Articles

A General and Fast Scoring Function for Protein−Ligand Interactions: A Simplified Potential Approach

Ingo Muegge*,† and Yvonne C. Martin

Pharmaceutical Products Division, Abbott Laboratories, Abbott Park, Illinois 60064-6100

Received September 24, 1998

A fast, simplified potential-based approach is presented that estimates the protein−ligand binding affinity based on the given 3D structure of a protein−ligand complex. This general, knowledge-based approach exploits structural information of known protein−ligand complexes extracted from the Brookhaven Protein Data Bank and converts it into distance-dependent Helmholtz free interaction energies of protein−ligand atom pairs (potentials of mean force, PMF). The definition of an appropriate reference state and the introduction of a correction term accounting for the volume taken by the ligand were found to be crucial for deriving the relevant interaction potentials that treat solvation and entropic contributions implicitly. A significant correlation between experimental binding affinities and computed score was found for sets of diverse protein−ligand complexes and for sets of different ligands bound to the same target. For 77 protein−ligand complexes taken from the Brookhaven Protein Data Bank, the calculated score showed a standard deviation from observed binding affinities of 1.8 log K_i units and an R^2 value of 0.61. The best results were obtained for the subset of 16 serine protease complexes with a standard deviation of 1.0 log K_i unit and an R^2 value of 0.86. A set of 33 inhibitors modeled into a crystal structure of HIV-1 protease yielded a standard deviation of 0.8 log K_i units from measured inhibition constants and an R^2 value of 0.74. In contrast to empirical scoring functions that show similar or sometimes better correlation with observed binding affinities, our method does not involve deriving specific parameters that fit the observed binding affinities of protein−ligand complexes of a given training set. We compared the performance of the PMF score, Böhm's score (LUDI), and the SMOG score for eight different test sets of protein−ligand complexes. It was found that for the majority of test sets the PMF score performs best. The strength of the new approach presented here lies in its generality as no knowledge about measured binding affinities is needed to derive atomic interaction potentials. The use of the new scoring function in docking studies is outlined.

Introduction

The fast and accurate prediction of binding free energies of putative protein−ligand complexes is crucial for lead discovery and optimization in structure-based drug design.[1] It is apparent that improved scoring functions are needed to more accurately identify or design biologically active molecules that fit a macromolecular target of known 3D structure.[2] Several methods have been developed to calculate binding free energies; they vary in accuracy and consume computer time between hours and less than seconds per complex. Force field based methods calculate binding affinity using energy functions developed for 3D structure refinements and molecular dynamics calculations. Of these, free energy perturbation (FEP) methods,[3] linear response approximation (LRA) in conjunction with FEP or semimicroscopic methods,[4−6] and discretized continuum (DC) methods[7] are capable of calculating absolute and relative binding free energies within an error range of about 1−4 log K_i units. Although these methods consume minutes to hours of computing time per complex and are therefore not fast enough to screen a large number of compounds against a given protein target, they also do not offer generally robust predictions since these methods still depend on critical parametrizations (e.g., dielectric constants in DC or van der Waals weighting coefficients in LRA). As a result, so-called empirical scoring functions have been developed. They are based on a model that includes calculating the most important physical properties of protein−ligand interaction, converting them into computable terms, and optimizing their coefficients by fitting the derived function to observed binding constants of a training set of protein−ligand complexes with known 3D structure.[8−16] Empirical scoring functions as well as force field based scoring functions have been used in several docking approaches[17−20] as well as de novo design programs.[9,10] The advantage of multivariate regression methods that are used to derive empirical scoring functions lies in its capability to reconcile many parameters in a model and

* To whom correspondence should be addressed. Phone: (203) 812-5139. Fax: (203) 812-2650. E-mail: ingo.mugge.b@bayer.com.
† Present address: Bayer Corp., 400 Morgan Lane, West Haven, CT 06516.

10.1021/jm980536j CCC: $18.00
Published on Web 02/24/1999

make good predictions within the model range. The main weakness of empirical scoring functions lies in the fact that it is unclear to what extent they can be applied to protein−ligand complexes that were not represented in the training set. Although Welch and co-workers reported a docking study where it was found that biotin docked to streptavidin scored highest in a pool of 80 000 compounds of the ACD,[21] this encouraging result appears to be an exception. Using empirical scoring functions in blind docking studies, it is usually found that if the correct binding mode of a ligand is among the sampled conformations, it is often not ranked the highest. It appears that the highest ranked configurations often have a root-mean-square (rms) deviation of 2 Å or more from the binding mode found in the crystal structure.[22]

A third way to derive a scoring function of protein−ligand association uses statistical, knowledge-based methods that have already been shown to be successful in protein folding studies.[23−29] A coarse-grained model (SMOG) was derived and used in the de novo design of new compounds.[30] Verkhivker and co-workers reported a free energy function that included ligand−protein interaction patterns of HIV-1 protease complexes as part of an energy function that also included terms for protein/ligand−water interaction, desolvation, isomerization, and entropy.[31] However, while the calculated binding free energies for seven test complexes agreed well with the observed binding constants, no direct correlation between the PMF contribution and the measured binding affinities of the test complexes was found. Very recently, a knowledge-based scoring function was proposed that uses atom pair potentials similar to those derived in this article combined with an additional term for changes in solvent-accessible surface area upon protein−ligand binding.[32] However, a performance study of this scoring function is not published yet.

Our previous work focused on a wide range of methods of calculating relevant biological free energies in proteins (e.g., binding free energies, redox potentials, and pK_a values) using methods such as free energy perturbation,[33,34] discretized continuum,[34,35] linear response approximation in combination with semimicroscopic approaches,[6,36−38] and quantitative structure−activity relationship (QSAR) studies.[39−42] Here we tackle the long-standing problem of defining a general, reliable, and fast scoring function for protein−ligand interaction by using the entire Brookhaven Protein Data Bank[43] as a knowledge base and treating solvation and entropic effects implicitly. The new scoring function is built upon protein−ligand atom pair interaction potentials. They can be calculated from structural data alone because an observed crystallographic complex represents the optimum placement of the ligand atoms relative to the protein atoms, subject of course to covalent constraints. If one considers hundreds of such complexes, there are literally millions of observed distances between ligand and protein atoms. The observed distance distribution of specific atom-type interactions is the fundamental basis of the scoring method described in this report. In contrast to empirical methods, it does not rely on fitting to observed affinities. Nevertheless, the potentials of mean force (PMF) score can be calculated just as quickly as the potential interaction energy using a standard energy function, and that takes only fractions of a second per complex. A large cutoff for deriving the PMF, a large number of atom pair types, an appropriate reference state, and a suitable ligand volume correction are key elements of the new scoring function.

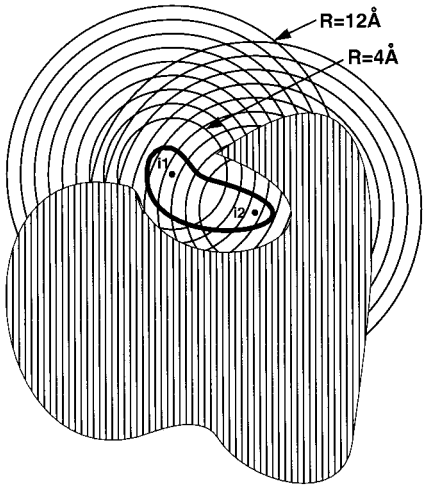

Figure 1. Schematic dissection of the protein−ligand−solvent system into spherical shells centered at two given ligand atoms $i1$ and $i2$. The shells are used to derive the volume correction terms $f^{j}_{\mathrm{Vol\_corr}}(r)$ (eq 14). They also illustrate how different the volume in a sphere with a radius of 12 Å around a specific ligand atom can be with respect to the ratio of taken protein and solvent volume. This ratio directly scales the reference bulk density term in eq 13 and therefore gives rise to implicitly taking solvent contributions of the ligand−protein interaction into account. The protein is shown as filled with vertical lines.

Method

If one chooses to dissect the protein−ligand binding free energy into all its physically meaningful contributions and then tries to evaluate them explicitly (e.g., desolvation, entropy, electrostatics, and hydrophobic interactions) one faces the difficult task of balancing large energy terms that add up to small binding free energies. Moreover, in order to get reliable estimates for the total binding free energy, it is important to evaluate all the terms rather than just a subset.[44] The problem then becomes that some of them are quite hard to quantify using available approaches (e.g., entropic contributions) or are time-consuming (e.g., reorganization energies). We circumvent this complicated task by deriving simplified potentials from known structural data to directly estimate the total protein−ligand binding free energy. Thus we treat all relevant contributions to the binding free energy implicitly.

The crystal and solution 3D structures of protein−ligand complexes collected in the Brookhaven Protein Database (PDB) as well as in proprietary databases provide a tremendous amount of information about protein−water systems as well as the specific interaction between proteins and ligands. Hundreds of different protein−inhibitor structures are available from the PDB alone. As the number of solved protein−ligand complexes grows rapidly it becomes possible to use this structural information to develop an energy function for protein−ligand interactions and ultimately an energy function that is able to predict the binding constants of ligands to proteins or to provide a ranking of binding constants of different ligands bound to a protein target. There are two basic assumptions that allow us to derive energy functions from a set of known structures using concepts of statistical mechanics. First, it is assumed that the protein−ligand complex is in a state of thermodynamic equilibrium and that this state retains the global minimum of the free energy of the complex. Second, the distribution of the molecules in their microscopic states

Table 1. Ligand Atom Types

CF	nonpolar carbon sp^3 aliphatic	OC	negatively charged oxygen (e.g., carboxylate)
CP	polar sp^3 carbon bonded to an atom other than carbon or hydrogen	OA	oxygen as hydrogen bond acceptor (e.g., keto, amide oxygen)
cF	nonpolar carbon aromatic	OE	oxygen in an ether bond
cP	polar carbon aromatic	OR	oxygen in planar ring
C3	nonpolar carbon sp^2 not aromatic	OS	oxygen bonded to atoms other than carbon or hydrogen
CW	polar carbon sp^2 not aromatic (e.g., bonded to carbonyl oxygen)	OD	oxygen bonded to hydrogen, except water
CO	carbon bonded to a negatively charged oxygen	P	phosphorus
CN	carbon bonded to a positively charged nitrogen	SA	sulfur as hydrogen bond acceptor
C0	sp carbon	SD	sulfur as hydrogen bond donor
NC	positively charged nitrogen (e.g., NH_3+ or guanidino group)	HL	hydrogen
NP	planar nitrogen bonded to 2 or 3 carbons but not to a hydrogen (can occur in a nonaromatic ring)	Zn	zinc
		CL	chlorine
NA	nitrogen as a hydrogen bond acceptor, not in a ring	Mn	manganese
ND	nitrogen as a hydrogen bond donor, not in a ring (e.g., amide nitrogen)	Mg	magnesium
NR	planar nitrogen in a ring structure (e.g., pyridine)	F	fluorine
N0	sp nitrogen bound to 1 carbon	Fe	iron
NS	nitrogen bound to atoms other than carbon or hydrogen and not of type ND	Br	bromine
		V	vanadium

Table 2. Protein Atom Types

CF	nonpolar aliphatic carbon (e.g., Cβ)
CP	polar aliphatic sp^2 or sp^3 carbon bonded to atoms other than carbon or hydrogen (e.g., backbone C or Cα)
cF	nonpolar carbon aromatic
cP	polar carbon aromatic
CO	carbon bonded to a negatively charged oxygen
CN	carbon bonded to a positively charged nitrogen
NC	positively charged nitrogen
ND	nitrogen as a hydrogen bond donor (e.g., backbone N, TRP NE, ASN ND)
NR	nitrogen in a planar ring structure (e.g., HIS ND, HIS NE)
OC	negatively charged oxygen
OA	oxygen as hydrogen bond acceptor (e.g., backbone O, ASN OD, GLN OE)
OD	oxygen as hydrogen bond donor (e.g., TYR OH, SER OG, THR OG)
OW	water oxygen
SA	sulfur as hydrogen bond acceptor (MET SD)
SD	sulfur as hydrogen bond donor (CYS SG)
HH	hydrogen

obeys Boltzmann's law. That is, there is a connection between the potential energy of the system and a probability to find a molecule or a complex of molecules in a certain microscopic state. Here we use only collected observations of protein−ligand atom distances to derive a new scoring function: we neglect all other structural information such as, for instance, hydrogen bond angles. In analogy to theoretical models used in protein folding studies,[45] it is possible to derive Helmholtz free energies of protein−ligand atom pair interactions, or PMF, from a database of protein−ligand complexes in aqueous solution at atomic resolution.

We are primarily interested in reliably ranking molecules according to their binding affinity to a biological target rather than predicting absolute binding energies. Therefore, we make the additional assumption that the sum over all protein−ligand atom pair interaction energies can be used as a good ranking measure for protein−ligand binding constants. Note, however, that there is no unique thermodynamic path to derive a total binding free energy of a protein−ligand complex based on the sum of protein−ligand atom pair interaction free energies. This is a conceptual disadvantage of our approach. We showed, however, that a very significant correlation between the sum of atom pair potentials and total binding free energy exists, and the sum is therefore a good measure for estimating binding affinities.

The idea of deriving appropriate PMF has been used before in protein folding studies by Sippl and co-workers for characterizing arbitrary atom pair interactions of protein atoms,[24] and we followed here Sippl's derivation with a few but essential changes due to the intermolecular nature of the protein−ligand interaction (see Appendix).

The protein−ligand interaction free energy (PMF) between a protein atom of type i and a ligand atoms of type j can be written as

$$A_{ij}(r) = -k_B T \ln\left[f^{j}_{\text{Vol\_corr}}(r) \frac{\rho^{ij}_{\text{seg}}(r)}{\rho^{ij}_{\text{bulk}}} \right] \quad (1)$$

where k_B is the Boltzmann constant, T is the absolute temperature, and r is the atom pair distance. $f^{j}_{\text{Vol\_corr}}(r)$ is the ligand volume correction factor, $\rho^{ij}_{\text{seg}}(r)$ is the number density of pairs of type ij in a structural database that occur in a certain radius range that is indicated by "seg", and ρ^{ij}_{bulk} represents the distribution of i and j when no interaction between i and j occurs. The quotient $\rho^{ij}_{\text{seg}}(r)/\rho^{ij}_{\text{bulk}}$ in eq 1 designates the pair correlation (radial distribution) function of a protein atom of type i paired with a ligand atom of type j in a structural database of protein−ligand complexes. The derivation of eq 1 and especially the definition of the ligand volume correction factor are given in the Appendix.

The scoring function is defined as the sum over all interatom interactions of the protein−ligand complex as

$$\text{PMF\_score} = \sum_{\substack{kl \\ r < r^{ij}_{\text{cut-off}}}} A_{ij}(r) \quad (2)$$

where $r^{ij}_{\text{cut-off}}$ is the cutoff radius for the atom type pair ij and we sum over all protein−ligand atom pairs kl in a database.

A conceptual advantage of our simplified potential approach is that solvation and entropic terms are treated implicitly. It is arguable, however, to what extent relevant desolvation of protein active site and of ligand, as well as the change of conformational entropy upon complex formation, are captured in this approach. Ligands bound to a protein binding site are usually not totally buried in the protein. A considerable part of the ligand is often exposed to the solvent. We attempt to take advantage of this fact by converting the degree of ligand penetration in the protein into implicit recognition of solvation effects. Since this is an important point of the approach we illustrate this idea considering the ligand atom $i2$ in Figure 1. We find that within a radius of 6 Å atom $i2$ is engulfed by either ligand atoms or protein atoms. That is, $i2$ cannot "feel" how deep it sits in the binding site within this radius. If we increase the radius, the space available for solvent molecules is increased. The ratio between the volume occupied by protein atoms and solvent atoms affects the density ρ^{ij}_{bulk} and therefore the depth of the PMF for each atom type pair ij (eq 1). The PMF become more pronounced (larger extrema) with a larger portion of $V(R)$ occupied by the solvent. As a result, a particular protein−ligand interaction at the protein surface is on average rated higher than one inside the protein. The overall PMF curve flattens with decreasing cutoff radius R. The effect of the solvent exposure is, of course, only taken into

account in an average sense, since a particular interaction of type *ij* can occur at the surface of the protein or inside the protein matrix as well. That is, the solvation effect in the *ij* potential arises from its average exposure to the solvent as compared to the same measure for all other *ij* pairs. Hence, the argument that solvation is included in the potentials arises from the use of a large cutoff of atom pair interactions that accounts for the presence or absence of protein atoms around the ligand. In other words, the small cutoff radii used by Verkhivker and co-workers[31] do not effectively capture solvation effects in the PMF.

As the PMF reflect Helmholtz free energies, entropic contributions are also captured implicitly. This includes folding entropy of protein and ligand in the complex conformation. However, it is not clear to what extent rotational and translational entropy contributions of the ligand are captured in our approach and whether there is still a need for additional solvation and/or entropy terms. Since it is crucial to include all relevant contributions to free energies in computational approaches,[44] we tried adding additional terms to our approach. We added solvation terms using the semi-microscopic protein dipole langevin dipole method[4] and terms that compensate for the loss of conformational freedom of the ligand upon complex formation.[46,47] We found that the added terms did not improve the correlation between calculated and observed binding free energies. Therefore, we omitted all additional terms and used a scoring function solely built on atom pair interaction PMF as given in eqs 1 and 2. Note that very recently a similar scoring function was proposed that combines PMF with an additional term for changes in the solvent-accessible surface area upon ligand binding.[32] However, results are not published yet and therefore cannot be compared to our scoring function at this time.

To relate the calculated PMF score (eq 2) to an absolute binding free energy we define a scaling factor ϵ that stands for all the different terms that are treated implicitly in our model,

$$\Delta G_{\text{bind}} = \text{PMF\_score}/\epsilon \qquad (3)$$

Implementation of the Method. A set of 16 protein and 34 ligand atom types was defined. The number of types is a compromise between detailed atom characterization and the ability to obtain reliable statistics to derive the potentials of mean force (PMF). The ligand and protein atom types are given in Tables 1 and 2. We chose entries from the Brookhaven protein database (PDB) for deriving the potentials of mean force (eq 1) by applying the following criteria. We searched for the keyword "complex" in the header, and then removed all entries that had a resolution below 2.5 Å, were complexed with RNA or DNA, were theoretical models, included ligands covalently bound to the protein, or were peptide inhibitors. This yielded 1269 entries. We took only the first of all those that had the same second and third letter of the four letter identification code of the PDB entry because they contain similar protein structures. This yielded 697 entries. (Comparing the potentials of mean force derived from the set of 1269 structures and those of the set of 697 structures did not show significant differences.) We then determined which atoms belonged to the protein and which to the ligand, using an automated process. Ligand and protein atom types were automatically assigned by an algorithm similar to that in BALI[48] (I. Muegge, unpublished data). If more than one protein was present and they were overlaid, the structure was discarded. If the ligand bond order could not be properly identified, the structure was discarded. If no atom pair of type *ij* was found in a certain segment "seg", the corresponding potential of mean force of this segment was set to 3 kcal/mol. This arbitrary value is slightly above the calculated PMF for atom pairs found at any distance. If the total number of occurrences of atom pairs of type *ij* in all segments was smaller than 1000, we set $A_{ij}(r) = 0$ kcal/mol in all segments. That is, we ignored the contributions of a particular pair type if it had statistically insufficient data.

The 697 protein structures from the PDB that were used for the derivation of the PMF are listed in the Supporting Information. The derived potentials of mean force are available as Supporting Information as well.

Results

Details of the Potentials of Mean Force. Table 3 shows the statistics of all protein−ligand atom pair types in detail. The number of atom pair occurrences varies greatly between the atom types. The largest number of atom pairs (250 217) occurred for the interaction between polar aliphatic carbons. For several ligand atom types there were not enough observations to derive reliable PMF. This includes spN (N0), spC (C0), Cl, Br, SD, and metals. To assess the quality of statistics of the PMF we compared the potentials for the interaction of *i* and *j* with those of *j* and *i* in case of a salt bridge. Figure 2 shows for a statistically moderate case that the curves for NCOC (the first two letters refer to the protein atom type; the last two to the ligand atom type) and OCNC are similar and that the minima occur at the same atom pair distance. The NCOC potential is slightly deeper than the OCNC potential. The occurrence of salt bridges is relatively rare compared to N−O hydrogen bonds or carbon−carbon interactions: the number of NCOC occurrences was 4258 and that of OCNC only 2788 compared with 32835 NDOA occurrences and 38354 cFcF occurrences. PMF curves of other protein−ligand pair types with higher occurrence rates and those where protein and ligand atom types were interchanged are even more indistinguishable (data not shown). Therefore, we conclude that the statistics are sufficient for the evaluation of the pair interaction energies. At any rate, it was found that the degree of deviation of the complementary PMF is the same as described by Sippl and co-workers.[24] Therefore we did not apply sparse data correction or smoothing algorithms.

Figure 2 shows examples of the potentials of mean force as derived from 697 complexes initially chosen from the Brookhaven protein database (about half of them were discarded due to different reasons as outlined in the Method section). There are several notable properties of the functions that can be observed from the figure. For the protein nitrogen donor−ligand oxygen acceptor interactions (NDOA), we see a valley of favorable interaction at a distance of ∼3 Å followed by a barrier of ∼1 kcal/mol at ∼4 Å. This pattern corresponds to the established hydrogen bond between these atom types. The apparent barrier reflects the preference for a hydrogen bond at 3 Å, which leaves few observed interactions at 4 Å. On the other hand, a salt bridge between a charged nitrogen and charged oxygen (NCOC) does not exhibit a barrier in the potential, reflecting the conventional wisdom that electrostatic interactions are not as sensitive to distances compared to hydrogen bonds. The minimum for the NCOC interaction lies at 3 Å whereas for the protein water−aliphatic carbon (OWCF) hydrophobic interaction we observe repulsive contributions up to a distance of 3.4 Å and a small favorable interaction at 4 Å. If one compares aromatic and aliphatic carbon−carbon interactions, one finds that aromatic carbon interactions tend to be favorable at shorter distances and their interaction energies are more favorable than those of aliphatic

Table 3. Logarithm of Protein−Ligand Atom Pair Occurrences in the Database[a]

ligand atom types	protein atom types															
	CP	CF	ND	OA	cF	OW	cP	HH	OC	OD	NC	CO	NR	CN	SA	SD
CP	5.4	5.2	5.1	5.1	4.8	4.8	4.4	4.6	4.3	4.3	4.1	4.0	3.9	3.9	3.3	3.2
CF	5.3	5.1	5.0	5.0	4.8	4.5	4.3	4.8	4.1	4.1	3.9	3.8	3.8	3.7	3.3	3.1
cF	5.1	5.0	4.8	4.8	4.6	4.3	4.1	4.5	4.0	4.0	3.8	3.7	3.6	3.5	3.1	3.0
OD	5.1	4.9	4.8	4.8	4.5	4.5	4.2	4.2	4.1	4.0	3.8	3.8	3.6	3.6	3.0	2.9
OS	5.1	4.9	4.8	4.7	4.4	4.4	4.1	3.9	4.0	4.0	3.9	3.7	3.6	3.7	3.0	3.0
cP	5.1	4.9	4.8	4.7	4.5	4.3	4.0	4.2	3.9	3.9	3.8	3.6	3.5	3.6	3.1	2.9
CW	4.9	4.7	4.6	4.5	4.3	4.1	3.9	4.3	3.7	3.7	3.6	3.4	3.4	3.4	2.9	2.7
OA	4.8	4.6	4.5	4.5	4.3	4.1	3.8	4.2	3.7	3.7	3.5	3.4	3.4	3.3	2.8	2.7
OC	4.8	4.6	4.5	4.4	4.2	4.1	3.8	4.0	3.7	3.7	3.6	3.4	3.4	3.4	2.9	2.9
HL	4.8	4.7	4.5	4.5	4.2	3.9	3.8	5.0	3.7	3.6	3.5	3.4	3.2	3.3	2.6	2.4
C3	4.7	4.6	4.4	4.4	4.3	3.8	3.8	4.3	3.5	3.6	3.3	3.2	3.3	3.1	2.8	2.6
NR	4.7	4.6	4.4	4.4	4.2	3.9	3.7	3.8	3.5	3.5	3.4	3.2	3.2	3.2	2.7	2.5
CO	4.6	4.4	4.3	4.2	3.9	3.9	3.6	3.7	3.4	3.4	3.4	3.1	3.2	3.2	2.6	2.6
OE	4.6	4.4	4.3	4.2	4.0	4.0	3.6	3.6	3.5	3.5	3.3	3.2	3.0	3.1	2.4	2.4
CN	4.6	4.4	4.3	4.2	4.0	3.8	3.6	3.3	3.5	3.4	3.3	3.2	3.0	3.1	2.6	2.6
NC	4.6	4.4	4.3	4.2	4.0	3.8	3.5	3.3	3.4	3.4	3.2	3.1	3.0	3.1	2.6	2.6
P	4.5	4.3	4.2	4.2	3.9	3.9	3.5	3.2	3.4	3.4	3.3	3.1	3.1	3.2	2.5	2.4
ND	4.4	4.3	4.1	4.1	3.9	3.7	3.5	4.2	3.3	3.2	3.2	3.0	3.0	3.0	2.5	2.3
NP	4.5	4.3	4.2	4.1	3.8	3.8	3.4	3.4	3.3	3.3	3.2	3.0	2.9	3.0	2.4	2.5
NS	3.9	3.7	3.6	3.5	3.4	3.0	3.0	3.0	2.7	2.8	2.5	2.4	2.6	2.2	1.5	1.7
F	3.5	3.3	3.2	3.1	3.0	2.7	2.4	2.8	2.5	2.3	1.8	2.2	2.0	1.5	1.9	1.4
OR	3.4	3.3	3.1	3.1	2.7	2.7	2.2	2.0	2.2	2.2	2.0	1.9	1.7	1.8	1.3	1.1
SA	3.1	3.0	2.8	2.8	2.5	2.3	2.2	1.0	2.1	1.9	1.7	1.8	1.7	1.5	0.7	1.0
NA	3.1	2.9	2.8	2.8	2.5	2.3	1.8	1.9	2.1	1.9	1.6	1.8	0.8	1.4	1.5	0.7
Br	2.9	2.6	2.6	2.5	2.3	2.1	1.9	0.0	1.7	1.9	1.4	1.4	1.3	1.1	1.0	1.5
C0	2.4	2.4	2.2	2.1	2.3	1.6	1.9	0.0	1.3	1.4	1.1	1.0	1.4	0.9	0.0	1.1
V	2.1	2.0	1.9	1.9	1.6	1.4	1.2	2.6	1.1	0.5	1.3	0.8	0.9	1.0	0.5	0.7
N0	2.2	2.0	1.9	1.8	2.0	1.3	1.4	0.0	1.1	1.1	0.8	0.8	0.9	0.5	0.3	0.0
Fe	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
CL	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Mn	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
SD	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Zn	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Mg	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

[a] $\log_{10}$ of the atom pair occurrences in the set of complexes used to derive the PMF. If the $\log_{10}$ is less than 3.0, the statistics for the atom type pair are considered insignificant and the interaction is ignored in calculating the score. Atom types are ordered by declining occurrences.

interactions. This indicates that an interaction of an aromatic ring is entropically favored over an aliphatic chain interacting with the protein. The interaction between two nitrogen hydrogen bond donors (NDND) was found to be insignificant, whereas that between aromatic polar carbon and charged oxygen (cPOC) is favorable at ∼3.6 Å.

A crucial part of our new approach is the introduction of the volume correction term (Appendix, eq 14). To show the large effect of this correction, Figure 3 shows the volume correction factor $f^{j}_{\mathrm{Vol\_corr}}(r)$ as a function of the atom pair distance of five different ligand atom types to an arbitrary protein atom type. For short distances (<4 Å) the correction factor is in the range 2−8. The effect of a correction factor greater than 1 is that the PMF become deeper compared to longer distances where $f^{j}_{\mathrm{Vol\_corr}}(r)$ approaches 1.0. That is, $f^{j}_{\mathrm{Vol\_corr}}(r)$ affects only the shape of the PMF and usually not the location of the extrema.

To include solvation effects implicitly we introduced a relatively large cutoff of 12 Å in deriving the PMF. As argued in the Method section, the solvation effect is captured in the ratio of protein to solvent volume occupied in a sphere of 12 Å radius around a given ligand atom *j*. The cutoff affects the reference state and therefore the deepness of the PMF. However, when we use the PMF to calculate a binding score we do not necessarily have to use the same cutoff. As can be seen from Figure 2, the statistics of different pair interactions differ. For instance, the NCOC potential becomes slightly positive for $r > 9$ Å. Since the number of interaction partners increases with r^3, this small deviation from zero would result in large positive contributions even though it is not quite clear how significant these contributions are. Hence, for scoring we use another cutoff that seemed to be optimal at 9 Å for the interaction of all non-carbon protein−ligand interactions. There is again no reason to apply the same cutoff radius to carbon−carbon interactions, and one could argue that hydrophobic interactions are not long-ranged. Therefore, we use a cutoff of 6 Å for carbon−carbon atom type interactions. Note that these cutoffs are only statistically motivated ("noise suppression") as the relevant PMF are shown to have reached approximately zero for all the different interactions at the given cutoff radii. Choosing different cutoffs up to 12 Å for all the PMF still showed reasonable but not as good correlation with the experiment for all the test cases described below.

Scoring Protein−Ligand Complexes. We applied the scoring function of eq 1 to six test sets of 77 different protein−ligand complexes taken from the Brookhaven PDB and an additional test set of 33 inhibitors that had been modeled into the binding site of the same protein. The results are presented in Table 4 and Figures 4−10. The complexes used are referenced in the figure by the PDB four-letter code (or reference number in case of HIV-1 protease). If there is more than one compound bound to the protein, we always used the one that was

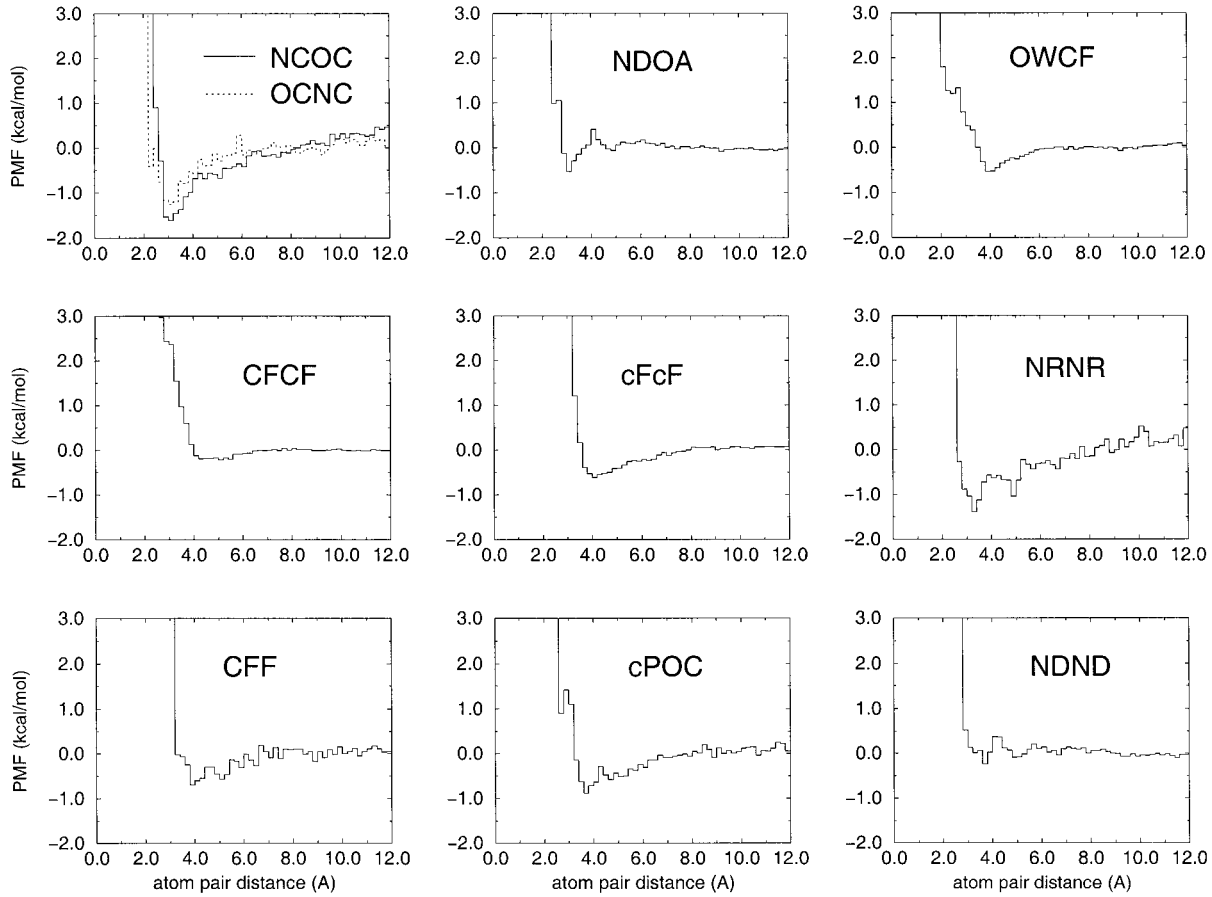

Figure 2. An arbitrary set of 12 PMF (out of a total of several hundred different PMF) is shown. The four-letter code refers to the atom pair interaction types, where the first two letters refer to the protein and the last two letters to the ligand (see Tables 1, 2, and the text). To demonstrate the statistical robustness of our method we show for the case of a salt bridge both NC as protein atom type and OC as ligand atom type (NCOC) and vice versa (OCNC). See text for further details.

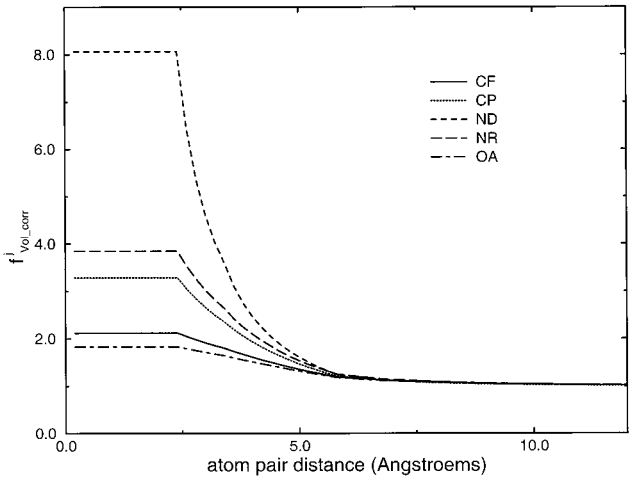

Figure 3. The volume correction term $f^{j}_{\mathrm{Vol\_corr}}(r)$ (eq 14) is shown for five indicated ligand atom types as a function of protein−ligand atom pair distance.

indicated in the given reference. Note that interactions involving hydrogen atoms and water molecules were omitted. In several cases the dependency of the results on the interaction with water molecules (as part of the protein) was included and found to be negligible. The influence of including hydrogen atoms was not analyzed in a systematic way as very few of the PDB complexes have explicit hydrogen atoms attached. Note that the neglect of hydrogen atoms is a strength of our approach as it reduces configurational sensitivity as elaborated in the Discussion section.

The first test set (set 1 in Table 4) contains 16 thrombin and trypsin complexes (Figure 4). The best scoring complex, 1tmt, is an outlier. When 1tmt is removed from the test set, the statistics are significantly improved: $R^2 = 0.92$, the standard deviation from the observed binding affinity (SD) becomes 0.73 log K_i units and the largest deviation (LD) is 1.5 log K_i units. Even with the outlier included, the statistics of set 1 are still the best we found, with $R^2 = 0.86$.

Set 2 (Figure 5, Table 4) contains 14 carboxypeptidase A and thermolysin complexes and 1 neutrophil collagenase complex. The removal of one outlier (the neutrophil collagenase complex 1mnc) improved the statistics dramatically. R^2 increased from 0.58 to 0.78, and the SD improved from 2.3 to 1.5 log K_i units.

Set 3 (Figure 6, Table 4) contains 18 L-arabinose binding protein complexes. Each of the nine crystal structures contains two different conformations of the sugar ligand that were treated separately. The correlation with observed data is not as high as for sets 1 and 2, but the SD is only about 0.86 log K_i units. This is a reflection of the fact that set 3 includes a narrower log K_i range of about 3 log K_i units compared with 7 and 10 log K_i units for sets 1 and 2, respectively.

Set 4 (Figure 7, Table 4) contains 11 endothiapepsin complexes that show a very poor correlation. One reason might be that the size of the inhibitors is relatively large

Table 4. Correlation between Experiment and PMF Score for Eight Test Sets[a]

no.	test set	figure	no. of complexes	R^2	SD[a]	LD	log K_i range	ϵ[b]
1	serine protease	4	16	0.87	0.96	2.65	7	10.8
1a	serine protease w/o 1 outlier		15	0.92	0.72	1.50	7	10.1
2	metalloprotease	5	15	0.58	2.31	6.42	10	10.0
2a	metalloprotease w/o 1 outlier		14	0.78	1.47	2.85	10	10.9
3	L-arabinose binding prot.	6	18 (9)[c]	0.48	0.86	1.43	3	5.5
4	endothiapepsin	7	11	0.22	1.89	3.33	4	6.3
5	others[d]	8	17	0.69	1.56	2.32	8	12.2
6	sets 1−5	9	77	0.61	1.84	5.01	12	12.8
6a	sets 1−5 w/o 1 outlier		76	0.64	1.70	4.64	12	13.3
6b	sets 1, 2, 3, 5 w/o 1 outlier		65	0.77	1.34	3.82	12	12.7
7	Böhm's training and test sets[e]		39	0.48	2.83	8.95	12	9.3
7a	Böhm's training and test sets[e] w/o 1 outlier		38	0.64	1.91	3.90	12	11.6
8	HIV-1 protease[f]	10	33	0.74	0.85	2.12	6	3.0

[a] The standard deviations (SD) and largest deviations (LD) from the observed affinities are given in log K_i units. [b] ϵ was calculated using eq 3. It represents a scaling factor of the PMF score and does not affect the statistics. ϵ serves to convert the PMF score into binding free energies. As such, ϵ represents all the contributions to the binding free energy that are treated implicitly in our simplified potential approach. With respect to the generality and predictability of absolute binding constants of our approach, it is especially instructive to compare the derived ϵ for different test sets. See discussion in the text. [c] The crystal structures of nine L-arabinose binding protein complexes contain two ligand conformations that were treated separately. [d] This set contains complexes of the combined training and test sets of Böhm's paper[10] that were not found in test sets 1−4 and show a resolution of 2.5 Å or better. [e] All available protein−ligand complexes of Böhm's training (30) and test (9) sets were used.[10] The remaining 15 complexes of his training set were modeled by Böhm and therefore were not available to us. [f] The 33 inhibitors taken from Holloway and co-workers[50] were placed and minimized in the active site of the crystal structure of the L-689,502-inhibited HIV-1 protease by using the Merck force field.[50]

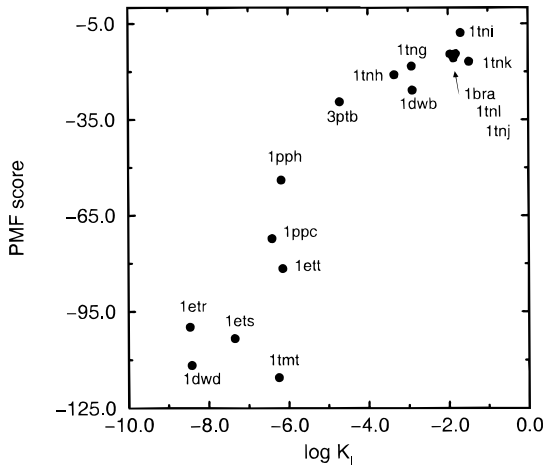

Figure 4. Calculated PMF score as a function of the observed binding affinity. A set of 16 serine protease complexes taken from the PDB is shown. The experimental data are taken from Tables 2 and 6 of ref 14. The results of the statistical analysis are given in Table 4 (set 1).

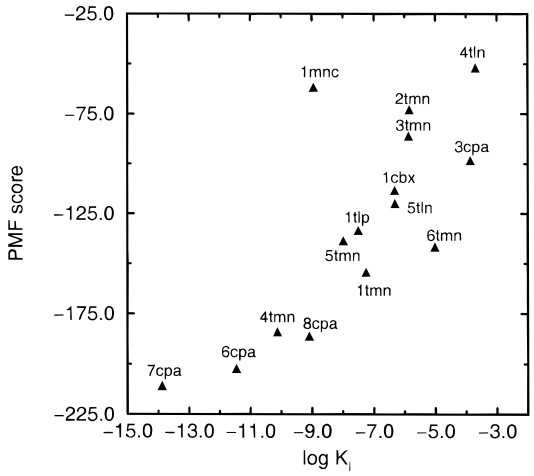

Figure 5. Calculated PMF score as a function of the observed binding affinity. A set of 15 metalloprotease complexes taken from the PDB is shown. The experimental data are taken from Table 3 of ref 14. The results of the statistical analysis are given in Table 4 (set 2).

and differs considerably. In general, a significant correlation was found between the number of heavy atoms of the ligand and the reciprocal PMF score that allows for the high correlation between the PMF score of ligands of different size and observed binding constants.[49] However, for large inhibitors the absolute difference in the number of heavy atoms is also large and this may lead to unreliable PMF scoring.

Set 5 (Figure 8, Table 4) consists of 17 protein−inhibitor complexes containing cytochrome P450, FKBP, HIV−protease, renin, DHFR, galactose binding protein, thymidylate synthase, retinol binding protein, TIM, myoglobin, and concanavalin as protein targets. They were chosen by taking those complexes of the combined training and test sets of Böhm[10] that were not present in our sets 1−4 and that have a resolution of 2.5 Å or better. This diverse set exhibits a good correlation between calculated PMF score and observed binding affinities. If we include also those complexes that have crystal resolutions below 2.5 Å (data not shown), we find them to fit nicely into the plot of Figure 8 with one drastic outlier, biotin bound to streptavidin (1stp), that scored about 6 log K_i units less than the observed binding affinity. This is due to the fact that the calculations were done by using the 1stp structure from the PDB that contains only a single streptavidin with a biotin. Streptavidin, however, is a tetramer, and interactions with a second subunit increase the binding of biotin by 8 orders of magnitude.

To compare our scoring function to the most widely used empirical scoring function we analyzed all 39 complexes of the combined training and test sets of Böhm[10] (set 7 of Table 4) that were available to us (his training set contains 15 modeled complexes unavailable to us). It should be pointed out here that we were not aware of an improved scoring function by Böhm[16] that was published only recently and therefore could not be taken into consideration. We found $R^2 = 0.48$ and $R^2 = 0.64$, neglecting the biotin/streptavidin outlier (data not shown). The latter correlation is only slightly worse than

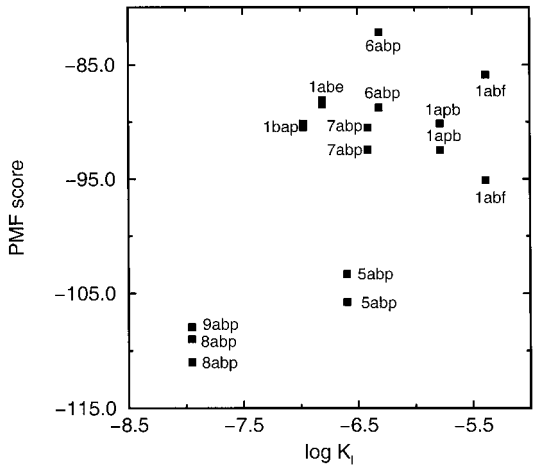

Figure 6. Calculated PMF score as function of the observed binding affinity. A set of 18 L-arabinose binding protein complexes taken from the PDB is shown. The experimental data are taken from Table 3 of ref 14. The results of the statistical analysis are given in Table 4 (set 3).

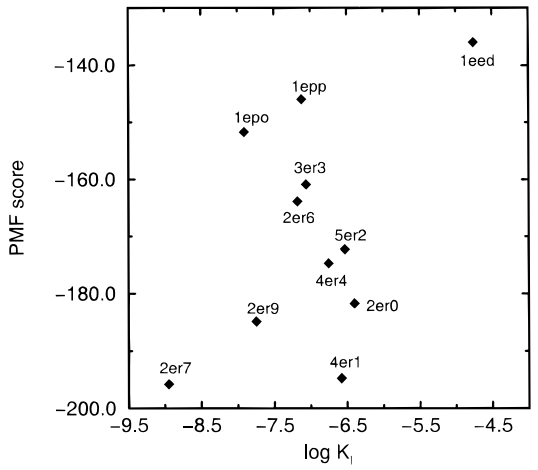

Figure 7. Calculated PMF score as function of the observed binding affinity. A set of 11 endothiapepsin complexes taken from the PDB is shown. The experimental data are taken from Table 6 of ref 14 and Table 1 of ref 13. The results of the statistical analysis are given in Table 4 (set 4).

the one found by using Böhm's scoring function ($R^2 = 0.69$) for these 39 complexes that include 30 complexes of Böhm's training set. Since our approach does not involve any fitting to known binding affinities we find this result very encouraging. The comparison between Böhm's scoring function and the PMF score is further discussed below.

To further evaluate the generality of our scoring function with respect to diverse sets of complexes, we combined sets 1−5 into set 6 of 77 diverse protein−ligand complexes (Figure 9, Table 4). Figure 9 shows a surprisingly high correlation considering the fact that the heavy atom numbers of the ligands range between 5 (1mbi) and 94 (4phv) and, for instance, 1fkf with 57 heavy ligand atoms is still predicted to bind better than 4phv in agreement with experiment. The overall results suggest that one should not use per heavy atom scores as used in additive statistical methods[30] since the effect of having more atoms in a ligand is compensated for here by implicit solvation effects due to the large cutoff in deriving the PMF. Although per heavy atom scores correlate with the experimental binding constants in our

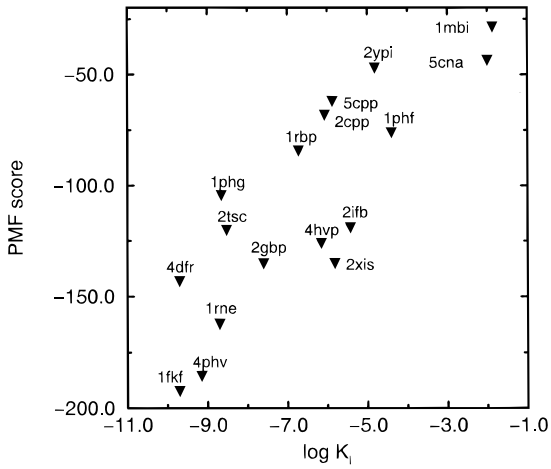

Figure 8. Calculated PMF score as a function of the observed binding affinity. A set of 17 different protein−ligand complexes taken from the PDB is shown. This set contains all compounds from the combined training and test sets of Böhm[10] that are not included in the sets 1−4 (Figures 4−7) and have a crystal structure resolution less than 2.5 Å. The experimental data are taken from Tables 1 and 3 of ref 10. The results of the statistical analysis are given in Table 4 (set 5).

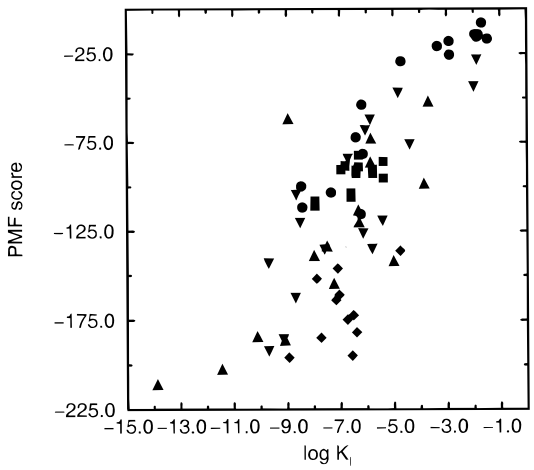

Figure 9. Calculated PMF score as function of the observed binding affinity. The combined sets of Figures 4−8 are shown (77 protein−ligand complexes taken from the PDB). The symbols correspond to those used in Figures 4−8. That is, ● depicts serine protease complexes, ▲ depicts metalloprotease complexes, ■ depicts L-arabinose binding proteins, ◆ depicts endothiapepsin complexes, and ▼ depicts the complexes of set 5. The results of the statistical analysis are given in Table 4 (set 6).

test sets, the correlation is found to be generally much lower (R^2 lies between 0.2 and 0.4).

Set 8 (Figure 10, Table 4) consists of 33 ligands taken from Table 1 of a study by Holloway and co-workers modeled into the crystal structure of the L-689,502-inhibited HIV-1 protease.[50] A significant statistical correlation between observed binding affinities and PMF scores was found ($R^2 = 0.74$) with a low standard deviation of 0.8 log K_i units from observed binding constants. Set 8 contains 2 subsets of inhibitors that vary at only one substitution site each. The results suggest that the PMF score is especially reliable for screening similar inhibitors against the same target.

The last column of Table 4 shows the scaling constant ϵ for all the test sets. ϵ was iterated using a regula falsi method such that the calculated ΔG_{bind} (eq 3) and the

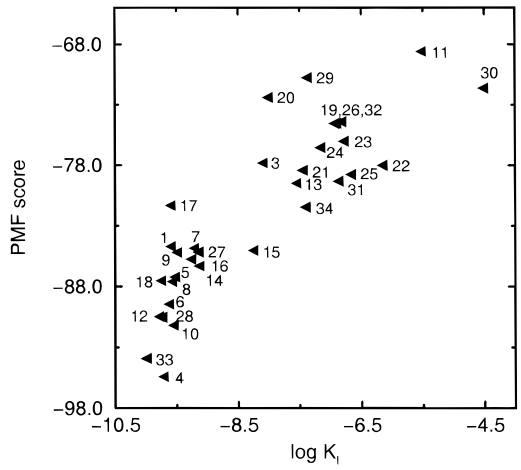

Figure 10. Calculated PMF score as a function of the observed binding affinity. A set of 33 ligands was placed in the active site of the L-689,502-inhibited HIV-1 protease structure and minimized by the using the Merck force field.[50]

observed binding affinities showed the same SD and LD from the linear regression line for the given test set. While ϵ was found to be very similar for sets 1, 2, 5, 6, and 7, it was found to be quite different from those of sets 3, 4, and 8. The performance of the scoring function on sets 3 and 4 is poor combined with narrow log K_i ranges of the test sets. This leads to unreliable regression lines and may explain the large deviations of ϵ for these two test sets compared to ϵ of the test sets 1, 2, 5, 6, and 7 that show similar ϵ. However, this does not explain the different ϵ for test set 8. Possible reasons are that the HIV-1 protease inhibitors were modeled into the protein and not placed optimally with respect to the PMF score. They also represent smaller structural changes between the different ligands. Although the correlation between measured and predicted log K_i's is good in test set 8, this finding indicates that the generality of the approach may be limited. In other words, the scoring function is expected to be more reliable for the screening of different ligands against the same target or for comparing different ligand (and/or protein) configurations in docking studies as suggested above. Without knowledge of any observed binding affinities, the scoring function is able to rank protein−ligand binding affinities but not reliable in giving estimates for absolute binding free energies.

In Table 5 we compare our scoring function to the empirical scoring function of Böhm[10] and the knowledge-based scoring function of SMOG.[30] The PMF score performs best for most of the test sets. This finding indicates the PMF score function is more general than the empirical scoring function of Böhm. It also indicates that a distance-dependent knowledge-based potential, as compared to the coarse-grained potential of SMOG, is a powerful feature of our approach that leads to better correlation between observed and calculated binding constants. These results are discussed in more detail below.

Discussion

The generality of the simplified potential approach presented above is relative, of course, because the Brookhaven database can be considered as the "training set" for the scoring function. The most important difference from training sets of other scoring functions (besides its size) is, however, that structural information is converted into free energies without any knowledge of binding affinities and without any fitting procedures. The only assumption made is that all the complexes contain bound ligands crystallized from water, and this is met since only experimentally measured structures from the PDB were used.

Although hydrogen atom types can be used in our approach they were omitted in all test sets. The reason for that is twofold. First, most of the complexes in the PDB contain either no hydrogen atoms or mainly polar hydrogen atoms necessitating that most hydrogen atoms would have to be modeled. Second, hydrogen atoms are expected to be much more sensitive to position changes. Specifically, scoring functions having penalty terms for hydrogen bond angles, such as that of Böhm,[10] are found to be sensitive to hydrogen positions and sometimes require minimization procedures to achieve the appropriate configurations. To what extent this insensitivity of our approach might lead to problems in finding the right configuration in docking studies remains to be seen. It should be pointed out, however, that docking studies using the PMF scoring function seem to be more reliable than, for instance, force field scoring in DOCK4. The PMF approach presented here has been implemented as an optional scoring function in DOCK4.[18] This docking program is especially created to allow the user to implement customized scoring functions. One apparent advantage lies in the implementation of a simplex algorithm that is not based on the evaluation of gradients for the configurational optimization. Therefore, our scoring function can be directly implemented in DOCK4 in a straightforward way to account for all intermolecular interactions. Note that for docking purposes the PMF at close distances are replaced by van der Waals potentials to avoid unfavorable steric interactions in the active site. Also, intramolecular van der Waals interactions of the ligand are added for docking purposes. However, the use of these additions to the scoring function are not needed for scoring purposes of reasonable protein−ligand complexes and are therefore not discussed here in detail but presented elsewhere.[49]

The idea of using simplified potentials for the prediction of binding constants is not new. There is at least one example reported where PMF are used as part of a scoring function.[31] In principle one can derive PMF from the much larger body of small-molecule data stored in the Cambridge database. One of our findings is that the definition of an appropriate reference state and the use of large cutoff radii are pivotal to derive physically meaningful Helmholtz free energies (PMF) that treat solvent and entropic effects implicitly. The definition of such a reference state for crystal structures of small molecules is not obvious and maybe not possible due to crystal contacts at relatively short distances. The need of introducing a volume correction factor that accounts for the ligand volume (intraligand interactions are omitted) was only acknowledged recently[51] and was found here independently. Taking these crucial elements of a PMF approach into account leads to results with significant correlation to experimental data as demonstrated in this paper. On the other hand, the use of crystallographic data of small molecules is useful in

Table 5. Comparison between PMF Score, Böhm's Score, and SMOG[a]

			PMF score[b]		Böhm's score		SMOG[d]	
no.	test set	no. of complexes	R^2	SD[d]	R^2	SD	R^2	SD
1	serine protease	16	0.87	0.96	0.76	1.39	0.76	1.34
1a	serine protease w/o 1 outlier	15	0.92	0.72	0.87	0.95	0.80	1.25
2	metalloprotease	15	0.58	2.31	0.41	3.27	0.58	2.29
2a	metalloprotease w/o 1 outlier	14	0.78	1.47	0.71	1.39	0.59	2.21
3	L-arabinose binding prot.	18	0.48	0.86	0.00	69.7	0.04	4.06
4	endothiapepsin	11	0.22	1.89	0.39	1.26	0.05	4.18
5	others	17	0.69	1.56	0.53[e]	2.21[e]	0.25	4.05
6	sets 1−5	77	0.61	1.84	0.30[f]	3.47[f]	0.21	4.43
6a	sets 1−5 w/o 1 outlier	76	0.64	1.70	0.34	2.91	0.23	4.41
6b	sets 1, 2, 3, 5 w/o 1 outlier	65	0.77	1.34	0.35	3.04	0.23	4.46
7	Böhm's training and test sets	39	0.48	2.83	0.69[e]	1.85[e]	0.21	5.27
7a	Böhm's training and test sets w/o 1 outlier	38	0.64	1.91	0.71[e]	1.77[e]	0.29	4.32
8	HIV-1 protease	33	0.74	0.85	0.46	1.62	0.08	4.98

[a] Sets listed as in Table 4. Standard deviations (SD) of calculated score from observed binding constants are given in log K_i units. [b] PMF score results are taken from Table 4. [c] Böhm's scoring function as implemented in the Ligand design module of the program package INSIGHT (MSI) was used unless indicated otherwise. [d] The SMOG scoring function was implemented using the Supporting Information of ref 30. Note that we report here the *total* score as calculated with SMOG. The score per heavy atom (as suggested and used in the original paper) yields results totally uncorrelated with observed binding constants for all test sets (data not shown). [e] These data were calculated using the reported scores from Böhm's paper.[10] [f] Since the ligand design score calculated for sets 1−4 is different from the score reported for set 5, we had to convert the scores of sets 1−4 with respect to the score of set 5. To do so we used eq 3, replaced the PMF score with Böhm's score, and calculated a factor $1/\epsilon$ using a regula falsi method such that the standard deviations of the observed binding constants and calculated scores were equalized.

deriving optimal interaction distances between atom types; in fact this was used for deriving empirical atom pair interaction potentials in programs such as GRID.[52] Indeed there might be a way to combine the advantages of crystallographic data of small molecules and protein−ligand complexes for designing a better scoring function, but this is, again, beyond the scope of this work.

The work of Verkhivker and co-workers[31] is closest to our approach. There are, however, major differences. Verkhivker and co-workers[31] reported a free energy function including knowledge-based ligand−protein interaction patterns of HIV-1 protease complexes based on ideas of using potentials of mean force (PMF) in studies of protein folding.[23−29] They derived atom interaction PMF using a small set of 30 HIV and SIV protein−ligand complexes and only 12 different atom pair types. Since they sought to treat solvation effects explicitly, they used a small cutoff of 6−7 Å for the PMF. Interactions that were found to be occluded by water molecules were not taken into account. There was no volume correction reported. By combining these pair interaction potentials with other explicit entropic and solvation terms they found good correlation between calculated and observed binding free energies for 7 HIV-1 inhibitors. Comparing the PMF alone with the observed binding free energies of these seven complexes showed no positive correlation. Verkhivker and co-workers argued that the predictive power of the PMF approach decreases with increasing number of parameters (different atom pair interaction types). We found, however, that with a relatively large number of parameters (several hundred) derived from a large database the predictive power of the PMF approach is actually enhanced compared to using a smaller set of atom pair interaction types (data not shown). This suggests that an optimal number of parameters for a certain problem exists and that the number is larger than 100. The PMF approach may be improved by employing different atom type definitions. The number of different atom types can probably not be increased, however, since the statistics already exclude several atom pair types due to an insufficient number of occurrences in the PDB.

One encouraging finding is that with smaller changes in the inhibitor and with smaller ranges of log K_i for the screened compounds, the statistical predictability of our scoring function increases. That is, this scoring function is capable of identifying micromolar leads by screening large databases. This is much harder to accomplish than distinguishing nanomolar and micromolar compounds in a database. In this respect, this result is of importance in applying the scoring function as part of a docking program in virtual screening.[49]

Table 5 shows a direct comparison between the PMF score, Böhm's scoring function, and the SMOG scoring function. Taking the standpoint that a correlation between calculation and observation is significant if $R^2 > 0.5$, five of eight sets show significant correlation to observed binding constants, for Böhm's scoring function only three of eight, and for SMOG only two of eight. If we consider all subsets of Table 5, we find significance in 10 of 13, 6 of 13, and 4 of 13 sets for the three scoring functions, respectively. In terms of the standard deviations of the calculated from the observed binding constants, we find that in five of eight sets the PMF score shows the lowest SD, in two sets Böhm's scoring function shows the lowest SD, and in one set the SMOG scoring function gives the lowest SD. Böhm's scoring function performs best for set 7 that contains 30 of 39 complexes to which it is fitted. Although Böhm's scoring function also shows the highest correlation in case of 11 endothiapepsin complexes (set 4), compared to the other scoring functions this correlation is statistically insignificant. For a set of 33 modeled HIV-1 protease inhibitors that differ only at two substitution sites, Böhm's scoring function performs significantly worse than the PMF score. Set 6, which contains 77 different protein−ligand complexes, gives the best estimate of the generality of the different scoring functions. Only the PMF score shows a significant correlation in this case while Böhm's method and SMOG fail. Another important test set is number 8, which shows how reliably different modeled ligands are scored. Again, the PMF score is the only scoring function that shows a significant correlation to observed binding constants. While

Böhm's scoring function almost reaches significance, SMOG totally fails. With respect to the argument that one can fit a scoring function like Böhm's to a given protein, it is instructive to comment again on set 6 which contains 30 of the 39 complexes Böhm's scoring function was fitted to. Here, Böhm's scoring function is the only one showing significant correlation with experiment; the PMF score almost reaches significance, and SMOG fails. However, removing one extreme outlier (1stp) in the PMF score (set 7a) shows a much improved correlation of the PMF score (Note that the streptavidin/biotin outlier can be explained by the fact that 1stp has only 1 streptavidin/biotin whereas streptavidin comes as a tetramer and the biotin interaction with other parts of the tetramer increases the binding affinities by 8 orders of magnitude). In fact, the results are comparable to Böhm's scoring function, in both correlation and standard deviation.

The predictability of the PMF score was found to be on the order of 1−2 log K_i units for diverse subsets of 77 different protein−ligand complexes. In comparison, standard deviations from observed binding constants on the order of 1 log K_i unit are reported for the training sets of the best empirical scoring functions.[12−14] However, since these studies do not present large test sets we do not exactly know how to compare our findings to theirs. At any rate, however, a standard deviation of more than 1 log K_i unit would probably not be good enough for scanning a database in search for micromolar leads. Fortunately, for a test set of 33 HIV-1 protease inhibitors modeled to the same protein, the predictability was significantly less than 1 log K_i unit. This result is very encouraging, and we hope it is representative for the screening of large sets of compounds against a particular protein target.

Changing the number of protein−ligand complexes in the protein database the PMF were derived from did not change the PMF significantly. Also including or excluding the numerous peptidic ligands in the PDB did not change the PMF significantly. However, it is not clear to what extent the peptidomimetic character of many ligands in the PDB biases the derived potentials. The quality of the scoring function can probably not be improved significantly within a few years by simply waiting for more structures to be deposited in the PDB or proprietary databases. There is, however, room for improvement coming from other directions. First, since the number of atom type occurrences varies over several orders of magnitude (Table 3), the choice of atom types assigned to the protein and ligand atoms is obviously statistically not optimal. That is, merging, creating, and deleting atom types may improve the scoring function. Second, there are several ligand atom types that do not occur with sufficient frequency to provide a significant potential of mean force (Table 3). In these cases one could derive empirical potentials from structural data of small molecules as is done in the GRID program.[52] However, this was not attempted here and needs further study.

There are other problematic points of the PMF score. For instance, there is no measure for the directionality of hydrogen bonds. Only the distances between the atoms govern the shape of the respective PMF. However, it is possible to derive PMF not only for distances but for certain angles as well (hydrogen bond directionality). We did not pursue this idea since it becomes more complicated to combine all the different types of PMF to a meaningful measure of binding constants in which one might be forced to use multivariate fitting methods (as in empirical scoring functions) to fit these terms to known binding affinities. Our goal, however, was to derive a scoring function solely built on structural information.

Conclusion

A knowledge-based simplified potential approach is presented to rank the binding affinities of protein−ligand complexes of given 3D structure. It is able to estimate the binding energies of sets of different ligands bound to the same protein target with good correlation to observed binding constants. The correlation of calculated binding scores of diverse sets of protein−ligand complexes is also significant and comparable to those reported for empirical scoring functions (e.g., Böhm[10]). Comparing the newly introduced PMF score for eight test sets containing 77 different protein ligand complexes of the Brookhaven protein database and 33 modeled HIV-1 protease inhibitors to Böhm's score and the SMOG score,[30] the PMF score performed best. This finding is very encouraging since, in contrast to empirical scoring functions, the simplified potential approach does not involve any fitting of parameters of the scoring function to observed binding affinities. Thus, the presented scoring function eliminates the major uncertainty of empirical scoring functions as to what extent the latter can be applied to protein−ligand complexes that are not represented in the training set used to derive the empirical scoring function.

Our scoring function is as fast as empirical scoring functions; therefore, it can be used for docking large sets of compounds into a protein binding site. The design of fast and reliable scoring functions presents a real challenge for docking programs today and appears to be the bottleneck for getting reliable conformations of protein−ligand complexes with given 3D protein structure. We hope that our scoring function will prove to be very helpful in solving this long-standing issue. We have implemented our scoring function into the DOCK4 program.[18] Preliminary studies show that the PMF score performs better than force field scoring in DOCK4 in the case of docking a database of small molecules with measured binding affinities to the FK506 binding protein.[49]

Acknowledgment. The authors thank Dr. Mark G. Bures for helpful discussions and critically reading the manuscript.

Appendix

Volume-Corrected Atom Pair Interaction Free Energies. The Helmholtz free energy $A(r)$ can be calculated from pair distribution functions $g(r)$ by[53]

$$A(r) = -k_B T \ln g(r) \tag{4}$$

where k_B is the Boltzmann constant, T is the absolute temperature, and r is the atom pair distance. The pair distribution function $g_{ij}(r)$ for a protein−ligand atom pair of type ij can be calculated from the number density

$\rho_{ij}(r)$ of occurrences of pairs of type ij at a certain distance r in a database of protein−ligand complexes

$$\rho_{ij}(r) = \sum_{pkl} \delta_{ij}(r - r_{pkl}) \tag{5}$$

where we sum over all protein−ligand complexes p of the database and all protein−ligand atom pairs kl of type ij. δ_{ij} designates the δ function. The pair distribution function of a protein atom of type i paired with a ligand atom of type j reads

$$g_{ij}(r) = \rho_{ij}(r)/\rho^{ij}_{\text{bulk}} \tag{6}$$

The bulk density ρ^{ij}_{bulk} represents the distribution of i and j when no interaction occurs. The implementation of number densities in our computational approach is done by evaluating the following expressions

$$\rho^{ij}_{\text{bulk}} = \sum_{pkl} \frac{n^{ij}_{\text{bulk}}}{V(R)} \quad \text{and} \quad \rho^{ij}_{(r,r+\Delta r)}(r) = \sum_{pkl} \frac{n^{ij}_{(r,r+\Delta r)}}{V_{(r,r+\Delta r)}(r)} \tag{7}$$

where $\rho^{ij}_{(r,r+\Delta r)}(r)$ is the number density $\rho_{ij}(r)$ in a spherical shell of thickness Δr that stretches between radii r and $r + \Delta r$. n^{ij}_{bulk} and $n^{ij}_{(r,r+\Delta r)}$ are the numbers of protein−ligand atom pair occurrences of type ij in the sphere with radius R and volume $V(R) = {}^4/_3\pi R^3$ and in the spherical shells with volume $V_{(r,r+\Delta r)}(r) = {}^4/_3\pi((r + \Delta r)^3 - r^3)$, respectively.

In order to calculate the distribution of j if no interaction with i occurs, one would have to treat the protein in its complex conformation but internally relaxed to its unbound state. This hypothetical protein state is not accessible in a reasonable way. Therefore, we calculate ρ^{ij}_{bulk} as in eqn 7, assuming that no interactions occur. This is reasonable since the most relevant short-range interactions are also the most infrequent ones and the bulk of the ij pairs stem from larger distances ($V \propto r^3$) where i and j can be considered to be noninteracting. ρ^{ij}_{bulk} serves as the reference density that guarantees that the potentials of mean force will approach zero with increasing distance between the atoms. To calculate ρ^{ij}_{bulk}, the volume $V(R)$ of the sample has to be given. Since the number density $\rho_{ij}(r)$ is calculated by analyzing a library of protein−ligand structures up to a cutoff distance R, the volume would be that of a sphere with radius R (eq 7). However, we focus on intermolecular interactions between the protein and the ligand. For a given ligand atom the interaction energy to all other ligand atoms is not considered. Therefore, the spherical reference volume needs to be corrected by eliminating the volume of the ligand itself. (In principle, we could take the intraligand interactions into account, but it is not clear how to include bonded interactions between ligand atoms. If one decides to exclude them, one would face conceptually the same problem of excluding parts of the ligand. Therefore, it is better to exclude the entire intraligand interaction and to introduce a reference volume correction that is computationally easy to implement, as outlined below.) This leads to significant changes in the potentials of mean force, similar to the very recently reported effect of solute volume corrections on pair distribution functions in solution.[51]

In order to derive an atom pair distance and ligand atom type dependent volume correction factor $f^{j}_{\text{Vol\_corr}}(r)$ averaged over all protein−ligand complexes in the database, we define the following terms. The sphere with radius R is dissected into spherical shells $s(r)$ (s = 1, 2, .., $R - 3$), where the first shell is actually a sphere with radius 4 Å and the other shells have a thickness of 1 Å (Figure 1). This coarse separation ensures that we have a reasonable number of ligand atoms in each occupied shell. The average numbers of non-hydrogen protein atoms in the volume $V(R)$ and in every shell $s(r)$ are defined as $\langle n^{\text{P}}_{\text{bulk}}\rangle$ and $\langle n^{\text{P}}_{s}\rangle$, respectively. $\langle n^{\text{P}}_{\text{bulk}}\rangle$ was evaluated by scanning the protein database and averaging over different origins of a sphere with radius R and different proteins, making sure that the entire sphere was filled with protein atoms each time. The $\langle n^{\text{P}}_{s}\rangle$ were then calculated by the volume ratios of $s(r)$ and $V(R)$ (Note that $\langle n^{\text{P}}_{s}\rangle$ is usually not an integer). $n^{\text{L}}_{s}(pl)$ and $n^{\text{L}}_{\text{bulk}}(pl)$ designate the number of ligand atoms l of the protein−ligand complex p in the shell s and in the sphere with radius R, respectively. The volume correction terms v^{pl}_{s} and v^{pl}_{bulk} were calculated for every ligand atom l of each complex p in the shell s and in the sphere with radius R, respectively,

$$v^{pl}_{s} = \frac{\langle n^{\text{P}}_{s}\rangle - n^{\text{L}}_{s}(pl)}{\langle n^{\text{P}}_{s}\rangle} \quad \text{and} \quad v^{pl}_{\text{bulk}} = \frac{\langle n^{\text{P}}_{\text{bulk}}\rangle - n^{\text{L}}_{\text{bulk}}(pl)}{\langle n^{\text{P}}_{\text{bulk}}\rangle} \tag{8}$$

In order to calculate the volume correction terms for specific ligand atoms of type j, we build the average over all ligand atoms l that are of type j and write

$$v^{pj}_{s} = \frac{1}{N_{l_j}} \sum_{l_j} v^{pl}_{s} \quad \text{and} \quad v^{pj}_{\text{bulk}} = \frac{1}{N_{l_j}} \sum_{l_j} v^{pl}_{\text{bulk}} \tag{9}$$

with l_j indicating that l is of type j. We average over all the protein−ligand structures p and get a mean ligand atom type volume correction in every shell s and in the bulk, respectively,

$$v^{j}_{s} = \frac{1}{N_p} \sum_{p} v^{pj}_{s} \quad \text{and} \quad v^{j}_{\text{bulk}} = \frac{1}{N_p} \sum_{p} v^{pj}_{\text{bulk}} \tag{10}$$

In order to smooth the coarse shell model the following simple procedure was applied. We define another set of spherical shells called segments seg(r) (seg = 1, 2, .., R/m) with thickness m consecutively separating our sphere with radius R into R/m segments. m was chosen to be 0.2 Å. We assign the v^{j}_{s} to the corresponding v^{j}_{seg} and write

$$\hat{v}^{j}_{\text{seg}} = \frac{1}{17} \sum_{\text{seg}-8}^{\text{seg}+8} v^{j}_{\text{seg}} \tag{11}$$

The volume-corrected bulk number densities of protein−ligand atom pair interactions and the number density in a segment are

$$\hat{\rho}^{ij}_{\text{bulk}} = \frac{1}{v^j_{\text{bulk}}} \frac{\int_0^R \rho_{ij}(r)\, \mathrm{d}r}{\int_0^R 4\pi r^2\, \mathrm{d}r} = \frac{1}{v^j_{\text{bulk}}} \rho^{ij}_{\text{bulk}} \quad \text{and}$$

$$\hat{\rho}^{ij}_{\text{seg}}(r) = \frac{1}{\hat{v}^j_{\text{seg}}} \frac{\int_{r=(\text{seg}-1)m}^{\text{seg}m} \rho_{ij}(r)\, \mathrm{d}r}{\int_{r=(\text{seg}-1)m}^{\text{seg}m} 4\pi r^2\, \mathrm{d}r} = \frac{1}{\hat{v}^j_{\text{seg}}} \rho^{ij}_{\text{seg}}(r) \quad (12)$$

respectively. Using eqs 6 and 12 the volume-corrected pair distribution functions are

$$g_{ij}(r) = \frac{\hat{\rho}^{ij}_{\text{seg}}(r)}{\hat{\rho}^{ij}_{\text{bulk}}} = \frac{v^j_{\text{bulk}}}{\hat{v}^j_{\text{seg}}} \frac{\rho^{ij}_{\text{seg}}(r)}{\rho^{ij}_{\text{bulk}}} = f^{\,j}_{\text{Vol\_corr}}(r) \frac{\rho^{ij}_{\text{seg}}(r)}{\rho^{ij}_{\text{bulk}}} \quad (13)$$

and we find the distance and ligand atom type dependent volume correction factors of the pair distribution functions to be

$$f^{\,j}_{\text{Vol\_corr}}(r) = v^j{}_{\text{bulk}}/\hat{v}^j_{\text{seg}} \quad (14)$$

Using eqs 4 and 13 the protein−ligand interaction free energy between atom types i and j is

$$A_{ij}(r) = -k_{\text{B}}T \ln \left[f^{\,j}_{\text{Vol\_corr}}(\text{r}) \frac{\rho^{ij}_{\text{seg}}(\text{r})}{\rho^{ij}_{\text{bulk}}} \right] = -k_{\text{B}}T \ln \rho^{ij}_{\text{seg}}(r) - k_{\text{B}}T \ln f^{\,j}_{\text{Vol\_corr}}(r) + k_{\text{B}}T \ln \rho^{ij}_{\text{bulk}} \quad (15)$$

Note that while the normalization constants ρ^{ij}_{bulk} have no effect on energy differences between two states, the introduction of the volume correction factors $f^{\,j}_{\text{Vol\_corr}}(r)$ changes the actual shape of the free energy function. $\ln f^{\,j}_{\text{Vol\_corr}}(r)$ approaches zero with increasing r.

References

(1) Greer, J.; Erickson, J. W.; Baldwin, J. J.; Varney, M. D. Application of the three-dimensional structures of protein target molecules in structure-based drug design. *J. Med. Chem.* **1994**, *37*, 1035−1054.

(2) Martin, Y. C. Challenges and prospects for computational aids to molecular diversity. *Perspect. Drug Discovery Des.* **1997**, *7/8*, 159−172.

(3) Kollman, P. Free energy calculations-applications to chemical and biological phenomena. *Chem. Rev.* **1993**, *7*, 2395−2417.

(4) Lee, F. S.; Chu, Z. T.; Warshel, A. Microscopic and semimicroscopic calculations of electrostatic energies in proteins by the POLARIS and ENZYMIX programs. *J. Comput. Chem.* **1993**, *14*, 161−185.

(5) Aqvist, J.; Medina, C.; Samuelsson, J. E. New method for predicting binding affinity in computer-aided drug design. *Prot. Eng.* **1994**, *7*, 385−391.

(6) Muegge, I.; Schweins, T.; Langen, R.; Warshel, A. Electrostatic control of GTP and GDP binding in the oncoprotein p21ras. *Structure* **1996**, *4*, 475−489.

(7) Sharp, K. A.; Honig, B. Electrostatic interactions in macromolecules. *Annu. Rev. Biophys. Biophys. Chem.* **1990**, *19*, 301−332.

(8) Williams, D. H.; Cox, J. P. L.; Doig, A. J.; Gardner, M.; Gerhard, U.; Kaye, P. T.; Lai, A. R.; Nicholls, I. A.; Salter, C. J.; Mitchell, R. C. Toward the semiquantitative estimation of binding constants. Guides for peptide-peptide binding in aqueous solution. *J. Am. Chem. Soc.* **1991**, *113*, 7020−7030.

(9) Böhm, H.-J. LUDI: Rule-based automatic design of new substituents for enzyme inhibitor leads. *J. Comput.-Aided Mol. Des.* **1992**, *6*, 593−606.

(10) Böhm, H.-J. The development of a simple empirical scoring function to estimate the binding constant for a protein−ligand complex of known three-dimensional structure. *J. Comput.-Aided Mol. Des.* **1994**, *8*, 243−256.

(11) Wallqvist, A.; Jernigan, R. L.; Covell, D. G. A preference-based free energy parametrization of enzyme−inhibitor binding. Application to HIV-1-protease inhibitor design. *Protein Sci.* **1995**, *4*, 1881−1903.

(12) Jain, A. N. Scoring noncovalent protein−ligand interactions: A continouus differentiable function tuned to compute binding affinites. *J. Comput.-Aided Mol. Des.* **1996**, *10*, 427−440.

(13) Head, R. D.; Smythe, M. L.; Oprea, T. L.; Waller, C. L.; Green, S. M.; Marshall, G. M. VALIDATE: A New method for the receptor-based prediction of binding affinites of novel ligands. *J. Am. Chem. Soc.* **1996**, *118*, 3959−3969.

(14) Eldridge, M. D.; Murray, C. W.; Auton, T. R.; Paolini, G. V.; Mee, R. P. Empirical scoring functions: I. The development of a fast empirical scoring function to estimate the binding affinity of ligands in receptor complexes. *J. Comput.-Aided Mol. Des.* **1997**, *11*, 425−445.

(15) Murray, C. W.; Auton, T. R.; Eldridge, M. D. Empirical scoring functions. II. The testing of an empirical scoring function for the prediction of ligand−receptor binding affinities and the use of Bayesian regression to improve the quality of the model. *J. Comput.-Aided Mol. Des.* **1998**, *12*, 503−519.

(16) Böhm, H.-J. Prediction of binding constants of protein ligands: A fast method for the prioritization of hits obtained from de novo design or 3D database search programs. *J. Comput.-Aided Mol. Des.* **1998**, *12*, 309−323.

(17) Kuntz, I. D. Structure-based strategies for drug design and discovery. *Science* **1992**, *257*, 1078−1082.

(18) Ewing, T.; Kuntz, I. D. Critical evaluation of search algorithms for automated molecular docking and database screening. *J. Comput. Chem.* **1997**, *18*, 1175−1189.

(19) Rarey, M.; Kramer, B.; Lengauer, T.; Klebe, G. A fast felxible docking method using an incremental construction algorithm. *J. Mol. Biol.* **1996**, *261*, 470−489.

(20) Jones, G.; Willett, P.; Glen, R. C.; Leach, A. R.; Taylor, R. Development and validation of a genetic algorithm for flexible docking. *J. Mol. Biol.* **1997**, *267*, 727−748.

(21) Welch, W.; Ruppert, J.; Jain, A. N. Hammerhead: fast, fully automated docking of flexible ligands to protein binding sites. *Chem. Biol.* **1996**, *3*, 449−462.

(22) Dixon, J. S. Evaluation of the CASP2 docking section. *Proteins: Struct., Funct., Genet., Suppl.* **1997**, *1*, 198−204.

(23) Sippl, M. J. Calculation of conformational ensembles from potentials of mean force. *J. Mol. Biol.* **1990**, *213*, 859−883.

(24) Sippl, M. J.; Ortner, M.; Jaritz, M.; Lackner, P.; Flöckner, H. Helmholtz free energies of atom pair interactions in proteins. *Folding Des.* **1996**, *1*, 289−298.

(25) Jones, D. T.; Taylor, W. R.; Thornton, J. M. A new approach to protein fold recognition. *Nature* **1992**, *358*, 86−89.

(26) Godzik, A.; Kolinski, A.; Skolnick, J. Topology fingerprint approach to the inverse protein folding problem. *J. Mol. Biol.* **1992**, *227*, 227−238.

(27) Ouzounis, C.; Sander, C.; Scharf, M.; Schneider, R. Prediction of protein structure by evaluation of sequence-structure fitness. *J. Mol. Biol.* **1993**, *232*, 805−825.

(28) Bahar, I.; Jernigan, R. L. Interresidue potentials in globular proteins and the dominance of highly specific hydrophilic interactions at close separation. *J. Mol. Biol.* **1997**, *266*, 195−214.

(29) Bahar, I.; Jernigan, R. L. Coordination geometry of nonbonded residues in globular proteins. *Folding Des.* **1996**,*1*, 357−370.

(30) DeWitte, R. S.; Shakhnovich, E. I. SMoG: de novo design method based on simple, fast, and accurate free energy estimates. 1. Methodology and supporting evidence. *J. Am. Chem. Soc.* **1996**, *118*, 11733−11744.

(31) Verkhivker, G.; Appelt, K.; Freer, S. T.; Villafranca, J. E. Empirical free eenrgy calculations of ligand-protein crystallographic complexes. I. Knowledge-based ligand-protein interaction potentials applied to the prediction of human immunodeficiency virus 1 protease binding affinity. *Protein Eng.* **1995**, *8*, 677−691.

(32) Gohlke, H.; Hendlich, M.; Klebe, G. Knowledge based scoring function to predict protein−ligand interactions. Poster presented at the 12th European QSAR Symposium in Copenhagen, Denmark, September 22−27, 1998.

(33) Muegge, I.; Ermler, U.; Fritzsch, G.; Knapp, E. W. Free energy of cofactors at the quinone-QA site of the photosynthetic reaction center of Rhodobacter spheroides calculated by minimizing the statistical error. *J. Phys. Chem.* **1995**, *99*, 17917−17925.

(34) Apostolakis, J.; Muegge, I.; Ermler, U.; Fritzsch, G.; Knapp, E. W. Free energy computations on the shift of the special pair redox potential − mutants of the reaction center of Rhodobacter spheroides. *J. Am. Chem. Soc.* **1996**, *118*, 3743−3752.

(35) Muegge, I.; Apostolakis, J.; Ermler, U.; Fritzsch, G.; Lubitz, W.; Knapp, E. W. Shift of special pair redox potential: electrostatic energy computations of mutants of the reaction center of Rhodobacter spheroides. *Biochemistry* **1996**, *35*, 8359−8370.

(36) Muegge, I.; Tao, H.; Warshel, A. A fast estimate of electrostatic group contributions to the free energy of protein-inhibitor binding. *Protein Eng.* **1997**, *10*, 1363−1372.

(37) Muegge, I.; Qi, X. P.; Wand, A. J.; Chu, Z. T.; Warshel, A. The reorganization energy of cytochrome *c* revisited. *J. Phys. Chem. B* **1997**, *101*, 825−836.

(38) Muegge, I.; Schweins, T.; Warshel, A. Electrostatic contributions to protein−protein binding affinities: Application to rap/raf-interaction. *Proteins: Struct., Funct., Genet.* **1998**, *30*, 407−423.
(39) Martin, Y. C.; Lin, C. T.; Wu, J. Application of CoMFA to the design and structural optimization of D1 dopaminaergic agonists. In *3D QSAR in Drug Design. Theory Methods and Applications*; Kubinyi, H. Ed.; ESCOM: Leiden, 1993; pp 643−660.
(40) Martin, Y. C.; Lin, C. T.; Hetti, C.; DeLazzer, J. PLS analysis of distance matrixes detects nonlinear relationships between biological potency and molecular properties. *J. Med. Chem.* **1995**, *38*, 3009−3015.
(41) Martin, Y. C.; Lin, C. T. Three-dimensional quantitative structure−activity relationships: D2 dopamine agonists as an example. In *The Practice of Medical Chemistry*; Wermuth, C. G., Ed.; Academic Press: London, 1996; pp 459−483.
(42) Kim, K. H.; Martin, Y. C.; Brooks, C. Quantitative structure activity relationship of 5-lipoxygenase inhibitors − inhibitory potency of triazinone analogues in a broken cell. *Quant. Struct. Act. Relat.* **1996**, *15*, 491−497.
(43) Bernstein, F. C.; Koetzle, T. F.; Williams, G. J. B.; Jr., E. F. M.; Brice, M. D.; Rodgers, J. R.; Kennard, O.; Shimanouchi, T.; Tasumi, M. The Protein Data Bank. A computer-based archival file for macromolecular structures. *Eur. J. Biochem.* **1977**, *80*, 319−324.
(44) Warshel, A.; Papazyan, A.; Muegge, I. Microscopic and semimacroscopic redox calculations: What can and cannot be learned from continuum methods. *J. Biol. Inorg. Chem.* **1997**, *2*, 143−152.
(45) Sippl, M. J. Boltzmann's principle, knowledge-based mean fields and protein folding. An approach to the computational determination of protein structures. *J. Comput.-Aided Mol. Des.* **1993**, *7*, 473−501.
(46) Searle, M. S.; Williams, D. H. The cost of conformational order: Entropy changes in molecular associations. *J. Am. Chem. Soc.* **1992**, *114*, 10690−10697.
(47) Searle, M. S.; Williams, D. H.; Gerhard, U. Partitioning of free energy contributions in the estimate of binding constants: Residual motions and consequences for amide-amide hydrogen bond strengths. *J. Am. Chem. Soc.* **1992**, *114*, 10697−10704.
(48) Hendlich, M.; Rippmann, F.; Barnickel, G. BALI: Automatic-assignment of bond and atom types for protein ligands in the Brookhaven Protein Databank. *J. Chem. Inf. Comput. Sci.* **1996**, *37*, 774−778.
(49) Muegge, I.; Martin, Y. C.; Hajduk, P. J.; Fesik, S. W. Evaluation of PMF-scoring in docking weak ligands to the FK506 binding protein. submitted.
(50) Holloway, M. K.; Wai, J. M.; Halgren, T. A.; Fitzgerald, P. M. D.; Vacca, J. P.; Dorsey, B. D.; Levin, R. B.; Thompson, W. J.; Chen, L. J.; deSolms, S. J.; Gaffin, N.; Ghosh, A. K.; Giuliani, E. A.; Graham, S. L.; Guare, J. P.; Hungate, R. W.; Lyle, T. A.; Sanders, W. M.; Tucker, T. J.; Wiggins, M.; Wiscount, C. M.; Woltersdorf, O. W.; Young, S. D.; Darke, P. L.; Zugay, J. A. A priori prediction of activity for HIV-1 protease inhibitors employing energy minimization in the active site. *J. Med. Chem.* **1995**, *38*, 305−317.
(51) Astley, T.; Birch, G. G.; Drew, M. G. B.; Rodger, P. M.; Wilden, G. R. H. Effect of available volumes on radial distribution functions. *J. Comput. Chem.* **1998**, *19*, 363−367.
(52) Goodford, P. J. A computational procedure for determining energetically favorable binding sites on biologically important macromolecules. *J. Med. Chem.* **1985**, *28*, 849−857.
(53) McQuarrie, D. A. *Statistical Mechanics*; Harper Collins Publishers: New York, 1976.

JM980536J