PROTEINS: Structure, Function, and Genetics 30:407–423 (1998)

Electrostatic Contributions to Protein–Protein Binding Affinities: Application to Rap/Raf Interaction

Ingo Muegge, Thomas Schweins, and Arieh Warshel*
Department of Chemistry, University of Southern California, Los Angeles, California

ABSTRACT **The challenge of evaluating absolute binding free energies of protein–protein complexes is addressed using the scaled Protein Dipoles Langevin Dipoles (PDLD/S) model in combination with the Linear Response Approximation (LRA). This is done by taking the complex between Rap1A (Rap) and the p21^{ras} binding domain of c-Raf (Raf-RBD) (Nassar et al., Nature 375:554–560, 1995) as a model system. Several formulations and different thermodynamic cycles are explored taking advantage of the LRA method and considering the protein reorganization during complex formation. The performance of different approximations is examined by comparing the calculated and observed absolute binding energies for the native complex and some of its mutants. The evaluation of the contributions of individual residues to the binding free energy, which is referred to here as *group contributions* is also examined. Special attention is paid to the role of the "*dielectric constant*," ϵ_{in} which is in fact a scaling factor that represents the contributions that are treated implicitly. It is found that explicit consideration of protein relaxation is crucial for obtaining reasonable results with small values of ϵ_{in}, but it is also found that such a treatment of protein–protein interactions is very challenging and does not always give stable results. This indicates that more advanced explicit calculations should be based on experimentally determined structures of both the complex and the isolated proteins. Nevertheless, it is demonstrated that the qualitative trend of the effect of mutations can be reproduced by considering the effect of protein reorganization implicitly, using $\epsilon_{in} \sim 25$ for ionized residues and $\epsilon_{in} \sim 4$ for polar residues. Thus, it is concluded that an explicit treatment of solvent relaxation (which is common to current continuum models) does not provide sufficient compensation for turning off the charges of ionized residues on the interaction surface of the Raf-RBD/Rap complex. Representing the missing contribution by large ϵ_{in} can, of course, reproduce the observed effect of ionized residues, but now the contribution of uncharged residues will be largely underestimated. Regardless of these conceptual problems, it is established that a very simple nonrelaxed approach, where the relaxation of both the protein and the solvent are considered implicitly, can provide an effective qualitative way for evaluating group contributions, using large and small values for ϵ_{in} of ionized and neutral residues, respectively. As much as the actual system studied is concerned we find that more residues than generally assumed play a role in Raf-RBD/Rap interaction. This includes residues that are not located at the protein–protein interaction surface. These residues contribute to the binding energy through direct charge–charge interaction without leading to drastic structural changes. The overall contribution of the surface residues is quite significant since Raf and Rap are positively and negatively charged, respectively, and their charges are distributed along the interaction site between the two proteins. Proteins 30:407–423, 1998.** © 1998 Wiley-Liss, Inc.

Key words: binding free energy; electrostatics; group contributions; thermodynamic cycle; solvation

INTRODUCTION

Specific recognition of biological ligands by proteins and nucleic acids and the formation of stable complexes is crucial for many biologically functioning systems.[1] The growing number of available crystal or solution structures of protein-inhibitor complexes, as well as antibody-antigen or enzyme-ligand complexes, challenges one to study the energetics of biological recognition in a quantitative way. While significant progress has been made in studies of protein–ligand interactions,[2,3,4,5,6] much less progress has been made in studies of protein–protein interactions. The number of residues involved in the protein–protein recognition is usually

Contract grant sponsor: DOE; Contract grant number: DE-FG03-94ER61945; Contract grant sponsor: Tobacco research grant; Contract grant number: 4RT-0002.

Dr. Muegge's present address is Pharmaceutical Product Division, Abbott Laboratories, Abbott Park, IL 60064-3500.
Dr. Schweins' present address is Hoechst Research & Technology, Biotechnology, D-65926 Frankfurt, Germany.
*Correspondence to: Arieh Warshel, Department of Chemistry, University of Southern California, Los Angeles, CA 90089-1062.

Received 7 April 1997; Accepted 18 September 1997

relatively small[7] and the binding free energy lies between 9–18 kcal/mol.[8] In principle, the formation of protein–protein complexes obeys the same physical rules as the protein–ligand interactions. However, protein–protein interactions are expected to involve more delicate balance between entropic and enthalpic contributions and this balance might not be easily reproduced by computer simulation approaches. Furthermore, the reorganization of the proteins upon complex formation (especially of the residues at the interaction surfaces) should be quite significant and has to be addressed adequately. Despite these difficulties, it is clearly important to advance our understanding of the nature of protein–protein interactions.

Calculations of free energies play a major role in correlating the structure and function of proteins.[2,5,9,10,11,12] Methods such as Molecular Dynamics/Free Energy Perturbation (MD/FEP),[2,11] Linear Response Approximation (LRA),[4,6,13] Protein Dipole Langevin Dipole (PDLD) approaches,[3,6] or Discretized Continuum approaches (DC)[14,15,16] and related methods, have become useful tools in calculations of relative and absolute protein–ligand binding affinities. However, the related issue of protein–protein interaction has not been studied in a systematic way using these methods. A possible reason might be the need to deal with a relatively large solute (namely, one of the proteins). Another reason is associated with the fact that the reorganization of the protein structure and the surrounding solvent upon complex formation cannot be neglected. The treatment of the protein reorganization effects presents a major challenge that has not been addressed in most previous studies.

The present work addresses the challenge of estimating protein–protein binding free energies by computer modeling approaches: by considering a thermodynamic cycle that reflects conformational changes upon docking. This is done by the implementation of a PDLD/S-LRA method that provides a reasonable estimate of protein–protein interactions. Our method is examined by calculating the binding affinity between the Ras-related Rap1A and the Ras binding domain (RBD) of the Ras effector molecule c-Raf1, a serine/threonine specific kinase, whose crystal structure was solved recently by Nassar et al.[17] This system is not only interesting as a model system for protein–protein interaction, but is also of great biological importance: that is, Rap1A is a GTP-binding protein which has the same effector region as the protein Ras that plays a major role in signal transduction and is involved in human cancer.[18] The elucidation of the process of complex formation between Ras, and effectors like the GTPase activating proteins (GAPs), Raf, or other kinases that bind at the same region, is crucial for learning more about the mechanism of signal transduction in the cell.

The Simulation Methods section considers several options for evaluating protein–protein binding energies, and develops the corresponding PDLD/S-LRA expression. The section starts by introducing the relevant thermodynamic cycles and the LRA treatment for the somewhat less challenging case of ligand binding to proteins (which was already considered in our previous works). Several related strategies are then developed for studies of protein–protein interactions. This section also derives the relevant expressions for the evaluation of group contributions to binding free energy. The Results and Discussion section examines the approaches developed in The Simulation Methods section by considering their performance in evaluating Rap/Raf affinity and the effect of mutations on this affinity. Special emphasis is placed on the relationship between the treatment of the protein reorganization and the optimal dielectric constant. This section also examines the dependence of the results on the model used. Finally, the results are discussed in the Concluding Remarks section, emphasizing the role of the protein reorganization and the need to consider this crucial effect either implicitly or explicitly.

SIMULATION METHODS

Calculation of Protein–Protein Binding Affinities

As a background for our treatment of protein–protein interactions, we start by considering our treatment of protein–ligand interactions[6] and some of its recent modifications. Here, as in studies of free energies of most biological processes[19,20] it is important to define the relevant thermodynamic cycle. For example, the cycle considered by Lee et al.[6] was used successfully in the past to calculate absolute antibody–antigen binding energies.[6] A closely related approach is used in constructing the inner cycle of Figure 1 (a, d, e, h) to describe the electrostatic contribution ($\Delta G^{elec}_{bind,l}$) to the absolute binding free energy of a ligand (l) to a protein (p), where the configurations of a protein/ligand complex are kept at a single configuration(s) in the bound state and single configuration (s′) at the dissociated state, but the solvent is allowed to relax at each step of the cycle. In this case we can express the total binding free energy as

$$\begin{aligned}(\Delta G_{bind})_{s'\to s} &= (\Delta G^{elec}_{bind})_{s'\to s} \\ &\quad + (\Delta G_{hyd} + \Delta G_{vdw} - T\Delta S')_{s'\to s} \\ &= (\Delta G^{p}_{elec,l})_s - (\Delta G^{w}_{elec,l})_{s'} + (\Delta G_{elec,l'})_{s'\to s} \\ &\quad + (\Delta G_{hyd} + \Delta G_{vdw} - T\Delta S')_{s'\to s} \end{aligned} \quad (1)$$

where $(\ )_s$ indicates that the corresponding form is evaluated at a single configuration of the protein ligand complex. The term $\Delta G^{w}_{elec,l}$ is the solvation energy of the ligand in water (without the van der Waals and hydrophobic contributions), $\Delta G^{p}_{elec,l}$ is the change of the electrostatic contribution to the solvation energy of the protein–ligand complex upon charging the ligand (from an artificial uncharged state where all atomic partial charges are set to zero, to the state where all charges including partial

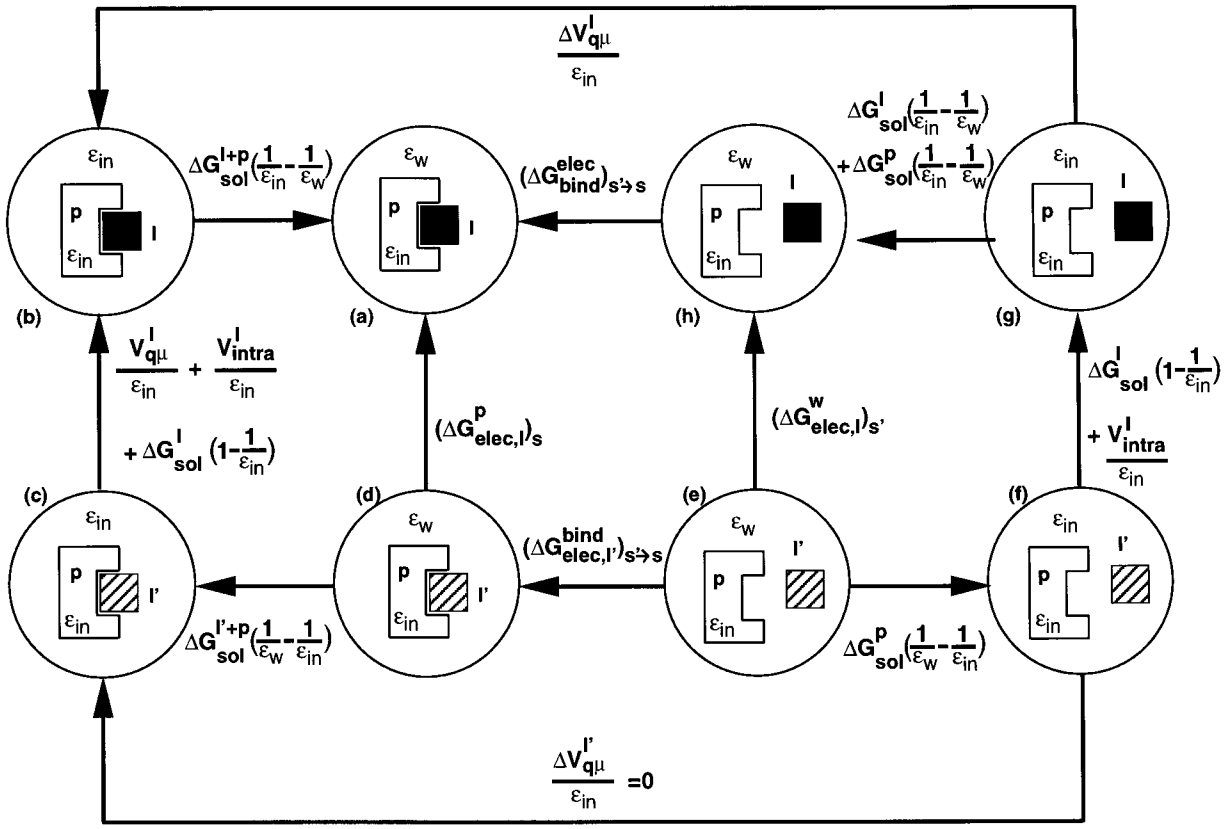

Fig. 1. Thermodynamic cycles for studies of the binding of a ligand (l) to a protein (p). The cycles consider only the electrostatic contributions to binding free energies. This is done by changing the charges of the ligand from their actual value (black diagram) to zero (cross hatched diagram). The figure only represents a single configuration (s) for the complex state and a single configuration (s′) for the unbound state. The proper average over the configuration of the system, on the charged and uncharged state of the ligand, is performed by the LRA approach as described in the text. The inner cycle (a, b, d, h) describes the binding process on a macroscopic level. The outer cycles (a, d, c, b) and (h, e, f, g) are used to provide the semimacroscopic estimates of the relevant electrostatic energies. This is done by using the PDLD/S three steps cycle where the solvent "dielectric constant" is changed from ϵ_w to ϵ_{in}, the ligand charges are changed, and finally the dielectric constant is changed back from ϵ_{in} to ϵ_w. The relevant PDLD/S energy contributions are given in the figure and more details about these evaluations are given elsewhere.[6]

charges are switched on) and $\Delta G^{bind}_{elec,l'}$ is the electrostatic contribution to the binding free energy of the uncharged ligand, l'. ΔG_{hyd} and ΔG_{vdw} are the hydrophobic, and van der Waals contributions to binding and $-T\Delta S'$ represents the nonelectrostatic entropic contribution which is associated with the binding of the uncharged ligand.

The free energy of Eq. (1) can be evaluated conveniently by the semimacroscopic PDLD/S model,[6] replacing the microscopic ΔG_{elec} terms by their semimacroscopic counterparts. This is achieved by considering the extra cycles (h, e, f, g) and (a, b, c, d) where the dielectric constant of the solvent around the protein is changed from that of water, ϵ_w, to a value that corresponds to the assumed protein "dielectric constant," ϵ_{in} (this parameter represents the implicit contribution of the protein model, and its true nature is explained in detail elsewhere.[11,21] The PDLD/S contributions of the different steps in the cycles are considered in detail elsewhere,[3,6] and are also given in Figure 1. Since we deal with a single protein/ligand configuration, we consider the sum of the PDLD/S terms as an effective potential (a potential that should be averaged to obtain the proper free energy) and write:

$$
\begin{aligned}
(\Delta G^{p}_{elec,l})_s &= U^{p}_{elec,l} \\
&= \left((\Delta G^{l+p}_{sol} - \Delta G^{l'+p}_{sol})\left(\frac{1}{\epsilon_{in}} - \frac{1}{\epsilon_w}\right)\right. \\
&\quad \left. + \Delta G^{l}_{sol}\left(1 - \frac{1}{\epsilon_{in}}\right) + \frac{V^{l}_{q\mu}}{\epsilon_{in}} + \frac{V^{l}_{intra}}{\epsilon_{\mathrm{in}}}\right)_s \\
(\Delta G^{w}_{elec,l})_{s'} &= U^{w}_{elec,l} = \left(\Delta G^{l}_{sol}\left(\frac{1}{\epsilon_{in}} - \frac{1}{\epsilon_w}\right)\right. \\
&\quad \left. + \Delta G^{l}_{sol}\left(1 - \frac{1}{\epsilon_{in}}\right) + \frac{V^{l}_{intra}}{\epsilon_{\mathrm{in}}}\right)_{s'} \qquad (2)
\end{aligned}
$$

where ΔG_{sol} denotes the electrostatic contribution to the solvation free energy of the indicated group in water. To be more precise, ΔG_{sol} should be scaled by $1/(1 - 1/\epsilon_w)$, but this small correction is neglected here. $V_{q\mu}$ is the electrostatic interaction between the indicated groups in vacuum (this is a standard PDLD notation), and in the present case, $V^{l'}_{q\mu} = 0$. V^{l}_{intra} is the intramolecular electrostatic interaction between the atoms of the ligand. Since 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.,[21] and a more consistent treatment should involve the use of the LRA or related approaches[6,21]). This approach provides a reasonable approximation for the actual free energy by using[6]:

$$\Delta G^{p}_{elec,l} = 1/2[\langle U^{p}_{elec,l}\rangle_{l'+p} + \langle U^{p}_{elec,l}\rangle_{l+p}]$$
$$\Delta G^{w}_{elec,l} = 1/2[\langle U^{w}_{elec,l}\rangle_{l'} + \langle U^{w}_{elec,l}\rangle_{l}] \quad (3)$$

where $\langle\ \rangle_l$ and $\langle\ \rangle_{l'}$ designate an MD average over a force field that correspond to the ligand in its charged and uncharged form, while $\langle\ \rangle_{l+p}$ and $\langle\ \rangle_{l'+p}$ designate an MD average over a force field of the protein-ligand system, where the bound ligand is in its charged and uncharged forms, respectively. It is important to realize that the average of Eq. (3) is always done where both contributions to the relevant U_{elec} are evaluated at the same configurations. That is, the PDLD/S energies of the charged and uncharged states are always evaluated at each average step at the same structure, but these structures are generated by MD simulations with the potential surface of the charged and uncharged states. Eqs. (2) and (3) give now:

$$\begin{aligned}\Delta G^{p}_{elec,l} - \Delta G^{w}_{elec,l} &\approx \frac{1}{2}\left[\left\langle(\Delta G^{l'+p\rightarrow l+p}_{sol})\left(\frac{1}{\epsilon_{in}} - \frac{1}{\epsilon_w}\right) + \frac{V^{l}_{q\mu}}{\epsilon_{in}}\right\rangle_{(l'+p)}\right.\\ &\left. + \left\langle(\Delta G^{l'+p\rightarrow l+p}_{sol})\left(\frac{1}{\epsilon_{in}} - \frac{1}{\epsilon_w}\right) + \frac{V^{l}_{q\mu}}{\epsilon_{in}}\right\rangle_{(l+p)}\right]\\ &- \frac{1}{2}\left[(\langle\Delta G^{l}_{sol}\rangle_{l'} + \langle\Delta G^{l}_{sol}\rangle_{l})\left(\frac{1}{\epsilon_{in}} - \frac{1}{\epsilon_w}\right)\right.\\ &- (\langle\Delta G^{l}_{sol}\rangle_{l+p} + \langle\Delta G^{l}_{sol}\rangle_{l'+p} - \langle\Delta G^{l}_{sol}\rangle_{l}\\ &\left. - \langle\Delta G^{l}_{sol}\rangle_{l'})\left(1 - \frac{1}{\epsilon_{in}}\right)\right] + \Delta G^{l}_{intra} \quad (4)\end{aligned}$$

where ΔG^{l}_{intra} is the contribution to the cycle due to structural relaxation of V^{l}_{intra}. Here we use the fact that $\Delta G^{l'}_{sol} = 0$, since the hydrophobic term is evaluated separately. Also note that ΔG^{p}_{sol} cancels out in the cycle. The last two terms of Eq. (4) can be written as:

$$-\frac{1}{2}\left[(\langle\Delta G^{l}_{sol}\rangle_{l} + \langle\Delta G^{l}_{sol}\rangle_{l'})\left(1 - \frac{1}{\epsilon_w}\right) - (\langle\Delta G^{l}_{sol}\rangle_{l+p} + \langle\Delta G^{l}_{sol}\rangle_{l'+p})\left(1 - \frac{1}{\epsilon_{in}}\right)\right] + \Delta G^{l}_{intra}. \quad (5)$$

Frequently it is advantageous to keep the ligand structure unchanged in the cycles used to evaluate Eq. (4) and to evaluate the energetics of its structural relaxation in water by a separate microscopic step. In this case, the last two terms in Eq. (4) reduce to $-\langle\Delta G^{l}_{sol}\rangle_{l+p}[(1/\epsilon_{in}) - (1/\epsilon_w)]$. For additional details, we refer the reader to references 3, 6 and 21.

Once the terms in Eq. (4) are evaluated, we can finally express the overall PDLD/S-LRA binding free energy as:

$$\begin{aligned}\Delta G_{bind} &= \Delta G^{p}_{elec,l} - \Delta G^{w}_{elec,l} + (\Delta G^{bind}_{elec,l'})'\\ &\quad + (\Delta G^{bind}_{relax,l'})' + \Delta G_{hyd} + \Delta G_{vdw} - T\Delta S'\\ &= \Delta G^{elec}_{bind} + \Delta G_{hyd} + \Delta G_{vdw} - T\Delta S' \quad (6)\end{aligned}$$

where $\Delta G^{bind}_{elec,l'}$ is now divided to two contributions. The $(\Delta G^{bind}_{elec,l'})'$ term represents the contribution to $\Delta G^{bind}_{elec,l'}$ from the process of binding p and l' while keeping each at the same geometry it assumes in the bound complex. This term is evaluated by PDLD/S calculations of $p + l'$ and of separate p and l', averaging in each case over a few configurations of p and l' generated at the bound state (this procedure is not a rigorous LRA procedure since the energies of the reactant and product are evaluated at different structures). The $(-\Delta G^{bind}_{relax,l'})'$ contribution represents the reorganization of p and l' once they are separated and allowed to relax from their configurations at the complex. Now the $-T\Delta S'$ does not include the entropic contributions to the $(\Delta G^{bind}_{relax,l'})'$ term, but only the translation and rotation terms and also the configurational entropy contributions of the relaxation step that are not captured in the finite LRA calculations of the relaxation step. The $(\Delta G^{bind}_{relax,l'})'$ term is neglected in our calculations of ligand binding to proteins assuming that the LRA treatment of the reorganization during the charging process (the evaluation of these contributions to $\Delta G^{p}_{elec,l}$) accounts for most of the protein reorganization effect. Unfortunately, this approximation is not always fully justified since the relaxation upon binding of the uncharged ligand might be significant.

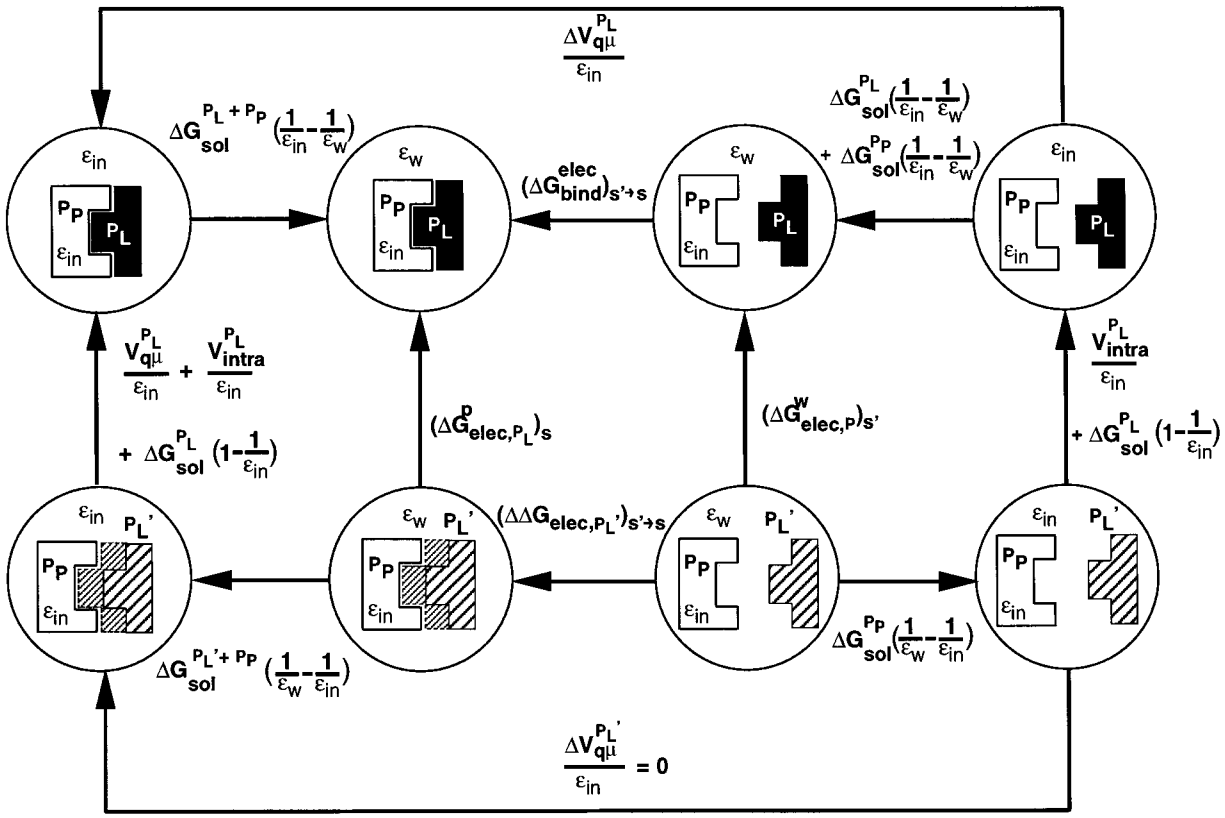

Fig. 2. Thermodynamic cycles used to study the interaction between two proteins, P_L and P_P. The cycles are almost identical to the cycles of Figure 1, except that now the ligand is replaced by another protein (P_L) and that the uncharging of P_L is used to consider the relaxation of P_P upon separation from P_L. The figure considers only the electrostatic free energy and this is done for a single configuration (s) of the $P_P + P_L$ system and for a single configuration (s′) of this system in the dissociated state. Thus it cannot be used to visualize the relaxation effect. This effect is captured by using the LRA approach where the electrostatic free energy $\Delta G^{p}_{elec,P_L}$ is evaluated for configurations generated by MD simulations on both states when P_L is charged and when P_L is uncharged. This includes configurations where the complex of P_P with the uncharged P_L is forced to partially dissociate. A partially dissociated configuration is also drawn in the figure, although it does not mean that the cycle involves any actual structural change; as stated above, we use the cycle for the same configurations in both states, but the LRA sampling then considers different configurations for the complex.

More specifically, when we consider the $a \to d \to e$ steps in Figure 1, it is assumed that most of the relaxation of the $a \to e$ step occurs during the uncharging process ($a \to d$), or in other words, that the relaxation effect is captured in $\Delta G^{p}_{elec,l}$. However, although the change in solvation of the protein upon binding is calculated in $(\Delta G^{bind}_{elec,l'})'$, we leave in the neglected $\Delta G^{bind}_{relax,l'}$ term the additional penetration of water upon relaxation of p from its ($p + l'$) configuration. Furthermore, we neglect relaxation of the protein sidechains upon dissociation and the effect of this relaxation might be significant.

Despite the above concern, it seems that the neglect of the $(\Delta G^{bind}_{relax,l'})'$ term is a reasonable approximation in case of the ligand binding to proteins, since the binding site is relatively small and its shape is often kept in absence of a ligand. The reorientational freedom of the residues, and particularly their sidechains upon ligand docking, is rather confined. This is not the case, however, when one deals with protein–protein binding. This case involves relatively large number of surface groups of two proteins that have to reorient in a very specific way to form hydrogen bond networks that finally build the interaction surface between the two proteins. In the case of the separated proteins, the interaction surfaces are fully exposed to aqueous solution and the hydrophilic sidechains are likely to reorient in water to maximize their solvation. In other words, the cycle of Figure 1 does not allow the surface groups of the first protein to fully relax since they are blocked by the surface groups of the second protein, which is considered as the ligand.

In order to circumvent the above problem, we introduce the thermodynamic cycle of Figure 2, which can be considered as a variant of the cycle of Figure 1. The new cycle considers in a somewhat artificial way one protein as a "ligand" (P_L), and the

second as a protein (P_P). Beyond this aspect, the main difference between the cycle of Figure 1 and that of Figure 2 is that the latter cycle considers simultaneously the uncharging of (P_L) and its separation from (P_P). Now the electrostatic contributions to the binding energy are given by:

$$\Delta G^{elec}_{bind} = \Delta G^{p}_{elec,P_L} - \Delta G^{w}_{elec,P_L} + \Delta\Delta G^{bind}_{elec,P_{L'}}. \quad (7)$$

$\Delta G^{w}_{elec,P_L}$ corresponds to the solvation of P_L and is formally identical to the corresponding term in Eq. (6). The difference is that the $\Delta G^{p}_{elec,P_L}$ term contains now the relaxation of P_P upon partial dissociation of the complex, while the term $\Delta\Delta G^{bind}_{elec,P_{L'}}$ represents the effect of separating the proteins from their finite separation to infinity. Since it is arbitrary which protein is considered P_L and which P_P, one can perform the calculation twice, treating the proteins in either way. Both results can be compared and used for estimating the systematic error in our approach.

The PDLD/S expression for the energy of Eq. (7) is formally identical to the expression of Eq. (1). The only difference is that the LRA treatment now considers a partially dissociated complex. In order to determine these terms we start by evaluating the PDLD/S effective potentials using the cycles of Figure 2. This gives:

$$\begin{aligned}(U^{p}_{elec,P_L})_s = &\left[(\Delta G^{P_L+P_P}_{sol} - \Delta G^{P_{L'}+P_P}_{sol})\left(\frac{1}{\epsilon_{in}} - \frac{1}{\epsilon_w}\right)\right.\\ &\left. + \Delta G^{P_L}_{sol}\left(1 - \frac{1}{\epsilon_{in}}\right) + \frac{V^{P_L}_{q\mu}}{\epsilon_{in}} + \frac{V^{P_L}_{intra}}{\epsilon_{in}}\right]_s\\ (U^{w}_{elec,P_L})_{s'} = &\left[\Delta G^{P_L}_{sol}\left(\frac{1}{\epsilon_{in}} - \frac{1}{\epsilon_w}\right)\right.\\ &\left. + \Delta G^{P_L}_{sol}\left(1 - \frac{1}{\epsilon_{in}}\right) + \frac{V^{P_L}_{intra}}{\epsilon_{in}}\right]_{s'} \quad (8)\end{aligned}$$

where P_L and $P_{L'}$ represent the charged and uncharged forms of P_L. The $V^{P_L}_{intra}$ term represents the intramolecular interaction of P_L. We also keep in mind that these potential corresponds to a single configuration of P_L and P_P. Next, we use the LRA approach allowing the proteins to reach a finite separation (R_2) during the relaxation process and obtain:

$$\begin{aligned}\Delta G^{elec}_{bind} \approx\ & 1/2[\langle U^{p}_{elec,P_L}\rangle_{(P_L+P_P,R_1)} + \langle U^{p}_{elec,P_L}\rangle_{(P_{L'}+P_P,R_2)}]\\ &- 1/2[\langle U^{w}_{elec,P_L}\rangle_{P_L(\infty)} + \langle U^{w}_{elec,P_L}\rangle_{P_{L'}(\infty)}] \quad (9)\end{aligned}$$

where R_1 denotes the complex configuration.

Here we assume that the free energy $\Delta\Delta G_{elec,P_{L'}}$, of separating the proteins from the finite distance R_2 to infinity is negligible and write:

$$\begin{aligned}\Delta G^{elec}_{bind} \approx\ & \frac{1}{2}\left[\left\langle(\Delta G^{P_L+P_P}_{sol} - \Delta G^{P_{L'}+P_P}_{sol})\left(\frac{1}{\epsilon_{in}} - \frac{1}{\epsilon_w}\right)\right.\right.\\ &\left.\left. + \Delta G^{P_L}_{sol}\left(1 - \frac{1}{\epsilon_{in}}\right) + \frac{V^{P_L}_{q\mu}}{\epsilon_{in}}\right\rangle_{(P_L+P_P,R_1)}\right]\\ &+ \frac{1}{2}\left[\left\langle(\Delta G^{P_L+P_P}_{sol} - \Delta G^{P_P}_{sol})\left(\frac{1}{\epsilon_{in}} - \frac{1}{\epsilon_w}\right)\right.\right.\\ &\left.\left. + \Delta G^{P_L}_{sol}\left(1 - \frac{1}{\epsilon_{in}}\right) + \frac{V^{P_L}_{q\mu}}{\epsilon_{in}}\right\rangle_{(P_{L'}+P_P,R_2)}\right]\\ &- \frac{1}{2}\left[\left\langle\Delta G^{P_L}_{sol}\left(\frac{1}{\epsilon_{in}} - \frac{1}{\epsilon_w}\right) + \Delta G^{P_L}_{sol}\left(1 - \frac{1}{\epsilon_{in}}\right)\right\rangle_{P_L(\infty)}\right]\\ &- \frac{1}{2}\left[\left\langle\Delta G^{P_L}_{sol}\left(\frac{1}{\epsilon_{in}} - \frac{1}{\epsilon_w}\right)\right.\right.\\ &\left.\left. + \Delta G^{P_L}_{sol}\left(1 - \frac{1}{\epsilon_{in}}\right)\right\rangle_{P_{L'}(\infty)}\right] + \Delta G^{P_L}_{intra} \quad (10)\end{aligned}$$

where $\Delta G^{P_L}_{intra}$ is the contribution of the ΔV_{intra} term. Here we used the fact that the electrostatic contribution to $\Delta G^{P_{L'}}_{sol}$ is zero and the fact that $\Delta G^{P_{L'}+P_P}_{sol} \approx \Delta G^{P_{L'}}_{sol} + \Delta G^{P_P}_{sol}$, since $V^{P_{L'}}_{q\mu} = 0$ when $P_{L'}$ and P_P are separated even to a rather small distance. We can then write

$$\begin{aligned}\Delta G^{elec}_{bind} \approx\ & \frac{1}{2}\left[\left\langle(\Delta G^{P_L+P_P}_{sol} - \Delta G^{P_{L'}+P_P}_{sol})\left(\frac{1}{\epsilon_{in}} - \frac{1}{\epsilon_w}\right)\right.\right.\\ &\left.\left. + \frac{V^{P_L}_{q\mu}}{\epsilon_{in}}\right\rangle_{(P_L+P_P,R_1)}\right]\\ &+ \frac{1}{2}\left[\left\langle(\Delta G^{P_L+P_P}_{sol} - \Delta G^{P_P}_{sol})\left(\frac{1}{\epsilon_{in}} - \frac{1}{\epsilon_w}\right)\right.\right.\\ &\left.\left. + \frac{V^{P_L}_{q\mu}}{\epsilon_{in}}\right\rangle_{(P_{L'}+P_P,R_2)}\right] + \frac{1}{2}\left[(\langle\Delta G^{P_L}_{sol}\rangle_{(P_L+P_P,R_1)}\right.\\ &\left. + \langle\Delta G^{P_L}_{sol}\rangle_{(P_{L'}+P_P,R_2)}\left(1 - \frac{1}{\epsilon_i}\right)\right] - \frac{1}{2}\left[(\Delta G^{P_L}_{sol}\rangle_{P_L(\infty)}\right.\\ &\left. + \langle\Delta G^{P_L}_{sol}\rangle_{P_{L'}(\infty)}\right)\left(1 - \frac{1}{\epsilon_w}\right)\right] + \Delta G^{P_L}_{intra}. \quad (11)\end{aligned}$$

The derivation of Eq. (11) assumes that the PDLD/S effective free energies can be described within the LRA, using harmonic Marcus' parabolas (Fig. 3). In the present case, we describe by parabolas the free energy

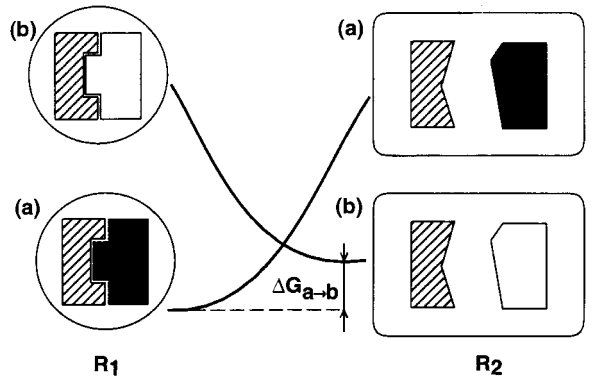

Fig. 3. A schematic illustration of the evaluation of $\Delta G^{p}_{elec,P_L}$ (Fig. 2) using the LRA/S approach. a and b represent, respectively, the charged and uncharged states of P_L. R_1 represents the conformation of the protein–protein complex, and R_2 represents the conformation of the two proteins at infinite separation. Thus, the two parabolas represent the free energy of states a and b as a function of the structure of the system along the $R_1 \rightarrow R_2$ reaction coordinate.

functions for separation and reorganization of P_L and P_P in state a (charged P_L) and state b (uncharged P_L). This parabolic description is an implicit assumption of the LRA treatment. Thus, wherever this assumption is valid there is no need to evaluate the change in intramolecular energy of the protein or solvent but only the change in interaction between the protein and ligand upon protein reorganization. The LRA has been found to be valid in studies of solvation processes[22,23,24,25] and in ligand protein interactions.[4,6,26,27] Here we assumed that this treatment can be extended to a combined process of uncharging and dissociation, at least when the dissociation is limited.

We can also reformulate the present derivation assuming that the LRA is valid for uncharging and a direct separation to infinity rather than to R_2. When this is done, one can write

$$\Delta G^{elec}_{bind} = \frac{1}{2}\left[\left\langle (\Delta G^{P_L+P_P}_{sol} - \Delta G^{P_{L'}+P_P}_{sol})\left(\frac{1}{\epsilon_{in}} - \frac{1}{\epsilon_w}\right) + \frac{V^{P_L}_{q\mu}}{\epsilon_{in}}\right\rangle_{(P_L+P_P,R_1)}\right] - \frac{1}{2}\left[\left\langle \Delta G^{P_L}_{sol}\left(\frac{1}{\epsilon_{in}} - \frac{1}{\epsilon_w}\right)\right\rangle_{P_L(\infty)}\right] - \frac{1}{2}\left[(\langle \Delta G^{P_L}_{sol}\rangle_{P_L(\infty)} - \langle \Delta G^{P_L}_{sol}\rangle_{P_L+P_P,R_1})\left(1 - \frac{1}{\epsilon_{in}}\right)\right] + \Delta G^{P_L}_{intra}. \quad (12)$$

Here we exploit the cancellation of the $\Delta G^{P_L}_{sol}$ contributions when $R_2 \rightarrow \infty$ and the fact that $V_{q\mu} = 0$ at infinite separation. We also allowed P_L to relax to its structure at infinite separation only at $P_L(\infty)$ rather than at $P_{L'}(\infty)$. Now we can write Eq. (12) in a way that the reorganization energy of P_L upon complex formation is separated from the other terms. This leads to

$$\Delta G^{elec}_{bind} = \frac{1}{2}\left[\left\langle (\Delta G^{P_L+P_P}_{sol} - \Delta G^{P_{L'}+P_P}_{sol})\left(\frac{1}{\epsilon_{in}} - \frac{1}{\epsilon_w}\right) + \frac{V^{P_L}_{q\mu}}{\epsilon_{in}}\right\rangle_{(P_L+P_P,R_1)} - \langle \Delta G^{P_L}_{sol}\rangle_{(P_L+P_P,R_1)}\left(\frac{1}{\epsilon_{in}} - \frac{1}{\epsilon_w}\right) + \langle \Delta\Delta G^{P_L}_{sol}\rangle_{(P_L(\infty)\rightarrow(P_L+P_P,R_1))}\left(1 - \frac{1}{\epsilon_w}\right)\right] + \Delta G^{P_L}_{intra} \quad (13)$$

where $\langle \Delta\Delta G^{P_L}_{sol}\rangle_{(P_L(\infty)\rightarrow(P_L+P_P,P_1))} = \langle \Delta G^{P_L}_{sol}\rangle_{(P_L+P_P,R_1)} - \langle \Delta G^{P_L}_{sol}\rangle_{(P_L(\infty)}$. The fact that Eq. (13) does not contain the average over $\langle\ \rangle_{P_{L'}}$ is a result of extending the LRA assumption to the process of separating the proteins from their initially relaxed state to infinity (see below). This results in the factor 1/2 which plays the same role as the factor 1/2 in continuum treatments such as Born's equation.[28] Furthermore, the LRA assumption leads to the elimination of the $\Delta G^{P_P}_{sol}$ term. This means that the reorganization energy of P_P is only given implicitly in the factor 1/2 of Eq. (13). Note, however, that since it is arbitrary which protein is considered as P_L and which as P_P, one uses both proteins as P_L and P_P, respectively, in separate simulations, and averages over the results. That is, the reorganization of both proteins is once treated explicitly and once implicitly. The deviation of the results gives us a rough estimate of the systematic error due to the assumption that the LRA still holds for separating both proteins to infinity.

The actual calculation of the $\Delta\Delta G_{sol} + \Delta G_{intra}$ terms is very challenging since it requires a reasonable estimate of the protein structural reorganization. This involves large opposing interactions. One option is to treat this P_L-reorganization in a separate cycle as drawn in Figure 4. The free energy of this cycle is given by

$$(\Delta\Delta G^{sol}_{bind})_{P_L+P_L(\infty))} \approx \Delta\Delta G^{w}_{sol}\frac{(1 - 1/\epsilon_{in})}{(1 - 1/\epsilon_w)} + \frac{\Delta V_{qq}}{\epsilon_{in}} = \Delta\Delta G^{in}_{sol} + \frac{\Delta V_{qq}}{\epsilon_{in}} \quad (14)$$

where $\Delta\Delta G^{in}_{sol} = (\Delta G^{w}_{sol})_{r_2} - (\Delta G^{w}_{sol})_{r_1}$ and V_{qq} is the internal electrostatic interaction of the protein (charge–charge, charge–dipole, dipole–dipole). Eq. (14) can be further manipulated by using the generalized Born approximation[3,11] and writing

$$\Delta\Delta G^{in}_{sol} + \Delta V_{qq} = \frac{\Delta V_{qq}}{\epsilon_{in}} \quad (15)$$

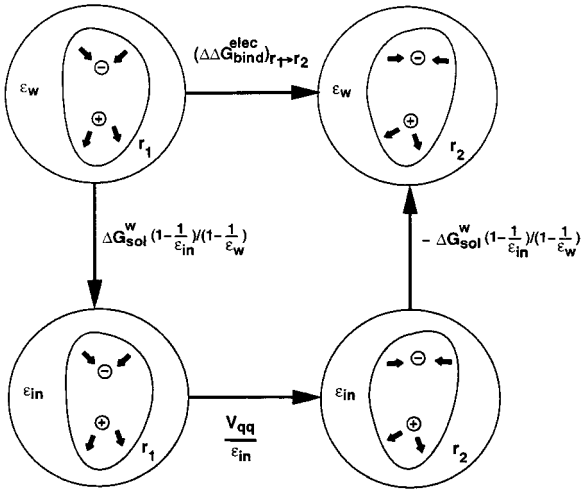

Fig. 4. A thermodynamic cycle that can be used to estimate the reorganization energy of P_L. The cycle describes schematically a structural change in a protein where the orientation of dipoles and the position of charges is changed. The corresponding reorganization free energy $\Delta\Delta G^{elec}_{bind}$ can be obtained by considering a PDLD/S cycle where the solvent dielectric is changed from ϵ_w to ϵ_{in}, the protein is allowed to rearrange and the dielectric is changed back to ϵ_w.

which leads finally to

$$(\Delta\Delta G^{elec}_{bind})_{(P_L \to P_L(\infty))} = -\Delta\Delta G^{w}_{sol}\left(1-\frac{1}{\epsilon_{in}}\right)\left(1-\frac{1}{(1-1/\epsilon_{in})\epsilon_{in}}\right) = -\Delta\Delta G^{w}_{sol}\left(1-\frac{2}{\epsilon_{in}}\right). \quad (16)$$

This expression gives the effect of the protein reorganization in terms of the corresponding change in solvation free energy. Thus we can rewrite Eq. (13) as

$$\Delta G^{elec}_{bind} \approx \frac{1}{2}\left\langle(\Delta\Delta G^{P_{L'}+P_P \to P_L+P_P}_{sol} - \Delta G^{P_L}_{sol}) \cdot\left(\frac{1}{\epsilon_{in}}-\frac{1}{\epsilon_w}\right)+\frac{V^{P_L}_{q\mu}}{\epsilon_{in}}\right\rangle_{(P_L+P_P,R_1)} + \left\langle\Delta\Delta G^{P_L}_{sol}\left(1-\frac{2}{\epsilon_{in}}\right)\right\rangle_{(P_L(\infty)\to P_L+P_P,R_1)}. \quad (17)$$

After estimating the electrostatic contributions to binding we can finally write

$$\Delta G_{bind} = \Delta G^{elec}_{bind} + \Delta G_{hyd} + \Delta G_{vdw} - T\Delta S' \quad (18)$$

where the last three terms have the same meaning as in Eq. (6). In addition to the approaches considered above, we also examine an alternative approach which ignores the protein reorganization. This is done by considering the cycle on the upper part of Figure 2 without any protein relaxation. This cycle gives

$$\Delta G^{elec}_{bind} = \left\langle(\Delta G^{P_L+P_P}_{sol} - \Delta G^{P_L}_{sol} - \Delta G^{P_P}_{sol}) \cdot\left(\frac{1}{\epsilon_{in}}-\frac{1}{\epsilon_w}\right)+\frac{\Delta V^{P_L}_{q\mu}}{\epsilon_{in}}\right\rangle_{(P_L+P_P,R_1)}. \quad (19)$$

Eq. (19) should provide results which should be quite close to the corresponding results of current continuum approaches (where reorganization energy is not considered explicitly).

Details of the Simulations

The present work involves MD simulations of the Raf-RBD/Rap complex, as well as of Raf-RBD and Rap in aqueous solution using the program ENZYMIX.[3] The crystal structure of the complex was provided by Wittinghofer.[17] The initial structures of the separated proteins were taken from the complex structure, then relaxed separately in a special procedure. This relaxation procedure presents a major challenge and was accomplished by a special iterative strategy that will be outlined below (after describing the PDLD/S procedure).

The PDLD/S energy of the relaxed structure was evaluated by placing Raf-RBD plus its interaction surface with Rap inside a water sphere of a 26.5 Å radius. This protein–water sphere was completed to a radius of 28 Å by surrounding it with a grid of Langevin dipoles. The region beyond the 28 Å radius was treated as a continuum. A small harmonic potential of the form $V' = \Sigma_i A(\vec{r}_i - \vec{r}^{\,0}_i)^2$, with A = 0.01 kcal mol^{-1}, was applied in order to keep the protein atoms near the corresponding observed positions. Protein atoms outside this sphere were held fixed and their electrostatic effect excluded from the model.[3] The ENZYMIX calculations were performed using the Surface Constrained All Atomic Solvent (SCAAS) boundary conditions and the Local Reaction Field method that provides a very efficient way of solving the problem of a proper treatment of long-range electrostatic interactions.[29] The PDLD/S-LRA calculations involved a series of four MD simulations of two picoseconds each for the charged and uncharged forms of P_L, and an automatic averaging of the corresponding PDLD/S effective potential. The calculations considered the GTP^{4-} and the Mg^{2+} ions in their protein site and treated all ionizable residues in the ionization state expected from them in water at pH = 7. Thus, all lysine, arginine, aspartic acid, and glutamic acid residues were treated at their charged states. This approximation could be removed by proper calculations of the pK_α in the actual protein site[21] but such a study is out of the scope of the present work (see Discussion section).

Some versions of our approach require us to evaluate the solvation of the complex as well as of the

separated proteins. Therefore, it is not sufficient to simulate the reorganization of a part (e.g., the interaction surface) of the complex. It is rather necessary to evaluate the relaxation of the protein surface in a way that would represent the force field of the entire protein. This is particularly crucial since the initial structures of the separated proteins were generated by using MD relaxation with the ENZYMIX force field rather than from experimentally determined structures (that were only available for the complex but not for the separated proteins). Thus, we had to design a special strategy that will allow the entire complex to relax properly in the complex state (this is also important in studying the global relaxation due to mutations). In designing such a strategy, we had to address the challenge of treating the relaxation consistently while using a manageable simulation sphere (such a spherical system allows for proper treatment of long-range effects). In order to accomplish this we developed an iterative simulation procedure where the center of our 26.5 Å simulation sphere is shifted to different regions of the protein. More specifically, we started placing the center of the simulation sphere in the center of the interaction surface and run the first 2 ps MD trajectory at a temperature of 300 K and 1 fs time step. In the next step, we displaced the center of the simulation sphere along the axis of the narrowest cylinder that contained all atoms of the complex, so that Rap and the interaction surface would be inside the simulation sphere. The remaining Raf-RBD residues that stuck out of the simulation sphere were held fixed. The new region, covered by the simulation sphere, was then treated by the same type of MD protocol described above. Next, we switched back to the first simulation region and repeated the whole procedure five times. The conformations generated at the end of the last four MD runs were used in the binding calculations.

In order to evaluate the solvation of the infinitely separated proteins, Raf and Rap structures underwent MD relaxation for 50 ps at a temperature of 300 K each (these systems are sufficiently small, so they did not require the above-mentioned iterative procedure). The heavy atoms were only weakly constrained to their positions with a harmonical force constant of A = 0.01 kcal mol^{-1}. In the case of Rap that contains the highly charged GTP^{4-}-Mg^{2+}-complex, the GTP and all heavy atoms in a sphere with radius 6 Å around the Mg^{2+}-ion were constrained with an harmonical force constant of A = 50.0 kcal mol^{-1} (as was also done in the MD simulations of the complex). The uncontrolled relaxation of this highly charged region that includes the Mg^{2+}-ion coordination sphere may result in a substantial displacements from the observed structure, our focus here is on energetics rather than on reproducing exact structures (see related discussion of Langen et al.[26]). The r.m.s deviations were 1.1 and 0.7 Å for Raf and Rap, respectively.

TABLE I. Contributions to the Free Energy of Raf-RBD/Rap Binding*

Term	Energy (kcal/mol)
$\langle \Delta G_{sol}^{(P_L+P_P,R_1)} \rangle_{(P_L+P_P,R_1)}$	−1122
$\langle \Delta G_{sol}^{(P_L+P_P,R_1)} \rangle_{(P_L+P_P,R_1)}^{P_L=Raf}$	−1066
$\langle \Delta G_{sol}^{(P_L+P_P,R_1)} \rangle_{(P_L+P_P,R_1)}^{P_L=Rap}$	−562
$\langle \Delta G_{sol}^{P_L} \rangle_{(P_L+P_P,R_1)}^{P_L=Raf}$	−747
$\langle \Delta G_{sol}^{P_L} \rangle_{(P_L+P_P,R_1)}^{P_L=Rap}$	−1219
$\langle V_{q\mu}^{P_L} \rangle_{(P_L+P_P,R_1)}$	−868
$\langle \Delta\Delta G_{sol}^{P_L} \rangle_{(P_L(\infty)\rightarrow P_L+P_P,R_1)}^{P_L=Raf}$	(42)
$\langle \Delta\Delta G_{sol}^{P_L} \rangle_{(P_L(\infty)\rightarrow P_L+P_P,R_1)}^{P_L=Rap}$	(−61)
$-T\Delta S'$	(10)
ΔG_{hyd}	(14)
ΔG_{vdw}	(−23)

*The calculations evaluate the various contributions to Eqs. (17) and (18). The meaning of the different terms is explained in the text.

Defining and Calculating Group Contributions

In order to correlate structure and function of our complex, it is useful to decompose the binding energy into residue-specific contributions (these contributions are called "group contributions"). This decomposition allows us to identify residues that play an essential role in protein complex formation. As argued before, the evaluation of group contributions is not unique and depends on the definition used.[13,27] In particular, we do not see a way to define a thermodynamic cycle that can be used to uniquely decompose the total binding energy of two interacting proteins (or a ligand and a protein) into its group contributions. The values of the relevant group contributions depend on their sequential occurrence in the thermodynamic cycle. In principle, one can take an operational definition and define group contribution as the free energy of introducing a new residue (glycine) in the given position. This definition allows one to compare calculated and observed free energies directly and to assess the reliability of the calculations. The magnitude of the group contributions depends largely on whether the other residues and the solvent are allowed to relax during the evaluation of the given contribution. The relaxation of the local environment upon mutation can be captured conveniently by performing MD calculations for the respective mutants. However, these MD simulations are very time-consuming. Nevertheless, such mutational calculations were performed here for the mutants Rap(D38A), Raf(R89L), Raf(R59A), Raf(K84A), and Raf(R67A).

Our starting point for evaluating group contributions is Eq. (10). Considering the contribution of a

TABLE II. Absolute Binding Energies of Raf-RBD/Rap Wild Type and Mutants*

		ΔG^{calc}_{bind} (kcal/mol)						ΔG^{obs}_{bind} † (kcal/mol)
		$\epsilon_{in}=2$	$\epsilon_{in}=3$	$\epsilon_{in}=4$	$\epsilon_{in}=5$	$\epsilon_{in}=6$	$\epsilon_{in}=25$	
Wild type	P_L = Raf	−48.6	−33.8	−26.4	−22.0	−19.1	−7.9	−8.2
Wild type	P_L = Rap	−56.4	−39.0	−30.2	−25.0	−21.5	−8.3	−8.2
Raf(R89L)	P_L = Rap	−43.0	−30.0	−23.4	−19.6	−17.0	−7.1	>−2.7
Rap(D38A)	P_L = Raf	−37.6	−26.4	−20.8	−17.5	−15.3	−6.7	−8.1
Raf(R59A)	P_L = Rap	−26.9	−19.1	−15.3	−12.9	−12.4	−5.9	−6.2
Raf(K84A)	P_L = Rap	−2.0	−2.9	−3.3	−3.5	−3.7	−6.0	−6.9
Raf(R67A)	P_L = Rap	−73.2	−49.7	−37.9	−30.8	−26.1	−8.3	−6.9

*The calculations involve the use of Eqs. (17) and (18) while neglecting $\langle \Delta\Delta G^{P_L}_{sol} \rangle_{(P_L(\infty)\rightarrow P_L+P_P,R_1)}$ and assuming that the last three terms of Eq. (15) cancel each other.
†Since the binding free energy of Rap(D38A) to Raf is not known, we give here the binding affinity between Raf and the mutant of $p21^{ras}$.[35] The binding affinity of the wild type $p21^{ras}$ and Rap is −10.7 kcal/mol. The changes in binding energy of the Rap/Raf complex upon mutations are very similar to the corresponding changes in Ras/Raf.[35,34,36] Therefore, we compare the calculated relative binding affinity of Rap(D38A)/Raf with the observed relative binding affinity of Ras(D38A)/Raf.

TABLE III. Relative Binding Energies of Raf-RBD/Rap Mutants*

		$\Delta\Delta G^{calc}_{bind}$ (kcal/mol)						$\Delta\Delta G^{obs}_{bind}$ † (kcal/mol)
		$\epsilon_{in}=2$	$\epsilon_{in}=3$	$\epsilon_{in}=4$	$\epsilon_{in}=5$	$\epsilon_{in}=6$	$\epsilon_{in}=25$	
Raf(R89L)	P_L = Rap	13.4	9.0	6.8	5.4	4.5	1.2	>5.5
Rap(D38A)	P_L = Raf	11.0	7.4	5.6	4.5	3.8	1.2	2.6
Raf(R59A)	P_L = Rap	29.5	19.9	14.9	12.1	10.1	2.4	2.0
Raf(K84A)	P_L = Rap	54.4	36.1	26.9	21.5	17.8	2.3	1.5
Raf(R67A)	P_L = Rap	−16.8	−9.3	−7.7	−5.8	−4.6	0.0	1.3

*The calculated values are based on the results given in Table II and the binding free energies are given relative to the wild type binding free energies.
†Taken from Nassar et al.[34] (Raf-mutants) and Herrmann et al.[35] (Ras mutant D38A).

single residue which is located at P_P we can write:

$$\Delta\Delta G^{i}_{bind} \simeq \Delta^{(i)}\left\langle \frac{1}{2}\left\langle (\Delta\Delta G^{P_L'+P_P\rightarrow P_L+P_P}_{sol}) \cdot \left(\frac{1}{\epsilon_{in}}-\frac{1}{\epsilon_w}\right)+\frac{V^{P_L}_{q\mu}}{\epsilon_{in}}\right\rangle_{P_L+P^{(i)}_P,R_1)}\right\rangle \quad (20)$$

where $\Delta^{(i)}$ designates the change of the function in brackets upon mutating the ith residue of P_P. Here we assume that the change in structure of P_L upon mutation of P_P is small so that the corresponding change in $\Delta\Delta G_{sol}$ is negligible. Furthermore, we also neglected the change in the last three terms of Eq. (18).

Eq. (20) can be further simplified by several approximations. The first one, which is referred to as the solvent-relaxed approximation treats the changes in the structure of the complex upon mutation implicitly by choosing the proper ϵ_{in}. In this way, we only consider the effect of changing the charge of the ith residue on $\Delta\Delta G_{sol}$ (this represents the reorientation of the solvent dipoles) and the changes in $V_{q\mu}$.

The second approximation, which is referred to as nonrelaxed approximation considers also the solvent relaxation implicitly, thus leading to

$$\Delta\Delta G^{i}_{bind} \simeq \frac{1}{2}\left\langle \frac{V^{P_L}_{q\mu}}{\epsilon_i}\right\rangle^{(i)} = \left\langle \frac{V^{P_L}_{q\mu}}{\epsilon_{eff}}\right\rangle^{(i)}. \quad (21)$$

It should be mentioned here that the present nonrelaxed treatment is somewhat different from its previous versions[27] which did not include the implicit relaxation associated with the factor 1/2 and the cycle of Figure 1. In general, we find in studies of electrostatic energies in proteins that $\epsilon_{eff}=4$ provides reasonable estimates of the contribution of neutral residues in case of nucleotide binding in $p21^{ras}$ [27] and heme "solvation" in cytochrome c.[13] This approach does not (and should not) work for charged residues, as their effect would be drastically overestimated. For charged residues, a scaling factor $2\epsilon_{in}=\epsilon_{eff}=40$ is appropriate. This approximation has been repeatedly validated (e.g. refs. 11 and 27).

RESULTS AND DISCUSSION

Rap/Raf-Binding Affinity

The binding affinity of Raf-RBD to Rap was calculated using the PDLD/S-LRA approach as described in the Simulation Method section using Eqs. (16) and (17). The absolute binding energies were calculated twice using Raf-RBD or Rap as P_L, respectively. The calculated contributions of all the terms in Eqs. (17) and (18) are given in Table I. These terms provide the raw data for the evaluation of the binding energy of the native complex. The comparison of the individual terms that contribute to the binding energy

TABLE IV. Binding Energies of Raf-RBD/Rap Mutants Evaluated Without Considering Any Protein Reorganization*

	$\Delta\Delta G_{bind}$ (kcal/mol)						$\Delta\Delta G_{bind}^{obs}$ †
	$\epsilon_{in}=2$	$\epsilon_{in}=3$	$\epsilon_{in}=4$	$\epsilon_{in}=5$	$\epsilon_{in}=6$	$\epsilon_{in}=25$	(kcal/mol)
Raf(R89L)	4.9	3.6	2.9	2.5	2.2	1.2	>5.5
Rap(D38A)	83.4	55.2	41.1	32.7	27.0	5.6	2.6
Raf(R59A)	58.5	39.3	29.7	24.0	20.1	5.6	2.0
Raf(K84A)	120.9	80.2	60.0	47.7	39.5	8.6	1.5
Raf(R67A)	−64.3	−41.6	−30.3	−23.5	−19.0	−1.8	1.3

*The calculations used Equation 19. The calculated binding free energies are given relative to the corresponding free energy of the wild type protein, evaluated using $\epsilon_{in}=4$. This binding free energy is −17 kcal/mol. See Table II for more information.
†Taken from Nassar et al.[34] (Raf-mutants) and Herrmann et al.[35] (Ras mutant D38A).

TABLE V. The Effect of ϵ_{in} and Different Models on Relative Binding Free Energies Obtained for Charged and Uncharged Residues*

	Nonrelaxed† V_{qu}/ϵ_{in} kcal/mol		Solvent-relaxed‡ $\Delta\Delta G^{i}_{bind}$ kcal/mol			Locally§ relaxed protein $\Delta\Delta G^{i}_{bind}$ kcal/mol		Observed $\Delta\Delta G_{bind}$
Construct	$\epsilon_{in}=4$	$\epsilon_{in}=25$	$\epsilon_{in}=4$	$\epsilon_{in}=25$	$\epsilon_{in}=40$	$\epsilon_{in}=4$	$\epsilon_{in}=25$	kcal/mol
Raf(R59A)	20.5	3.3	14.7	3.0	2.1	12.7	2.2	2.0
Raf(K65A)	2.0	0.3	1.0	0.2	0.2	−3.2	0.0	1.1
Raf(R67A)	17.5	2.8	13.1	2.8	2.0	13.0	2.5	1.5
Raf(K84A)	17.5	2.8	13.0	2.6	1.9	14.1	2.8	1.3
Raf(R89L)	20.5	3.3	18.4	3.6	2.5	15.0	3.1	>5.5
Raf(N64A)	2.1		1.7	0.3	0.2			0.5
Raf(Q66A)	0.7		0.2	0.1	0.0			1.6
Raf(T68A)	3.1		2.8	0.5	0.3			1.5
Raf(V69A)	0.3		0.7	0.1	0.0			1.1

*Relative binding energies of mutants of Raf-RBD/Rap.[34] The upper and lower parts of the table involve charged and uncharged residues, respectively.
†The calculations used Eq. (21).
‡The calculations used Eq. (20) treating the changes in the protein structure upon mutation implicitly.
§The calculations used Eqs. (17) and (18). However, in contrast to the results of Tables II and III, where the entire protein was free to relax upon mutation, here only atoms in a small sphere with 8 Å radius around the mutated residue were allowed to relax during the MD relaxations of the mutated protein. The special treatment of shifting the simulation sphere over the complex was therefore unnecessary and not used. Instead, the relevant regions of the complex and the unbound proteins underwent 2 ps MD relaxation, respectively. Therefore, this model is called "locally" relaxed.

reveals that these terms are two orders of magnitude larger than the actual binding energy. Thus, we have to deal with large numbers that have to compensate for each other. Particularly, the Coulomb interaction between both proteins fluctuates strongly (±10%) with small conformational changes. This leads to relatively large statistical errors, although this error is reduced by configurational averaging.

The relaxation terms, $\langle \Delta\Delta G_{sol}^{P_L} \rangle_{P_L(\infty)\rightarrow P_L+P_P,R_1}$, appear to be unstable and unreliable giving 42 and −61 for the cases with $P_L = Raf$ and $P_L = Rap$, respectively. This is not so unexpected since the ability of MD simulations to actually predict the reorganization of the entire protein is at present very limited. Obviously, it will be interesting to repeat the calculations when the structure of the isolated proteins is known, but at present it is better to ignore this term. In principle we have to account for the last three terms of Eq. (18), and we have reasonable estimates of these quantities. The ΔG_{hyd} and ΔG_{vdw} can be evaluated by the standard PDLD/S procedure[3] using a field-dependent hydrophobic term for ΔG_{hyd} and an estimate of the compensation between solute and solvent van der Waals interactions. Similarly, we can estimate the free energy term $(-T\Delta S')$ associated mainly with the change of translational and rotational entropy upon binding (the electrostatic contributions and part of the configurational entropy is captured in the LRA calculations of the electrostatic free energy). Although there are no rigorous studies of this contribution, it is estimated to be between 10 and 15 kcal/mol in protein–protein complexes.[8,30,31] Of course, the estimate is not certain and, furthermore, the field dependent hydrophobic term is probably not sufficiently reliable since we are dealing with highly charged surfaces and our approach has not been examined or refined for this case. Apparently, as seen from Table I, the sum of the last three terms of Eq. (18) is approximately zero, so it might be reasonable to assume that these terms cancel each other out or that their sum is rather small and can be neglected in the present treatment.

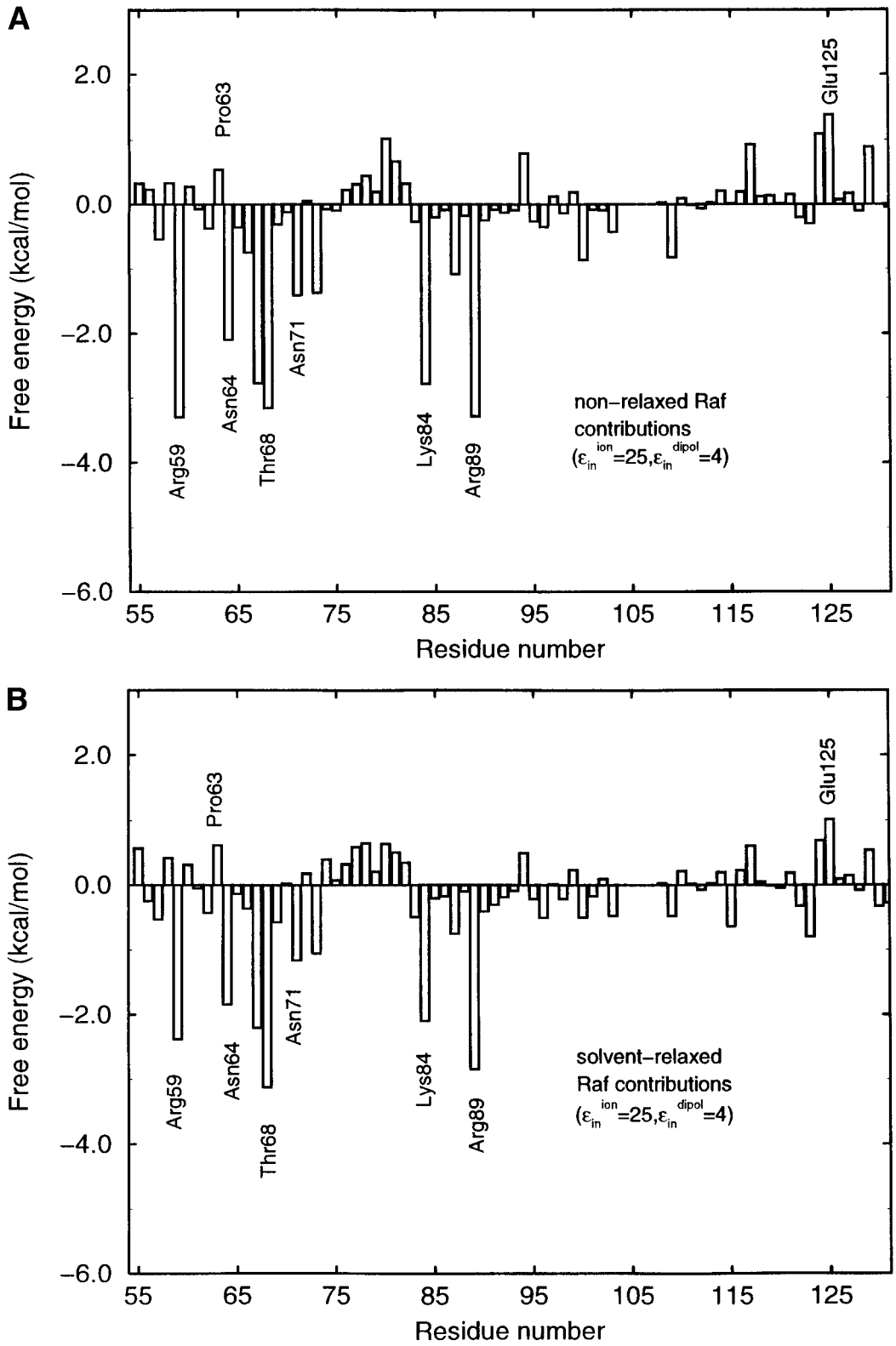

Fig. 5. Electrostatic contributions of individual residues of Raf-RBD to Raf-RBD/Rap binding. **A:** Nonrelaxed group contributions with $\epsilon_{in}^{ion} = 25$ and $\epsilon_{in}^{dipol} = 4$. **B**: Solvent-relaxed group contributions with $\epsilon_{in}^{ion} = 25$ and $\epsilon_{in}^{dipol} = 4$.

The calculated absolute binding energy of the wild type complex is presented in the first two rows of Table II for different values of ϵ_{in}. The calculations with $\epsilon_{in} = 25$ seem to reproduce the observed binding free energy. This, however, cannot be considered as a definite conclusion, considering the uncertainties in $\langle \Delta\Delta G^{P_L}_{sol} \rangle_{P_L(\infty)\rightarrow P_L+P_P,R_1}$ that led to its neglect and the last three terms of Eq. (18). Thus, this part of our study cannot be considered as a systematic study of protein–protein binding energy, but as an exploratory study of the state-of-the-art in such calculations and the problems involved in different strategies. In this respect our calculations indicate that the structural relaxation needed for the evaluation of $\langle \Delta\Delta G^{P_L}_{sol} \rangle_{P_L(\infty)\rightarrow P_L+P_P,R_1}$ cannot be reproduced in a reliable way by MD simulations. More reliable results should be produced by knowing the crystal structures of the complex and the isolated proteins (note in this respect that we only need the three observed structures and do not reassume any information about the dynamics of the structural change). Such a study, however, is not within the scope of the present work.

Calculating Absolute Binding Free Energies for Some Mutants

A significant part of the uncertainty in our calculations can be reduced in evaluating the effect of mutations, where some of the errors are expected to cancel out. For instance, since we are free in choosing which protein serves as P_L and which as P_P, we can arbitrarily place the mutation site on P_P. That leads to cancellation of the $\langle \Delta\Delta G^{P_L}_{sol}[1 - (2/\epsilon_{in})] \rangle_{P_L(\infty)\rightarrow P_L+P_P,R_1}$ term of Eq. (17), which is one of the major sources of the instability in our calculations (see Table I and the above discussion).

In order to examine the performance of our strategies in evaluating the effect of mutations, as well as the validity of our conclusions about ϵ_{in}, we considered several important mutants on the Rap and Raf-RBD surfaces. The calculations involve evaluation of the *absolute* binding energies of the mutant and the wild type rather than by applying a different cycle that calculates relative binding energies indirectly by using 'mutational approaches' that perform the mutation in the complex and in the dissociated proteins. Obviously, such a mutational approach would provide more reasonable results for *changes* in binding free energies upon mutation. Here, however, we address the much more challenging issue of the overall interaction between two proteins (we still study the effect of mutations, but this is done mainly as a tool for validation of our different approaches).

The results of the mutational calculations are summarized in Tables II and III. As seen from the tables, the effect of some mutations can only be reproduced with large values of ϵ_{in}, while the effect of others can be reproduced with $\epsilon_{in} = 4$. The need to use large ϵ_{in} most probably reflects the effect of incomplete relaxation and indicates that direct information about the mutant structures in the complex and the isolated proteins might be crucial for quantitative results. Nevertheless, as will be shown below, we can use our approach in qualitative studies of the contributions of individual residues to the overall binding process and for addressing the biological "importance" of different residues. It is also quite clear from inspection of Table III that $\epsilon_{in} = 25$ gives on average much better results than those obtained with smaller values of ϵ_{in}. This is significant, although the conclusion about the need of large ϵ_{in} is also obtained from the calculations of the absolute binding energies (Table II). That is, the results reported in Table I and Eq. (18) could have been used to reproduce the observed binding free energy with small ϵ_{in}, assuming that the positive estimate of $\langle \Delta\Delta G^{P_L}_{sol} \rangle_{P_L(\infty)\rightarrow P_L+P_P,R_1}$ reflects the correct trend and the reorganization of P_L. On the other hand, considering several mutations leads to much more unique conclusion about the need to use large ϵ_{in}.

In order to present a more complete analysis, we also consider in Table IV the result obtained by the approach of Eq. (19), which does not allow for protein relaxation. The results of this approach are definitely less quantitative than those represented in Table III, but they clearly establish that small values of ϵ_{in} cannot provide realistic values for ΔG_{bind}.

It is quite obvious from the above discussion that the protein relaxation should be taken into account either explicitly or implicitly. Yet, as much as the explicit calculations are concerned, it appears that the present approach (and probably any current alternative approach) cannot provide a quantitative estimate of the reorganization of the entire protein. Thus, we consider here one more variant of the variation procedure of Eqs. (17) and (18), which is termed the "locally relaxed" approach and is presented in Table V. This approach only allows for protein relaxation within an 8 Å radius from the mutated residue, thus reducing the uncertainty associated with the MD-generated global relaxation. That is, formally, the local relaxation approach is less rigorous than the "complete" global relaxation, but as long as we cannot fully trust the treatment it might be preferable to use the local relaxation approach in practical studies. As much as the present study is concerned, the conclusions from comparing Table III and the local relaxation results of Table V are not unique. Overall, the results of the local relaxation model with $\epsilon_{in} = 25$ appear to be more stable than those obtained by the global relaxation approach. Yet our aim is to capture the relaxation explicitly rather than to do so implicitly by using a large ϵ_{in}. In this respect it is instructive that the largest observed effect (>5.5 kcal/mol) for Raf(R89L) is captured by the global relaxation with $\epsilon_{in} = 4$, while the local relaxation model yields too large effect (~15 kcal/mol) with the same ϵ_{in}. On the other

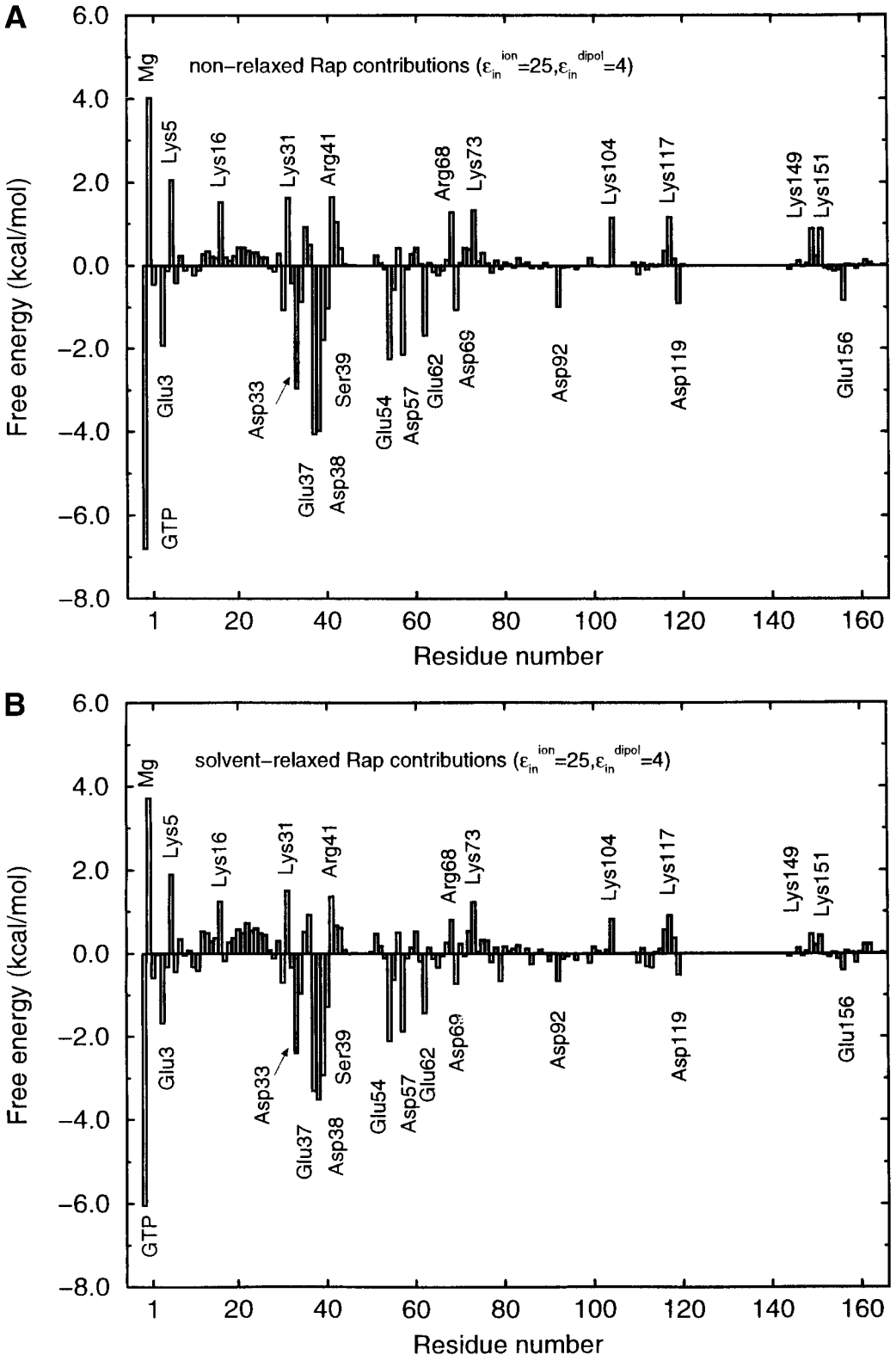

Fig. 6. Electrostatic contributions of individual residues of Rap to Raf-RBD/Rap binding. **A:** Non-relaxed group contributions with $\epsilon_{in}^{ion} = 25$ and $\epsilon_{in}^{dipol} = 4$. **B**: Solvent-relaxed group contributions with $\epsilon_{in}^{ion} = 25$ and $\epsilon_{in}^{dipol} = 4$.

hand, in the case of the Raf(K84A) mutation, it is found that the local relaxation model with $\epsilon_{in} = 4$ gives smaller effect than that obtained with the global relaxation approach.

In order to make the present analysis more complete, and determine the optimal ϵ_{in} for evaluation of group contributions, we also examined in Table V much simpler models that do not consider explicitly any protein relaxation. Evidently, none of the models considered reproduced the observed compensation of charge–charge interaction without using large ϵ_{in}, although the global relaxation model does it in some cases. It is significant that the effect of the solvent relaxation does not provide sufficient compensation. Thus a major part of the experimentally observed compensation would not be captured by current continuum models with small ϵ_{in}. Experimental observation of relatively small charge–charge interactions in the range of 0–5 kcal/mol, rather than the corresponding gas-phase values of $V_{q\mu} > 100$ kcal/mol, forces one to use a Coulomb's law with a large effective dielectric constant that is, of course, supposed to represent all the contributions that are not included explicitly.[11] Trying to treat the system more explicitly should allow one to use smaller values of ϵ_{in}. However (in contrast to common belief), using $\epsilon_{in} = 4$ and a reasonably realistic model of the effect of the *solvent* do not reproduce the experimentally observed compensation effect. A significant part of this effect is due to the protein reorganization, as was pointed out in our early discussion of dielectric effects in proteins (e.g., see Figure 3 of ref. 32).

Table V also considers the group contributions of uncharged residues. Here the picture is very different from that of charged residues. That is, in the case of uncharged residues, it is possible to reproduce the observed trend with $\epsilon_{in} = 4$. This confirms our previous point that the proper dielectric constant for charged and uncharged residues is very different (see e.g., Table 1 in ref. 11 or ref. 27), as long as one is not treating the entire system explicitly. This also means that recent attempts to upgrade ϵ_{in} to ~20 in DC calculations[33] probably accounts for the overestimate of charge–charge interaction (as pointed out much earlier in our studies[32]) but at the same time would lead to underestimation of dipole–charge interactions.

Evaluating Group Contributions

After establishing that our simple solvent-relaxed model, and even the simpler nonrelaxed model, can give reasonable estimate of group contributions to binding, we may use these models to examine biologically relevant problems. This was already done in our previous study of GTP and GDP binding in $p21^{ras}$ [27] and it is done here in Figs. 5 and 6 using $\epsilon_{in} = 25$ and $\epsilon_{in} = 4$ for ionized and neutral residues, respectively. Figure 5 depicts the nonrelaxed and solvent-relaxed contributions of Raf-RBD. As expected, positively charged residues stabilize the complex, whereas negatively charged residues destabilize the protein–protein binding. In particular, Asn64, Gln66, Arg67, and Thr68 that participate in forming the interaction surface between Raf-RBD and Rap exhibit contributions in our calculations that increase the binding energy. This is not surprising since the interaction is mostly due to hydrogen bonds and polar interaction of charged residues. However, it is interesting to note that the strong peak of Arg59 is due to closer contact (3.9 Å) to Rap-Glu37 that has been formed during the MD relaxation. The polar contact (2.9 Å distance between Oδ of Rap-Glu37 and NH_2 of Arg69 was widen to 3.7 Å). The large contribution of Arg89 that is located in the center of the interaction surface is consistent with the large effect of the R89L mutation.[34] The large contribution of Lys84 is due to a 4 Å contact to Rap-Asp33, and it provides a significant contribution to the binding energy.

Figure 6 considers the group contributions of Rap residues to the Raf-RBD/Rap binding. The largest peaks are due to the contributions of GTP^{4-} and MG^{2+}, as they carry the largest charges. The contribution to binding energy of the residues at positions 26, 30, and 31 (that are different in Rap and Ras) is rather small except for the destabilizing effect of the positive charge of Lys31 (Fig. 6). (In fact, the binding affinity of the Rap(K31E) mutant to Raf-RBD, relative to Rap wild type, is increased by a factor of 20 (1.8 kcal/mol)[34]). The effector region of Rap that comprises residues 32–40 shows a stabilizing electrostatic contributions to the protein–protein binding. Especially, Asp33, Glu37, and Asp38 provide the largest contributions. Astonishingly, Arg41 that forms a hydrogen bond with Raf-Asn64 at the edge of the interaction surface does not appear to provide electrostatic stabilization to the complex. This is probably due to its repulsive electrostatic interaction with the positively charged Raf-Lys65 sidechain that lies within 5 Å of the also positively charged Arg41 side chain. The stabilization coming from Glu3 is due to the interaction with Raf-Lys65. The charged sidechains lie within 6 Å. Glu54 that shows also a large stabilizing contribution is located in the middle of the interaction surface and its direct contact to Raf is mediated by a water molecule. Although Asp 57 is not located at the interaction surface but at least 7 Å away, it lies on the normal vector through the center of the interaction surface. The negative charge of this residue has therefore a combined stabilizing influence on binding of all the positive charges of Raf that are located at the interaction surface. This is an example of how buried residues can influence binding. A similar example is provided by the interaction of Glu62, especially with Raf-Arg59 more than 8 Å away. Although the absolute value of those contributions can be wrong by as much as a factor of two, this analysis shows that not only residues that form the

contact surface between the proteins are responsible for the binding force but also those that are buried in the protein matrix.

CONCLUDING REMARKS

The evaluation of binding free energies of protein–protein complexes presents a significant challenge to current computer modeling approaches. This challenge is addressed here by using the PDLD/S approach combined with the LRA treatment. This is done while considering different strategies with alternative thermodynamic cycles and different treatments of the protein reorganization. The Raf-RBD/Rap complex is taken as a test case, thus combining the somewhat formal computational study with a preliminary analysis of a system with significant biological importance.

It is found that the calculations can produce the absolute binding energy of the native complex and several of its mutants on a qualitative level. This is quite encouraging since such calculations are much more challenging than the more standard calculations of the effect of mutations on binding processes that are usually done by indirect thermodynamic cycles.[2] This is particularly true in cases of complexation of two proteins with many ionizable surface groups where the final binding free energy reflects compensation of vary large opposing contributions. The ability of different approaches to reproduce the observed effect of ionized residues depends on the treatment of the relaxation of the solvent and the protein and the value used for ϵ_{in}.

The most rigorous approach that involves global protein relaxation reproduces the observed trend of some mutations with $\epsilon_{in} = 4$, but the effect of other mutations can only be reproduced with $\epsilon_{in} = 25$. Considering explicitly only the relaxation of the solvent gives the best results with ϵ_{in} somewhat larger than 25. Apparently, the use of the global relaxation approach involves significant instabilities which reflect the difficulties of reproducing major protein reorganization by MD simulations. Nevertheless, the use of this approach indicates that explicit consideration of protein reorganization should allow one to use small ϵ_{in}. As stated before, ϵ_{in} does not represent the actual local dielectric of the protein, but merely the factors that are not treated explicitly. Thus, the need to use a large ϵ_{in} in most of our treatments reflects in part the difficulties in accurately reproducing the protein reorganization upon complexation by MD simulations. Here it will be very beneficial to have the crystal structures of the bound and unbound proteins (from which the latter are not yet available) or examine systems where such structures are available, but this is out of the scope of the present work. Another possibility is that some of the protein ionizable groups change their ionization states during the dissociation process. As stated in the Method section, we kept ionized groups at their ionization states in water at pH = 7. We could have calculated the actual pK_a in the protein[21] at each form of the protein, but such calculations are again out of the scope of the present work. At any rate, the implicit treatment of the possible change in ionization state also accounts in part for the need of using large ϵ_{in}. Interestingly, it is found that the effects of mutations of polar residues is best reproduced with $\epsilon_{in} = 4$. This result is instructive since the recent suggestion of using large ϵ_{in} as a uniform dielectric constant[33] in discretized continuum calculations would lead to a drastic underestimation of the charge–dipole and dipole–dipole interactions involved in the contribution of polar residues to protein–protein interactions. On the other hand, the use of large ϵ_{in} exclusively for charge–charge interactions is consistent with our original suggestion.[32]

Regardless of the above difficulties of obtaining quantitative estimates of protein–protein interactions with small ϵ_{in}, we establish here a practical way of evaluating the effects of mutation on such interactions, using $\epsilon_{in} \approx 25$ for mutations of ionized residues and $\epsilon_{in} \approx 4$ for mutations of polar residues. Furthermore, we also established that the results of our expensive LRA approach can be approximated quite effectively using the nonrelaxed group contributions, considering only the average of the vacuum Coulomb interactions with $\epsilon_{in} \approx 25$ for charged residues and $\epsilon_{in} \approx 4$ for polar residues. This very simple approach provides quite a useful way of correlating electrostatic energies with biological function. In particular, such an approach can provide an electrostatic fingerprint that locates these residues that are responsible for binding.

ACKNOWLEDGMENTS

We would like to thank A. Wittinghofer for providing the crystal structure of the Raf-RBD/Rap complex. This work has been supported in part by DOE Grant DE-FG03-94ER61945 and the Tobacco research grant 4RT-0002. I.M. gratefully acknowledges support by a fellowship of the Deutsche Forschungsgemeinschaft. T.S. greatly acknowledges support by an Otto-Hahn-Fellowship of the Max-Planck-Gesellschaft.

REFERENCES

1. Chothia, C., Janin, J. Principles of protein–protein recognition. Nature 256:705–708, 1975.
2. Kollman, P. Free-energy calculations—applications to chemical and biological phenomena. Chem. Rev. 7:2395–2417, 1993.
3. Lee, F.S., Chu, Z.-T., Warshel, A. A microscopic and semimicroscopic calculations of electrostatic energies in proteins by the POLARIS and ENZYMIX programs. J. Comp. Chem. 14:161–185, 1993.
4. Åqvist, J., Medina, C., Samuelsson, J.E. New method for predicting binding-affinity in computer-aided drug design. Protein Engineering 7:385–391, 1994.
5. Straatsma, T.P., McCammon, J.A. Computational Alchemy. Ann. Rev. Phys. Chem. 43:407–435, 1992.
6. Lee, F.S., Chu, Z.-T., Bolger, M.B., Warshel, A. Calculations

of antibody–antigen interactions: Microscopic and semi-microscopic evaluation of free energies of binding of phosphorylcholine analogs to McPC603. Protein Engineering 5:215–228, 1992.
7. Janin, J., Chothia, C. The structure of protein–protein recognition sites. J. Biol. Chem. 265:16027–16030, 1990.
8. Janin, J. Elusive affinities. Proteins 21:30–39, 1995.
9. Warshel, A. Electrostatic basis of structure-function correlation in proteins. Acc. Chem. Res. 14:284–290, 1981.
10. Warshel, A. "Computer Modeling of Chemical Reactions in Enzymes and Solutions." New York: John Wiley & Sons, Inc., 1991.
11. Warshel, A., Åqvist, J. Electrostatic energy and macromolecular function. Ann. Rev. Biophys. Biophys. Chem. 20:267–298, 1991.
12. van Gunsteren, W.F., Mark, A.E. On the interpretation of biochemical data by molecular dynamics simulation. Eur. J. Biochem. 204:947–961, 1992.
13. Muegge, I., Qi, X.P., Wand, A.J., Chu, Z.T., Warshel, A. The reorganization energy of cytochrome C revisited. J. Phys. Chem. B 101:825–836, 1997.
14. Sharp, K.A., Honig, B. Electrostatic interactions in macromolecules. Ann. Rev. Biophys. Biophys. Chem. 19:301–332, 1990.
15. Gilson, M., Davis, M., Luty, B., McCammon, J. Computation of electrostatic forces on solvated molecules using the Poisson-Boltzmann equation. J. Phys. Chem. 97:3591–3600, 1993.
16. Sharp, K.A. Electrostatic interactions in macromolecules. Current Opinion in Structural Biology 4:234–239, 1994.
17. Nassar, N., Horn, G., Herrmann, C., Scherer, A., McCormick, F., Wittinghofer, A. The 2.2 Å crystal structure of the Ras-binding domain of the serine/threonine kinase c-Rafl in complex with Rap1A and a GTP analogue. Nature 375:554–560, 1995.
18. Barbacid, M. *ras* genes. Ann. Rev. Biochem. 56:779–827, 1987.
19. Warshel, A. Calculations of enzymic reactions: Calculations of pK_a, proton transfer reactions, and general acid catalysis reactions in enzymes. Biochemistry 20:3167, 1981.
20. Wong, C.F., McCammon, J.A. Dynamics and design of enzymes and inhibitors. J. Am. Chem. Soc. 108:3830–3832, 1986.
21. Sham, Y.Y., Chu, C.T., Warshel, A. Consistent calculations of pK_a's of ionizable residues in proteins: Semi-microscopic and microscopic approaches. J. Phys. Chem. B 101:4458, 1997.
22. Warshel, A. Dynamics of reactions in polar solvents. Semiclassical trajectory studies of electron transfer and proton transfer reactions. J. Phys. Chem. 86:2218, 1982.
23. Hwang, J.K., Warshel, A. Microscopic examination of free-energy relationships for electron transfer in polar solvents. J. Am. Chem. Soc. 109:715–720, 1987.
24. Kuharski, R.A., Bader, J.S., Chandler, D., Spirk, M., Klein, M.L., Impey, R.W. Molecular model for aqueous ferrous–ferric electron transfer. J. Chem. Phys. 89:3248–3257, 1988.
25. King, G., Warshel, A. Investigation of the free energy function for electron transfer reactions. J. Chem. Phys. 93:8682–8692, 1990.
26. Langen, R., Schweins, T., Warshel, A. On the mechanism of guanosine triphosphate hydrolysis in ras p21 proteins. Biochemistry 31:8691–8696, 1992.
27. Muegge, I., Schweins, T., Langen, R., Warshel, A. Electrostatic Control of GTP and GDP binding in the oncoprotein $p21^{ras}$. Structure 4:475–789, 1996.
28. Warshel, A., Russell, S.T. Calculations of electrostatic interactions in biological systems and in solutions. Q. Rev. Biophys. 17:283–422, 1984.
29. Lee, F.S., Warshel, A. A local reaction field method for fast evaluation of long-range electrostatic interactions in molecular simulations. J. Chem. Phys. 97:3100–3107, 1992.
30. Searle, M.S., Williams, D.H. The cost of conformational order: Entropy changes in molecular associations. J. Am. Chem. Soc. 114:10690–10697, 1992.
31. Searle, M.S., Williams, D.H., Gerhard, U. Partitioning of free energy contributions in the estimation of binding constants: Residual motions and consequences for amide-amide hydrogen bond strength. J. Am. Chem. Soc. 114: 10697–10704, 1992.
32. Warshel, A., Russell, S.T., Churg, A.K. Macroscopic models for studies of electrostatic interactions in proteins: Limitations and applicability. Proc. Natl. Acad. Sci. USA 81:4785–4789, 1984.
33. Antosiewicz, J., McCammon, J.A., Gilson, M. Prediction of pH-dependent properties of proteins. J. Mol. Biol. 238:415–436, 1994.
34. Nassar, N., Horn, G., Herrmann, C., Block, C., Janknecht, R., Wittinghofer, A. Ras/Rap effector specificity determined by charge reversal. Nature Struct. Biol. 3:723–729, 1996.
35. Herrmann, C., Martin, G.A., Wittinghofer, A. Quantitative analysis of the complex between $p21^{ras}$ and the Ras-binding-domain of the human Raf-1 protein kinase. J. Biol. Chem. 270:2901–2905, 1995.
36. Block, C., Janknecht, R., Herrmann, C., Nassar, N., Wittinghofer, A. Quantitative structure-activity analysis correlating Ras/Raf interaction in vitro to Raf activation in vivo. Nature Struct. Biol. 3:244–251, 1996.