PROTEINS: Structure, Function, and Genetics 34:395–402 (1999)

What Determines the van der Waals Coefficient β in the LIE (Linear Interaction Energy) Method to Estimate Binding Free Energies Using Molecular Dynamics Simulations?

Wei Wang,[1] Jian Wang,[2] and Peter A. Kollman[2]*
[1]*Graduate Group in Biophysics, University of California, San Francisco, California*
[2]*Department of Pharmaceutical Chemistry, University of California, San Francisco, California*

ABSTRACT **Recently a semiempirical method has been proposed by Åqvist et al.[1,2,3] to calculate absolute and relative binding free energies. In this method, the absolute binding free energy of a ligand is estimated as $\Delta G_{bind} = \alpha\langle V^{el}_{bound} - V^{el}_{free}\rangle + \beta\langle V^{vdw}_{bound} - V^{vdw}_{free}\rangle$, where V^{el}_{bound} and V^{vdw}_{bound} are the electrostatic and van der Waals interaction energies between the ligand and the solvated protein from an molecular dynamics (MD) trajectory with ligand bound to protein and V^{el}_{free} and V^{vdw}_{free} are the electrostatic and van der Waals interaction energies between the ligand and the water from an MD trajectory with the ligand in water. A set of values, $\alpha = 0.5$ and $\beta = 0.16$, was found to give results in good agreement with experimental data. Later, however, different optimal values of β were found in studies of compounds binding to P450cam[4] and avidin.[5] The present work investigates how the optimal value of β depends on the nature of binding sites for different protein-ligand interactions. By examining seven ligands interacting with five proteins, we have discovered a linear correlation between the value of β and the weighted non-polar desolvation ratio (WNDR), with a correlation coefficient of 0.96. We have also examined the ability of this correlation to predict optimal values of β for different ligands binding to a single protein. We studied twelve neutral compounds bound to avidin. In this case, the WNDR approach gave a better estimate of the absolute binding free energies than results obtained using the fixed value of β found for biotin-avidin. In terms of reproducing the relative binding free energy to biotin, the fixed-β value gave better results for compounds similar to biotin, but for compounds less similar to biotin, the WNDR approach led to better relative binding free energies. Proteins 1999;34:395–402.**

Key words: molecular dynamics; ligand-protein interactions; van der Waals energies; electrostatic energies; desolvation

INTRODUCTION

Free energy perturbation (FEP) and thermodynamic integration (TI) methods are rigorous approaches to evaluate binding free energies of ligands to a receptor. However, sampling and convergence problems associated with these approaches prevent them from being widely used in structure-based drug design.[4,5,6] Thus, development of fast and accurate methods to be used for structure-based drug design is still very much needed.[4,5,6] Åqvist et al. have recently proposed a semiempirical molecular dynamics method, termed the linear interaction energy (LIE) approximation, for the estimation of absolute binding free energies.[1,2,3] It is faster than FEP or TI because it does not sample any intermediate state between the initial and final states. The method has been applied to several different systems and the results are in good agreement with experimental data.

The LIE method is based on linear response assumptions.[1,2,3] It divides the interaction between the ligand and its environment into electrostatic and van der Waals parts. The binding free energy is estimated as

$$\Delta G_{bind} = \Delta G^{el}_{bind} + \Delta G^{vdw}_{bind} \approx \alpha\langle V^{el}_{bound} - V^{el}_{free}\rangle + \beta\langle V^{vdw}_{bound} - V^{vdw}_{free}\rangle \quad (1)$$

where V^{el}_{bound} and V^{vdw}_{bound} are the electrostatic and van der Waals interaction energies between the ligand and the solvated protein from an molecular dynamics (MD) trajectory with ligand bound to protein and V^{el}_{free} and V^{vdw}_{free} are the electrostatic and van der Waals interaction energies between the ligand and the water from an MD trajectory with the ligand in water, $< >$ denotes an ensemble average, and α and β are two empirical parameters.

For several protein systems,[1,2,3] Åqvist et al. have found that $\alpha = 0.5$ and $\beta = 0.16$ gave calculated binding free energies in good agreement with experimental data. Paulsen and Ornstein, however, found that $\alpha = 0.5$ and $\beta = 1.043$ resulted in good estimation of the binding free energies of 11 substrates binding to cytochrome P450cam.[7] The difference between the two sets of parameters was rationalized as perhaps due to different force fields, GROMOS and CVFF respectively, used in the two studies.[7] Wang et al. [8] applied this method to calculate binding free

Grant sponsor: National Institutes of Health; Grant numbers: GM-56531 and GM-56609.
*Correspondence to Peter A. Kollman, Department of Pharmaceutical Chemistry, University of California, San Francisco, San Francisco, CA 94143. E-mail: pak@cgl.ucsf.edu
Received 10 August 1998; Accepted 19 October 1998

energies of 14 compounds binding to avidin using the Cornell et al. force field.[9] Their results showed that $\alpha = 0.5$ and $\beta = 1.0$ gave reasonable estimates of the binding free energies with respect to the corresponding experimental results.

These studies raise an interesting question: can one set of α and β be used in different protein-ligand complexes to give reasonable estimates of binding free energies? Although Wang et al. used the Cornell et al. force field,[9] they found similar values of α and β as did Åqvist et al. for the trypsin-benzamidine complex.[8] This suggests that the use of different force fields can not explain the difference in α and β found in different simulations.

What leads to a common value of 0.5 for α, albeit Åqvist has shown α may be reduced from this value for ligands containing OH groups?[25] The value of $\alpha = 0.5$ came from the first order approximation of electrostatic contribution to the binding free energy.[4] This 0.5 also appears in other semiempirical methods, such as Generalized Born model (GB). It has been shown that this first order approximation is reasonable.[19] Thus, the use of $\alpha = 0.5$ has a physical justification. On the other hand, there is no similar argument to obtain β, which has been derived empirically. It is reasonable to think that the value of β is binding-site dependent since it is the scaling factor for van der Waals interactions. Thus, the question is, is there a factor to describe the nature of binding sites and can one relate this factor to the value of β?

In the present work, we have applied the LIE method to a variety of protein-ligand systems and have tried to answer the above question. In our study, α has been kept as 0.5 (as discussed above) and β is adjustable. We defined a ratio factor that we termed as the weighted non-polar desolvation ratio (WNDR) (described below). We studied seven different complexes whose binding affinities were known experimentally and examined the correlation between β and WNDR. β was optimized separately for each complex to reproduce the experimental binding free energy. WNDR of these seven complexes were also calculated. For these seven complexes we have observed a linear correlation between the value of β and WNDR.

We then used this observed linear correlation to calculate the binding free energies for 12 neutral compounds bound to avidin. WNDR was calculated for each compound and used to pick a separate β value for each ligand. In this case, the WNDR approach gave a better estimate of absolute binding free energies than the results obtained using the fixed value of β found for biotin-avidin. In terms of relative free energies of binding to biotin, the fixed β gave better results for compounds similar to biotin than the WNDR approach. On the other hand, for dissimilar compounds, better relative binding free energies were obtained using the WNDR approach.

METHODS

All calculations presented in this work were performed using the AMBER5.0[10] simulation package and the Cornell et al. force field[9] with TIP3P[11] water model. RESP[12] charges were used for all ligand atoms. For each system a pair of simulations was performed, one with the ligand in a 20-Å sphere of waters, the other with the ligand bound to the protein with a cap of waters around the complex. The cap of waters around the complexes were filled up to 20 Å from the center of mass of the ligand. All simulations were carried out at 300 K. The SHAKE[13] procedure was employed to constrain all bonds connecting hydrogen atoms. The time step of the simulations was 1.5 fs with a dual cutoff of 10 Å and 17 Å for the nonbonded interactions. The nonbonded pairs were updated every 30 steps. All atoms within 18 Å of the center of mass of the ligand were allowed to move. Atoms between 18 Å and 20 Å were restrained by a 20- kcal/mol/Å^2 harmonic force. A 100-kcal/mol/Å^2 harmonic position restrain was applied on the center of mass of the ligand in each simulation. Electrostatic and van der Waals interaction energies between the ligand and its environment, i.e. water molecules in the ligand-free state or protein residues and water molecules in the ligand-bound state, within the 20-Å sphere centered at the center of mass of the ligand were calculated using the CARNAL and ANAL modules of AMBER.[10]

Since adding counter ions to the system leads to slow convergence of the simulations,[5] we followed Åqvist et al.'s approach to maintain a neutral protein system in the simulations by changing the charge state of some charged residues.[5] Specifically, the farthest charged residues from the center of mass of the ligand were turned off to keep the 20-Å sphere of the protein neutral. The protonation states of charged residues beyond 20 Å were also adjusted to neutralize the entire system.

Prior to MD simulations, a series of minimizations were carried out using a protocol in which all heavy atoms were restrained with 5,000, 1,000, 100, and 10 kcal/mol/Å^2 harmonic forces respectively. The maximum number of minimization steps was 50,000 steps and the convergence criterion for the energy gradient was 0.5 kcal/mol/Å. Data collection was performed after a 50 ps equilibration. It took 100 to 300 ps to obtain converged average energies. Our criteria for convergence was that the average energies in the two halves of the trajectory differ by less than 5 kcal/mol. The convergence was checked for every simulation (see below). Solvent accessible surfaces (SAS) were calculated using program MSMS[14] after all hydrogen atoms were removed from the PDB files.

When protein and ligand bind to each other, the solvent accessible surface (SAS) of the complex is smaller than the sum of SAS of protein and ligand before binding, because part of the protein and ligand which are exposed to water in the free states are buried upon binding. We termed the loss of the SAS due to binding as the total desolvation SAS (tdSAS). It is calculated as following:

$$\text{total desolvation SAS(tdSAS)} = \text{SAS}_{complex} - \text{SAS}_{protein} - \text{SAS}_{ligand}. \quad (2)$$

Obviously, the total desolvation solvent accessible surface

TABLE I. Surface Tension Parameters Taken From Eisenberg and McLachlan's Work[17]

Atom type	σ (kcal/mol/Å^2)
C	16
S	21
N	−6
N^+	−50
O	−6
O^-	−24

of atom i ($tdSAS^i$) due to binding was calculated as:

$$\text{total desolvation } SAS^i\ (tdSAS^i) = SAS^i_{complex} - SAS^i_{protein} - SAS^i_{ligand} \quad (3)$$

where i = C, N, O, N^+,O^- and S, following the atom classes defined by Eisenberg and McLachlan.[17]

The non-polar desolvation ratio (NDR) is defined as the ratio of total desolvation SAS of all nonpolar groups, carbon and sulfur atoms in this work, to the total desolvation SAS (tdSAS).

$$NDR = (tdSAS^c + tdSAS^S)/tdSAS \quad (4)$$

where $tdSAS^C$ and $tdSAS^S$ represent total desolvation SAS of carbon and sulfur atoms respectively. NDR roughly reflects how hydrophobic the binding site is.

In order to get a more accurate representation of the hydrophobicity of the binding site, different groups on the surface of the binding site should not be treated equally. For example, burial of charged groups is more unfavorable than burial of polar groups. So the desolvation SAS for different groups have been weighted differently.

The total weighted desolvation SAS (twdSAS) due to binding was calculated as:

$$\text{total weighted desolvation SAS(twdSAS)} = \sum_i (\sigma(i) \times dSAS_i) \quad (5)$$

where $\sigma(i)$ is surface tension parameters taken from Eisenberg and McLachlan's work.[17] These values are listed in Table I.

The weighted nonpolar desolvation ratio (WNDR) was defined as the ratio of all nonpolar groups' weighted desolvation SAS to total weighted desolvation SAS:

$$WNDR = ((\sigma(C) \times dSAS^C) + (\sigma(S) \times dSAS^S))/twdSAS \quad (6)$$

It should be pointed out that WNDR can be greater than 1 since the weights for polar groups are negative.

Initially, simulations on seven ligands binding to five proteins were done. All these simulations started from crystal structures taken from Protein Data Bank. Their PDB entries are 3ptb, 1dwc, 1dwd, 1aaq, 5hvp, 1avd, and 2cpp. The value of α was kept as 0.5 and the values of adjustable parameter β were optimized to fit the experimental data.

The simulations on twelve neutral compounds bound to avidin were performed in a similar way. A rationale for the use of a neutral COOH rather than the COO^- present in biotin and its analogs was presented in ref. 8. The structures for the twelve complexes were obtained using a docking algorithm.[21] The computational details were reported elsewhere.[8]

RESULTS AND DISCUSSION

The average ligand-solvent electrostatic and van der Waals interaction energies for ligand bound and free states are shown in Table II together with the calculated binding free energies. The convergence errors estimated by averaging over the first and second half of the simulation trajectories are also listed in the Table II. The small errors indicate that the averaged results are stable.

Since Åqvist et al. and Ornstein et al. obtained different values of β, we calculated the binding free energies on the same systems as they did, namely complexes of trypsin-benzamidine (PDB entry 3ptb)[6] and camphor-P450cam (PDB entry 2cpp).[7] We found two different values of β, 0.14 and 0.81, which gave a good fit to the experimental data respectively. The two values are far from each other and neither can give satisfactory results for both systems. We used the Cornell et al. force field in these two systems while Åqvist et al. used the GROMOS force field for trypsin-benzamidine and Ornstein et al. used the CVFF force field for camphor-P450cam. The β values we found for these two systems are each close to what Åqvist et al. (β = 0.16) and Ornstein et al. (β = 1.043) obtained. This further emphasizes what we noted above - the force field is not the determining factor in why β is so different in trypsin from that in P450cam. Different force fields may give different values of interaction energies in the ligand bound and free states. However, taking the interaction energy differences between the two states may lead to some cancellation of the differences between the different force fields.

We applied this method to other five protein-ligand complex systems and the optimal values of β were also different for each. As discussed above, we keep α = 0.5 in all our simulations (see Introduction). One fixed β could not give the results, which are in good agreement with experimental values in all cases. This is in contrast to previous studies by Åqvist et al. in which one universal value of β was good for different systems.

The dependence of β on ligand-protein system is not unreasonable, given that the nature of binding sites is different for different proteins: some binding sites have many non-polar residues while other binding sites are mostly composed by polar residues. From the previous study of biotin binding to streptavidin by Miyamoto and Kollman,[14,15] it was found that the binding is dominated by the difference between the van der Waals interaction energies in ligand-bound and free state. It appears to be determined by the dispersion attraction in the bound state and the hydrophobic effect (repulsion term) in the ligand-free state. On the other hand in the complex of N-acetyl

TABLE II. Calculated Binding Free Energies, WNDR and β Values of Seven Systems

Protein	PDB entry	VDW[d] (kcal/mol)	EL[e] (kcal/mol)	WNDR[f]	β	ΔG_{cald} (Kcal/mol)	ΔG_{expt} (kcal/mol)
Trypsin	3ptb						
In water[a]		−8.20 ± 0.18	−104.44 ± 0.92	1.361	0.14	−6.55	−6.54
In protein[b]		−20.35 ± 0.69	−114.14 ± 1.92				
Difference[c]		−12.15	−9.70				
HIVP	1aaq						
In water		−39.76 ± 0.5	−111.48 ± 1.50	1.42	0.195	−7.61	−7.60
In protein		−64.31 ± 0.25	−117.12 ± 1.02				
Difference	−24.55	−5.64					
HIVP	5hvp						
In water		−35.44 ± 0.82	−252.18 ± 2.99	1.48	0.11	−10.30	−10.30
In protein		−65.15 ± 0.63	−266.24 ± 0.80				
Difference	−29.71	−14.06					
Thrombin	1dwc						
In water		−23.82 ± 0.08	−229.80 ± 1.03	1.24	0.61	−10.76	−10.67
In protein		−45.21 ± 0.31	−225.24 ± 1.04				
Difference		−21.39	+4.56				
Thrombin	1dwd						
In water		−34.76 ± 0.28	−209.07 ± 1.98	1.20	0.61	−11.59	−11.63
In protein		−60.99 ± 0.18	−200.25 ± 0.70				
Difference	−26.228	+8.82					
Avidin	1avd						
In water		−19.53 ± 0.06	−52.68 ± 0.16	1.10	0.87	−20.37	−20.40
In protein		−37.57 ± 0.30	−62.40 ± 0.16				
Difference		−18.04	−9.36				
P450-cam	2cpp						
In water		−14.97 ± 0.22	−12.87 ± 0.24	1.05	0.81	−7.85	−7.90
In protein		−26.69 ± 0.14	−9.58 ± 0.06				
Difference		−11.72	+3.29				

[a]Ligand in the free state, i.e. in water.
[b]Ligand in the bound state, i.e. ligand bound to solvated protein.
[c]The difference of van der Waals and electrostatic interaction energies between the ligand free and bound states.
[d,e]van der Waals and electrostatic interaction energies between the ligand and its environments, which are water in the free state and solvated protein in the bound state respectively.
[f]Weighted non-polar desolvation ratio (WNDR), see text for definition.

tryptophan with α-chymotrypsin, the binding free energy has a larger contribution from the electrostatic term. This suggested that there might be some correlation between the value of β and the hydrophobicity of the binding site.

Using Eq. (1) to calculate binding free energies, the contributions of binding free energy due to electrostatic and van der Waals interactions between the ligand and the protein are considered. The contributions of binding free energies due to desolvation, entropy loss due to reduced conformation freedom, etc, are implicitly included by empirically optimizing the value of β.

In order to study how the value of β depends on the nature of the binding site, as the first step, we examined the non-polar desolvation ratio (NDR), i.e. the ratio of desolvation SAS of non-polar groups to the total desolvation SAS (tdSAS). It is a rough representation of how hydrophobic the binding site is. We plotted the NDR versus the values of β in Figure 1. The linear correlation coefficient is 0.89.

Many previous studies have shown that there is a correlation between the change of solvent accessible surfaces and solvation and binding free energies. Empirical atomic solvation parameters have been obtained for different atom types[17,18] and these parameters have been success-

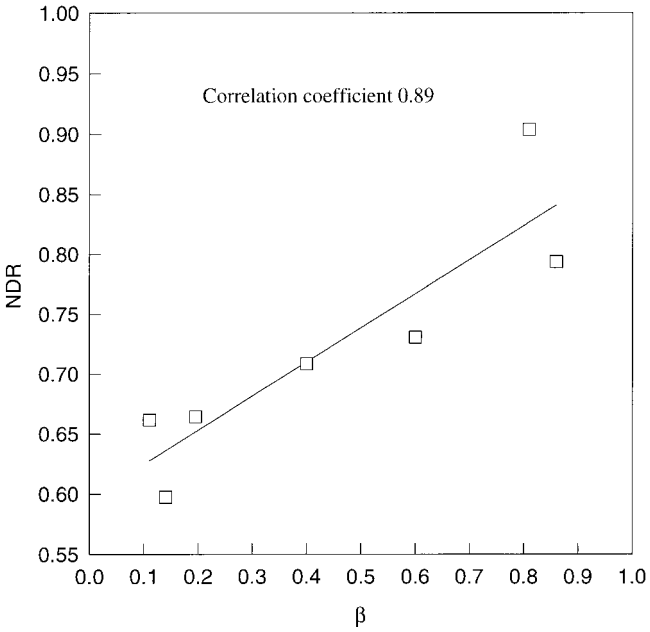


Fig. 1. β value versus non-polar desolvation ratio (NDR) for the seven calibration systems.

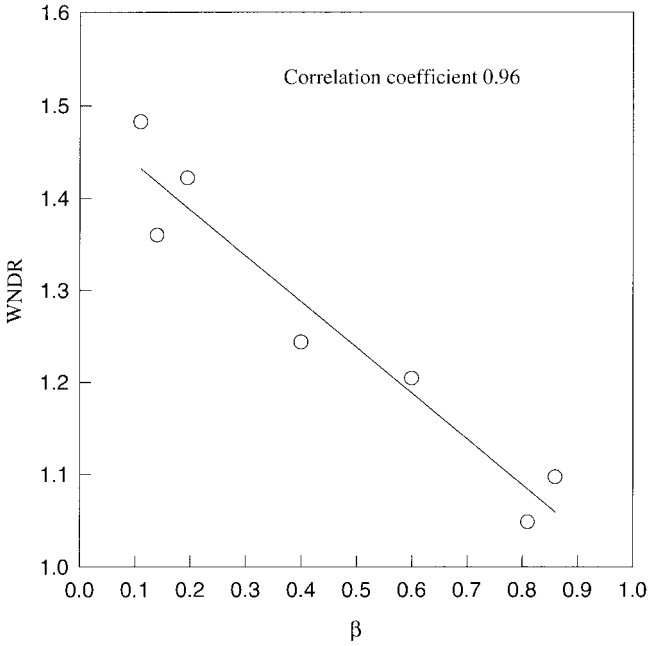


Fig. 2. β value versus weighted non-polar desolvation ratio (WNDR) for the seven calibration systems.

fully applied to studying protein stability and protein ligand binding.[19,20] Encouraged by the good correlation between the NDR and β, we used a weighted non-polar desolvation ratio (WNDR see definition in Methods) to discriminate contributions to binding from different groups. It is natural to use the corresponding solvation parameter of each group as its weight. Thus, non-polar groups have positive weights and polar and charged groups have negative weights (see Table I). In the present work, we used the solvation parameters published by Eisenberg and McLachlan.[17] We calculated the WNDR and the optimal β for the seven protein-ligand complexes. The correlation between WNDR and values of β is shown in Figure 2 and the correlation coefficient is 0.96.

It is clear that the Åqvist Eq. (1) is a simplification of a more general attempt to represent ΔG_{bind} in terms of components. For example, Eq. (7) can be considered as such a general approach.

$$\Delta G_{bind} = \Delta G_{el} + \Delta G_{vdw} + \Delta G_{desolv} + \Delta G_{t/r} + \Delta G_{intra} \quad (7)$$

where ΔG_{el} and ΔG_{vdw} are contributions to the binding free energy from electrostatic and van der Waals interactions between the ligand and the protein respectively, ΔG_{desolv} is the free energy change due to the desolvation effect, i.e. the loss of solvent accessible surface due to complexation, $\Delta G_{t/r}$ is the translational/rotational free energy change upon complexation, and ΔG_{intra} includes free energy contributions from the conformational entropy loss and the changes of intramolecular energies of the ligand and the protein upon binding.

The Åqvist Eq. (1) attempts to "fold in" the last three components of Eq. (7) into Eq. (1) by using parameters of α and β. Of those three terms that are in Eq. (7) but not in Eq. (1), $\Delta G_{t/r}$ stands out as not being very molecule dependent and should be roughly constant for typical ligand-protein complexes. Thus, we examined the ability of the following equation to represent the binding data in the seven ligand-protein complexes studied above.

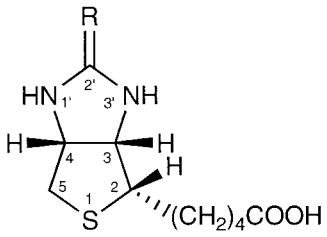
$$\Delta G_{bind} = \alpha\langle V^{el}_{bound} - V^{el}_{free}\rangle + \beta\langle V^{vdw}_{bound} - V^{vdw}_{free}\rangle + \Delta G_{t/r}. \quad (8)$$

We tested values of $\Delta G_{t/r}$ in the range of 7–11 kcal/mol [26,27] and derived β value for each system . Fitting of these constant values of $\Delta G_{t/r}$ still led to a good fit between the WNDR and β, with correlation coefficient r varying between 0.98 ($\Delta G_{t/r}$ = 7 kcal/mol) to 0.95 ($\Delta G_{t/r}$ = 11 kcal/mol), compared to the 0.96 we found before. Thus, the use of Eq. (8) rather than Eq. (1) appear to be equally efficient for molecular dynamics modeling of the free energies of ligand-protein complexes.

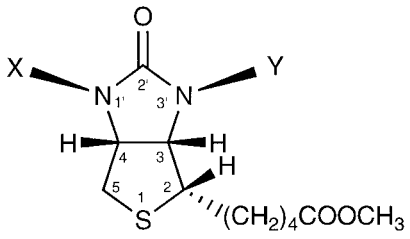
All of the above simulations started from crystal structures. In real drug design, people are interested in docking novel ligands into binding sites. In another word, this method would be more useful if it could be used in estimating binding free energies for structures obtained from docking. In addition, the linear correlation shown in Figure 2 was obtained from examining different protein-ligand systems. It is of interest to examine this relationship for different ligands binding to the same protein. We thus have applied this correlation to predict optimal values of β for twelve neutral compounds bound to avidin (compounds 2–13 in Figure 3).[8] No crystal structure of any of these complexes was available. The compounds (see Fig. 3) were docked into the binding site of avidin using a docking program developed by Wang et al.[21] Different values of β were assigned based on WDNR of each complex structure after the MD simulation. As a comparison, a fixed-β value, β = 0.87, was applied to calculate the binding free energies too (see Table III). The reason to use β = 0.87 is that it gave best estimate of binding free energy of biotin (compound 1) binding to avidin.

From Figure 4, we can see, compared with a fixed-β value, nine points out of twelve are in equal or better agreement with experiment using the WDNR to determine a different β for each complex. For a fixed-β value, the binding free energies were almost always underestimated. However, using the WDNR values, this systematic error was reduced in most cases. In three complexes, the binding free energies were overestimated. This implies that there were random errors involved instead of systematic ones. The error may come from the correlation itself since we only had seven points in our fit set. Another possible source may be errors of the structures that were obtained from docking ligands into the binding site. The worst case is ligand four. The reason for this is discussed elsewhere.[8]

We also examined the relative binding free energies between biotin (compound 1) and other compounds binding to the avidin (Table IV and Figure 5). From Figure 3, we can see that compounds 2, 3, 6, 7 are more similar to biotin than compounds 5, 8–13. The desolvation effect of compounds 2, 3, 6, 7 should be also similar to biotin. In

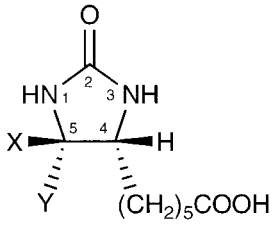


1 (R=O) **2** (R=S) **3** (R=NH)



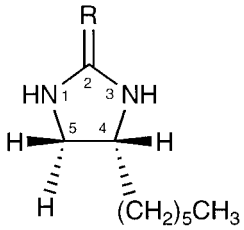
4 (X=CH_3O and Y=H)

5 (X=H and Y=CH_3O)

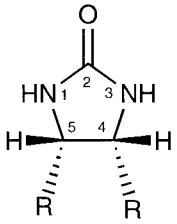


6 (X=H and Y=CH_3)

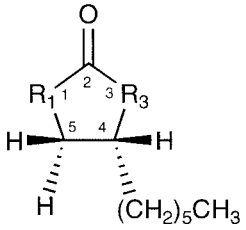
7 (X=CH_3 and Y=H)



8 (R=O) **9** (R=NH)



10 (R=H)
11 (R=CH_3)



12 (R_1=O and R_3=NH)
13 (R_1=NH and R_3=O)

Fig. 3. Biotin analogues used for comparing the WNDR approach and the fixed-β approach.

Table IV, a fixed β gave better results for compounds 2, 3, 6, 7 than WNDR. The reason is that errors introduced by second term in Eq. (1) using a fixed β cancelled each other when we calculate relative binding free energies between similar compounds. However, for compounds 5, 8–13, WNDR gave better results than fixed β since the desolvation effect of compounds 5, 8–13 are different from biotin.

After observing the correlation between the WNDR and the value of β, we want to understand the physical meaning of this correlation. As we noted above in Eq. (7), the value of β has in it implicit contributions from a number of terms.

Jorgensen and coworkers have included a specific $\Delta G_{desolve}$ term, a solvent accessible surface term, to Eq. (1).[23,24] They calculated the change of solvent accessible surface due to complexation and introduced another parameter γ in addition to α and β. By optimizing the three parameters, they obtained some improvements compared with results obtained using Eq. (1). However they did not discriminate the different contributions of polar and non-polar groups to the binding free energies.[23,24] It is known that burial of non-polar groups is favorable for binding, while burial of polar groups contributes unfavorably to binding free energies.

From our own experience and previous studies,[1–3,8,23,24] van der Waals interactions are always favorable for binding. The contribution to binding free energy of desolvation effect depends on the components of the binding site. The higher percentage of non-polar groups in the binding site, the higher percentage non-polar groups are desolvated due to binding, apparently, the more favorable contribution to the binding free energy due to the desolvation effect. Since

TABLE III. Calculated Binding Free Energies for 12 Compounds Binding to Avidin With WNDR and β = 0.87 Approaches

Number	ΔG (kcal/mol) (expt)	WNDR	β	ΔG (kcal/mol) (cald)[a]	ΔG (kcal/mol) (cald)[b]
2	−16.9	1.070	0.88	−17.09	−16.93
3	−14.3	1.130	0.76	−11.58	−13.65
4	−8.8	1.070	0.88	−18.03	−17.81
5	−12.2	1.100	0.82	−16.47	−17.29
6	−14.0	1.150	0.72	−11.98	−13.94
7	−16.5	1.145	0.73	−14.53	−16.82
8	−11.1	1.117	0.78	−12.64	−13.66
9	−7.4	1.150	0.72	−9.97	−11.91
10	−4.5	1.240	0.53	−6.93	−8.78
11	−6.4	1.170	0.67	−8.00	−9.38
12	−5.0	1.105	0.81	−10.24	−11.15
13	−7.4	1.096	0.83	−12.60	−13.19

[a]Results obtained by using β which was estimated from the correlation between WNDR and β.
[b]Results calculated using β = 0.87.

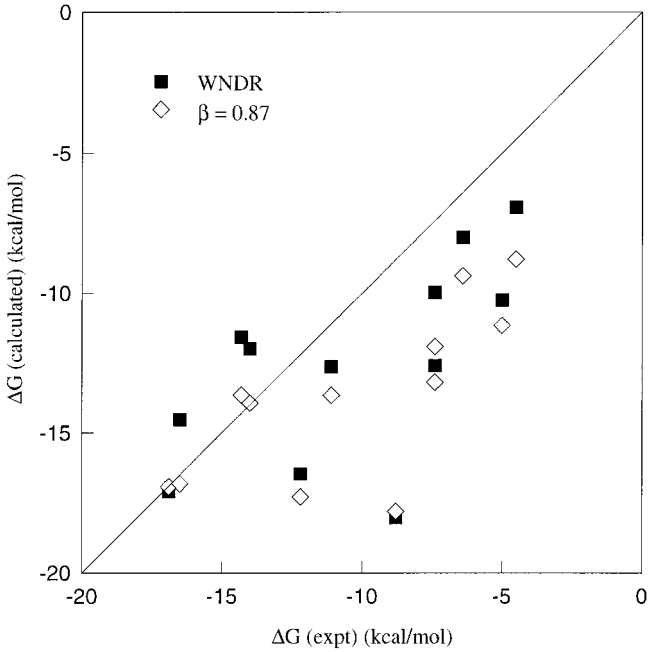


Fig. 4. Observed versus calculated binding free energies for the 12 compounds binding to avidin using β predicted from the correlation obtained from Figure 2 and β = 0.87, respectively.

TABLE IV. Relative Binding Free Energies Between Biotin (Compound 1) and Other Compounds Binding to Avidin

Relative	ΔΔG (kcal/mol) (expt)	ΔΔG (kcal/mol)[a]	ΔΔG (kcal/mol)[b]
2-1	3.5	3.3	3.5
3-1	6.1	8.8	6.8
4-1	11.6	2.4	2.6
5-1	8.2	3.9	3.1
6-1	6.4	8.4	6.5
7-1	3.9	5.9	3.6
8-1	9.3	7.8	6.7
9-1	13.0	10.4	8.5
10-1	15.9	13.5	11.6
11-1	14.0	12.4	11.0
12-1	15.4	10.2	9.3
13-1	13.0	7.8	7.2

[a]Results obtained by using β which was estimated from the correlation between WNDR and β.
[b]Results calculated using β = 0.87.

the weights for polar groups are negative, the more polar groups buried, the smaller the denominator of Eq. (6), thus, the larger the WNDR. Conversely, the smaller the WNDR, the more hydrophobic a binding site, the larger the value of β. Therefore, the correlation between WNDR and β makes sense.

CONCLUSION

In this work, we present a correlation between the weighted desolvation non-polar ratio (WNDR) and values of β in the linear interaction energy method. β in the LIE method is a factor to describe how much van der Waals interactions and desolvation effect contribute to binding free energies. The WNDR represents the hydrophobicity of the binding sites. This correlation was found by studying different systems. It suggests that the value of β is predictable by calculating the WNDR of the system, especially for systems in which very different ligands bind to the same protein or ligands bind to different binding sites of the same protein. We applied this correlation to 12 complexes of avidin whose complex structures were obtained using a docking algorithm. The results are promising, but the further investigation is needed to examine if the correlation is found in other systems.

The correlation between the WNDR and the value of β should be useful in drug design. It is easy to calculate the WNDR for any ligand-protein complex. The value of β can be obtained from the correlation presented here and this should allow a more accurate estimate of binding free energy to be obtained from the linear interaction energy (LIE) method. The further role of LIE method in drug design is to refine the leads found by docking algorithms using database screening or de novo design. With the development of increased computer power, molecular dy-

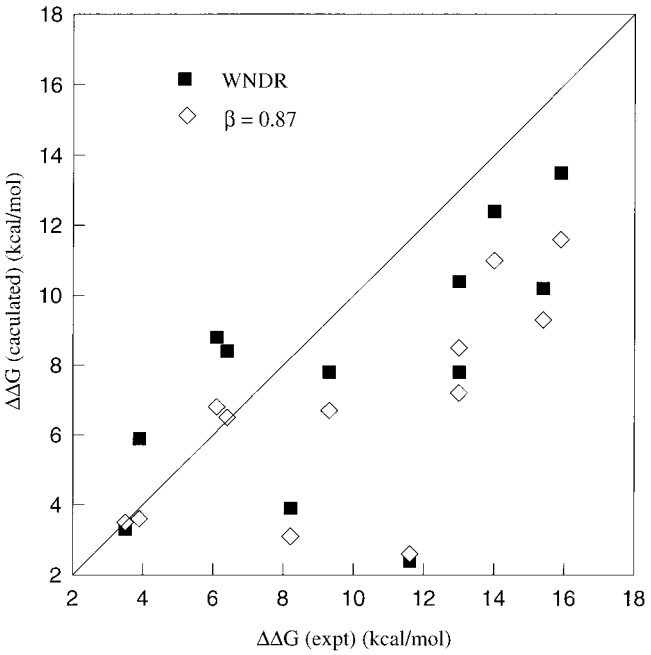


Fig. 5. Observed versus calculated binding free energies between biotin (compound 1) and other compounds binding to avidin using β predicted from the correlation obtained from Figure 2 and β = 0.87, respectively.

namics can become more useful. Thus the LIE method can be expected to be more widely used in finding novel leads for ligands that bind tightly to macromolecules.

ACKNOWLEDGEMENT

This research was supported in part by NIH grants GM-56531 to P. Ortizde Montellano, P.I. and GM-56609 (E. Arnold, P.I.). Part of this investigation was completed using the facilities in the UCSF Computer Graphics Lab, T. Ferrin, director, supported by RR-1081 from the NIH. W.W is grateful to Drs. Lu Wang and Yong Duan and Jed Pitera for their stimulating discussions, and a Regents fellowship of UCSF for support. J.W. thanks the Natural Science and Engineering council of Canada for a postdoctoral fellowship to carry out part of this research.

REFERENCES

1. Åqvist J, Medina C, Samuelsson J. New method for predicting binding affinity in computer-aided drug design. Protein Eng 1994;7:385–391.
2. Hansson T, Åqvist J. Estimation of binding free energies for HIV proteinase inhibitors by molecular dynamics simulations. Protein Eng 1995;8:1137–1144.
3. Åqvist J. Calculation of absolute binding free energies for charged ligands and effects of long-range electrostatic interactions. J Comput Chem 1996;17:1587–1597.
4. Kollman PA. Free energy calculations: Applications to chemical and biochemical phenomena. Chem Rev 1993;93:2395–2417.
5. Beveridge DL, Dicapua FM. Free energy via molecular simulations: Application to chemical and biochemical system. Annu Rev Biophys Biophys Chem 1989;18: 431–492.
6. van Gunsteren WF. Methods for calculation of free energies and binding constants: Successes and problems. In: van Gunsteren WF, Weiner PK, editors. Computer simulation of biomolecular systems. Leiden: ESCOM; 1989. p 27–59.
7. Paulsen MK, Ornstein R. Binding free energy calculations for P450cam-substrate complexes. Protein Eng 1996;9:567–571.
8. Wang J, Dixon R, Kollman PA. Ranking ligand binding affinities with avidin: A molecular dynamics based interaction energy study. Proteins 1999;34:in press.
9. Cornell WD, Cieplak P, Bayly CI, et al. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. J Am Chem Soc 1995;117:5179–5197.
10. Pearlman DA, Case DA, Caldwell JW, et al. AMBER, a package of computer programs for applying molecular mechanics, normal mode analysis, molecular dynamics and free energy calculations to simulate the structural and energetic properties of molecules. Comp Phys Comm 1995;91:1–41.
11. Jorgensen WL, Chandrasekhar J, Madura J, Impey RW, Klein ML. Comparison of simple potential functions for simulating liquid water. J Chem Phys 1983;79:926–935.
12. Bayly CI, Cieplak P, Cornell WD, Kollman PA. A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges - the RESP model. J Phys Chem 1993;97: 10269–10280.
13. Ryckaert JP, Ciccotti G, Berendsen HJC. Numerical integration of the cartesian equations of motion of a system with constraints: Molecular dynamics of n-alkanes. J Comput Phys 1977;23: 327–341.
14. Sanner MF, Olson AJ, Spehner J. Reduced surface - an efficient way to compute molecular surfaces. Biopolymers 1996;38: 305–320.
15. Miyamoto S, Kollman PA. Absolute and relative binding free energy calculations of the interaction of biotin and its analogs with strepavidin using molecular dynamics free energy perturbation approaches. Proteins 1993;16:226–245.
16. Miyamoto S, Kollman PA. What determines the strength of noncovalent association of ligands to proteins in aqueous solution. Proc Natl Acad Sci USA 1993;90:8402–8406.
17. Eisenberg D, McLachlan AD. Solvation energy in protein folding and binding. Nature 1986;319:199–203.
18. Williams RL, Vila J, Perrot G, Scheraga HA. Empirical solvation models in the context of conformational energy searches—application to bovine pancreatic trypsin inhibitor. Proteins 1992; 14:110–119.
19. Still WC, Tempcyk A, Hawley RC, Hendrickson T. Semianalytical treatment of solvation for molecular mechanics and dynamics. J Am Chem Soc 1990;112:6127–6129.
20. Cramer CJ, Truhlar DG. General parameterized SCF model for free energies of solvation in aqueous solution. J Am Chem Soc 1991;113:8305–8311.
21. Wang J, Kollman PA, Kuntz ID. Flexible ligand docking: A systematic approach. Proteins 1999; in press.
22. Ajay, Murcko MA. Computational methods to predict binding free energy in ligand-receptor complexes. J Med Chem 1995;38:4953–4967.
23. Jones-Hertzog DK, Jorgensen WL. Binding affinities for sulfonamide inhibitors with human thrombin using Monte Carlo simulations with a linear response method. J Med Chem 1997;40:1539–1549.
24. Carlson HA, Jorgenson WL. An extended linear response method for determining free energies of hydration. J Phys Chem 1995;99: 10667–10673.
25. Marelius J, Hansson T, Åqvist J. Calculation of ligand binding free energies from molecular dynamics simulations. Int J Quant Chem 1998;69:77–88.
26. Williams DH, Cox JPL, Doig AJ, et al. Toward the semiquantitative estimation of binding constants—guides for peptide peptide binding in aqueous solution. J Am Chem Soc 1991;113:7020–7030.
27. Searle MS, Williams DH, Gerhard U. Partitioning of free energy contributions in the estimation of binding constants—residual motions and consequences for amide-amide hydrogen bond strengths. J Am Chem Soc 1992;114:10697–10704.