1pw6], (24) FRH [1py2], (25) FRI [1qvn], and (26) IL-2Rα [1z92]. In addition, MM-PB/SA single trajectory binding effective energies are depicted for the complexes of IL-2 with (27) FRG [1m48], (28) CMM [1m49], (29) FRB [1pw6], (30) FRH [1py2], (31) FRI [1qvn], and (32) IL-2Rα [1z92]. The range of the ordinate values is identical in all plots (1) – (32). For reasons of clarity, MM-PB/SA single trajectory binding effective energies are depicted again with a magnified ordinate scale for the complexes of IL-2 with (33) FRG [1m48], (34) CMM [1m49], (35) FRB [1pw6], (36) FRH [1py2], (37) FRI [1qvn], and (38) IL-2Rα [1z92]. Figures were generated by gnuplot.[26]

```
wt   APTSSSTKKTQLQLEHLLLDLQMILNGINNYKNPKLTRMLTFKFYMPKKATELKHLQCLEEELKPLE
1m47 -----STKKTQLQLEHLLLDLQMILNGINNYKNPKLTRMLTFKFYMPKKATELKHLQCLEEELKPLE
1m4C -----STKKTQLQLEHLLLDLQMILNGINNYKNPKLTRMLTFKFYMPKKATELKHLQCLEEELKPLE
1m48 ---SSSTKKTQLQLEHLLLDLQMILNGINNYKNPKLTRMLTFKFYMPKKATELKHLQCLEEELKPLE
1m49 ---SSSTKKTQLQLEHLLLDLQMILNGINNYKNPKLTRMLTFKFYMPKKATELKHLQCLEEELKPLE
1pw6 ----SSTKKTQLQLEHLLLDLQMILNGINNYKNPKLTRMLTFKFYMPKKATELKHLQCLEEELKPLE
1py2 -----STKKTQLQLEHLLLDLQMILNGINNYKNPKLTRMLTFKFYMPKKATELKHLQCLEEELKPLE
1qvn ---SSSTKKTQLQLEHLLLDLQMILNGINNYKNPKLTRMLTFKFYMPKKATELKHLQCLEEELKPLE
1z92 -----STKKTQLQLEHLLLDLQMILNGINNYKNPKLTRMLTFKFYMPKKATELKHLQCLEEELKPLE
          ************************************************************

wt   EVLNLAQSKNFHLRPRDLISNINVIVLELKGSETTFMCEYADETATIVEFLNRWITFCQSIISTLT
1m47 EVLNLAQ--NFHLRPRDLISNINVIVLELKG----FMCEYADETATIVEFLNRWITFCQSIISTLT
1m4C EVLNLA---------RDLISNINVIVLELKGS---FMCEYADETATIVEFLNRWITFCQSIISTL-
1m48 EVLNLAQSK----NFRDLISNINVIVLELKGSETTFMCEYADETATIVEFLNRWITFCQSIISTLT
1m49 EVLNLAQ------RPRDLISNINVIVLELKGSETTFMCEYADETATIVEFLNRWITFCQSIISTL-
1pw6 EVLNLAQSKNFHLRPRDLISNINVIVLELKGSETTFMCEYADETATIVEFLNRWITFCQSIISTLT
1py2 EVLNLAQ------RPRDLISNINVIVLELKG-ETTFMCEYADETATIVEFLNRWITFCQSIISTL-
1qvn EALNLAQ------RPRDLISNINVIVLELKGSETTFMCEYADETATIVEFLNRWITFCQSIISTL-
1z92 EVLNLA-------RPRDLISNINVIVLELKGSETTFMCEYADETATIVEFLNRWITFCQSIISTLT
     *.****         ***************    ******************************
```

Figure S9: Multiple sequence alignment of sequences of IL-2 crystal structures (PDB codes: `1m47`, `1m4c`, `1m48`, `1m49`, `1pw6`, `1py2`, `1qvn`, and `1z92`). Residues that have not been resolved are indicated by a dash (`-`) and were modeled using MODELLER 7v7[2] to match the full length wild-type sequence (`wt`). Ala69 of one of the crystal structures (PDB code: `1qvn`) was mutated to alanine using MODELLER 7v7 to match the wt sequence. The multiple sequence alignment was created using CLUSTAL-W.[27]

References

(1) 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* (1), 235-242.

(2) Sali, A.; Blundell, T. L., Comparative Protein Modeling by Satisfaction of Spatial Restraints. *J. Mol. Biol.* **1993,** *234* (3), 779-815.

(3) 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.

(4) Schuettelkopf, A. W.; van Aalten, D. M. F., PRODRG: a tool for high-throughput crystallography of protein-ligand complexes. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2004,** *60*, 1355-1363.

(5) 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 an empirical binding free energy function. *J. Comput. Chem.* **1998,** *19* (14), 1639-1662.

(6) Case, D. A.; Darden, T. A.; Cheatham, T. E., III; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Merz, K. M.; Pearlman, D. A.; Crowley, M.; Walker, R. C.; Zhang, W.; Wang, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Wong, K. F.; Paesani, F.; Wu, X.; Brozell, S.; Tsui, V.; Gohlke, H.; Yang, L.; Tan, C.; Mongan, J.; Hornak, V.; Cui, G.; Beroza, P.; Mathews, D. H.; Schafmeister, C.; Ross, W. S.; Kollman, P. A., AMBER 9, University of California, San Francisco. **2006**.

(7) Wang, J. M.; Cieplak, P.; Kollman, P. A., How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? *J. Comput. Chem.* **2000,** *21* (12), 1049-1074.

(8) 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.

(9) Wang, J. M.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A., Development and testing of a general amber force field. *J. Comput. Chem.* **2004,** *25* (9), 1157-1174.

(10) Cieplak, P.; Cornell, W. D.; Bayly, C.; Kollman, P. A., Application of the Multimolecule and Multiconformational Resp Methodology to Biopolymers - Charge Derivation for DNA, Rna, and Proteins. *J. Comput. Chem.* **1995,** *16* (11), 1357-1377.

(11) 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* (2), 926-935.

(12) 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* (12), 10089-10092.

(13) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C., Numerical-Integration of Cartesian Equations of Motion of a System with Constraints - Molecular-Dynamics of N-Alkanes. *J. Comput. Phys.* **1977,** *23* (3), 327-341.

(14) Miyamoto, S.; Kollman, P. A., Settle - an Analytical Version of the Shake and Rattle Algorithm for Rigid Water Models. *J. Comput. Chem.* **1992,** *13* (8), 952-962.

(15) Gohlke, H.; Hendlich, M.; Klebe, G., Knowledge-based scoring function to predict protein-ligand interactions. *J. Mol. Biol.* **2000,** *295* (2), 337-356.

(16) Sotriffer, C. A.; Gohlke, H.; Klebe, G., Docking into knowledge-based potential fields: A comparative evaluation of DrugScore. *J. Med. Chem.* **2002,** *45* (10), 1967-1970.

(17) Pfeffer, P.; Gohlke, H., DrugScore(RNA) - Knowledge-based scoring function to predict RNA-ligand interactions. *J. Chem. Inf. Model.* **2007,** *47* (5), 1868-1876.

(18) 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-RaIGDS complexes. *J. Mol. Biol.* **2003,** *330* (4), 891-913.

(19) 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.

(20) 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.

(21) Deng, N. J.; Cieplak, P., Insights into affinity and specificity in the complexes of alpha-lytic protease and its inhibitor proteins: binding free energy from molecular dynamics simulation. *Phys. Chem. Chem. Phys.* **2009,** *11* (25), 4968-4981.

(22) 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-p-MHC complex. *Proteins* **2007,** *67* (4), 1026-47.

(23) Tuncel, A.; Kavakli, I. H.; Keskin, O., Insights into subunit interactions in the heterotetrameric structure of potato ADP-glucose pyrophosphorylase. *Biophys. J.* **2008,** *95* (8), 3628-3639.

(24) R Development Core Team, *A Language and Environment for Statistical Computing*, 2.6.2; Vienna, Austria. **2008**.

(25) DeLano, W. L. *PyMOL Molecular Graphics System*, 0.99rc6; DeLano Scientific LLC: 2006.

(26) Williams T.; Kelley C.; Broeker H.B.; E.A., M., *gnuplot - An Interactive Plotting Program*. **2007**.

(27) Thompson, J. D.; Higgins, D. G.; Gibson, T. J., Clustal-W - Improving the Sensitivity of Progressive Multiple Sequence Alignment through Sequence Weighting, Position-Specific Gap Penalties and Weight Matrix Choice. *Nucleic Acids Res.* **1994,** *22* (22), 4673-4680.