PROTEINS: Structure, Function, and Genetics 39:393–407 (2000)

Examining Methods for Calculations of Binding Free Energies: LRA, LIE, PDLD-LRA, and PDLD/S-LRA Calculations of Ligands Binding to an HIV Protease

Yuk Yin Sham, Zhen Tao Chu, Holly Tao, and Arieh Warshel*
Chemistry Department, University of Southern California, Los Angeles, California

ABSTRACT **Several strategies for evaluation of the protein-ligand binding free energies are examined. Particular emphasis is placed on the Linear Response Approximation (LRA) (Lee et. al., Prot Eng 1992;5:215–228) and the Linear Interaction Energy (LIE) method (Åqvist et. al., Prot Eng 1994;7:385–391). The performance of the Protein Dipoles Langevin Dipoles (PDLD) method and its semi-microscopic version (the PDLD/S method) is also considered. The examination is done by using these methods in the evaluating of the binding free energies of neutral C2-symmetric cyclic urea-based molecules to Human Immunodeficiency Virus (HIV) protease. Our starting point is the introduction of a thermodynamic cycle that decomposes the total binding free energy to electrostatic and non-electrostatic contributions. This cycle is closely related to the cycle introduced in our original LRA study (Lee et. al., Prot Eng 1992;5:215–228). The electrostatic contribution is evaluated within the LRA formulation by averaging the protein-ligand (and/or solvent-ligand) electrostatic energy over trajectories that are propagated on the potentials of both the polar and non-polar (where all residual charges are set to zero) states of the ligand. This average involves a scaling factor of 0.5 for the contributions from each state and this factor is being used in both the LRA and LIE methods. The difference is, however, that the LIE method neglects the contribution from trajectories over the potential of the non-polar state. This approximation is entirely valid in studies of ligands in water but not necessarily in active sites of proteins. It is found in the present case that the contribution from the non-polar states to the protein-ligand binding energy is rather small. Nevertheless, it is clearly expected that this term is not negligible in cases where the protein provides preorganized environment to stabilize the residual charges of the ligand. This contribution can be particularly important in cases of charged ligands. The analysis of the non-electrostatic term is much more complex. It is concluded that within the LRA method one has to complete the relevant thermodynamic cycle by evaluating the binding free energy of the "non-polar" ligand, ℓ', where all the residual charges are set to zero. It is shown that the LIE term, which involves the scaling of the van der Waals interaction by a constant β (usually in the order of 0.15 to 0.25), corresponds to this part of the cycle. In order to elucidate the nature of this non-electrostatic term and the origin of the scaling constant β, it is important to evaluate explicitly the different contributions to the binding energy of the non-polar ligand, $\Delta G_{\mathbf{bind},\ell'}$. Since this cannot be done at present (for relatively large ligands) by rigorous free energy perturbation approaches, we evaluate $\Delta G_{\mathbf{bind},\ell'}$ by the PDLD approach, augmented by microscopic calculations of the change in configurational entropy upon binding. This evaluation takes into account the van der Waals, hydrophobic, water penetration and entropic contributions, which are the most important free energy contributions that make up the total $\Delta G_{\mathbf{bind},\ell'}$. The sum of these contributions is scaled by a factor θ and it is argued that obtaining a quantitative balance between these contributions should result in θ = 1. By doing so we should have a reliable estimate of the value of the LIE β and a way to understand its origin. The present approach gives θ values between 0.5 and 0.73, depending on the approximation used. This is encouraging but still not satisfying. Nevertheless, one might be able to use our PDLD approach to estimate the change of the LIE θ between different protein active sites.**

It is pointed out that the LIE method is quite similar to our original approach where the electrostatic term was evaluated by the LRA method and the non-electrostatic term by the PDLD method (with its vdw, solvation, and hydrophobic contributions). The practical difference is that the LIE method approximates the non-electrostatic term by the average of the van der Waals interaction, while our LRA method evaluates this term by the PDLD method. This point is illustrated by the fact that our LRA approach gives results of similar quality to those obtained by the LIE approach. Finally it is

Grant sponsor: National Institutes of Health; Grant number: GM24498.

Yuk Yin Sham's present address is Laboratory of Experimental and Computational Biology, NCI-FCRDC, Building 469, Room 151, Frederick, MD 21702.

Holly Tao's present address is Double Twist, Inc, 1999 Harrison Street Suite 1100, Oakland, CA 94612.

*Correspondence to: Arieh Warshel, Chemistry Department, University of Southern California, Los Angeles, CA 90089-1062.

Received 31 August 1999; Accepted 6 January 2000

found that results of similar quality are obtained by the PDLD/S method and the LRA method. This is significant since the PDLD/S method is much faster than the LRA and LIE methods. However, more studies of the relative accuracy on other systems are needed in order to establish their relative merits. Proteins 2000;39:393–407.

Key words: ligand binding; Linear Response Approximation; binding entropy; drug design

INTRODUCTION

The development of a reliable method for the evaluation of absolute binding free energies of ligand protein complexes is of significant practical and fundamental interest. An effective method should provide a guide for rational drug design and help in providing a deeper understanding of structure function correlation in proteins. Reliable estimates of binding free energies should help in identifying and refining "lead compounds" in computer-aided drug design. Different strategies can be used in studies of absolute binding energies. This include formally rigorous approaches such as the Free Energy Perturbation (FEP) method,[1-4] which is usually restricted to "mutations" of small parts of the ligand[2] and the all-atom Linear Response Approximation (LRA)[4] and its variants,[5,6] which allows one to calculate the absolute binding energy of relatively large ligands. More approximated and significantly faster approaches, which frequently focus on electrostatic energies, are also quite effective.[7] These include the scaled Protein Dipoles Langevin Dipoles (PDLD/S) method (e.g., Lee et al.[4] and Madura et al.[8]) and the LRA version of this approach (PDLD/S-LRA)[9] as well as other approaches (e.g., Madura et al.[8] and Froloff et al.[8,10]).

Despite the availability of several options for the calculations of binding free energy, the actual performance of different approaches has not been fully established. We do not yet have benchmarks for comparative examination of different strategies and for a systematic assessment of the corresponding error range. This work presents a comparative study of different approaches by using a benchmark of several inhibitors of HIV protease. The approaches considered will be the PDLD/S-LRA, PDLD-LRA, and several versions of all-atom LRA method, including the LIE approach.

METHODS

LRA Calculations

We start this section by considering the LRA method. This will serves as a general framework for other approaches. The validity of the LRA method has been established by simulation charge transfer (CT) processes in solutions [11–13] and more recent related studies (e.g., King and Warshel[14] and Kuharski et al.[15]). Subsequent studies have noted the applicability of this approach as an approximation to Free Energy Perturbation (FEP) calculations.[16] The applicability of the LRA to calculating binding free energies was pointed out and demonstrated first by Lee et. al.[4]

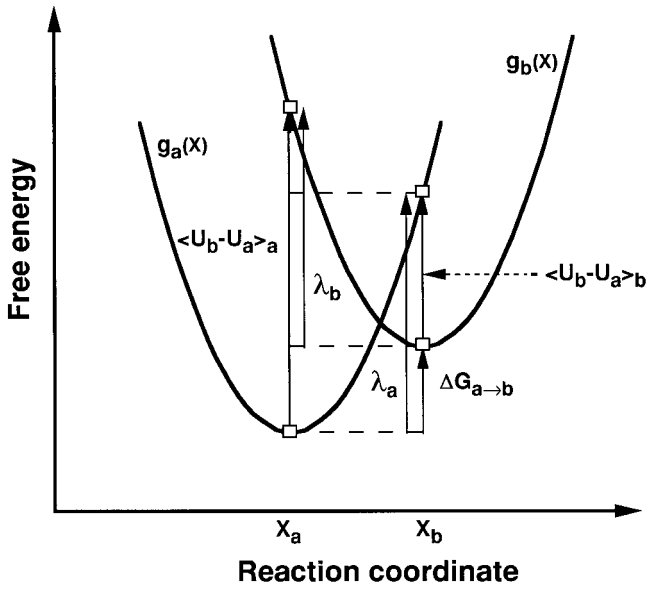

Fig. 1. Showing the free energy functions of the reactant and product states, $[g_a(X)$ and $g_b(X)]$, as two Marcus' type parabolas of equal curvature. The solvent reorganization energies at the minima of g_a and g_b are given by $\lambda_b = \langle U_b - U_a\rangle_a - \Delta G_{a\to b}$ and $\lambda_a = \langle U_a - U_b\rangle_b + \Delta G_{a\to b}$. By assuming $\lambda_a = \lambda_b$ one can obtain the LRA estimate of $\Delta G_{a\to b}$ [see Eqs. (4) and (5)].

Our derivation of the LRA method is based on studying the functions that describe the free energies of the reactant (a) and product (b) states. These free energy functions (the g_α's of Fig. 1) are parabolas of equal curvature in the macroscopic continuum model where the LRA is assumed, rather than obtained. In this limit the g_α' are the macroscopic Marcus' parabolas.[17] In microscopic molecular systems the g_α's are defined by[11,14]

$$g_a(X) = -\beta^{-1} \ln P_a(X) \tag{1}$$

$$g_b(X) = -\beta^{-1} \ln P_b(X) + \Delta G_{a\to b}$$

where X is the reaction coordinate and $P_a(X)$ is the probability of finding a specific value X while running trajectories on the potential surface of state a (see King and Warshel[14]). $\beta = 1/k_BT$ (where k_B is the Boltzmann constant and T is the absolute temperature).

Our simulation studies[11,12,14] established that the microscopic g_a and g_b for charge transfer reactions in water are indeed parabolas with very similar curvature, so that the LRA is valid. In this case we can consider the system as a collection of quasi-harmonic oscillators where we have the relationship[13]

$$\lambda_b = \langle U_b - U_a\rangle_a - \Delta G_{a\to b} \tag{2}$$

here U_a and U_b are the potential surface of state a and b, respectively, while $\langle\rangle_a$ designates an average over trajectories which are propagated on the potential U_a. λ_b is the so-called "solvent reorganization energy," which is usually defined [in addition to the definition of Eq. (2)] by $\lambda_b = g_b(X_a) - g_b(X_b)$, where X_a and X_b are the minima of g_a and

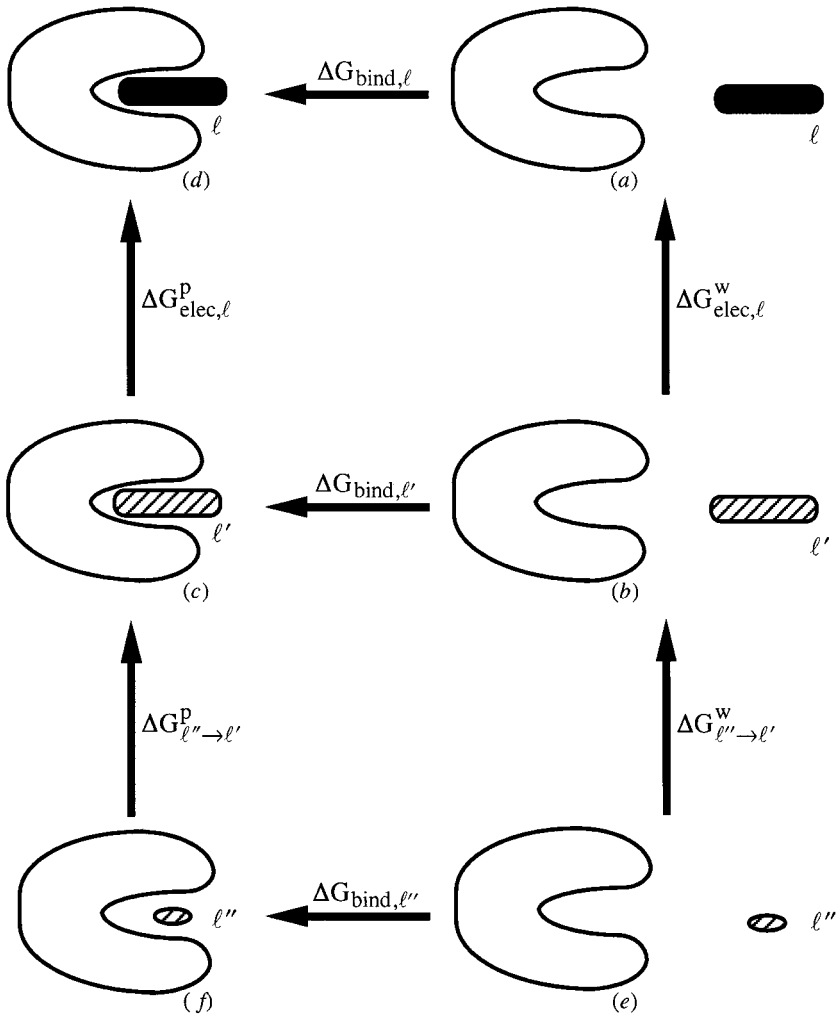

Fig. 2. A thermodynamic cycle that describes the binding of a ligand to a protein and provides a clear definition of the LRA and LIE approximations. The electrostatic contributions to the binding free energy, $\Delta G^{w}_{elec,\ell}$ and $\Delta G^{p}_{elec,\ell}$, can be calculated reliably by using the LRA method for the process of "uncharging" the residual charges of the polar ligand to zero in water (a→b) and in protein (c→d), respectively. The estimation of the $\Delta G_{bind,\ell'}$ term considers the lower cycle (b→e→f→c) and can be carried out by a shrinking process describe in the text.

g_b, respectively (see Fig. 1). Similarly we evaluate λ_a and obtain

$$\lambda_a = \langle U_a - U_b \rangle_b + \Delta G_{a\to b} \tag{3}$$

Using the fact that $\lambda_a = \lambda_b$ when g_a and g_b have the same curvature, we obtain by equating Eq. (2) to Eq. (3):

$$\langle U_a - U_b \rangle_b + \Delta G_{a\to b} = \langle U_b - U_a \rangle_a - \Delta G_{a\to b} \tag{4}$$

This gives our general LRA expression

$$\Delta G_{a\to b} = \frac{1}{2}\left[\langle U_b - U_a \rangle_a + \langle U_b - U_a \rangle_b\right] \tag{5}$$

An alternative derivation based on expanding the Free Energy Perturbation (FEP) expressions are given elsewhere.[16]

Applying the LRA to a binding process requires one to formulate the relevant thermodynamic cycles and to access what steps can be approximated reliably by the LRA method. Here we consider the binding cycle of Figure 2.

In the upper cycle ($a \to b \to c \to d$), we start by taking the ligand (ℓ) in water and "uncharging" it by setting all the residual charges to zero. Next we take the non-polar ligand (ℓ') and bind it to the protein. Finally we recharge (ℓ') in the protein active site. The binding free energy is now given by

$$\begin{aligned}\Delta G_{\mathrm{bind},\ell} &= \Delta G_{a\to b} + \Delta G_{b\to c} + \Delta G_{c\to d} \\ &= -\Delta G^{w}_{\mathrm{elec},\ell} + \Delta G_{\mathrm{bind},\ell'} + \Delta G^{p}_{\mathrm{elec},\ell}\end{aligned} \tag{6}$$

where $\Delta G_{\mathrm{bind},\ell'}$ is the free energy of binding the non-polar ligand to the protein, while $\Delta G^{w}_{elec,\ell}$ and $\Delta G^{p}_{elec,\ell}$ are the electrostatic free energies associated with charging the ligand in water (w) and in protein (p), respectively. These electrostatic contributions can be evaluated by the LRA method using

$$\begin{aligned}\Delta G^{w}_{\mathrm{elec},\ell} &= \frac{1}{2}\left[\langle \Delta U^{w}_{\mathrm{elec},\ell'\to\ell}\rangle_{\ell} + \langle \Delta U^{w}_{\mathrm{elec},\ell'\to\ell}\rangle_{\ell'}\right] \\ &= \frac{1}{2}\left[\langle U^{w}_{\mathrm{elec},\ell}\rangle_{\ell} + \langle U^{w}_{\mathrm{elec},\ell}\rangle_{\ell'}\right] \\ \Delta G^{p}_{\mathrm{elec},\ell} &= \frac{1}{2}\left[\langle \Delta U^{p}_{\mathrm{elec},\ell'\to\ell}\rangle_{\ell} + \langle \Delta U^{p}_{\mathrm{elec},\ell'\to\ell}\rangle_{\ell'}\right]\end{aligned} \tag{7}$$

TABLE I. Calculated Electrostatic Contributions to the LRA and LIE Binding Free Energies for Cyclic Urea-Based HIV Protease Inhibitors†

Inhibitor[a]	$\langle U^p_{\mathrm{elec},\ell}\rangle_\ell$	$\langle U^p_{\mathrm{elec},\ell}\rangle_{\ell'}$	$\langle U^w_{\mathrm{elec},\ell}\rangle_\ell$	$\langle U^w_{\mathrm{elec},\ell}\rangle_{\ell'}$	LRA $\Delta G^p_{\mathrm{elec},\ell}$	LRA $\Delta G^w_{\mathrm{elec},\ell}$	LIE $\Delta G^p_{\mathrm{elec},\ell}$	LIE $\Delta G^w_{\mathrm{elec},\ell}$	LRA $\Delta\Delta G^{w\to p}_{\mathrm{elec},\ell}$	LIE $\Delta\Delta G^{w\to p}_{\mathrm{elec},\ell}$
ℓ_1	−78.4	−1.5	−69.5	−0.3	−40.0	−34.9	−39.2	−34.7	−5.0	−4.5
ℓ_2	−42.2	−3.4	−35.5	−0.3	−22.8	−17.9	−21.1	−17.8	−4.9	−3.3
ℓ_3	−50.5	−1.6	−39.9	−0.2	−26.0	−20.1	−25.2	−20.0	−6.0	−5.3
ℓ_4	−44.5	−2.7	−36.3	−0.2	−23.6	−18.3	−22.2	−18.1	−5.3	−4.1
ℓ_5	−43.6	−2.2	−35.1	−0.2	−22.9	−17.7	−21.8	−17.5	−5.2	−4.3

†Electrostatic free energies (in kcal/mol) are evaluated from the corresponding averages $\langle\ \rangle$ over trajectories generated by the potential surface of the indicated state. ℓ and ℓ' indicate the polar and the non-polar states of the ligand, while p and w designate protein and water respectively. The LRA and LIE estimates correspond to the use of Eqs. (8) and (12), respectively.
[a]The notation for the inhibitors is defined in Figure 5.

$$= \frac{1}{2}\left[\langle U^p_{\mathrm{elec},\ell}\rangle_\ell + \langle U^p_{elec,\ell}\rangle_{\ell'}\right]$$

where $U_{\mathrm{elec},\ell}$ designates the electrostatic interactions between the ligand and its surrounding. Here $\Delta U_{\mathrm{elec},\ell'\to\ell} = U_{\mathrm{elec},\ell} - U_{\mathrm{elec},\ell'}$ and we make use of the fact that $U_{\mathrm{elec},\ell'} = 0$. Now we obtain from Eqs. (6) and (7):

$$\Delta G^{LRA}_{\mathrm{bind},\ell} = \frac{1}{2}\left[\langle U^p_{\mathrm{elec},\ell}\rangle_\ell + \langle U^p_{\mathrm{elec},\ell}\rangle_{\ell'} - \langle U^w_{\mathrm{elec},\ell}\rangle_\ell - \langle U^w_{\mathrm{elec},\ell}\rangle_{\ell'}\right] + \Delta G_{\mathrm{bind},\ell'} = (\Delta\Delta G^{w\to p}_{\mathrm{elec},\ell})^{LRA} + \Delta G_{\mathrm{bind},\ell'} \quad (8)$$

$\Delta G_{\mathrm{bind},\ell'}$ can be obtained by the lower cycle of Figure 2 (b $\to$ e $\to$ f $\to$ c). In this cycle we shrink the non-polar ligand, ℓ', to nothing, ℓ'', in water, then bind the zero size ligand to the protein, and finally grow it back to its original size in the protein. This gives

$$\Delta G_{\mathrm{bind},\ell'} = \Delta G^p_{\ell''\to\ell'} - \Delta G^w_{\ell''\to\ell'} + \Delta G_{\mathrm{bind},\ell''} \quad (9)$$

The terms $\Delta G^w_{\ell''\to\ell'}$ and $\Delta G^p_{\ell''\to\ell'}$ can be evaluated, at least in principle, by FEP approaches or related treatments (see PDLD-LRA and PDLD/S-LRA Calculations section). The values of all the free energy contributions to $\Delta G_{\mathrm{bind},\ell''}$ are zero except the value of an entropic contribution that is related to the difference between the volume of the binding site and the molar volume (which is available for the ligand at 1 M concentration in water). Thus we can write

$$\Delta G_{\mathrm{bind},\ell''} = -RT\ln\left(\frac{v_0}{v^{site}_{eff}}\right) \quad (10)$$

where v^{site}_{eff} is the effective volume of the active site and v_0 is the volume occupied by one molecule at 1M concentration (v_0 = 1,660 Å^3 at 300 K).[18,19]

Several approximations can be applied to Eq. (8). One useful approach involves the neglect of $\langle U_{\mathrm{elec},\ell}\rangle_{\ell'}$. This is justified for a charging process in water where the water dipoles are oriented randomly around the non-polar state and the solvation energy is given by $1/2\langle U^w_{elec,\ell}\rangle_{\ell'}$.[20] In proteins, however, it is not fully justified to neglect $\langle U^p_{elec,\ell}\rangle_{\ell'}$, since the protein environment can remain preoriented even in the presence of the non-polar ligand (see $\langle U^p_{elec,\ell}\rangle_{\ell'}$ in Table I). At any rate one may try to use the approximation:

$$\Delta G_{\mathrm{bind},\ell} \cong \frac{1}{2}\left[\langle U^p_{\mathrm{elec},\ell}\rangle_\ell - \langle U^w_{\mathrm{elec},\ell}\rangle_\ell\right] + \Delta G_{\mathrm{bind},\ell'} \quad (11)$$

Before discussing our approach for the evaluation of Eqs. (8) and (11), it is useful to introduce the closely related LIE approach. The LIE approximation expresses the binding energy as

$$\Delta G^{LIE}_{\mathrm{bind},\ell} = \alpha[\langle U^p_{\mathrm{elec},\ell}\rangle_\ell - \langle U^w_{\mathrm{elec},\ell}\rangle_\ell] + \beta[\langle U^p_{\mathrm{vdw},\ell}\rangle_\ell - \langle U^w_{\mathrm{vdw},\ell}\rangle_\ell] \quad (12)$$

where α and β are scaling constants while $U_{\mathrm{elec},\ell}$ designates the electrostatic interaction between the ligand and its surrounding. The rational for this has been provided in Åqvist et al.[5] by a somewhat less rigorous way than that used in deriving Eqs. (8) and (11). That is, Åqvist et al.[5] considered the electrostatic part of the binding process as a transition from a state (state A) with the isolated ligand in the gas phase and a protein with a ready made cavity (in the shape of the ligand) to a state (state B) where the ligand is bound to the protein. The electrostatic contribution of this process was estimated by the LRA approximation. Although such a process is conceptually valid, it is difficult to treat the potential U_A in actual calculations without introducing an actual non-polar ligand (or a related constraint) in the cavity. This problem was removed in a recent work,[21] which essentially used the cycle introduced earlier by Lee et. al. (the a $\to$ b $\to$ c $\to$ d cycle of Fig. 2 of the present work and the upper cycle of Fig. 1 of Lee et al.[4]) to justify the LIE approach. The cycle of Marelius et al.,[21] however, keeps the intramolecular interactions within the ligand, upon annihilating the ligand-protein electrostatic interaction, while Lee et. al. also annihilated the intramolecular interactions within the ligand. Now the LIE approach evaluates the $\Delta\Delta G_{\mathrm{bind},\ell'}$ term using the vdw interactions of the polar rather than non-polar ligand (this is needed since the LIE does not involve simulations of the non-polar ligand).

Our approximation of Eq. (8) combines the approach used by Lee et. al.[4] in the first implementation of the LRA treatment in calculations of binding free energies with a microscopic estimate of the entropic term. This approach retains the rigorous LRA terms of Eq. (8) but approximates $\Delta G_{\mathrm{bind},\ell'}$ by

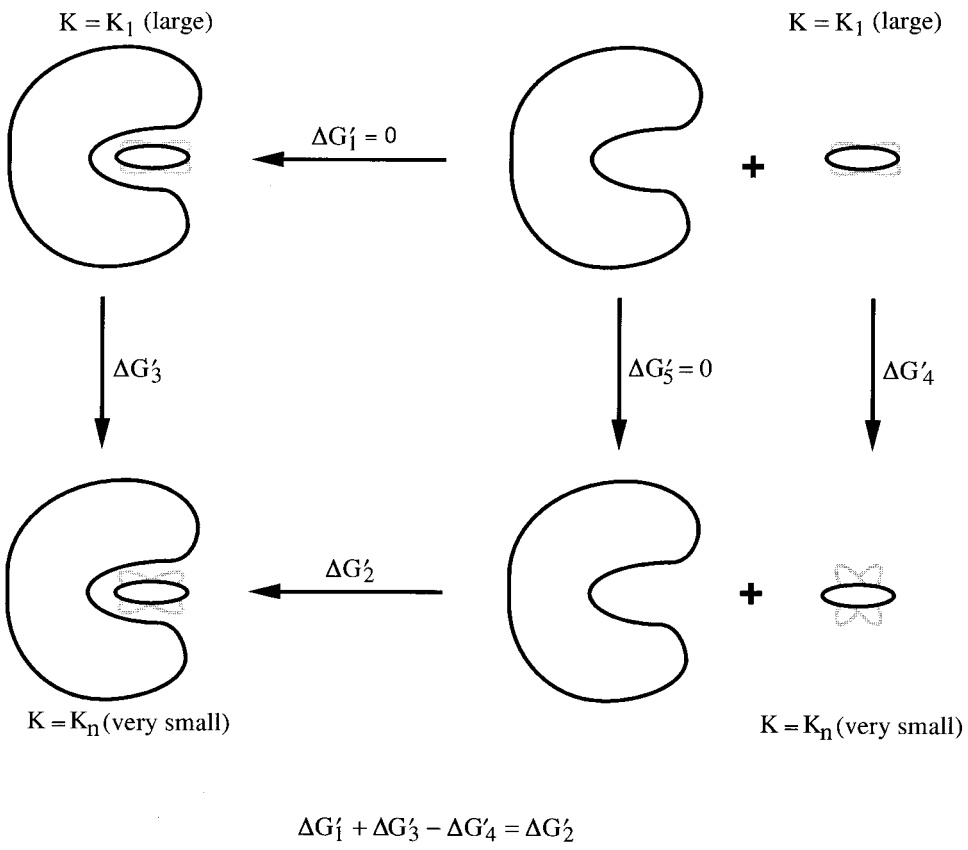

Fig. 3. A thermodynamic cycle for estimating of the configurational entropy contribution to the binding free energy. This calculation is carried out using a FEP approach with a pertubative reduction of a position restraint which is imposed on each atom of the ligand in its specific environment (protein and water). An additional constraint is used to define the effective volume available for the ligand (see text). The arrows in steps a→b and d→c indicates that the corresponding free energies are evaluated in several FEP steps. The gray traces of the ligand indicate schematically the motion of the ligand relative to its reference position.

$$\Delta G^{LRA}_{\mathrm{bind},\ell'} = \Delta G^{w\to p}_{\mathrm{vdw},\ell'} + \Delta G^{w\to p}_{\mathrm{hyd},\ell'} + \Delta G^{w\to p}_{\mathrm{elec},\ell'} - T(\Delta\Delta S^{w\to p}_{\ell'})' = \Delta G^{PDLD}_{\mathrm{bind},\ell'} - T(\Delta\Delta S^{w\to p}_{\ell'})' \quad (13)$$

where $\Delta G^{w\to p}_{vdw,\ell'}$ is the van der Waals contribution associated with moving ℓ' from water to protein, $\Delta G^{w\to p}_{hyd,\ell'}$ is the hydrophobic contribution, $\Delta G^{w\to p}_{elec,\ell'}$ is the electrostatic contribution (that involves the effect of water penetration to the active site). Finally $-T\Delta\Delta S'$ is the entropic contribution associated with the change in the configurational space of the ligand upon transfer from water to the protein active site. The first three terms of Eq. (13) are evaluated now by the PDLD treatment (see PDLD-LRA and PDLD/S-LRA Calculations section). The $-T\Delta\Delta S'$ term was neglected in the original work of Lee et. al. This term can be evaluated by direct simulations, which involve a release of a restraint potential using a FEP approach (see Entropy Calculations section). Now the LRA expression is given by

$$\Delta G^{LRA}_{\mathrm{bind},\ell} = \frac{1}{2}[\langle U^{p}_{\mathrm{elec},\ell}\rangle_{\ell} + \langle U_{\mathrm{elec},\ell}\rangle_{\ell'} - \langle U^{w}_{\mathrm{elec},\ell}\rangle_{\ell} - \langle U^{w}_{\mathrm{elec},\ell}\rangle_{\ell'} + \theta[\Delta G^{PDLD}_{\mathrm{bind},\ell'} - \gamma[T(\Delta\Delta S^{w\to p}_{\ell'})']] \quad (14)$$

Here we scaled the contributions to $\Delta G_{\mathrm{bind},\ell'}$ to compensate for possible missing effects or for possible deficiencies in the model (this approach has already been used in our previous PDLD studies which involved calculation of the relevant hydrophobic and salvation energies[22]).

Entropy Calculations

The $\Delta\Delta S'$ term of Eq. (14) can be estimated by a method that is related to the approach introduced by Hermans and co-workers.[18] This approach involves a perturbative change in the restraints imposed on the ligand in its specific environment (protein or water). That is, the atoms of the ligand are subjected to a position restraint.

$$U'_{\mathrm{rest},j} = \frac{1}{2}K_j[\sum_i (r^{\ell}_i - \bar{r}^{\ell}_i)^2] \quad (15)$$

where $\bar{r}$ is the vector of the average coordinates of the ligands at the protein site. Now the restraint K is changed according to the cycle of Fig. 3 where the resulting free energy changes are designated by $\Delta G'$. In the upper part of the cycle a large restraint potential is applied so that the non-polar ligand is practically frozen in both the protein site and in water so that the corresponding absolute entropy is zero. Thus, increasing K_1 above a given limiting value will result in a zero contribution from the change in the configurational entropy ($\Delta G'_1 = -T\Delta S'_1 \approx 0$) for the binding free energy that would be obtained with this restraint. The left and right side of the cycle involve a change of the magnitude of K to a small value (K_n) in n-1 steps. The overall change in configurational entropy for this cycle can be written as

$$-(T\Delta S'_2 - T\Delta S'_1) = -(T\Delta S'_3 - T\Delta S'_4) \quad (16)$$

since ΔS_1 the entropy is a state function (so it can be evaluated by thermodynamic cycles), we have

$$-T\Delta S_2' = -(T\Delta S_3' - T\Delta S_4') \tag{17}$$

where $\Delta S_2'$ change in the configurational entropy of the unrestrained and non-polar ligand, which is the contribution we are looking for. Now our problem is to obtain $\Delta S_3'$ and S_4'. Here we assume that the free energies obtained by changing K_1 to K_n ($\Delta G_3'$ and $\Delta G_4'$) only reflect the configurational entropy contributions, which are associated with the effective volume available for the ligand. This assumption depends on the chosen $\bar{r}$ where the optimal $\bar{r}$ should give the smallest $\Delta G'$. This is so since the state with $K = 0$ is independent of r and thus the larger $\Delta G'$, the largest is its enthalpic component. The optimal $\bar{r}$ should be obtained by relaxing the system without the constraint and by trying several alternative $\bar{r}$'s. At any rate, with the above assumption we can write

$$-T\Delta S_2' = \Delta G_3' - \Delta G_4' \tag{18}$$

In evaluating $\Delta G_3'$ and $\Delta G_4'$, we find it convenient to use several consecutive FEP steps rather than one long step as is done in regular FEP procedures with equally spaced windows (e.g., Warshel[19]). In each of these steps we change K_j and K_{j+1}, where the new K_j is taken as the K_{j+1} of the previous FEP step. Thus our approach involves the use of the following expression:

$$\Delta G'(K_1 \rightarrow K_n) = \sum_{j=1}^{n-1} \Delta G'(K_j \rightarrow K_{j+1})$$

$$\Delta G'(K_j \rightarrow K_{j+1}) = = \sum_{m=0}^{N-1} \Delta\Delta G_j'(m \rightarrow m+1)$$

$$\Delta\Delta G_j'(m \rightarrow m+1) = -RT \ln\langle \exp\{-(U_{m+1}^j - U_m^j)/RT\}\rangle_{U_m^j} \tag{19}$$

where

$$U_m^j = (1-\lambda_m)U'_{\text{rest},j}(K = K_j) + \lambda_m U'_{\text{rest},j+1}(K = K_{j+1}) \tag{20}$$

and $\lambda_m$is changed from zero to one in N equal increments. Our approach is basically equivalent to the use of variable increments in λ_m.

Now we have to resolve the fact that $\Delta G_n'$ diverges (at least formally) for infinitely long simulation, when $K_n \rightarrow 0$. Fortunately we are interested in the entropic contribution for a 1 M standard state. This can be obtained in principle by choosing a simulation sphere of a volume which is equal to the molar volume (v_0 = 1,660 Å^3) while allowing K_n to approach zero. However, such an approach should encounter major convergence problems since the ligand is unlikely to sample the large simulation sphere in a reasonable simulation time. A faster convergence would be obtained by allowing the ligand to move in a smaller effective volume $v_{eff}^{(n)}$. Imposing such a restraint does not change the situation in the protein since $v_{eff}^{(n)}$ is chosen to be larger than the volume available for the ligand in the active site. In water, on the other hand, the ligand should be allowed to sample the available effective volume. Thus we perform the simulation while confining the ligand to $v_{eff}^{(n)}$ but we also consider the free energy associated with changing $v_{eff}^{(n)}$ to v_0. This free energy is simply given by[19]

$$\Delta G = -RT \ln(v_0/v_{eff}^{(n)}). \tag{21}$$

Thus we have

$$-T(\Delta S_2')_0 = \Delta G_3' - \Delta G_4' - RT \ln(v_0/v_{eff}^{(n)}). \tag{22}$$

where $\Delta s_2'$ designate the correspondence entropy in a standard state. The problem is, however, to introduce the proper volume restraint. That is, when the ligand is an isolated atom the restraint is easily introduced by letting the lowest value of K_n to correspond to v_{eff}[18]

$$v_{eff} = \left(\frac{2\pi RT}{K}\right)^{3/2}. \tag{23}$$

Here v_{eff} is taken as the classical partition function obtained with the given harmonic potential, while v_0 is the actual partition function for a single particle confine to a sphere with a volume v_0. Equation (23) is also valid for the center of gravity of a polyatomic molecule. In our case, however, the restraint is imposed on the Cartesian coordinates of each ligand atom and this involves a restriction of rotational and intramolecular entropy. Thus we have to find an alternative way of confining the ligand to v_{eff}. Our solution of this problem involves the use of two restraints, one for the mapping purpose and the other for restricting the system to the proper $v_{eff}^{(n)}$. That is, in addition to the mapping restraint we restrain the center of gravity of the ligand (or alternatively a central atom) by a Cartesian restraint (with $K = K_{\text{eff}}$) that determines the effective volume of the ligand through Eq. (23). This restraint is not allowed to change during the FEP mapping process. Thus we use a mapping potential of the form

$$U_m^j = (1-\lambda_m)U'_{cons,j}(K = K_j) + \lambda_m U'_{cons,j+1}(K = K_{j+1}) + \frac{1}{2}K_{eff}(r_c - r_c^0)^2 \tag{24}$$

where r_c is the coordinate of a central atom. Using this U_m' leaves v_{eff} unaffected by the change of λ_m. Now we can let K_n to approach zero without a divergence in $\Delta S'$ since the volume of the ligand is restricted by the K_{eff} term.

It might be instructive to note here the differences between our approach to that introduced by Hermans and Wang.[18] The approach of Hermans and Wang has been formulated in terms of the translational and rotational degrees of freedom of a rigid ligand while focusing on cases where the internal dynamics of the ligand is not perturbed. Our approach addresses the more general case where the change of internal degrees of freedom is also important. This requires us to introduce a Cartesian treatment that involves simultaneous freezing of all the substrate degrees of freedom. Our treatment also involves the introduction of the volume restraint (the K_{eff} term). Furthermore, Hermans and Wang[18] evaluated the effect of the constraint on the ligand in the gas phase and the protein site, while our

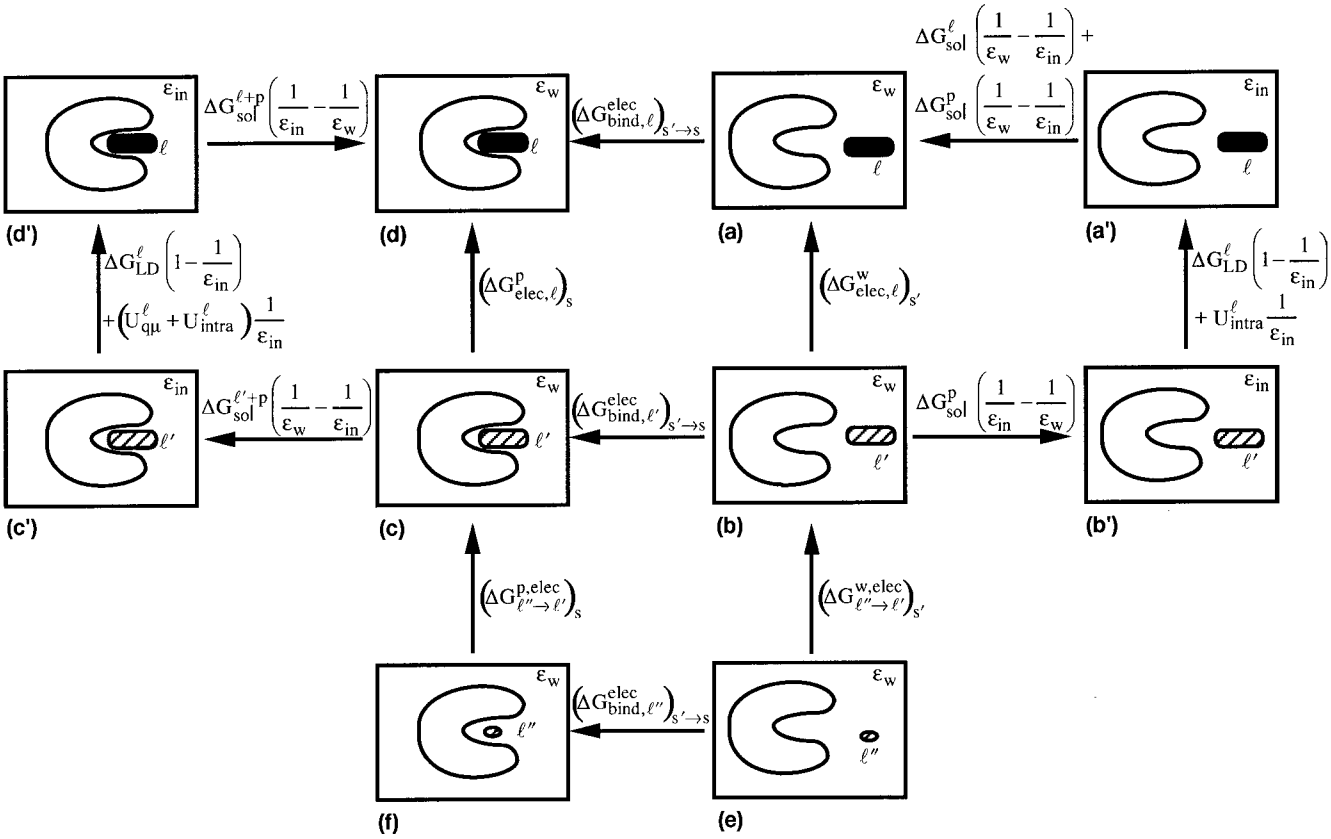

Fig. 4. A thermodynamic cycle for calculating binding energy using the PDLD/S method. The inner binding cycle (a→b→c→d) is equivalent to that of Figure 2 and to a part of the cycle used by Lee et al.[4] The electrostatic terms associated with a→b and c→d are evaluated by the external cycles in the figure (a→a′→b′→b and c→c′→d′→ d) and are evaluated by using Eq. (26) with ϵin = 4. The free energy ΔG bind,ℓ′ and the other contributions of ΔG bind,ℓ′ are evaluated by Eq. (30) using an approximation to the b→e→f→c cycle.

treatment considers the effect of moving the ligand from solution to the protein site, which corresponds directly to the binding process. Considering the entropic effect in the gas phase instead of in solution is usually a good approximation. However, this approach might miss the entropic effects associated with the conformational changes of polar ligand and these effects are evaluated consistently by our approach.

Our procedure has performed quite well in the test case of binding of benzene to T4 lysozyme (we obtained 8 kcal/mol entropic contribution to the binding energies as compared to the estimate of 8.4 kcal/mol by Hermans and Wang[18]). Similarly we obtained very encouraging results in evaluation of activation entropies (this work is now in press). This does not guarantees, however, that the present method will give reasonable results in case of large flexible ligands which involve major convergence problems.

PDLD-LRA and PDLD/S-LRA Calculations

One of the promising options for evaluating the absolute binding free energies is the PDLD/S-LRA method.[4,9,22–24] This approach has been described in details elsewhere (e.g., Muegge et al.[9]) but the present work introduces several modifications that are expected to increase the stability of the calculations. Thus, we describe below the main point of this method.

We start by considering the electrostatic contributions to the absolute binding free energy. This is done by using the thermodynamic cycle of Figure 4, which is in fact the PDLD/S equivalent of the cycle of Figure 2 and the upper cycle of our original LRA study.[4] Now the relevant PDLD/S-LRA electrostatic contributions are given by

$$\Delta G^{\text{elec}}_{\text{bind},\ell} = \Delta G^{p}_{\text{elec},\ell} - \Delta G^{w}_{\text{elec},\ell} + \Delta G^{\text{elec}}_{\text{bind},\ell'} \qquad (25)$$

In evaluating the first two terms of Eq. (25), we start by considering a single configuration (s) for the protein-ligand complex and a single configuration (s′) for the dissociated complex. The effective electrostatic potential for both the a→b and c→d steps are evaluated by the PDLD/S approach, where the results of the microscopic PDLD model are scaled by considering the external cycles c → c′ → d′ → d and a → a′ → b′ → b. These cycles involve a change of the dielectric constant of the solvent around the protein from that of water, ϵ_w, to the value of the so-called "protein dielectric", ϵ_{in} (this parameter represents the contributions which are not treated explicitly[25]). The overall potentials for the two cycles are given by

$$(\Delta G^{p}_{\mathrm{elec},\ell})_s = \overline{U}^{p}_{\mathrm{elec},\ell} = \left((\Delta G^{\ell+p}_{sol} - \Delta G^{\ell+p}_{\mathrm{sol}}) \left(\frac{1}{\varepsilon_{in}} - \frac{1}{\varepsilon_{w}} \right) + \Delta G^{\ell}_{sol} \left(1 - \frac{1}{\varepsilon_{in}} \right) + \frac{U^{\ell}_{q\mu}}{\varepsilon_{in}} + \frac{U^{\ell}_{intra}}{\varepsilon_{in}} \right)_s$$

$$(\Delta G^{w}_{\mathrm{elec},\ell})_{s'} = \overline{U}^{w}_{\mathrm{elec},\ell} = \left((\Delta G^{\ell}_{sol}) \left(\frac{1}{\varepsilon_{in}} - \frac{1}{\varepsilon_{w}} \right) + \Delta G^{\ell}_{sol} \left(1 - \frac{1}{\varepsilon_{in}} \right) + \frac{U^{\ell}_{intra}}{\varepsilon_{in}} \right)_{s'} \quad (26)$$

where we use $\overline{\mathrm{U}}$ rather than $\overline{\mathrm{U}}$ to designate the fact that $\overline{\mathrm{U}}$ is an effective potential rather than the fully microscopic potential of the previous sections. $\Delta \mathrm{G}_{\mathrm{sol}}$ denotes the electrostatic contribution to the solvation free energy of the indicated group in water (e.g., $\Delta \mathrm{G}^{\ell+p}_{sol}$ denotes the solvation of the protein-ligand complex in water). To be more precise, $\Delta \mathrm{G}_{\mathrm{sol}}$ should be scaled by $1/(1 - 1/\epsilon_{\mathrm{w}})$ but this small correction is neglected here. The values of the $\Delta \mathrm{G}_{\mathrm{sol}}$'s are evaluated by the Langevin dipole solvent model. $\mathrm{U}^{\ell}_{q\mu}$ is the electrostatic interaction between the charges of the ligand and the protein dipoles in vacuum (this is a standard PDLD notation). In the present case $\mathrm{U}^{\ell'}_{q\mu} = 0$. $\mathrm{U}^{\ell}_{intra}$ is the intramolecular electrostatic interaction of the ligand. Now the PDLD/S results obtained with a single protein-ligand configuration cannot capture properly the effect of the protein reorganization (see discussion in Sham et al.[26]) and a more consistent treatment should involve the use of the LRA or related approaches (e.g., Lee et al.[4] and Sham et al.[26]). This approach provides a reasonable approximation for the corresponding electrostatic free energies by using the equivalent of Eq. (5):

$$\Delta G^{p}_{\mathrm{elec},\ell} = \frac{1}{2} [\langle \overline{U}^{p}_{\mathrm{elec},\ell} \rangle_{\ell'} + \langle \overline{U}^{p}_{\mathrm{elec},1} \rangle_{\ell}]$$

$$\Delta G^{w}_{\mathrm{elec},\ell} = \frac{1}{2} [\langle \overline{U}^{w}_{\mathrm{elec},\ell} \rangle_{\ell'} + \langle \overline{U}^{w}_{\mathrm{elec},\ell} \rangle_{\ell}] \quad (27)$$

where the effective potential $\overline{\mathrm{U}}$ is defined in Eq. 26, $\langle\ \rangle_{\ell}$ and $\langle\ \rangle_{\ell'}$ designate an MD average over the coordinates of the ligand-complex in their polar and non-polar forms. It is important to realize that the average of Eq. (27) is always done where both contributions to the relevant $\overline{\mathrm{U}}_{\mathrm{elec}}$ are evaluated at the same configurations. That is, the PDLD/S energies of the polar and non-polar states are evaluated at each averaging step by using the same structure. However, we generate two set of structures one from MD runs on the polar state and one from MD runs on the non-polar state. This is basically the same approach used in the microscopic LRA but now with the effective potential $\overline{U}_{\mathrm{elec}}$.

Next we have to evaluate the binding free energy of the non-polar ligand, $\Delta \mathrm{G}_{\mathrm{bind},\ell'}$, of the lower cycle of Figure 4, where we mutate ℓ' to a shrinked ligand ℓ''. This mutation procedure gives more stable results than those obtained by our previous strategy that involved the direct b $\rightarrow$ c step, which requires a direct evaluation of the solvation energy of the isolated protein. In implementing step f $\rightarrow$ c, it is not so practical to use the LRA approach. The problem is that running trajectories with the potential of ℓ'' may lead to exponentially large contributions from the vdw energies of ℓ'. Thus, we approximate $\Delta \mathrm{G}_{\ell'' \rightarrow \ell'}$ by

$$\Delta G^{p,\mathrm{elec}}_{\ell'' \rightarrow \ell'} \cong \langle \overline{U}^{p}_{\mathrm{elec},\ell'} \rangle_{\ell'} - \langle \overline{U}^{p}_{\mathrm{elec},\ell''} \rangle_{\ell''} = \langle \Delta \overline{U}^{p}_{\mathrm{elec}\ell'' \rightarrow \ell'} \rangle_{\ell'} \quad (28)$$

where we average the PDLD contributions for ℓ' over trajectories obtained with the potential of ℓ' and average the contributions from ℓ'' over trajectories obtained with the potential of ℓ''. The effect of the protein reorganization upon change from ℓ'' to ℓ' is neglected, since it is assumed that the major part of the protein relaxation occurs in the charging step ($\ell \rightarrow \ell'$) (see however, the Concluding Remarks section). One may wonder why our electrostatic contribution is different than zero since the protein-ligand electrostatic interaction is zero and thus the corresponding FEP or LRA contribution will be zero. However, in an actual FEP procedure, that considers the $\ell' \rightarrow \ell''$ transformation, the non-zero van der Waals contribution depends on the reorganization of the system. This contribution reflects the change in protein-water (rather than protein-ligand) electrostatic interaction as a result of water penetration and this effect is approximated by our $\Delta \mathrm{G}^{p,elec}_{\ell'' \rightarrow \ell'}$. The same approximation used in the evaluation of $\Delta \mathrm{G}^{p,elec}_{\ell'' \rightarrow \ell'}$ is used for the vdw and hydrophobic term giving

$$\Delta G^{p,vdw}_{\ell'' \rightarrow \ell'} - \Delta G^{w,vdw}_{\ell'' \rightarrow \ell'} + \Delta G^{p,hyd}_{\ell'' \rightarrow \ell'} - \Delta G^{w,hyd}_{\ell'' \rightarrow \ell'} = \langle \overline{U}^{p}_{vdw,\ell'} \rangle_{\ell'} - \langle \overline{U}^{p}_{vdw,\ell''} \rangle_{\ell'} + \langle \overline{U}^{p}_{hyd,\ell'} \rangle_{\ell'} - \langle \overline{U}^{p}_{hyd,\ell''} \rangle_{\ell'} \quad (29)$$

where $\overline{U}_{\mathrm{vdw},\ell''}$ designates the PDLD vdw interaction and $\overline{U}_{\mathrm{hyd}}$ designates the field-dependent PDLD hydrophobic energy.[27] Now our final expression for $\Delta \mathrm{G}_{\mathrm{bind},\ell'}$ is

$$\begin{aligned} \Delta G_{\mathrm{bind},\ell'} &= \Delta G^{p}_{\ell'' \rightarrow \ell'} - \Delta G^{w}_{\ell'' \rightarrow \ell'} \\ &= \Delta G^{p,elec}_{\ell'' \rightarrow \ell'} + \Delta G^{p,vdw}_{\ell'' \rightarrow \ell'} + \Delta G^{p,hyd}_{\ell'' \rightarrow \ell'} \\ &\quad - \Delta G^{w,vdw}_{\ell'' \rightarrow \ell'} - \Delta G^{w,hyd}_{\ell'' \rightarrow \ell'} - T(\Delta \Delta S'_{\ell'})^{w \rightarrow p} \\ &= \Delta G^{PDLD/S}_{bind,\ell'} - T(\Delta \Delta S'_{\ell'})^{w \rightarrow p} \end{aligned} \quad (30)$$

The term $(\Delta \Delta \mathrm{S}'_{\ell'})^{\mathrm{w} \rightarrow \mathrm{p}}$ represents the loss of configurational entropy upon moving ℓ' from water to the protein active site rather than that associated with e $\rightarrow$ f step. This is, of course, not the case in FEP studies where the entropic contribution is obtained in the $\ell' \rightarrow \ell''$ steps. As in the case of Eq. (14), we scale the contributions to $\Delta \mathrm{G}_{\mathrm{bind},\ell'}$ by θ and use

$$\Delta G_{bind,\ell'} = \theta \{ \Delta G^{PDLD/S}_{bind,\ell'} - \gamma [T(\Delta \Delta S^{w \rightarrow p}_{\ell'})'] \} \quad (31)$$

In addition to the PDLD/S approach we also examine here the PDLD approach where each term in the PDLD/S expression is replaced by the corresponding PDLD term.

The results of the different approaches discussed above depend, of course, on the ionization state of different protein residues and thus on the pH of the surrounding solvent. This pH effect can be calculated by our PDLD/S-LRA approach. In this approach, we calculate first the intrinsic pK_{a} of each ionizable residue and then evaluate the effective interaction between these residues by a

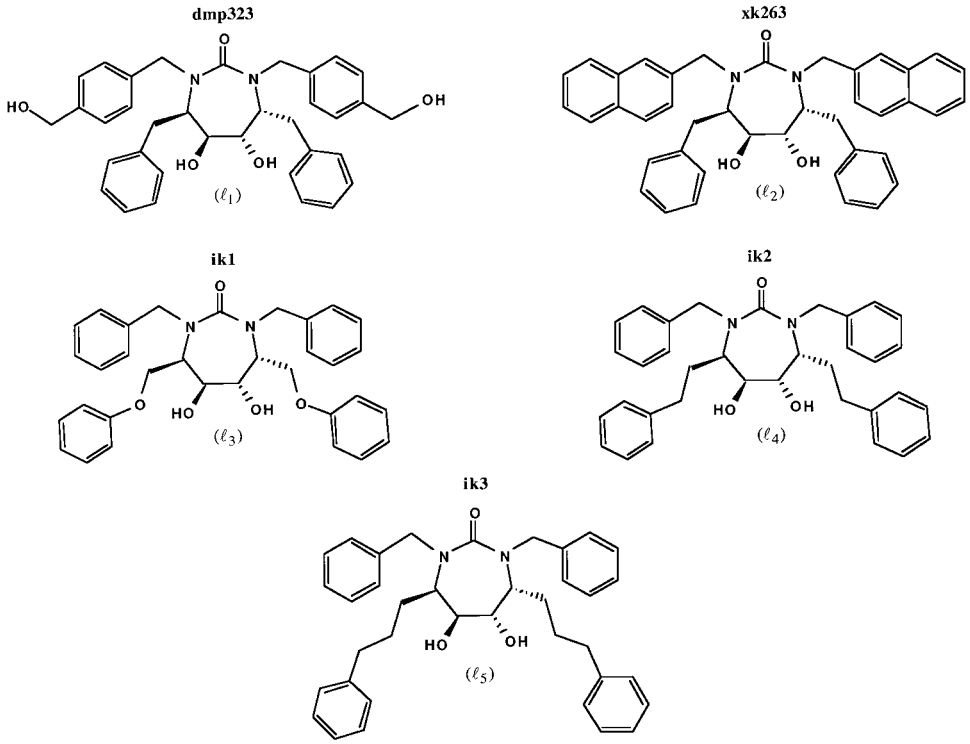

Fig. 5. The structure and notation of the ligands used in the present study. The ligand ℓ_1, ℓ_2, ℓ_3, and ℓ_4 correspond to DMP323 and its derivatives (1a, 2, and 3 according to the notation of Hultén et al.[31]) ℓ_5 corresponds to XK263.

hybrid approach (for details see Sham et al.[26]). The calculations include the ligand (ℓ) in its polar state.

RESULTS

Our benchmark consists of a class of neutral C2-symmetric cyclic urea based molecules shown in Figure 5. These molecules are potent inhibitors of the Human Immunodeficiency Virus (HIV) protease, which play an essential role in the maturation of the HIV virus.[28] The special feature of these inhibitors is the incorporation of a carbonyl oxygen groups onto a seven member cyclic urea ring that replace the commonly observed tetracoordinated crystallographic water.[29,30] This water molecule is a unique feature of the retroviral protease and the replacement of this water is presumed to be energetically favorable and should enhance the binding specificity of the inhibitor to the HIV protease. The effectiveness of the designed has been confirmed by high resolution X-ray crystallographic studies and the system has been subjected to theoretical studies that examine the predictive power of the LIE method.[31] Here we take this system as our benchmark in order to establish the relationship between the all-atom LIE method and the various approaches described in the previous section.

LRA and Related Models

The LRA simulations, which also provide the LIE results, were performed using the ENZYMIX program[27] with its Surface Constrained All-Atom Solvent (SCAAS) spherical boundary conditions.[32,33] This involve a 16 Å simulation sphere which is surrounded by a bulk continuum. The electrostatic interactions were treated by the Local Reaction Field (LRF) method. [34] The simulations utilized the unified-atom GROMOS force field (in order to simplify the comparison of the results of Hultén et al.[31]). The simulation involve 10 ps equilibration followed by a 4 ns data collection with 1 fs time step at 300°K. The binding free energies should depend on the ionization states of the protein residues at the pH used in the given experimental study. These ionization states can be evaluated by using the PDLD/S-LRA method (see Sham et al.[26]) for the protein with the ligand ℓ_1. Here, however, we used the ionization states used previously by Hultén et. al.,[31] which only considered the ionization of the nearest ionizable residues (Arg8, Arg8' Arg87, Arg87', Asp25, Asp29, Asp29', Asp30, and Asp30') with the exception of Asp25'.

The electrostatic energy contributions obtained by the simulations are summarized in Table I. As can be seen from table the averages obtained by running trajectories on the non-polar state $\langle\ \rangle_{\ell'}$ are practically zero in water (as it should be) and is also very small in the protein. This may justify the approximation of neglecting the $\langle\ \rangle_{\ell'}$ term in the LIE approach. However, in some cases where the protein increases the stabilization of a polar or charged ligand by a preoriented polar environment one would expect a large contribution from $\langle\ \rangle_{\ell'}$. This is clearly the situation in cases of ionizable residues that were considered in previous pK_a's studies.[26] Thus we conclude that in cases of non-polar ligands it may be reasonable to neglect the $\langle\ \rangle_{\ell'}$ term, but this approximation might not be fully justified in cases of very polar or charged ligands.

The non-electrostatic contributions to the LIE approach are given in Table II. The table presents the results of the original LIE treatment, which considers an average of the

TABLE II. Calculated LIE Non-Electrostatic Contributions to Binding Free Energies for Cyclic Urea-Based HIV Protease Inhibitors

	MD averaged over U_{ℓ}^{a}					MD averaged over $U_{\ell'}^{b}$				
Inhibitor	$\langle U_{vdw,\ell}^{p}\rangle_{\ell}$	$\langle U_{vdw,\ell}^{p}\rangle_{\ell}$	ΔG_{vdw}^{p}	ΔG_{vdw}^{p}	$\Delta\Delta G_{vdw}^{w\rightarrow p}$	$\langle U_{vdw,\ell'}^{p}\rangle_{\ell'}$	$\langle U_{vdw,\ell'}^{p}\rangle_{\ell'}$	ΔG_{vdw}^{p}	ΔG_{vdw}^{p}	$\Delta\Delta G_{vdw}^{w\rightarrow p}$
ℓ_1	−89.4	−41.7	−14.3	−6.7	−7.6	−99.4	−54.9	−16.9	−9.3	−7.6
ℓ_2	−99.2	−54.4	−15.9	−8.7	−7.2	−104.9	−59.4	−17.8	−10.1	−7.7
ℓ_3	−79.0	−45.8	−12.6	−7.3	−5.3	−85.4	−52.4	−14.5	−8.9	−5.6
ℓ_4	−83.9	−46.3	−13.4	−7.4	−6.0	−88.8	−53.6	−15.1	−9.1	−6.0
ℓ_5	−95.1	−51.5	−15.2	−8.2	−7.0	−94.9	−57.2	−16.1	−9.7	−6.4

[a]Non-electrostatic contributions (in kcal/mol) to the binding free energies, where the corresponding $\Delta\Delta G_{vdw}$'s are evaluated by scaling the average vdw potential ($<U_{vdw}>$) of the polar ligand in water (w) and in protein (p) by a factor β of 0.16 (i.e., $\Delta G_{vdw} = \beta\langle U_{vdw,\ell}\rangle_{\ell}$).
[b]Alternative estimation of the non-electrostatic contribution to the binding free energies, where the $\Delta\Delta G_{vdw}$'s are evaluated by scaling the average vdw potential ($<U_{vdw}>$) of the non-polar ligand (ℓ') in water (w) and in protein (p) by a factor β of 0.17 (i.e., $\Delta G_{vdw} = \beta\langle U_{vdw,\ell'}\rangle_{\ell'}$). In this way we approximate the binding of the non-polar ligand.

TABLE III. Calculated and Observed LRA and LIE Binding Free Energies (in kcal/mol) for Cyclic Urea Based HIV Protease Inhibitors

Inhibitor	ΔG_{bind}^{a}	ΔG_{bind}^{b}	ΔG_{bind}^{c}	ΔG_{bind}^{d}	ΔG_{bind}^{e}	ΔG_{bind}^{f}	ΔG_{bind}^{obs}
ℓ_1	−14.3	−12.1	−12.1	−11.7	−11.1	−10.0	<−12.8
ℓ_2	–	−10.5	−11.0	−11.7	−13.5	−14.0	−12.9
ℓ_3	−12.7	−10.6	−10.9	−10.9	−12.2	−12.2	−11.2
ℓ_4	−11.5	−10.1	−10.1	−10.6	−9.7	−9.3	−9.5
ℓ_5	−12.8	−11.3	−10.7	−10.9	−8.3	−5.8	−8.9

[a]Previously obtained LIE binding free energies.[29] These results were obtained by simulation that involve the SHAKE procedure.
[b]Calculated binding free energies obtained using LIE method [Eq. 12] with α = 0.5 and β = 0.16).
[c]Calculated binding free energies obtained using a modified LIE approach where the β term is evaluated from $<U_{vdw,\ell'}>_{\ell'}$. The scaling used is α = 0.5 and β = 0.17.
[d]Calculated binding free energies obtained using Eq. (8) where the electrostatic contributions are evaluated by the LRA method and the non-electrostatic term is evaluated by $\beta < U_{vdw,\ell'} >_{\ell'}$ with a scaling factor β = 0.15.
[e]Calculated binding free energies obtained using Eq. (14) where the electrostatic contributions are evaluated by the LRA method and the non-electrostatic contributions are evaluated by the PDLD approach without the entropy contribution and with a scaling factor θ = 0.49.
[f]Calculated binding free energies obtained using Eq. (14) where the electrostatic contributions are evaluated by the LRA method and the non-electrostatic contributions are evaluated by the PDLD approach and also includes an entropy contribution scaled by 0.25 (γ = 0.25). The overall scaling of the non-electrostatic term is θ = 0.61.

vdw energy over the force field of the polar ligand and our alternative treatment that involve an average over the force field of the non-polar ligand. This alternative treatment offers a clearer connection to the LRA cycle of Figure 2 and thus a simpler way of exploring the origin of the β term in the LIE approach.

The calculated binding LRA and LIE free energies are summarized in Table III. The table includes different expressions for ΔG_{bind}. This includes, among others, the ΔG_{bind}^{LIE} ($\Delta G_{bind}^{(b)}$) obtained from Eq. (12), the ΔG_{bind}^{LRA} ($\Delta G_{bind}^{(d)}$) obtained from the more rigorous LRA electrostatic energy [Eq. (8)] and the non-electrostatic term evaluated from $\beta\langle U_{vdw,\ell'}\rangle_{\ell'}$. The table also includes the results of several variants of Eq. (14), where the electrostatic terms are evaluated by the LRA approximation and $\Delta G_{bind,\ell'}$ is evaluated by the PDLD approximation with and without a scaled entropy term. The general agreement between the calculated and observed results is similar in the different approximations, although $\Delta G_{bind}^{(c)}$ and $\Delta G_{bind}^{(e)}$ (in the Table notation) may provide a somewhat better results. This is, however, not a central point of the present study. Our main points are the following: i) Quite different approaches gives similar results once they are scaled by β or a related parameter. ii) The rigorous LRA electrostatic term can be approximated by the corresponding term of Eq. (11), although this may not be the case in preorganized binding sites. iii) Taking the β term from the vdw interaction of the polar and non-polar ligands gives similar results. iv) Using Eq. (14) and evaluating the non-electrostatic contributions by the PDLD approach without the entropy contribution ($\Delta G_{bind}^{(e)}$) gives good results with θ ≈ 0.5. Adding the entropy term with a scaling γ = 0.25 also gives reasonable results with θ ≈ 0.6. Regardless of the convergence of the entropy term (see PDLD/S-LRA and PDLD-LRA Results section), it is clear that the PDLD estimate of the non-electrostatic cycle provides an effective way of completing the LRA cycle. This means that our original LRA treatment,[4] that used the PDLD energy for the lower cycle may be as effective as the subsequent LIE method.

Assuming that the electrostatic contributions are evaluated reliably by the LRA treatment we may conclude that the scaled non-electrostatic term gives a correct estimate of the magnitude of the correct value of $\Delta G_{bind,\ell'}$. Thus a treatment that involves minimal scaling (i.e., θ ≈ 1) is likely to capture correctly the nature of the contributions

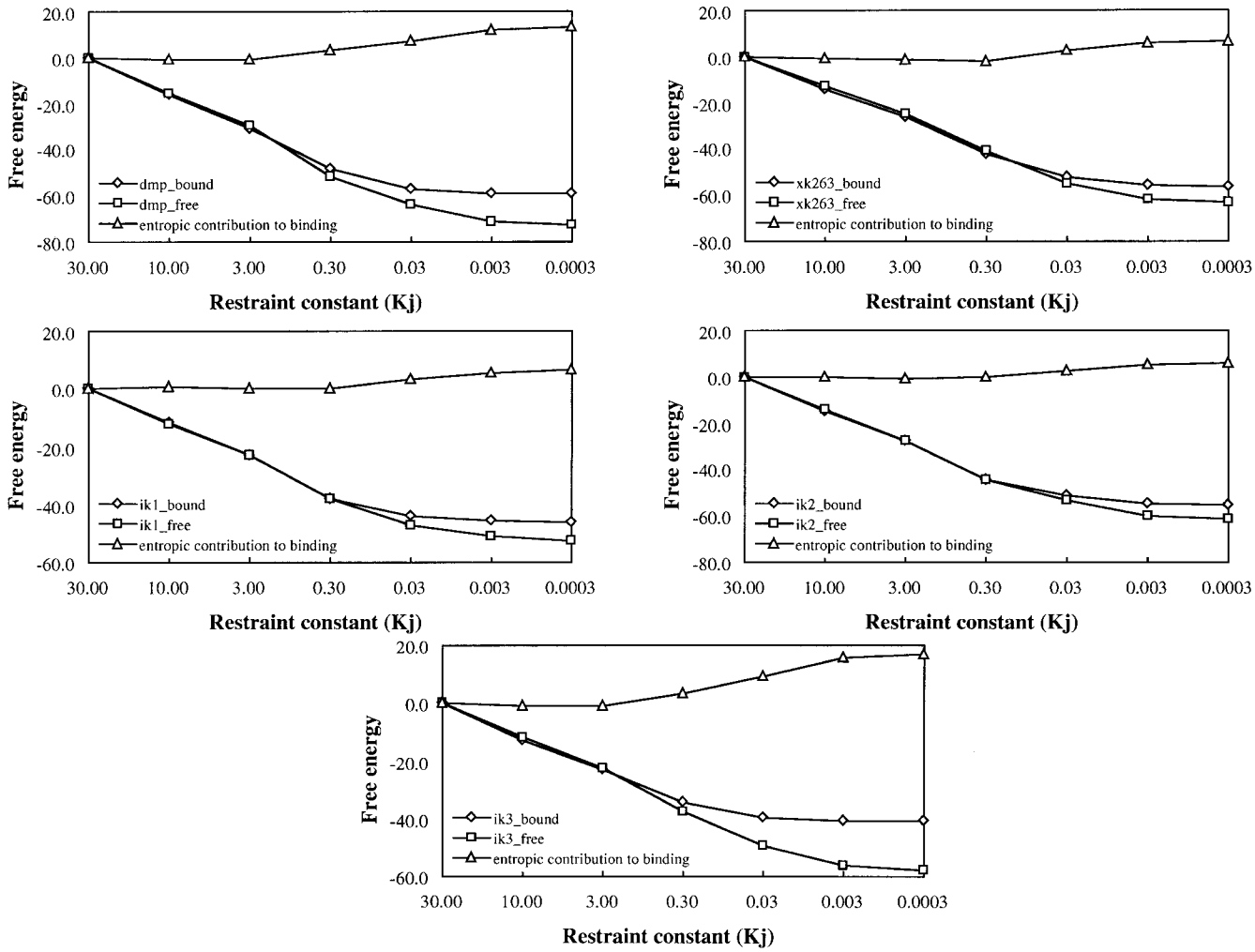

Fig. 6. The contributions to binding free energies from configurational entropies as a function of the restraint force constant (K_j). The figure gives the contributions for change from $K_1 = 30$ to $K = K_j$. The contributions obtained with the bound ligand ($\Delta G'_3$) and for the ligand in water ($\Delta G'_4$) are designated by diamonds and squares respectively. The difference between ($\Delta G'_3$) and ($\Delta G'_4$), which is the desire entropic contribution to binding energy, is designated by triangles. The figure does not include the effective volume correction of Eq. (21).

that gives $\Delta G_{\mathrm{bind},\ell'}$ and help in rationalizing the LIE approximation (see discussion in the concluding remarks section). Our most consistent attempt to evaluate the non-electrostatic term involves the use of the PDLD approach and Eq. (14). This approach considers the balance between the hydrophobic, vdw, solvation, and entropic contributions in the binding of the non-polar ligand. Obtaining the proper balance is, of course, a major challenge and the fact that β is relatively large is encouraging. Evidently the balance is not perfect, in particular, the calculated entropy is probably overestimated and thus requires a special scaling (see below).

PDLD/S-LRA and PDLD-LRA Results

The PDLD/S and the PDLD/S-LRA approaches (as well as the corresponding PDLD treatment) should provide an alternative to the more rigorous approaches of the previous section. Here we examine the performance of the PDLD/S-LRA and PDLD-LRA approaches considering the cycle $a \rightarrow b \rightarrow e \rightarrow f \rightarrow c \rightarrow d$ of Figure 2 where $\Delta G_{\mathrm{bind},\ell'}$ is calculated by the lower cycle. The actual electrostatic calculations involve the cycles of Figure 4 where the free energies of the inner cycles that corresponds to the cycles of Figure 2 are evaluated by the external cycles. These cycles represent the PDLD/S process of changing the dielectric of the surrounding solvent from that of water (ϵ_w) to the effective dielectric of the protein, ϵ_{in}, and back from ϵ_{in} to ϵ_w (see also PDLD-LRA and PDLD/S-LRA Calculations section). The PDLD/S contribution for each step of the cycle is obtained by the proper scaling of the corresponding microscopic estimate of the Langevin Dipoles (LD) free energy (see Fig. 4). The details of this procedure and the meaning of ϵ_{in} are discussed extensively in other works.[25,26] The PDLD/S-LRA calculation for the electrostatic contribution $\Delta G_{a\rightarrow b}$ and $\Delta G_{c\rightarrow d}$ involves PDLD/S averages over 20 configurations generated by the same MD simulations used in the LRA study of the previous section. The configuration selected correspond to

TABLE IV. Calculated Electrostatic Contributions to the Binding Free Energies of Cyclic Urea-Based HIV Protease Inhibitors†

	PDLD				PDLD/S			
Inhibitor	$\langle \overline{U}^{p}_{elec,\ell}\rangle_{\ell}$	$\langle \overline{U}^{p}_{elec,\ell}\rangle_{\ell'}$	$\langle \overline{U}^{w}_{elec,\ell}\rangle_{\ell}$	$\langle \overline{U}^{w}_{elec,\ell}\rangle_{\ell'}$	$\langle \overline{U}^{p}_{elec,\ell}\rangle_{\ell'}$	$\langle \overline{U}^{p}_{elec,\ell}\rangle_{\ell'}$	$\langle \overline{U}^{w}_{elec,\ell}\rangle_{\ell}$	$\langle \overline{U}^{w}_{elec,\ell}\rangle_{\ell'}$
ℓ_1	−66.3	−6.7	−20.9	−18.1	−34.3	−3.5	−13.7	−11.9
ℓ_2	−35.9	−6.0	−12.0	−10.1	−19.0	−3.2	−7.8	−6.6
ℓ_3	−40.3	−8.6	−17.0	−14.2	−21.4	−4.9	−11.1	−9.3
ℓ_4	−38.5	−7.7	−15.7	−12.9	−20.3	−4.5	−10.3	−8.4
ℓ_5	−45.9	−8.8	−16.3	−13.7	−23.8	−4.7	−10.7	−9.0

†The electrostatic contributions to the binding free energies (in kcal/mol). These contributions are defined by Eq. (27) and are calculated using the PDLD and the PDLD/S methods. The calculations correspond to the a → b → c → d cycle of Fig. 4. The relevant free energies are obtained from Eq. (27) using the average of the change in the electrostatic potential ($\overline{U}_{elec}$) over the charged (ℓ) and the uncharged (ℓ') states of the ligand in water (w) and in protein (p). The calculations were performed on 20 structures generated by MD simulation using the GROMOS force field with 1 fs time steps at 300 K and 20 ps intervals.

TABLE V. Calculated PDLD "Non-Electrostatic" Contributions to the Binding Free Energies of Cyclic Urea-Based HIV Protease Inhibitors†

Inhibitor	$(\Delta\Delta G^{elec}_{\ell''\to\ell'})^{w\to p}$	$(\Delta\Delta G^{vdw}_{\ell''\to\ell'})^{w\to p}$	$(\Delta\Delta G^{hyd}_{\ell''\to\ell'})^{w\to p}$	$-T(\Delta\Delta S'_{\ell'})^{w\to p}$
ℓ_1	20.3	−17.0	−15.8	17.1
ℓ_2	27.1	−29.2	−15.6	10.7
ℓ_3	15.4	−19.2	−9.0	10.3
ℓ_4	17.5	−16.9	−9.5	9.9
ℓ_5	25.6	−18.1	−13.7	20.9

†The terms $\Delta\Delta G_{\ell''\to\ell'}$ represent the corresponding differences (e.g., $(\Delta\Delta G^{hyd}_{\ell''\to\ell'})^{w\to p} = \Delta G^{hyd,p}_{\ell''\to\ell'} - \Delta G^{hyd,w}_{\ell''\to\ell'}$). $\Delta\Delta G^{elec}_{\ell''\to\ell'}$ is reffered to as a "non-electrostatic" term since it involves the nonpolar ligand. The ΔΔG terms are obtained (in kcal/mol) from Eqs. (28–30) using the PDLD method. Each of these terms is an averaged of 20 configurations for the non-polar state of the ligand. The entropy contribution, $-T(\Delta\Delta S'_{\ell'})^{w\to p}$, is estimated by the FEP approach described in Entropy Calculations section.

10 ps intervals. It is important to note that the PDLD/S calculations are quite stable and the similar results were obtained by PDLD/S average configurations generated with much shorter MD simulations.

As explained before, we evaluated $\Delta G_{bind,\ell'}$ by PDLD calculations and augmented this result by entropy calculations. The entropic contributions were calculated as depicted in Figure 6 according to approach of the Entropy Calculations section, where the restraint K was change from 30 to 0.003 kcal mol-1 Å^2 in a sequential FEP simulations each of ten mapping steps and total simulation time of 110 ps. The effective volume restraint, K_{eff}, was taken as 1 kcal mol^{-1} Å^2 which corresponds to an effective volume of 2.2 Å^3 and a volume correction of 3.9 kcal mol^{-1}. The final contributions (of $-T\Delta\Delta S'$) were 17.1, 10.7, 10.3, 9.9, and 20.9 kcal mol^{-1} for ℓ_1, ℓ_2, ℓ_3, ℓ_4, and ℓ_5. Although these results reproduce the fact ℓ_1 and ℓ_5 have the largest number of freely rotated bonds they do not appear to be reliable enough for our purpose. This is obvious from the fact (see Table VI) that the entropic contributions have to be scaled by 0.25 to reproduce a reasonable $\Delta G_{bind,\ell'}$. It is possible that if the problems are associated with some type of convergence problems, although longer simulations reproduce similar trend. We are probably dealing here with similar difficulties to those associated with direct attempts to calculate binding free energies by FEP approaches and with the fact that entropy calculations converge very slowly. At any rate the point of this paper is not to overcome the challenge of obtaining reliable microscopic FEP results by direct simulations nor to overcome the challenge of obtaining reliable entropic value. We only tried to explore the chance of obtaining some trend in the calculated entropy, which can be explored in improving the estimates of $\Delta G_{bind,\ell'}$.

The results of PDLD/S-LRA and PDLD-LRA simulations are summarized in Tables IV–VI. The LRA electrostatic contributions of the PDLD method are scaled by α = 0.25 rather than 0.5. This is due to the fact that the Langevin dipoles are treated like induced dipoles, which are instantly relaxed in the presence of any given charge set. In this way we get in water the similar contribution for average over state a and state b. This feature has been discussed for example in our study of reorganization energies in cytochrome c.[35] It is not hard to modify the PDLD treatment in a way that the Langevin dipoles polarized by state a will be kept in the same polarization when they interact with state b. Such a treatment is not likely, however, to change significantly the results of our calculation since the model was calibrated on solvation in water. In the PDLD/S model we used α = 0.5, which is consistent with the microscopic LRA scaling. It should be noted, however, that ϵ_{in} is a parameter that depends on the actual treatment used and it would probably have a different value if we used a LD model, which is more consistent with microscopic response of the solvent dipoles.

As to the scaling of the non-electrostatic part, we manage to have a relatively large value of θ (θ = 0.6 and

TABLE VI. Calculated Binding Free Energies of Cyclic Urea-Based HIV Protease Inhibitors[†]

	PDLD					PDLD/S					
Inhibitor	$\Delta\Delta G_{elec}^{w\to p}$	$(\Delta\Delta G_{non\text{-}elec}^{w\to p})^b$	$(\Delta\Delta G_{non\text{-}elec}^{w\to p})^c$	$(\Delta G_{bind}^{calc})^b$	$(\Delta G_{bind}^{calc})^c$	$\Delta\Delta G_{elec}^{w\to p}$	$(\Delta\Delta G_{non\text{-}elec}^{w\to p})^b$	$(\Delta\Delta G_{non\text{-}elec}^{w\to p})^c$	$(\Delta G_{bind}^{calc})^b$	$(\Delta G_{bind}^{calc})^c$	ΔG_{bind}^{obs}
ℓ_1	−8.5	−5.7	−4.8	−14.2	−13.3	−6.1	−7.2	−6.0	−13.4	−12.1	<-12.8
ℓ_2	−4.9	−8.1	−8.9	−13.1	−13.8	−3.8	−10.3	−11.0	−14.1	−14.8	−12.9
ℓ_3	−4.4	−5.8	−6.0	−10.3	−10.4	−2.9	−7.4	−7.4	−10.3	−10.4	−11.2
ℓ_4	−4.4	−4.1	−3.8	−8.5	−8.2	−3.0	−5.2	−4.7	−8.2	−7.7	−9.5
ℓ_5	−6.2	−2.8	−0.6	−9.0	−6.7	−4.4	−3.6	−0.7	−8.0	−5.1	−8.9

[†]The electrostatic and non-electrostatic contributions are calculated (in kcal/mol) by the PDLD and the PDLD/S methods. $\Delta\Delta G_{elec}^{w\to p}$ is the scaled electrostatic contribution [the scaling constant α is 0.25 for the PDLD and 0.5 for the PDLD/S (see discussion in the PDLD/S-LRA and PDLD-LRA Results section]. $(\Delta\Delta G_{non\text{-}elec}^{w\to p})^b$ is the sum of the scaled non-electrostatic contributions obtained without the $-T(\Delta\Delta S_{\ell'}^{w\to p})$ contribution (the scaling constant is 0.46 for the PDLD and 0.58 for the PDLD/S). $(\Delta\Delta G_{non\text{-}elec}^{w\to p})^c$ includes the entropic contributions of the $-T(\Delta\Delta S_{\ell'}^{w\to p})$ term which was sclaed by $\gamma = 0.25$ (see Eq. (14)). The scaling constant, θ, for the total non-electrostatic ocntribution was taken as 0.59 for the PDLD and 0.73 for the PDLD/S).

0.73 for the PDLD and PDLD/S models, respectively). This means, as discussed in the LRA and Related Models section, that the PDLD and PDLD/S models capture the compensation between the contributions to $\Delta G_{bind,\ell'}$. This compensation is not quantitative, however, and more studies will be needed to fully resolve the nature of the non-electrostatic term.

At any rate, the overall trend in the binding calculations is as good and perhaps better than that obtained by the LRA and LIE method. This does not mean that the PDLD/S is better than the LIE method but that it offers a practical and efficient alternative. Furthermore, as shown in the previous section the PDLD/S provides a way of obtaining the non-electrostatic term with well defined physical terms.

CONCLUDING REMARKS

This work examined several strategies for evaluating binding free energies with a particular emphasis on the relationship between the LRA and LIE approximations and the performance of the PDLD/S-LRA and PDLD-LRA methods. A clear thermodynamic cycle (which is directly related to our earlier cycle[4]) is defined for the LRA and LIE method. This cycle helps in studying the nature of the energy contributions that leads to the overall binding free energy. The performance of the different approximation is examined by calculating the binding free energies of cyclic urea-based ligands to HIV protease. Although the performance of the different approaches is similar in the present case, it is possible to deduce some useful conclusions about the relationship between these different treatments. It is shown that the electrostatic term of the LRA approach can be approximated in our specific test case by the corresponding LIE term, although the LIE approach neglects the LRA contribution from the average over the state with the non-polar ligand (the $\langle\ \rangle_{\ell'}$ term). However, we expect that the $\langle\ \rangle_{\ell'}$ term would not be negligible in cases when the protein provides preorganized environment that stabilizes the ligand residual charges and this problem might be significant in cases of the charged ligands.

While the electrostatic term can be determined quite accurately by the LRA approach, it is much harder to estimate the non-electrostatic term (the $\Delta G_{bind,\ell'}$ term). The LIE approach evaluates this contribution by using the β term of Eq. (12). Although this approach is quite effective, the origin of the scaling constant, β, is still not clear. It is tempting to determine the nature of β by using the lower cycle of Figure 2 and evaluating the $\Delta G_{\ell'\to\ell}$ term using the LRA approximation so that

$$\Delta G_{\ell'\to\ell''} \cong \theta[\langle U_{vdw,\ell'}\rangle_{\ell'} + \langle U_{vdw,\ell'}\rangle_{\ell''}] \approx \beta[\langle U_{vdw,\ell'}\rangle_{\ell''}] \quad (32)$$

The problem is that the LRA approach as well as the corresponding FEP approach involves large fluctuations of the contribution from U_{vdw} at ℓ'' ($\langle U_{vdw,\ell'}\rangle_{\ell''}$). One might still hope that the vdw contributions satisfies somehow the trend of the LRA relationship and that $\langle U_{vdw,\ell'}\rangle_{\ell''}$ has a similar trend to that of $\langle U_{vdw,\ell'}\rangle_{\ell'}$. Unfortunately we do not obtain at present convergent results for such an LRA treatment and, of course, we do not have an a priori way to determine β. This problem can be addressed by using our LRA approach [Eq. (14)] where the non-electrostatic term is evaluated using the vdw, hydrophobic, and water penetration PDLD terms as well as a microscopic estimate of the change in configurational entropy. Assuming that these are all the relevant contributions to $\Delta G_{bind,\ell'}$ (see below, however) and exploiting the fact that the LRA electrostatic contribution to the binding energy is relatively reliable, so it can be compared to the corresponding "observed" value (although this quantity cannot be measured), we can write

$$\Delta G_{bind,\ell'}^{obs'} \cong \Delta G_{bind}^{obs} - (\Delta\Delta G_{elec,\ell}^{w\to p})^{LRA} \quad (33)$$

where obs′ designates an estimate of the corresponding "observed" value. If we require that both the LRA and LIE non-electrostatic contributions will reproduce $\Delta G_{bind,\ell'}^{obs'}$, we obtain from Eqs (33), (12) and (13)

$$\beta[\langle U_{vdw,\ell'}^{p}\rangle_{\ell'} - \langle U_{vdw,\ell'}^{w}\rangle_{\ell'}] \cong \theta\{\Delta G_{bind,\ell'}^{PDLD} - \gamma[T(\Delta\Delta S_{\ell'}^{w\to p})']\} = \Delta G_{bind,\ell'}^{obs'} \quad (34)$$

This gives

$$\beta = \frac{\theta\{\Delta G_{bind,\ell'}^{PDLD} = \gamma[T(\Delta\Delta S_{\ell'}^{w\to p})']\}}{[\langle U_{vdw,\ell'}^{p}\rangle_{\ell'} - \langle U_{vdw,\ell'}^{w}\rangle_{\ell'}} \quad (35)$$

Here we retain the γ scaling of the entropic contribution since it is not clear if this term reaches a reasonable convergence. At any rate, obtaining the proper balance between the contribution of Eq. (34) should give $\theta = 1$. The present study gives θ values between 0.5 and 0.73 depending on the approximation used. This is encouraging but not sufficient to make a unique analysis of the nature of the LIE β or to predict what this β should be. We might be able, however, to use our PDLD approach and Eq. (35) to estimate the change in the LIE β between different proteins active sites.

It is interesting to consider at this stage the LRA approach introduced in our original work.[4] This approach used the LRA method for the electrostatic part of the thermodynamic cycle and evaluated the non-electrostatic part by the PDLD approach with its vdw, hydrophobic, and water penetration contributions. As shown in the present work this strategy gives as good results as the LIE approach. Apparently the thermodynamic cycles of the LRA and LIE methods are closely related (see the cycles in Lee et al.[4], Marelius et al.,[21] and Fig. 2 of this work as well as the discussion in the LRA Calculations section). This means that the real difference between the LIE and LRA methods is the fact that the LIE estimates the lower cycle by the scaled vdw interaction, while our LRA approach evaluates this contribution by the relevant PDLD terms. The fact that the LIE method neglects the $\langle\ \rangle_{\ell'}$ LRA contribution is not a fundamental difference.

Although we take into account the protein reorganization upon uncharging of the ligand by the LRA treatment of the electrostatic term, we do not take into account the protein reorganization upon binding of the non-polar ligand. This term could have been reflected in the calculations if an LRA treatment of the $\ell''\rightarrow\ell'$ step was valid.

As clarified in this work the LRA approach allows one to obtain from the average interaction energy the corresponding free energy (at least when the LRA is valid). However, it seems to us that this fact has not been perhaps fully appreciated by those who have examined and used the LIE approach, without considering its relation to the LRA approximation. For example, an interesting recent work[36] expressed the binding free energy as $\Delta G = -RT \ln\{\langle \exp(-E_{int}/RT)\rangle\}$ using $E_{int} = \alpha U_{elec} + \beta U_{vdw}$ where α and β are the LIE scaling factors. It was then suggested that such an expression can capture the entropic contributions to the binding energy, which is assumed to be missing in the LIE treatment. However, the LIE expression corresponds already to the free energy rather than just the interaction energy.

One might speculate here about the reason why the LIE scaling of $\langle U_{vdw}\rangle$ gives reasonable estimate of the sum of the different terms that make up $\Delta G_{bind,\ell'}$. When the ligand is in water, the $\langle U_{vdw}\rangle$ term is proportional to the hydrophobic and vdw terms in water. It thus may also be proportional to the corresponding terms in the protein. However, the possible proportionality to the water penetration and configurational entropies terms are harder to rationalize. It is possible, however, the change in solvation of the protein upon removal of the non-polar ligand is related to the ligand-protein vdw term, since this term is related to the volume available for water penetration. Similarly it is possible that $\langle U^{P}_{vdw}\rangle$ effect associated with the restriction of the ligand motion by the protein active site, since a larger restriction may result in larger $\langle U^{P}_{vdw}\rangle$. It might also tempting to use the argument presented in the derivation of Eq. (32). However, none of the above arguments is expected to provide a universal β.

Finally it is significant that the PDLD-LRA and PDLD/S-LRA give results of similar quality to those obtained by the LRA and LIE methods. This is important since the PDLD/S method is much faster than the more microscopic LRA and LIE approaches.

ACKNOWLEDGMENTS

We are grateful to Professor Johan Åqvist for providing us a demo that help in checking our version of the GROMOS force field in the Enzymix program.

REFERENCES

1. Warshel A, Sussman F, Hwang J-K. Evaluation of catalytic free energies in genetically modified proteins. J Mol Biol 1988;201:139–159.
2. Kollman P. Free energy calculations: applications to chemical and biochemical phenomena. Chem Rev 1993;93:2395–2417.
3. Merz KM, Jr, Rossi KA, Smith GM, Baldwin JJ. Application of the free energy perturbation method to human carbonic anhydrase II inhibitors. J Med Chem 1995;38:2061–2069.
4. Lee FS, Chu ZT, Bolger MB, Warshel A. Calculations of antibody-antigen interactions: microscopic and semi-microscopic evaluation of the free energies of binding of phosphorylcholine analogs to McPC603. Prot Eng 1992;5:215–228.
5. Åqvist J, Medina C, Samuelsson J-E. A new method for predicting binding affinity in computer-aided drug design. Prot Eng 1994;7:385–391.
6. Åqvist J, Hansson T. Estimation of binding free energy for HIV proteinase inhibitors by molecular dynamics simulations. Prot Eng 1995;8:1137–1144.
7. Bohm HJ, Klebe G. What can we learn from molecular recognition in protein-ligand complexes for the design of new drugs. Ang Chemie 1996;35:2589–2614.
8. Madura JD, Nakajima Y, Hamilton RM, Wierzbicki A, Warshel A. Calculations of the electrostatic free energy contributions to the binding free energy of sulfonamides to carbonic anhydrase. Struc Chem 1996;7:131–137.
9. Muegge I, Tao H, Warshel A, A fast estimate of electrostatic group contributions to the free energy of protein inhibitor binding. Prot Eng 1998;10:1363–1372.
10. Froloff N, Windemuth A, Honig B, On the calculation of binding free energies using continuum methods: application to MHC class-I protein-peptide interactions. Protein Sci 1997;6:1293–1301.
11. Warshel A, Dynamics of reactions in polar solvents: semiclassical trajectory studies of electron-transfer and proton-transfer reactions. J Phys Chem 1982;86:2218–2224.
12. Hwang J-K, Warshel A. Microscopic examination of free energy relationships for electron transfer in polar solvents. J Am Chem Soc 1987;109:715–720.
13. Warshel A, Hwang J-K. Simulation of the dynamics of electron transfer reactions in polar solvents: semiclassical trajectories and dispersed polaron approaches. J Chem Phys 1986;84:4938–4957.
14. King G, Warshel, A. Investigation of the free energy functions for electron transfer reactions. J Chem Phys 1990;93:8682–8692.
15. Kuharski RA, Bader JS, Chandler D, Sprik M, Klein M L, Impey RW. Molecular models for aqueous ferrous-ferric electron transfer. J Chem Phys 1988;89:3248.
16. Levy RM, Belhadj M, Kitchen DB. Gaussian fluctuation formuale for electrostatic free-energy changes in Soluton. J Chem Phys 1991;95:3627–3633.
17. Marcus RA. Chemical and electrochemical electron transfer theory. Annu Rev Phys Chem 1964;15:155.

18. Hermans J, Wang L. Inclusion of loss of translational and rotational freedom in theoretical estimates of free energies of binding: application to a complex of benzene and mutant T4 lysozyme. J Am Chem Soc 1997;119:2707–2714.
19. Warshel A. Computer modeling of chemical reactions in enzymes and solutions. New York: John Wiley & Sons; 1991.
20. Warshel A, Russell ST. calculations of electrostatic interactions in biological systems and in solutions. Q Rev Biophys 1984;17:283–421.
21. Marelius, J, Graffernorgberg, M, Hansson, T, Hallberg, A, and Åqvist J. Computation of affinity and selectivity-binding of 2,4-Diaminopteridine and 2,4-diaminoquinazoline inhibitors to dihydrofolate reductases. J Comp-Aided Mol Des 1998;12:119–131.
22. Tao H, Chu ZT, Warshel A. Quantitative studies of ligand-receptor interactions: a rapid evaluation of binding free energies of endothiapepsin to its inhibitors. In: Pacific symposium on biocomputing. Singapore: World Scientific Press; 1996.
23. Langen R, Brayer GD, Berghuis AM, McLendon G, Sherman F, Warshel A. Effect of the Asn52→Ile mutation on the redox potential of yeast cytochrome *c*. J Mol Biol 1992;224:589–600.
24. Muegge I, Schweins T, Langen R, Warshel A. Electrostatic control of GTP and GDP Binding in the oncoprotein p21 ras. Structure 1996;4:475–489.
25. King G, Lee FS, Warshel A, Microscopic simulations of macroscopic dielectric constants of solvated proteins. J Chem Phys 1991;95:4366–4377.
26. Sham YY Chu, ZT, Warshel A. Consistent calculations of pka's of ionizable residues in proteins: semi-microscopic and macroscopic approaches. J Phys Chem B 1997;101:4458–4472.
27. Lee FS, Chu ZT, Warshel, A. Microscopic and semimicroscopic calculations of electrostatic energies in proteins by the POLARIS and ENZYMIX programs. J Comp Chem 1993;14:161–185.
28. Lam PYS, Jadhav PK, Eyermann C, et al. Rational design of potent, bioavailable, nonpeptide cyclic ureas as HIV protease inhibitors. Science 1994;263:380–384.
29. Wlodawer A, Miller M, Jaskólski M, et al. Conserved folding of in retroviral proteases: crystal structure of synthetic HIV-1 protease. Science 1989;1989:616–621.
30. Lapatto R, Blundell T, Hemmings A, et al. X-ray analysis of HIV-1 proteinase at 2.7 Å resolution confirms structural homology among retroviral enzymes. Nature 1989;342:299–302.
31. Hultén J, Bonham NM, Nillroth U, et al. Cyclic HIV-1 protease inhibitors derived from Mannitol: synthesis, inhibitory potencies and computational predictions of binding affinities. J Med Chem 1997;40:885–897.
32. King G. Warshel A. A surface constrained all-atom solvent model for effective simulations of polar solutions. J Chem Phys 1989;91: 3647–3661.
33. Sham YY, Warshel A. The surface constrained all atom model provides size independent results in calculations of hydration free energies. J Chem Phys 1998;109:7940–7944.
34. Lee FS, Warshel A. A local reaction field method for fast evaluation of long-range electrostatic interactions in molecular simulations. J Chem Phys 1992;97:3100–3107.
35. Muegge I, Qi PX, Wand AJ, Chu ZT, Warshel A. The reorganization energy of cytochrome *c* revisited. J Phys Chem B 1997;101:825–836.
36. Wang J, Dixon R, Kollman PA. Ranking ligand binding affinities with Avidin: a molecular dynamic-based interaction energy study. Proteins: Struct Func Genet 1999;34:69–81.