Optimization of binding electrostatics: Charge complementarity in the barnase-barstar protein complex

LEE-PENG LEE^{1,2} AND BRUCE TIDOR¹

¹Department of Chemistry, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139-4307, USA

²Department of Physics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139-4307, USA

(RECEIVED September 19, 2000; ACCEPTED November 17, 2000)

Abstract

Theoretical and experimental studies have shown that the large desolvation penalty required for polar and charged groups frequently precludes their involvement in electrostatic interactions that contribute strongly to net stability in the folding or binding of proteins in aqueous solution near room temperature. We have previously developed a theoretical framework for computing optimized electrostatic interactions and illustrated use of the algorithm with simplified geometries. Given a receptor and model assumptions, the method computes the ligand-charge distribution that provides the most favorable balance of desolvation and interaction effects on binding. In this paper the method has been extended to treat complexes using actual molecular shapes. The barnase-barstar protein complex was investigated with barnase treated as a target receptor. The atomic point charges of barstar were varied to optimize the electrostatic binding free energy. Barnase and natural barstar form a tight complex ($K_d \sim 10^{-14}$ M) with many charged and polar groups near the interface that make this a particularly relevant system for investigating the role of electrostatic effects on binding. The results show that sets of barstar charges (resulting from optimization with different constraints) can be found that give rise to relatively large predicted improvements in electrostatic binding free energy. Principles for enhancing the effect of electrostatic interactions in molecular binding in aqueous environments are discussed in light of the optima. Our findings suggest that, in general, the enhancements in electrostatic binding free energy resulting from modification of polar and charged groups can be substantial. Moreover, a recently proposed definition of electrostatic complementarity is shown to be a useful tool for examining binding interfaces. Finally, calculational results suggest that wild-type barstar is closer to being affinity optimized than is barnase for their mutual binding, consistent with the known roles of these proteins.

Keywords: Electrostatic optimization; molecular recognition; protein electrostatics; binding affinity

Because of the large desolvation penalty incurred by polar and charged groups on protein folding or binding in aqueous solution, theoretical and experimental studies have found electrostatic interactions to be net destabilizing in a variety of contexts near room temperature (Hendsch and Tidor

1994; Waldburger et al. 1995; Wimley et al. 1996; Yang and Honig 1995; Wang et al. 1996; Misra et al. 1994a,b; Sharp 1996; Shen and Wendoloski 1996; Brucoleri et al. 1997; Novotny et al. 1997; Froloff et al. 1997; Hendsch and Tidor 1999). These interactions form because the hydrophobic effect is strongly stabilizing, outweighing the electrostatic contribution that is computed to be unfavorable and often large. Calculations show that, whereas salt bridges and hydrogen bonds formed across binding interfaces contribute a favorable *interaction* term to binding free energy, this is often outweighed by the *desolvation* term (because of the

Reprint requests to: Dr. Bruce Tidor, Department of Chemistry, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139-4307, USA; e-mail: tidor@mit.edu; fax: (617) 252-1816.

Article and publication are at www.proteinscience.org/cgi/doi/10.1110/ps.40001.

loss of electrostatic interaction of charged and polar groups with water and solvent ions on binding). Electrostatic interactions are believed to enhance specificity, however, because of the large penalty for burying but not compensating polar and charged groups (Tanford et al. 1960; Paul 1982; Hendsch and Tidor 1994; Sindelar et al. 1998).

The unfavorable nature of protein electrostatics is intriguing. Is this a consequence of physics or a choice of biology? In other words, is it possible for two appropriately designed molecular binding partners to have favorable intermolecular electrostatic interactions that compensate and surpass the desolvation penalty? We have developed a theoretical framework for use with the continuum electrostatic model that carries out electrostatic optimization. The charge distribution for a ligand is optimized to give the most favorable balance of these competing factors on binding a given receptor (Lee and Tidor 1997; Kangas and Tidor 1998, 1999; Chong et al. 1998). Electrostatic optimization, which is applicable beyond simple continuum models to general linear response models, is of special interest for three reasons. First, the ability to obtain optimal charge distributions provides a useful point of reference for understanding natural charge distributions as well as the effects of mutations. Second, optimization theory leads to an important measure of electrostatic complementarity shown here (Kangas and Tidor 1998). Third, it is likely that optimal-ligand computations will be a powerful guide to enhancing affinity through mutant and inhibitor design.

In previous work, the electrostatic optimization scheme was applied to a range of model receptors and complexes with idealized geometries of spheres and slabs (Lee and Tidor 1997; Kangas and Tidor 1998; Chong et al. 1998), including a spherical model for barnase (Chong et al. 1998). Results have shown that when the ligand-charge distribution is optimized, the electrostatic binding free energy can be favorable with reasonable values for point-charge magnitudes (Kangas and Tidor 1998; Chong et al. 1998). In fact, a proof has been completed showing, with a few simplifying assumptions, that the optimization produces favorable electrostatics of binding at zero ionic strength (Kangas and Tidor 1999). Taken together, these observations suggest that it might be possible to improve electrostatic interactions in natural complexes.

We have extended the electrostatic optimization method to treat the actual shapes of real molecules. Here we apply this optimization method to study the barnase-barstar complex, an extremely tight-binding pair with many polar and charged groups at the binding interface and a K_d of 10^{-14} M (Schreiber and Fersht 1993; Hartley 1993). Even so, the electrostatic binding free energy was calculated to be somewhat unfavorable at 14 kcal/mole (Chong et al., in prep.). The hydrophobic contribution is sufficient to overcome this term and drive binding. Most other complexes for which the calculations have been made show even more unfavorable

binding electrostatics, consistent with the observation that this is an especially tight complex. In this study, barnase was chosen to be the fixed receptor. Barstar atomic point charges were treated as variable and optimized to produce the net most favorable binding to barnase. The results show that sets of barstar charges (resulting from optimization with different constraints) can be found that give rise to computed improvements of 10 to 20 kcal/mole in electrostatic binding free energy.

In this paper we give a detailed description of our current electrostatic optimization method, including molecular shape, and investigate the nature of the optimal charge distributions. We extract principles for enhancing electrostatic interactions in molecular binding in aqueous solution. Moreover, we show the utility of a new measure of electrostatic complementarity (Kangas and Tidor 1998), and we also examine the evolutionary question of whether barnase (the enzyme) is more optimized to bind to barstar (the inhibitor), or whether the inhibitor is more optimized to bind the enzyme. In another paper we show that, although charge distributions can be found that give much higher binding affinity than barstar, with the constraint that limits residues to the 20 common amino acids, barstar is actually very well optimized (Lee and Tidor 2001). The implication of these findings for molecular design is exciting: Because laboratory chemistry is not limited to amino acids, novel molecules might be designed with substantially improved affinity.

Methods

In the continuum electrostatic model used here, a macromolecule is treated as a low dielectric ($\epsilon = 4$) cavity, defined by its molecular surface, embedded in a high dielectric ($\epsilon = 80$) continuum solvent at an ionic strength near physiological levels. The charge distribution of the macromolecule is a set of point partial charges, $\{Q_i\}$, at the atom centers, $\{\vec{r}_i\}$. In cases in which the potential is derived from the Poisson or the linearized Poisson-Boltzmann (PB) equation, the electrostatic free energy is $G = \frac{1}{2}\sum Q_i\Phi(\vec{r}_i)$, in which $\Phi(\vec{r}_i)$ is the potential at the position of the charge Q_i (Sharp and Honig 1990a). This free energy includes entropic contributions from the solvent, including ion reorganization in the case of the linearized Poisson-Boltzmann equation (Sharp and Honig 1990a; Overbeck 1990; Fogolari and Briggs 1997; Roux and Simonson 1999). The factor of $\frac{1}{2}$ exists because of the fact that the free energy arises from the interaction between the charges and their self-induced reaction field. In many cases of interest, $\Phi(\vec{r})$ is obtained by solving the linearized Poisson-Boltzmann equation,

$$\vec{\nabla} \cdot (\epsilon(\vec{r})\vec{\nabla}\Phi(\vec{r})) - \kappa^2(\vec{r})\Phi(\vec{r}) = -4\pi\sum_j Q_j\delta(\vec{r} - \vec{r}_j) \quad (1)$$

in which $\kappa^2(\vec{r}) = 8\pi e^2 I(\vec{r})/kT$ with $I(\vec{r})$ the bulk ionic strength, e the unit charge (magnitude of the charge of an electron), k the Boltzmann constant, T the absolute temperature, and $\epsilon(\vec{r})$ the spatially dependent dielectric determined by the molecular surface formed by rolling a probe water molecule of radius 1.4 Å over the macromolecule (Richards 1977; Connolly 1983). In cases in which

the molecular boundary is a simple shape (e.g., spheres, infinite cylinders, spheroids, and slabs) analytical solutions exist. Otherwise, the linearized PB equation is solved numerically. Here we used the actual molecular shapes of barnase and barstar, and numerically computed $\Phi(\vec{r})$ from $\{Q_i\}$, $\{\vec{r}_i\}$, $\kappa(\vec{r})$, and $\epsilon(\vec{r})$. The radii and charges were taken from the CHARMM PARAM19 parameter set (Brooks et al. 1983). A 2 Å Stern layer surrounded the molecule and represents an ion-exclusion region (Gilson and Honig 1987; Bockris and Reddy 1973) and the salt concentration was 0.145 M. Each calculation was done using a focusing technique, in which a low grid spacing calculation using 23% fill and Debye-Hückel boundary conditions (Klapper et al. 1986) was done to determine the potential at the grid boundary for a higher grid spacing calculation (using 92% fill). Ten translations relative to the grid (65 points in each Cartesian direction) were done and the average was used.

The electrostatic free energy of binding is the difference between the electrostatic free energy in the bound and the unbound states,

$$\Delta G_{\text{bind}} = G^{\text{bound}} - G^{\text{unbound}}. \quad (2)$$

Because the unbound and bound structures are very similar (Guillet et al. 1993; Buckley et al. 1994; Ratnaparkhi et al. 1998; Martin et al. 1999), we use a rigid binding model in which the bound state is taken from the X-ray crystal structure of the complex (C) and the unbound state is simply extracted from the complex of the ligand (L, barstar) and the receptor (R, barnase) (Fig. 1). A cap of 12 water molecules from the complex that appears more tightly associated with barnase is treated as part of the re-

ceptor in the bound and unbound states. A detailed description of the preparation of this structure, including the building of hydrogen atom positions, is given elsewhere (Chong et al., in prep.).

The free energy of the unbound state is given by,

$$G^{\text{unbound}} = \frac{1}{2} \sum_{i=1}^{n_L} Q_i^L \Phi^L(\vec{r}_i^L) + \frac{1}{2} \sum_{i=1}^{n_R} Q_i^R \Phi^R(\vec{r}_i^R) \quad (3)$$

in which $\Phi^L(\vec{r})$ is the potential from ligand charges in the unbound state, $\Phi^R(\vec{r})$ is the potential from receptor charges in the unbound state, and n_L and n_R are the number of partial atomic charges in the ligand and receptor, respectively.

The free energy of the bound state is given by,

$$G^{\text{bound}} = \frac{1}{2} \sum_{i=1}^{n_L} Q_i^L \Phi^C(\vec{r}_i^L) + \frac{1}{2} \sum_{i=1}^{n_R} Q_i^R \Phi^C(\vec{r}_i^R) \quad (4)$$

in which $\Phi^C(\vec{r})$ is the potential from both ligand and receptor charges in the complex. Because of the linearity of equation 1, the superposition principle holds and we can write $\Phi^C(\vec{r})$ as a sum of $\Phi^{CL}(\vec{r})$, the potential from ligand charges in the complex, and $\Phi^{CR}(\vec{r})$, the potential from receptor charges in the complex:

$$G^{\text{bound}} = \frac{1}{2} \sum_{i=1}^{n_L} Q_i^L (\Phi^{CL}(\vec{r}_i^L) + \Phi^{CR}(\vec{r}_i^L)) + \frac{1}{2} \sum_{i=1}^{n_R} Q_i^R (\Phi^{CL}(\vec{r}_i^R) + \Phi^{CR}(\vec{r}_i^R)). \quad (5)$$

Fig. 1. Ribbon drawing of the bound complex of barnase (blue) and barstar (yellow), including 12 interfacial water molecules (gray). Side chains found to be important for binding in mutagenesis experiments are highlighted (Schreiber and Fersht 1993, 1995; Schreiber et al. 1994; Hartley 1993). Figures 1, 2, and 3 were made with MOLSCRIPT (Kraulis 1991) and RASTER3D (Merritt and Bacon 1997).

We may rearrange these terms and substitute into equation 2 to obtain

$$\Delta G_{\text{bind}} = \underbrace{\frac{1}{2} \sum_{i=1}^{n_L} Q_i^L (\Phi^{\text{CL}}(\vec{r}_i^L) - \Phi^L(\vec{r}_i^L))}_{\Delta G_{\text{desolv,L}}} + \underbrace{\frac{1}{2} \left(\sum_{i=1}^{n_L} Q_i^L \Phi^{\text{CR}}(\vec{r}_i^L) + \sum_{i=1}^{n_R} Q_i^R \Phi^{\text{CL}}(\vec{r}_i^R) \right)}_{\Delta G_{\text{inter,L-R}}} + \underbrace{\frac{1}{2} \sum_{i=1}^{n_R} Q_i^R (\Phi^{\text{CR}}(\vec{r}_i^R) - \Phi^R(\vec{r}_i^R))}_{\Delta G_{\text{desolv,R}}} \quad (6)$$

In electrostatic optimization, the ligand-charge distribution is varied to optimize ΔG_{bind} (Lee and Tidor 1997; Kangas and Tidor 1998). Here this is achieved by varying the individual Q_i^L values. The potentials $\Phi^{\text{CR}}(\vec{r})$ and $\Phi^R(\vec{r})$, from the receptor charges, are computed numerically with a modified version of the computer program DELPHI (Sharp and Honig 1990a,b; Gilson et al. 1988). $\Phi^{\text{CL}}(\vec{r})$ and $\Phi^L(\vec{r})$, from the unknown ligand charges, can be expressed in terms of $\{Q_i^L\}$ and a set of potentials calculated numerically, as we now show. Using superposition, we decompose $\Phi(\vec{r})$ as (Kangas and Tidor 1998),

$$\Phi(\vec{r}) = \sum_j Q_j \hat{\Phi}_j(\vec{r}) \quad (7)$$

where $\hat{\Phi}(\vec{r})$ satisfies

$$\vec{\nabla} \cdot (\epsilon(\vec{r}) \vec{\nabla} \hat{\Phi}_j(\vec{r})) - \bar{\kappa}^2(\vec{r}) \hat{\Phi}_j(\vec{r}) = -4\pi\delta(\vec{r} - \vec{r}_j). \quad (8)$$

$\hat{\Phi}_j$ is then the potential from a unit charge at the position $\vec{r}_j$ (i.e., a Green function). We can now write the desolvation penalty of the ligand as

$$\Delta G_{\text{desolv,L}} = \frac{1}{2} \sum_{i=1}^{n_L} \sum_{j=1}^{n_L} Q_i^L Q_j^L (\hat{\Phi}_j^{\text{CL}}(\vec{r}_i^L) - \hat{\Phi}_j^L(\vec{r}_i^L)). \quad (9)$$

Furthermore, because of the principle of reciprocity, the two terms in the intermolecular interaction free energy in equation 6 are equal and we have

$$\Delta G_{\text{inter,L-R}} = \sum_{i=1}^{n_L} Q_i^L \Phi^{\text{CR}}(\vec{r}_i^L). \quad (10)$$

We have expressed ΔG_{bind} as a quadratic function of $\{Q_i\}$. In vector notation,

$$\Delta G_{\text{bind}} = \frac{1}{2} (\vec{Q}^L)^T \vec{A} \vec{Q}^L + (\vec{Q}^L)^T \vec{B} + C \quad (11)$$

where

$$A_{ij} = \hat{\Phi}_j^{\text{CL}}(\vec{r}_i^L) - \hat{\Phi}_j^L(\vec{r}_i^L) \quad (12)$$

$$B_i = \Phi^{\text{CR}}(\vec{r}_i^L) \quad (13)$$

and

$$C = \frac{1}{2} \sum_{i=1}^{n_R} Q_i^R (\Phi^{\text{CR}}(\vec{r}_i^R) - \Phi^R(\vec{r}_i^R)). \quad (14)$$

Equation 11 describes a paraboloid possessing a minimum corresponding to the optimized electrostatic binding free energy (Lee and Tidor 1997).

Allowing all ligand charges to change is a special case of the more general problem of leaving a subset of ligand charges fixed and the remainder variable. For instance, in this study we fix the backbone charges, and in a companion study we fix all charges except for a single side chain (Lee and Tidor 2001). Let $\{Q_i^{\text{L}^v}\}$ denote the set of variable ligand charges and $\{Q_i^{\text{L}^f}\}$ denote the set of fixed ligand charges. It is straightforward to show that ΔG_{bind} can be written analogously as a quadratic function of $\{Q_i^{\text{L}^v}\}$,

$$\Delta G_{\text{bind}} = \frac{1}{2} (\vec{Q}^{\text{L}^v})^T \vec{A} \vec{Q}^{\text{L}^v} + (\vec{Q}^{\text{L}^v})^T \vec{B} + C \quad (15)$$

where

$$A_{ij} = \hat{\Phi}_j^{\text{CL}}(\vec{r}_i^{\text{L}^v}) - \hat{\Phi}_j^L(\vec{r}_i^{\text{L}^v}) \quad (16)$$

$$B_i = \Phi^{\text{CR}}(\vec{r}_i^{\text{L}^v}) + \sum_{j=1}^{n_{\text{L}^f}} (\hat{\Phi}_j^{\text{CL}}(\vec{r}_i^{\text{L}^v}) - \hat{\Phi}_j^L(\vec{r}_i^{\text{L}^v})) Q_j^{\text{L}^f} \quad (17)$$

and

$$C = \frac{1}{2} \sum_{i=1}^{n_R} Q_i^R (\Phi^{\text{CR}}(\vec{r}_i^R) - \Phi^R(\vec{r}_i^R)) + \frac{1}{2} \sum_{i=1}^{n_{\text{L}^f}} Q_i^{\text{L}^f} (\Phi^{\text{CL}}(\vec{r}_i^{\text{L}^f}) - \Phi^L(\vec{r}_i^{\text{L}^f})) + \frac{1}{2} \sum_{i=1}^{n_{\text{L}^f}} Q_i^{\text{L}^f} \Phi^{\text{CR}}(\vec{r}_i^{\text{L}^f}) \quad (18)$$

Equations 16, 17, and 18 reduce to equations 12, 13, and 14 when $\{Q_i^{\text{L}^v}\}$ is $\{Q_i^L\}$, and $\{Q_i^{\text{L}^f}\}$ is the empty set.

In theory, the ligand desolvation matrix, $\vec{A}$, is symmetric (because of reciprocity) and positive definite (Kangas and Tidor 1998). All eigenvalues should be positive and $\Delta G_{\text{desolv,L}}$ should be positive (a desolvation penalty). ΔG_{bind} is a multidimensional paraboloid with a minimum $\Delta G_{\text{bind}}^{\text{opt}}$ at

$$\vec{Q}^{\text{opt}} = -\vec{A}^{-1} \vec{B}. \quad (19)$$

In practice, because of numerical errors inherent in the calculation, it is possible that $\vec{A}$ as computed is nonsymmetric; in such cases we symmetrize the matrix by averaging across the diagonal. The more problematic consequence of limited numerical accuracy is that $\vec{A}$ as computed might not be positive definite, which means the extremum computed is not a minimum but a saddle point. The size of the negative eigenvalues relative to positive eigenvalues is some indication of the extent of numerical inaccuracy. Here, there are negative eigenvalues that are quite small, the largest magnitude being 10^{-5} of the largest positive eigenvalue. Our remedy for this numerical problem is to use singular value decomposition (Strang 1993; Press et al. 1992). We construct the optimal ligand charges only from eigenvectors with eigenvalues greater than the absolute value of the most negative eigenvalue. To eliminate eigenvectors

with small eigenvalues, equation 19 must be used to reexpress $\vec{Q}^{\text{opt}}$ as a sum over eigenvectors of $\vec{A}$. Writing $\vec{A}$ as $\vec{V}\vec{\Lambda}\vec{V}^{-1}$, in which the columns of $\vec{V}$ are the n_{LV} independent eigenvectors $\{\vec{v}_i\}$ of $\vec{A}$ and $\vec{\Lambda}$ is a diagonal matrix with the corresponding eigenvalues $\{\lambda_i\}$ as its diagonal entries (Strang 1993), then

$$\vec{Q}^{\text{opt}} = (\vec{V}\vec{\Lambda}\vec{V}^{-1})^{-1}\vec{B} = \vec{V}\vec{\Lambda}^{-1}\vec{V}^{-1}\vec{B}. \quad (20)$$

With simple manipulations one can show that

$$\vec{Q}^{\text{opt}} = \sum_{i=1}^{n_{\text{LV}}} \beta_i \vec{v}_i \quad (21)$$

in which $\beta_i = (1/\lambda_i)D_i$ and $\vec{D} = \vec{V}^{-1}\vec{B}$. The optimized charge distribution is then constructed as a sum over the desired eigenvectors, indicated by Σ' ,

$$\vec{Q}^{\text{opt}} = \sum_{i=1}^{n_{\text{LV}}} \beta_i \vec{v}_i. \quad (22)$$

It is often useful to impose constraints on the variable charges $\{Q_i^{\text{LV}}\}$ during optimization. For instance, the size of the charges could be constrained to be chemically reasonable. In this study we limited individual charge values to between -0.85 and $+0.85$ (units are the magnitude of an electron charge). For some of the calculations we constrained the total ligand charge and the total charge of each side chain. These constraints are all linear in $\{Q_i^{\text{LV}}\}$, and because our objective function (i.e., the function to be optimized), ΔG_{bind} , is quadratic in $\{Q_i^{\text{LV}}\}$, the optimization problem is a quadratic programming problem, which has been extensively studied by the operations research community. We use the computer program package LOQO to solve this constrained optimization problem (Vanderbei 1997a,b, 1998). If $\vec{A}$ were positive definite, we would just use equation 15 as the objective function with the various linear constraints. However, with the removal of some eigenvectors, we have to rewrite the objective function. The variable charges $\{Q_i^{\text{LV}}\}$ are a linear combination of the eigenvectors that remain,

$$\vec{Q}^{\text{LV}} = \sum_{i=1}^{n_{\text{LV}}} c_i \beta_i \vec{v}_i. \quad (23)$$

With no constraints, all instances of c_i are 1 for the optimal charge set $\{Q_i^{\text{opt}}\}$. When constraints are imposed, the optimum achievable is worse than $\{Q_i^{\text{opt}}\}$ and the free energy deviation of $\vec{Q}$ from the unconstrained $\Delta G_{\text{bind}}^{\text{opt}}$ is (Lee and Tidor 1997),

$$\Delta \Delta G = \frac{1}{2} (\vec{Q} - \vec{Q}^{\text{opt}})^T \vec{A} (\vec{Q} - \vec{Q}^{\text{opt}}) \quad (24)$$

$$= \frac{1}{2} \sum_{i=1}^{n_{\text{LV}}} (1 - c_i) \beta_i \vec{v}_i^T \vec{A} (1 - c_i) \beta_i \vec{v}_i \quad (25)$$

$$= \frac{1}{2} \sum_{i=1}^{n_{\text{LV}}} (1 - c_i)^2 E_i \quad (26)$$

where

$$E_i = \beta_i \vec{v}_i^T \vec{A} \beta_i \vec{v}_i. \quad (27)$$

$\Delta \Delta G$ is the objective function. On optimization, a set of $\{c_i^{\text{opt}}\}$ are found such that $\Delta \Delta G$ is minimized and the optimum charge distribution is

$$\vec{Q}^{\text{opt}} = \sum_{i=1}^{n_{\text{LV}}} c_i^{\text{opt}} \beta_i \vec{v}_i. \quad (28)$$

Results and Discussion

Optimal charge distributions have enhanced binding electrostatics

The electrostatic optimization procedure was applied to selected sets of barstar atomic charges. Atomic positions and the dielectric boundary for barstar were taken from the X-ray crystal structure of the complex as were the geometry and point-charge distribution for barnase (Buckle et al. 1994). Optimizations were carried out with different sets of constraints on the atomic point charges, and the results are listed in Table 1. When actual (unoptimized) barstar point charges were used ($\{Q_i^{\text{wt}}\}$), the electrostatic binding free energy was computed to be unfavorable by about 14 kcal/mole. This reflects the observation made in a number of studies that most complexes pay more in electrostatic dehydration penalty than is recovered in attractive electrostatic interactions across the interface (Hendsch and Tidor 1994; Waldburger et al. 1995; Wimley et al. 1996; Yang and Honig 1995; Wang et al. 1996; Misra et al. 1994a,b; Sharp 1996; Shen and Wendoloski 1996; Brucoleri et al. 1997; Novotny et al. 1997; Froloff et al. 1997; Hendsch and Tidor 1999). The relatively small electrostatic free energy effect ($+14$ kcal/mole) for the relatively large buried surface area in the complex (803 and $878 \text{ \AA}^2$ from barnase and barstar, respectively) is consistent with this being a tight-binding pair. The hydrophobic effect is sufficient to overcome the unfavorable electrostatics as well as internal, translational, and rotational entropy to drive binding.

Table 1. Optimized charge distributions and energetics for barstar^a

Charge set	$\{Q_i^{\text{LV}}\}$	n_{LV}	Constraints	ΔG_{bind} (kcal/mol)
$\{Q_i^{\text{wt}}\}$	None	0	—	+14.2
$\{Q_i^{\text{opt1}}\}$	All barstar charges	839	$-0.85 \leq Q_i \leq 0.85$	-6.1
$\{Q_i^{\text{opt2}}\}$	Side chain charges	403	$-0.85 \leq Q_i \leq 0.85$	-0.6
$\{Q_i^{\text{opt3}}\}$	Side chain charges	403	$-0.85 \leq Q_i \leq 0.85$; $-1.5 \leq Q_n^{\text{sc}} \leq 1.5$	-0.1

^a $\{Q_i^{\text{wt}}\}$ is the unoptimized wild-type barstar charge distribution for comparison. The remaining three charge distributions represent optimizations with the indicated constraints. $\{Q_i^{\text{LV}}\}$ describes the set of variable ligand (barstar) charges in each optimization and n_{LV} is the number of charges in that set. $\{Q_n^{\text{sc}}\}$ is the total side-chain charge of the n th residue.

In all optimizations the value of each barstar atomic point charge was restricted to the range of -0.85 to $+0.85$. When this was the only constraint applied and all 839 barstar atomic charges in the model were optimized, the resulting charge distribution, ($\{Q_i^{\text{opt1}}\}$), had an electrostatic binding free energy of -6.1 kcal/mole. This value is favorable and represents a computed enhancement of about 20 kcal/mole over wild type. When the backbone charges were fixed at their standard values and only the 403 side chain atomic point charges were allowed to vary, the computed optimum ($\{Q_i^{\text{opt2}}\}$) had an electrostatic binding free energy of -0.6 kcal/mole, which is 5.5 kcal/mole less stable than ($\{Q_i^{\text{opt1}}\}$) but is still favorable. When the further constraint was added that the total charge magnitude of each side chain was <1.5 , the resulting charge distribution, ($\{Q_i^{\text{opt3}}\}$), had an electrostatic binding free energy of -0.1 kcal/mole. These results show that charge optimization produces atomic charge distributions with reasonable magnitudes and apparently extremely tight binding. Moreover, they show that electrostatics can, in principle, have a net favorable effect on binding (Kangas and Tidor 1999). It is of interest to examine these charge distributions in greater detail to understand what aspects are important and by what mechanisms electrostatic enhancements to binding might be achieved.

Optimization produces enhanced intramolecular and intermolecular effects on binding

Of fundamental interest are the mechanisms used by optimized ligand-charge distributions to improve the balance of desolvation and interaction effects on binding. To examine this, the linearity of the continuum treatment was used to dissect the energetics. There are three types of electrostatic terms that contribute to binding here: solvation, intermolecular (or direct), and intramolecular (or indirect) interactions (Hendsch and Tidor 1999; Chong et al. 1998; Caravella et al. 1999). Solvation represents the loss in electrostatic interaction of a functional group with solvent on binding, and intermolecular (or direct) terms represent screened coulombic interactions between a functional group and the binding partner in the bound state. Intramolecular (or indirect) terms represent electrostatic interactions between a functional group and other functional groups *within the same binding partner*. Even if the geometry of the interaction is unaltered on binding, an effect on the free energy of binding can be produced by differences in solvent screening in the bound and unbound states. Generally the effect is to strengthen intramolecular interactions on binding (be they favorable or unfavorable), and the effects are largest when at least one of the functional groups is at the binding interface. It is a striking result that these enhanced intramolecular interactions are computed to be significant in many natural systems, and there is some experimental evi-

dence for their importance as well (Hendsch and Tidor 1999).

The free energetic analysis for individual side chains of wild-type barstar and the side chain optimized charge distribution (opt3) binding to barnase is shown in Tables 2 and 3, respectively. The totals indicate that the optimum achieves binding enhancements through a larger dehydration penalty (ΔG_{solv} plus ΔG_{intra}) that is more than offset through a significantly larger improvement in intermolecular electrostatic interactions (ΔG_{inter}). Side chains making moderate and strong intermolecular interactions as well as those making enhanced intramolecular interactions (>1 kcal/mole) are indicated in the Tables and in Figure 2. The optimized charge distribution uses at least three mechanisms to improve binding. There are (i) enhanced intermolecular (direct) interactions (residues 37, 38, 42, 45, 76, and 77), (ii) reduced desolvation (35 and 39), and (iii) new favorable intramolecular (indirect) interactions (residues 41, 47, and 82). Strong direct intermolecular interactions were achieved by the optimum through the involvement of more side chains at the interface as well as through a small number of more distant electrostatic interactions. Side chains with enhanced intramolecular interactions generally lie just below the binding interface and make attractive electrostatic interactions with side chains at the interface. These interactions are largely screened by solvent in the unbound state but are enhanced on binding. Coupling of effects is present. For instance, on optimization, a number of side chains have worse intramolecular binding effects relative to wild type that are more than offset by improvements in desolvation or interaction.

Inventory of hydrogen-bonded contacts and ion pairs is only slightly increased through optimization

Examination of individual interactions across the interface for barstar and the ($\{Q_i^{\text{opt3}}\}$) optimized charge distribution reveals very similar types of interactions made in both (Tables 4 and 5). Decreased polarity causes the optimum to remove somewhat distant hydrogen bonds of Asn33 and Trp38 present in barstar. Increased polarity results in the strengthening of interactions in which a barstar side chain provides the hydrogen-bond acceptor. Finally, two new hydrogen-bond types of interaction are added. One is between Arg59 of barnase and Glu76 of barstar through redistributing charge such that all the atoms of the carboxylate bear a partial negative charge. The other is at most a weak hydrogen bond, between Wat29 (an interfacial water molecule) and Trp38 of barstar. Three ion pairing interactions remain similar in opt3, but are somewhat overpolarized relative to the wild type. Additionally, one new ion pair is added involving Arg59 of barnase and negative charge density added to Gln72 of barstar.

Table 2. *Component analysis for wild-type barstar binding barnase^a*

Side chain	ΔG_{solv}	ΔG_{intra}	ΔG_{inter}	ΔG_{total}	Side chain	ΔG_{solv}	ΔG_{intra}	ΔG_{inter}	ΔG_{total}
1	0.02	-0.20	0.41	0.24	46	0.24	0.13	-0.96	-0.59
2	0.00	-0.02	0.07	0.04	47	0.00	0.00	-0.08	-0.08
3	0.00	0.00	0.00	0.00	48	0.00	-0.02	0.04	0.03
4	0.00	0.00	0.00	0.00	49	0.00	0.00	0.00	0.00
5	0.00	0.00	0.00	0.00	50	0.00	0.00	0.00	0.00
6	0.00	0.00	0.00	0.00	51	0.00	0.00	0.00	0.00
7	0.00	0.00	0.00	0.00	52	0.00	0.02	-0.06	-0.04
8	0.00	0.01	-0.02	-0.01	53	0.00	0.00	-0.01	-0.01
9	0.00	0.00	0.00	0.00	54	0.00	-0.01	0.03	0.02
10	0.00	0.00	0.00	0.00	55	0.00	0.00	0.00	0.00
11	0.00	0.00	0.00	-0.01	56	0.00	0.00	0.00	0.00
12	0.00	0.00	0.00	0.00	57	0.00	0.01	-0.03	-0.02
13	0.00	0.00	0.00	0.00	58	0.00	0.00	0.01	0.00
14	0.00	0.00	0.00	0.00	59	0.00	0.01	-0.01	-0.01
15	0.00	0.02	-0.01	0.01	60	0.00	0.00	0.01	0.00
16	0.00	0.00	0.00	0.00	61	0.00	0.00	0.00	0.00
17	0.01	-0.04	0.26	0.23	62	0.00	0.00	0.00	0.00
18	0.00	0.00	0.00	0.00	63	0.00	0.00	0.00	0.00
19	0.00	-0.01	0.03	0.02	64	0.00	0.00	0.00	0.00
20	0.00	0.00	0.00	0.00	65	0.00	0.00	0.00	0.00
21	0.04	-0.12	-0.02	-0.10	66	0.00	0.00	0.00	0.00
22	0.00	-0.05	0.09	0.05	67	0.00	0.00	0.00	0.00
23	0.00	0.06	-0.15	-0.09	68	0.00	0.04	-0.14	-0.10
24	0.00	0.00	0.00	0.00	69	0.00	0.00	-0.04	-0.04
25	0.00	0.00	0.00	0.00	70	0.00	0.00	0.00	0.00
26	0.00	0.00	0.00	0.00	71	0.00	0.00	0.00	0.00
27	0.02	-0.20	0.47	0.29	72	0.01	-0.03	0.22	0.19
28	0.01	0.05	-0.02	0.04	73	0.00	0.00	0.00	0.00
29	0.96	0.70	-2.81	-1.15	74	0.00	0.00	0.00	0.00
30	0.01	-0.05	0.23	0.19	75	0.00	-0.06	0.20	0.14
31	0.00	0.00	0.00	0.00	76	<i>1.39</i>	0.26	-6.95	-5.30
32	0.09	-0.01	0.08	0.16	77	0.00	0.00	0.00	0.00
33	<i>1.65</i>	-0.47	-1.90	-0.72	78	0.00	-0.11	0.21	0.11
34	0.00	0.00	0.00	0.00	79	0.00	0.00	0.00	0.00
35	<i>13.30</i>	-2.88	-24.49	-14.08	80	0.93	0.55	-1.90	-0.42
36	0.00	0.00	0.00	0.00	81	0.00	0.00	0.00	0.00
37	0.00	0.00	0.00	0.00	82	0.00	0.00	0.00	0.00
38	0.83	-0.02	-0.78	0.03	83	0.01	0.13	-0.26	-0.12
39	<i>11.76</i>	<i>1.86</i>	-27.58	-13.96	84	0.00	0.00	0.00	0.00
40	0.00	0.00	0.00	0.00	85	0.00	-0.01	0.02	0.01
41	0.00	0.00	0.00	0.00	86	0.00	0.00	0.00	0.00
42	0.57	0.49	-3.15	-2.08	87	0.00	0.00	0.00	0.00
43	0.00	0.00	0.00	0.00	88	0.00	0.00	0.00	0.00
44	0.01	0.09	-0.17	-0.07	89	0.00	0.00	0.00	0.00
45	0.00	0.00	0.00	0.00	Total	31.85	0.11	-69.15	-37.20

^a All ΔG values are in kcal/mol. ΔG_{solv} is the desolvation penalty for the side chain atoms of the indicated residue, ΔG_{intra} is the intramolecular (indirect) contribution for the indicated side chain with the remainder of barstar, and ΔG_{inter} is the intermolecular (direct) electrostatic interaction between the indicated side chain and barnase. ΔG_{total} is the sum of the contributions. A barstar backbone contribution of -0.71 kcal/mol and a barnase dehydration contribution of 52.1 kcal/mol is to be added to the final ΔG_{total} to give the total electrostatic binding free energy. Bold and italic indicate favorable and unfavorable contributions, respectively, of magnitude exceeding 1 kcal/mol.

This enhanced binding affinity of the optimum relative to the wild type cannot simply be attributed to increases in the inventory of intermolecular interactions. The overpolarization and the new hydrogen bonds are estimated to contribute -5.8 kcal/mole and the new ion pair -1.4 kcal/mole. This is roughly half of the binding free energy improvement of opt3 over the wild type. The remaining -7.1 kcal/mole comes

from interactions from other side chains, which are not involved in short-range electrostatic interactions across the interface. Taken together, these results show that simple inventories of short-range interactions present in the bound state neglect important contributions to binding energetics that can be responsible for many orders of magnitude in a binding constant.

Table 3. Component analysis for an optimized charge distribution ($\{Q^{\text{opt3}}\})^a$

Side chain	ΔG_{solv}	ΔG_{intra}	ΔG_{inter}	ΔG_{total}	Side chain	ΔG_{solv}	ΔG_{intra}	ΔG_{inter}	ΔG_{total}
1	0.08	0.58	-1.22	-0.56	46	0.48	0.50	-1.96	-0.98
2	0.00	0.01	-0.01	-0.01	47	0.33	-1.97	2.75	<i>1.11</i>
3	0.00	0.04	-0.08	-0.04	48	0.04	0.24	-0.55	-0.26
4	0.00	0.00	-0.01	0.00	49	0.12	0.94	-1.97	-0.92
5	0.00	-0.02	0.05	0.03	50	0.00	0.01	-0.02	-0.01
6	0.00	0.00	0.00	0.00	51	0.00	0.07	-0.16	-0.09
7	0.00	0.00	0.00	0.00	52	0.00	0.00	0.00	0.00
8	0.00	0.00	0.00	0.00	53	0.01	-0.16	0.43	0.28
9	0.00	0.00	0.00	0.00	54	0.00	0.00	0.00	0.00
10	0.00	-0.01	0.03	0.03	55	0.00	0.00	0.00	0.00
11	0.00	0.00	0.00	0.00	56	0.00	-0.05	0.14	0.09
12	0.00	0.00	0.00	0.00	57	0.00	0.00	0.01	0.01
13	0.01	-0.01	0.04	0.04	58	0.00	0.00	0.00	0.00
14	0.05	-0.18	0.07	-0.05	59	0.00	-0.01	0.05	0.04
15	0.00	0.00	0.00	0.01	60	0.00	-0.01	0.04	0.03
16	0.02	-0.07	0.31	0.26	61	0.00	0.00	0.00	0.00
17	0.54	-0.04	0.62	<i>1.12</i>	62	0.00	0.00	0.00	0.00
18	0.03	0.03	0.00	0.06	63	0.01	0.03	0.01	0.04
19	0.00	-0.01	0.02	0.01	64	0.01	0.03	0.01	0.04
20	0.00	-0.07	0.20	0.14	65	0.01	0.03	0.01	0.04
21	0.10	0.08	-0.56	-0.38	66	0.00	0.00	0.00	0.00
22	0.01	0.07	-0.16	-0.08	67	0.00	-0.01	0.06	0.04
23	0.00	0.03	-0.08	-0.05	68	0.01	-0.08	0.22	0.14
24	0.05	0.52	-1.19	-0.63	69	0.17	-0.12	0.08	0.13
25	0.02	0.15	-0.28	-0.12	70	0.19	-0.37	<i>1.04</i>	0.86
26	0.03	-0.56	0.84	0.31	71	0.00	-0.05	0.12	0.07
27	0.13	-0.55	0.92	0.49	72	0.44	0.29	-1.99	-1.25
28	0.05	-0.05	-0.03	-0.04	73	0.26	-0.76	<i>1.02</i>	0.52
29	0.81	-0.02	-1.80	-1.01	74	0.10	0.76	-1.95	-1.09
30	0.49	-0.74	<i>1.34</i>	<i>1.09</i>	75	0.01	0.25	-0.57	-0.31
31	0.00	0.00	0.00	0.00	76	<i>4.92</i>	<i>1.01</i>	-12.30	-6.36
32	0.11	-0.19	0.07	-0.01	77	0.18	<i>1.19</i>	-2.73	-1.36
33	0.59	-0.70	-1.35	-1.46	78	0.00	-0.06	0.10	0.03
34	0.34	0.08	-1.57	-1.15	79	0.01	0.25	-0.53	-0.27
35	<i>11.93</i>	-2.42	-23.81	-14.30	80	0.33	0.91	-2.30	-1.06
36	0.00	0.02	0.01	0.03	81	0.00	0.00	0.00	0.00
37	0.93	0.47	-4.34	-2.94	82	0.17	-1.07	<i>1.46</i>	0.56
38	2.23	0.87	-6.67	-3.57	83	0.04	-0.39	0.53	0.18
39	<i>10.06</i>	2.20	-24.98	-12.72	84	0.00	0.05	-0.10	-0.05
40	0.04	0.37	-0.86	-0.46	85	0.00	-0.02	0.04	0.02
41	0.23	-1.92	<i>3.46</i>	<i>1.77</i>	86	0.00	0.05	-0.10	-0.05
42	2.59	<i>1.72</i>	-8.75	-4.44	87	0.00	0.00	0.00	0.00
43	0.00	0.00	0.00	0.00	88	0.00	-0.01	0.01	0.01
44	0.54	0.39	-2.05	-1.11	89	0.00	0.00	0.00	0.00
45	0.45	<i>1.78</i>	-4.08	-1.85	Total	40.29	3.27	-95.02	-51.46

^a All ΔG values are in kcal/mol. ΔG_{solv} is the desolvation penalty for the side chain atoms of the indicated residue, ΔG_{intra} is the intramolecular (indirect) contribution for the indicated side chain with the remainder of barstar, and ΔG_{inter} is the intermolecular (direct) electrostatic interaction between the indicated side chain and barnase. ΔG_{total} is the sum of the contributions. A barstar backbone contribution of -0.71 kcal/mol and a barnase dehydration contribution of 52.1 kcal/mol is to be added to the final ΔG_{total} to give the total electrostatic binding free energy. Bold and italic indicate favorable and unfavorable contributions, respectively, of magnitude exceeding 1 kcal/mol.

Optimized binding electrostatics are not directly sensitive to total charge or distant functional groups

The nominal charge on barnase is +1 and barstar is -6, assuming unshifted pK_{a} s at neutral pH. There is a general belief that complementarity is enhanced for cases of oppositely charged binding partners, even though a number of

natural counterexamples exist (Klapper et al. 1986; Allison et al. 1985, 1988; Marquart et al. 1983; Bajorath et al. 1991). Here we show that the electrostatic binding free energy is insensitive to the total molecular charge across a relatively broad range in the case of barstar, in which there are many solvent-exposed side chains distant from the binding interface. Table 6 gives the results of three further op-

Fig. 2. Ribbon drawing of the bound complex of barnase (blue) and barstar (yellow), indicating side chains with electrostatic interactions computed to be greater than 1 kcal/mole for wild type and optimized (opt3) barstar charge distributions. (A) ΔG_{inter} for $\{Q^{\text{wt}}\}$, (B) ΔG_{inter} for $\{Q^{\text{opt3}}\}$, (C) ΔG_{intra} for $\{Q^{\text{wt}}\}$, and (D) ΔG_{intra} for $\{Q^{\text{opt3}}\}$.

timizations in which side chain atomic charges were constrained to the range of -0.85 to $+0.85$, backbone charges were fixed at their ordinary values, and the total molecular charge of the optimum was constrained to be -10 , 0 , or $+10$. The electrostatic binding free energy was essentially the same value for each of these cases as that obtained when the total molecular charge was unconstrained ($\{Q_i^{\text{opt2}}\}$). The total molecular charge constraint, increasing or decreasing the net charge by 10 , was met by distributing extra charge to locations away from the interface and that had little effect on binding.

It can be shown that the binding free energy difference between an optimal and nonoptimal charge distribution binding to the same receptor with the same geometry is given by,

$$\Delta\Delta G = \frac{1}{2} (\vec{Q} - \vec{Q}^{\text{opt}})^T \vec{A} (\vec{Q} - \vec{Q}^{\text{opt}}) \quad (29)$$

in which $\vec{Q}$ is the nonoptimal and $\vec{Q}^{\text{opt}}$ is the optimal charge distribution (Lee and Tidor 1997; Kangas and Tidor 1998). If all point charges but one, Q_k , are optimal, the binding free energy difference is,

$$\Delta\Delta G_k = \frac{1}{2} A_{kk} (Q_k - Q_k^{\text{opt}})^2. \quad (30)$$

A_{kk} can be viewed as one measure of the sensitivity of the binding free energy to deviations from optimal for charge Q_k . It should be noted that A_{kk} is a diagonal term in the desolvation matrix and will be positive definite, largest for ligand point charges near the binding interface, and smaller for point charge locations away from the interface. In actual practice, multiple point-charge magnitudes will be nonoptimal and coupling between point charges will be affected by the entire row or column containing A_{kk} . Such off diagonal terms are especially small when Q_k is solvent exposed in the unbound and bound ligand so that its electrostatic inter-

Table 4. Ion pairs in wild-type and optimized (opt3) charge distribution^a

Ion pair ^b		Distance (Å)	Optimized residue charge	Atomic point charge		
				Atom	Q_i^{wt}	Q_i^{opt3}
Asp39b*	Arg83bn	1.6	−0.69	CB	−0.16	−0.36
	Arg87bn	1.9		CG	0.36	0.44
	Lys27bn	3.9		OD1	−0.60	−0.65
Glu76b*	Arg59bn	2.1	−1.15	OD2	−0.60	−0.85
				CB	0.00	0.45
				CG	−0.16	0.85
				CD	0.36	−0.85
				OE1	−0.60	−0.75
Asp35b*	Lys62bn	4.6	−0.90	OE2	−0.60	−0.85
				CB	−0.16	−0.20
				CG	0.36	0.85
				OD1	−0.60	−0.79
				OD2	−0.60	−0.75
Gln72b*	Arg59bn	4.8	−1.50	CB	0.00	0.49
				CG	0.00	−0.53
				CD	0.55	−0.85
				OE1	−0.55	0.08
				NE2	−0.60	−0.69
				HE21	0.30	−0.85
				HE22	0.30	0.85

^a Boldface type indicates a new electrostatic interaction in the optimized charge distribution. Normal type indicates interactions also present in wild type.

^b A suffix of b* indicates barstar and bn indicates barnase.

actions are largely screened by solvent. To summarize, large values of A_{kk} are expected to indicate the importance of optimizing partial charge Q_k . Small values of A_{kk} might not signal the nonimportance of optimizing Q_k because of context-dependent effects; but these effects should be minimal for solvent-exposed ligand point charges away from the binding site.

Figure 3A shows a ribbon drawing of barnase and a C_α trace of the ligand barstar, in which the radius of each barstar C_α atom represents the average A_{kk} for that residue's side-chain atoms. The largest values of A_{kk} are at the binding interface. Figure 3B is analogous, but the radius of each C_α represents the r.m.s. difference between ($\{Q_i^{\text{opt2}}\}$) and ($\{Q_i^{\text{opt2a}}\}$) (the unconstrained optimization and that constrained to a total charge of −10) for each side chain. The results show that the constraint was satisfied with very little change to the locations with large A_{kk} values and the biggest changes generally occurred at locations with very small A_{kk} values. As a further illustration of this principle, when all point charges in the barstar ligand were fixed to their wild-type value and only side chain point-charge locations with $A_{kk} \geq 0.1$ (138 atomic locations) were optimized in the range of −0.85 to +0.85, giving ($\{Q_i^{\text{opt2r}}\}$), the electrostatic binding free energy was only 1.6 kcal/mole worse than ($\{Q_i^{\text{opt2}}\}$). Figure 4 emphasizes that atoms with small values of A_{kk} tend to be the least restricted through optimization. This is

consistent with other studies that show that exposed charges away from the binding interface tend to have negligible effects on the binding free energy.

Improvements through optimization are not parameter restricted

One concern in applying a mathematical optimization procedure such as that used here is that the optimization might numerically fine-tune the interactions beyond our confidence in the model parameters. To test this we computed electrostatic binding free energies for barstar binding to barnase using three different charge and radius parameter sets and compared this to binding of the opt1 (optimized with CHARMM PARAM19). The results show binding enhancements of −20.3, −17.2, −15.1, and −10.6 kcal/mole with CHARMM PARAM19 (Brooks et al. 1983), OPLS (with 0 or 1.25 Å for hydrogen atom radius. All other atomic radii were taken as $2^{-5/6}$ (Jorgensen and Tirado-Rives 1988) and PARSE (Sitkoff et al. 1994), respectively. That a common optimized charge distribution produced large binding enhancements when used in the context of other parameter sets argues that much of the binding advantage is independent of numerical optimization within a single set of parameters.

The use of complementary electrostatic potentials to assess binding interfaces

A useful graphical representation of the electrostatic properties of a molecule is a plot of the screened coulombic potential at the molecular surface. Such representations are very common in reports of new macromolecular structures and can be conveniently computed and displayed with software packages such as GRASP (Nicholls et al. 1991). The images show the overall effect of the pattern of polar and charged residues at the surface. Figures 5A and B show this representation for the ligand barstar and also for the receptor barnase projected onto the barstar surface for ease of comparison. Whereas the surfaces are generally complementary in certain areas, the detailed distributions do not correspond well, particularly given that this is such a tight-binding complex.

However, unbound-state screened coulombic potentials could be unable to give a qualitative representation of electrostatic binding complementarity because they do not explicitly consider the desolvation penalty and the intermolecular interactions recovered in the bound state. We have previously proposed a measure of electrostatic complementarity based on electrostatic optimization theory (Kangas and Tidor 1998); the optimization condition can be expressed as,

$$\vec{A} \cdot \vec{Q}^{\text{opt}} + \vec{B} = 0 \quad (31)$$

Table 5. *Hydrogen bonds in wild-type and optimized (opt3) charge distribution^a*

Donor		Charge		Acceptor		Charge		Distance (Å)
Residue ^b	Atom	Q_i^{wt}	Q_i^{opt3}	Residue ^b	Atom	Q_i^{wt}	Q_i^{opt3}	
Lys27bn	HZ1	0.35		Thr42b*	OG1	-0.65	-0.85	1.94
Arg59bn	H	0.25		Asp35b*	OD1	-0.60	-0.79	1.98
Arg59bn	HH22	0.35		Glu76b*	OE1	-0.60	-0.75	2.07
Glu60bn	H	0.25		Asp35b*	OD2	-0.60	-0.75	2.45
Arg83bn	HH11	0.35		Gly43b*	O	-0.55		2.11
Arg83bn	HH12	0.35		Asp39b*	OD1	-0.60	-0.65	2.11
Arg83bn	HH12	0.35		Asp39b*	O	-0.55		2.93
Arg83bn	HH21	0.35		Asp39b*	OD1	-0.60	-0.65	1.58
Arg87bn	HH21	0.35		Asp39b*	OD2	-0.60	-0.85	1.94
His102bn	HE2	0.30		Asp39b*	OD2	-0.60	-0.85	1.84
Tyr29b*	HH	0.40	0.30	Arg83bn	O	-0.55		1.88
Gly31b*	H	0.25		His102bn	ND1	-0.40		2.12
Asn33b*	HD22	0.30	0.09	His102bn	O	-0.55		2.29
Leu34b*	H	0.25		Glu60bn	OE2	-0.60		1.86
Trp38b*	HE1	0.30	0.04	Wat48bn	OH2	-0.834		2.45
Trp38b*	HE1	0.30	0.04	Wat56bn	OH2	-0.834		2.63
Trp38b*	HE1	0.30	0.04	Wat60bn	OH2	-0.834		2.90
Wat14bn	H2	0.417		Asp35b*	O	-0.55		2.17
Wat22bn	H2	0.417		Asp35b*	OD2	-0.60	-0.75	1.87
Wat29bn	H2	0.417		Asp35b*	OD1	-0.60	-0.79	1.86
Wat33bn	H1	0.417		Asp39b*	OD1	-0.60	-0.65	2.29
Wat36bn	H1	0.417		Val45b*	O	-0.55		2.39
Wat36bn	H2	0.417		Val45b*	O	-0.55		2.14
Wat128bn	H2	0.417		Asp35b*	OD2	-0.60	-0.75	1.97
Wat155bn	H2	0.417		Gly43b*	O	-0.55		1.97
Arg59bn	HH12	0.35		Glu76b*	CD	0.36	-0.85	2.24
Wat29bn	H1	0.417		Trp38b*	CB	0.00	-0.24	2.93

^a Boldface type indicates a new electrostatic interaction in the optimized charge distribution. Normal type indicates interactions also present in wild type.

^b A suffix of b* indicates barstar and bn indicates barnase.

which can be obtained from equation 19. It represents a pair of potentials, one from the ligand-charge distribution ($\vec{A}$, Q^{L}) and one from the receptor (B) that are equal in magnitude and opposite in sign throughout the ligand volume at an optimum. The first term is the ligand desolvation potential and represents the bound- minus the unbound-state potential due only to the ligand-charge distribution; the second term is the receptor interaction potential and represents the bound-state potential from the receptor-charge distribution. These potentials are plotted at the barstar molecular surface in Figure 5C and D for actual barstar and barnase (not one of the computed optima). The potentials are remarkably complementary, with the pattern and intensity of blue and red being a near quantitative match between the pair. Moreover, the most intense features of the figures are from residues implicated experimentally as being especially important to binding (Schreiber and Fersht 1993, 1995; Schreiber et al. 1994; Hartley 1993). For example, the two most intense red regions in Figure 5C are from Asp35 and Asp39 of barstar, which are buried at the binding interface and make a series of interactions with positively charged side chains from barnase that are largely responsible for the corresponding intense blue region in Figure 5D. These observations

support the proposal that these potentials represent a useful measure of electrostatic complementarity. Scripts for use with the GRASP program (Nicholls et. al. 1991) to compute and display these potentials are available from <http://mit.edu/tidor>.

Schreiber and coworkers (1994) point out that there are small areas of negative charge on barnase (Asp54, Glu73, and Glu60) that are not matched by nearby positive charges on barstar, possibly leading to a reduction in binding affinity. The appropriate complementarity potentials for examining these residues in barnase are displayed in Figure 5E and F (see below). Detailed examination of the three-di-

Table 6. *Dependence of electrostatic binding free energy on total charge*

Charge set	Total charge	ΔG_{bind} (kcal/mol)
$\{Q_i^{\text{opt2}}\}$	-0.3 ^a	-0.56
$\{Q_i^{\text{opt2a}}\}$	-10.0	-0.52
$\{Q_i^{\text{opt2b}}\}$	0.0	-0.56
$\{Q_i^{\text{opt2c}}\}$	10.0	-0.52

^a Opt2 had no constraint on the total charge. The remaining charge sets had the total charge constrained to the listed value.

Fig. 3. Ribbon drawing of barnase and C_α trace of barstar, with (A) the radius of each barstar C_α atom representing the average A_{kk} for that side chain and (B) the radius of each barstar C_α atom representing the r.m.s difference between $\{Q_i^{\text{opt}2}\}$ and $\{Q_i^{\text{opt}2a}\}$ (the unconstrained optimization and that constrained to a total charge of -10) for that side chain.

mensional graphical objects used to create Figure 5E and F indicated no substantial lack of complementarity from these charged groups because of a relatively small desolvation potential. This resulted partially from their placement within the interface and partially from their intramolecular interactions with positively charged groups within barnase, which were enhanced on binding because of the accompanying reduction in solvent screening (Chong et al. 1998; Caravella et al. 1999; Hendsch and Tidor 1999). This stresses the observation that compensation of electrostatic groups on binding can come not only from introducing intermolecular partners but also from strengthening intramolecular effects on binding.

Relative optimization

An interesting feature of the potentials described above for assessing electrostatic complementarity is that they are not symmetric. Specifically, a difference of potentials (bound minus unbound) is plotted for the ligand, whereas simply a single potential is plotted for the receptor. Rather than being awkward, this reflects the fact that different improvements are expected when the enzyme is fixed and the inhibitor is optimized than when the inhibitor is fixed and the enzyme is optimized. Said another way, the enzyme might be further from its optimum than is the ligand. In principle such a situation could arise by chance, or it could be the result of different evolutionary pressures acting on each of the binding partners. For the purpose of illustration, here we describe a relative optimization analysis for barnase and barstar. Two measures were computed to judge which binding partner is closer to its optimum (or in other words, is more complementary to its partner). The first measure is simply to reverse the role of ligand and receptor and recompute the complementary potentials described above. These

are shown in Figures 5E and F and can be compared with Figures 5C and D. Whereas both sets are quite complementary, parts C and D are somewhat more complementary than parts E and F. This can be confirmed numerically (the correlation coefficient computed over the ligand atom centers is -0.97 for the former pair and -0.92 for the latter. Perfect complementarity would be -1).

The second measure involves substantially more computation but is more accurate. The charge distribution of barstar was fixed whereas that of barnase was optimized. Comparison of energetic improvements was made to calculations carried out with barnase fixed and barstar optimized. The results for a number of optimization procedures are given in Table 7. In each case, optimization of barnase leads

Fig. 4. Scatter plot of the A_{kk} versus the difference between $Q_k^{\text{opt}2}$ and $Q_k^{\text{opt}2r}$ (ΔQ_k). Only charges with small A_{kk} (desolvation matrix element) have significant ΔQ_k (change in atomic charge between optimizations that change the total ligand charge significantly).

Fig. 5. Potential plots displayed using GRASP (Nicholls et al. 1991). Screened coulombic potential for barstar (A) and barnase (B) on barstar's molecular surface; barstar's desolvation potential (C) and interaction potential (D) on barstar's molecular surface; and barnase's desolvation potential (E) and interaction potential (F) on barnase's molecular surface. Increasingly positive (negative) electrostatic potential is indicated by progressively darker shades of blue (red).

to greater binding enhancements than a similar optimization carried out for barstar. Because barstar has fewer atom centers than barnase in the model, care was taken to compare optimizations with the same number of variable charges. Both measures suggest that barstar, the inhibitor, is closer to being optimized for binding barnase than the other way around. This result is consistent with the known roles of these proteins. Barstar is an inhibitor of barnase and its primary function is presumably to fold and bind to barnase. It is reasonable to expect it to optimize binding to barnase. On the other hand, barnase is an enzyme with its own catalytic role. Optimal binding to barstar is expected to be less of a priority. It should be noted that there must be many other constraints and considerations regarding the roles of these molecules, some of which we might be unaware. Because of the need to rapidly and efficiently inhibit barnase, further optimization of the kinetics of folding and binding of barstar is also logical (Schreiber et al. 1994; Schreiber and Fersht 1996; Gabdoulline and Wade 1997; Nölting et al. 1997).

Importance of electrostatics

The calculational results reported here suggest that electrostatic interactions could be engineered to stabilize protein binding. Nevertheless, the magnitude of the stabilizing effect is rather modest for cases examined to date (e.g., be-

tween 0 and 6 kcal/mole for barnase). One might ask whether electrostatics, therefore, only play at best a minor role and that attention should rather be focused exclusively on energy terms that contribute more. The hydrophobic effect, for instance, contributes dozens of kcal/mole of favorable binding free energy to protein-protein complexes, and

Table 7. Improvement in ΔG_{bind} : A comparison between barnase and barstar optimization

Degree of freedom (Constraints) ^a	Barnase optimization		Barstar optimization	
	$n_{\text{L.v}}$	$\Delta\Delta G_{\text{bind}}^b$	$n_{\text{L.v}}$	$\Delta\Delta G_{\text{bind}}^b$
All charges	1071	-30.8	839	-20.3
Side chain charges	533	-27.9	403	-14.8
Side chain charges ($-1.5 \leq \sum_{\text{sidechain}} Q_i \leq 1.5$)	533	-27.0	403	-14.3
Side chain charges with $A_{kk} > 0.1$	179	-25.9	138	-13.3
100 side chain charges with the greatest A_{kk}	100	-21.4	100	-12.0
200 side chain charges with the greatest A_{kk}	200	-26.5	200	-14.6
300 side chain charges with the greatest A_{kk}	300	-27.8	300	-14.8

^a The constraints are in addition to the individual charge range constraint $-0.85 \leq Q_i \leq 0.85$.

^b $\Delta\Delta G_{\text{bind}}$ is the improvement of the optimized charge distribution over wild type in kcal/mol.

lost translational and rotational entropy can contribute dozens of kcal/mole unfavorably (Chothia and Janin 1975; Tidor and Karplus 1994). Should not one look to these or other large net interactions to understand differences in binding affinity for related complexes and to design tight-binding molecules? The answer is clearly No because what matters most in understanding binding differences and in design studies is not the terms that are large, but rather the terms whose magnitude can span a wide range. Certainly a full and accurate accounting of all terms is best, but this is beyond the scope of current technology (Froloff et al. 1997). Lost translations and overall rotations exert their effect through the logarithm of the mass and of the principal moments of inertia product (Hill 1986), so large perturbations in size (which change other interactions) are required to cause large changes in affinity. Improvements in shape complementarity (van der Waals packing) and the hydrophobic binding contribution are acceptable ways to produce tighter binding ligands, and a large number of efficient algorithms have been developed to exploit these effects (Desmet et al. 1992; Harbury et al. 1995; Dahiyat and Mayo 1997; Harbury et al. 1998). However, hydrophobic improvements tend to increase ligand size and hydrophobicity, which can be detrimental if the ultimate goal is a drug. Additionally, hydrophobic interactions alone might not impart sufficient specificity for many applications, and the magnitude of binding enhancement available through packing and van der Waals improvement is unclear. Here we show that, in principle, the difference between barstar (a ligand that appears to be making very good electrostatic interactions) and an optimized ligand is 10–20 kcal/mole. Thus, a focus on electrostatics is warranted because the potential gains through improving electrostatics can be quite large. Moreover, because electrostatics can impart substantial specificity (Sindelar et al. 1998), enhancements in affinity through improved electrostatics might also lead to enhancements in specificity (Dempster and Tidor, in prep.). Whereas interfaces deficient in shape complementarity can often be simply detected and remedies designed through visual inspection of structures, our results show that deficiencies in electrostatic complementarity can be more difficult to detect. The methodologies for analysis and design presented here will serve as useful tools in further studies of electrostatic complementarity.

Conclusion

Electrostatic optimization methods have been extended to deal efficiently with detailed molecular geometries through numerical computation of the necessary free energetic terms. Application to the barnase-barstar complex has revealed the predominance of relatively few barstar residues at the interface whose electrostatic charge distributions are computed to produce especially large effects on the binding

free energy. Experiments involving mutational analysis have implicated many of the same residues in binding. Analysis of optimized charge distributions suggests three general mechanisms for enhancing the net electrostatic effect on binding. First, improvements to the hydrogen-bond inventory across the interface improve binding. This includes not only increasing the number of interactions through the introduction of appropriately placed hydrogen-bond donors and acceptors, but also the elimination of ligand interfacial polar groups that are inappropriately placed to make good-geometry intermolecular interactions (deletion or hydrophobic replacement of unsatisfied donors and acceptors). Second, the enhancement of intramolecular electrostatic effects on binding through changes in solvent screening appears repeatedly in natural complexes as well as in electrostatic optima (Chong et al. 1998; Hendsch and Tidor 1999; Caravella et al. 1999). This frequently takes the form of introducing polar or charged ligand groups that interact favorably with interfacial ligand electrostatic groups. Solvent screening in the unbound state attenuates the interaction, which is correspondingly larger in the bound state because of displacement of high-dielectric solvent with low-dielectric protein. This effect, whose importance has not generally been recognized, suggests that attention must be focused on intramolecular as well as intermolecular interactions in the analysis and design of binding partners. Third, tuning of particular electrostatic groups can have significant effects on the overall electrostatic contribution to binding. Because the details depend on the geometry and charge distributions in the bound and unbound states, it is not yet possible to extract simple rules to predict the appropriate polarity to place at a site without carrying out a full calculation. The clear implication for combinatorial library design is that a range of polarities should be sampled at electrostatically critical regions of binding pockets. Further studies of electrostatic optimization, both theoretical and experimental, could lead to a focusing of these general guidelines as well as a catalog of particular functional-group motifs that are effective in enhancing electrostatic binding contributions in a broad class of circumstances. Our results suggest that complexes tighter than those generally observed for biomolecules might be achievable and is in agreement with recent suggestions by Kuntz and coworkers (1999).

A procedure for illustrating and evaluating electrostatic complementarity is a direct result of charge-optimization theory (Kangas and Tidor 1998). This measure of charge complementarity clearly reveals compensatory electrostatics at binding interfaces, which had generally not been readily apparent using previous methods. Scripts for computing and displaying this new measure with the GRASP (Nicholls et al. 1991) computer program are available at <http://mit.edu/tidor>. Our calculational analysis has shown that barstar is more electrostatically optimized than barnase

for binding in this complex, consistent with barnase having at least one other major function, namely to catalyze the hydrolysis of RNA.

The electrostatic calculations made here used the binding free energy of the prefolded inhibitor as the target function. In actual practice, barstar must also fold to a stable molecule, so affinity enhancing modifications would need to be selected with an eye toward maintaining barstar stability. Extensions to the methodology used here to allow dual optimization are possible, in which the stability of the isolated ligand and the affinity of its complex with a receptor are simultaneously optimized.

In the current work great latitude was permitted in the optimized charge distributions. Constraints were placed on individual atomic charge magnitudes, on side chain charges, and on molecular charge, but no requirements were enforced that the charge distributions be achievable by proteins. In this manner we have investigated general features of optimized ligand charge distributions and compared them to natural barstar to learn where improvements can arise, what form they might take, and what types of magnitudes one might expect. In other work we show that, given the barstar protein backbone and only the 20 common amino acid side chains, barstar is remarkably well optimized (Lee and Tidor 2001). Combined, the pair of studies suggests that substantial enhancements in binding affinity, in general, might be available through improvement of electrostatic interactions. This is likely to require synthetic chemistry that goes far beyond the common amino acids in producing a finely graded diversity of shapes and electrostatic charge distributions, which could be particularly useful in small molecule inhibitor design.

Acknowledgments

We thank Barry Honig for making the DELPHI and GRASP computer programs available; Martin Karplus for CHARMM; Robert J. Vanderbei for LOQO; and Mouni G. Bawendi, Charles L. Brooks III, Sylvia T. Ceyer, Irwin D. Kuntz, and members of our research group for helpful discussions. This work was supported by the National Institutes of Health (GM55758 and GM56552).

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC section 1734 solely to indicate this fact.

References

- Allison, S.A., Ganti, G., and McCammon, J.A. 1985. Simulation of the diffusion-controlled reaction between superoxide and superoxide dismutase. I. Simple models. *Biopolymers* **24**: 1323–1336.
- Allison, S.A., Bacquet, R.J., and McCammon, J.A. 1988. Simulation of the diffusion-controlled reaction between superoxide and superoxide dismutase. II. Detailed models. *Biopolymers* **27**: 251–269.
- Bajorath, J., Kitson, D.H., Kraut, J., and Hagler, A.T. 1991. The electrostatic potential of *Escherichia coli* dihydrofolate reductase. *Proteins* **11**: 1–12.
- Bockris, J.O'M. and Reddy, A.K.N. 1973. *Modern Electrochemistry*. Plenum, New York.
- Brooks, B.R., Bruccoleri, R.E., Olafson, B.D., States, D.J., Swaminathan, S., and Karplus, M. 1983. CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem.* **4**: 187–217.
- Bruccoleri, R.E., Novotny, J., Davis, M.E., and Sharp, K.A. 1997. Finite difference Poisson-Boltzmann electrostatic calculations: Increased accuracy achieved by harmonic dielectric smoothing and charge antialiasing. *J. Comput. Chem.* **18**: 268–276.
- Buckle, A.M., Schreiber, G., and Fersht, A.R. 1994. Protein-protein recognition: Crystal structural analysis of a barnase-barstar complex at 2.0 Å resolution. *Biochemistry* **33**: 8878–8889.
- Caravella, J.A., Carbeck, J.D., Duffy, D.C., Whitesides, G.M., and Tidor, B. 1999. Long-range electrostatic contributions to protein-ligand binding estimated using protein charge ladders, affinity capillary electrophoresis, and continuum electrostatic theory. *J. Am. Chem. Soc.* **121**: 4340–4347.
- Chong, L.T., Dempster, S.E., Hendsch, Z.S., Lee, L.-P., and Tidor, B. 1998. Computation of electrostatic complements to proteins: A case of charge stabilized binding. *Protein Sci.* **7**: 206–210.
- Chothia, C. and Janin, J. 1975. Principles of protein-protein recognition. *Nature* **256**: 705–708.
- Connolly, M.L. 1983. Analytical molecular surface calculation. *J. Appl. Crystallogr.* **16**: 548–558.
- Dahiyat, B.I. and Mayo, S.L. 1997. De novo protein design: Fully automated sequence selection. *Science* **278**: 82–87.
- Desmet, J., De Maeyer, M., Hazes, B., and Lasters, I. 1992. The dead-end elimination theorem and its use in protein side-chain positioning. *Nature* **356**: 539–542.
- Fogolari, F. and Briggs, J.M. 1997. On the variational approach to Poisson-Boltzmann free energies. *Chem. Phys. Lett.* **281**: 135–139.
- Froloff, N., Windemuth, A., and Honig, B. 1997. On the calculation of binding free energies using continuum methods: Application to MHC class I protein-peptide interactions. *Protein Sci.* **6**: 1293–1301.
- Gabdoulline, R.R. and Wade, R.C. 1997. Simulation of the diffusional association of barnase and barstar. *Biophys. J.* **72**: 1917–1929.
- Gilson, M.K. and Honig, B.H. 1987. Calculation of electrostatic potentials in an enzyme active site. *Nature* **330**: 84–86.
- Gilson, M.K., Sharp, K.A., and Honig, B.H. 1988. Calculating the electrostatic potential of molecules in solution: Method and error assessment. *J. Comput. Chem.* **9**: 327–335.
- Guillet, V., Laphorn, A., Hartley, R.W., and Mauguen, Y. 1993. Recognition between a bacterial ribonuclease, barnase, and its natural inhibitor, barstar. *Structure* **1**: 165–177.
- Harbury, P.B., Tidor, B., and Kim, P.S. 1995. Repacking protein cores with backbone freedom: Structure prediction for coiled coils. *Proc. Natl. Acad. Sci. USA* **92**: 8408–8412.
- Harbury, P.B., Plecs, J.J., Tidor, B., Alber, T., and Kim, P.S. 1998. High-resolution protein design with backbone freedom. *Science* **282**: 1462–1467.
- Hartley, R.W. 1993. Directed mutagenesis and barnase-barstar recognition. *Biochemistry* **32**: 5978–5984.
- Hendsch, Z.S. and Tidor, B. 1994. Do salt bridges stabilize proteins? A continuum electrostatic analysis. *Protein Sci.* **3**: 211–226.
- . 1999. Electrostatic interactions in the GCN4 leucine zipper: Substantial contributions arise from intramolecular interactions enhanced on binding. *Protein Sci.* **8**: 1381–1392.
- Hill, T.L. 1986. *An Introduction to Statistical Thermodynamics*. Dover, New York.
- Jorgensen, W.L. and Tirado-Rives, J. 1988. The OPLS potential function for proteins. Energy minimizations for crystals of cyclic peptides and crambin. *J. Am. Chem. Soc.* **110**: 1657–1666.
- Kangas, E. and Tidor, B. 1998. Optimizing electrostatic affinity in ligand-receptor binding: Theory, computation, and ligand properties. *J. Chem. Phys.* **109**: 7522–7545.
- . 1999. Charge optimization leads to favorable electrostatic binding free energy. *Phys. Rev. E* **59**: 5958–5961.
- Klapper, I., Hagstrom, R., Fine, R., Sharp, K., and Honig, B. 1986. Focusing of electric fields in the active site of Cu-Zn superoxide dismutase: Effects of ionic strength and amino-acid modification. *Proteins* **1**: 47–59.
- Kraulis, P.J. 1991. MOLSCRIPT: A program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.* **24**: 946–950.
- Kuntz, I.D., Chen, K., Sharp, K.A., and Kollman, P.A. 1999. The maximal affinity of ligands. *Proc. Natl. Acad. Sci. USA* **96**: 9997–10002.
- Lee, L.-P. and Tidor, B. 1997. Optimization of electrostatic binding free energy. *J. Chem. Phys.* **106**: 8681–8690.

- . 2001. Barstar is electrostatically optimized for tight-binding to barnase. *Nat. Struct. Biol.*, in press.
- Marquart, M., Walter, J., Deisenhofer, J., Bode, W., and Huber, R. 1983. The geometry of the reactive site and of the peptide groups in trypsin, trypsinogen and its complexes with inhibitors. *Acta. Crystallogr. Ser. B* **39**: 480–490.
- Martin, C., Richard, V., Salem, M., Hartley, R., and Mauguen, Y. 1999. Refinement and structural analysis of barnase at 1.5 Å resolution. *Acta. Crystallogr. Ser. D* **55**: 386–398.
- Merritt, E.A. and Bacon, D.J. 1997. Raster3D: Photorealistic molecular graphics. *Method. Enzymol.* **277**: 505–524.
- Misra, V.K., Hecht, J.L., Sharp, K.A., Friedman, R.A., and Honig, B. 1994a. Salt effects on protein-DNA interactions: The λ cl repressor and EcoRI endonuclease. *J. Mol. Biol.* **238**: 264–280.
- Misra, V.K., Sharp, K.A., Friedman, R.A., and Honig, B. 1994b. Salt effects on ligand-DNA binding: Minor groove binding antibiotics. *J. Mol. Biol.* **238**: 245–263.
- Nicholls, A., Sharp, K.A., and Honig, B. 1991. Protein folding and association: Insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins* **11**: 281–296.
- Nölting, B., Golbik, R., Neira, J.L., Soler-Gonzalez, A.S., Schreiber, G., and Fersht, A.R. 1997. The folding pathway of a protein at high resolution from microseconds to seconds. *Proc. Natl. Acad. Sci. USA* **94**: 826–830.
- Novotny, J., Bruccoleri, R.E., Davis, M., and Sharp, K.A. 1997. Empirical free energy calculations: A blind test and further improvements to the method. *J. Mol. Biol.* **268**: 401–411.
- Overbeek, J.T.G. 1990. The role of energy and entropy in the electrical double layer. *Colloid. Surf.* **51**: 61–75.
- Paul, C.H. 1982. Building models of globular protein molecules from their amino acid sequences. I. Theory. *J. Mol. Biol.* **155**: 53–62.
- Press, W.H., Teukolsky, S.A., Vetterling, W.T., and Flannery, B.P. 1992. *Numerical Recipes in C: The Art of Scientific Computing*. Cambridge University Press, Cambridge, 2nd ed.
- Ratnaparkhi, G.S., Ramachandran, S., Udgaonkar, J.B., and Varadarajan, R. 1998. Discrepancies between the NMR and X-ray structures of uncomplexed barstar: Analysis suggests that packing densities of protein structures determined by NMR are unreliable. *Biochemistry* **37**: 6958–6966.
- Richards, F.M. 1977. Areas, volumes, packing, and protein structure. *Annu. Rev. Biophys. Bioeng.* **6**: 151–176.
- Roux, B. and Simonson, T. 1999. Implicit solvent models. *Biophys. Chem.* **78**: 1–20.
- Schreiber, G. and Fersht, A.R. 1993. Interaction of barnase with its polypeptide inhibitor barstar studied by protein engineering. *Biochemistry* **32**: 5145–5150.
- . 1995. Energetics of protein-protein interactions: Analysis of the barnase-barstar interface by single mutations and double mutant cycles. *J. Mol. Biol.* **248**: 478–486.
- . 1996. Rapid, electrostatically assisted association of proteins. *Nat. Struct. Biol.* **3**: 427–431.
- Schreiber, G., Buckle, A.M., and Fersht, A.R. 1994. Stability and function: Two constraints in the evolution of barstar and other proteins. *Structure* **2**: 945–951.
- Sharp, K. 1996. Electrostatic interactions in hirudin-thrombin binding. *Biophys. Chem.* **61**: 37–49.
- Sharp, K.A. and Honig, B. 1990a. Calculating total electrostatic energies with the nonlinear Poisson-Boltzmann equation. *J. Phys. Chem.* **94**: 7684–7692.
- . 1990b. Electrostatic interactions in macromolecules: Theory and applications. *Annu. Rev. Biophys. Biophys. Biom.* **19**: 301–332.
- Shen, J. and Wendoloski, J. 1996. Electrostatic binding energy calculation using the finite difference solution to the linearized Poisson-Boltzmann equation: Assessment of its accuracy. *J. Comput. Chem.* **17**: 350–357.
- Sindelar, C.V., Hendsch, Z.S., and Tidor, B. 1998. Effects of salt bridges on protein structure and design. *Protein Sci.* **7**: 1898–1914.
- Sitkoff, D., Sharp, K.A., and Honig, B. 1994. Accurate calculation of hydration free energies using macroscopic solvent models. *J. Phys. Chem.* **98**: 1978–1988.
- Strang, G. 1993. *Introduction to Linear Algebra*. Wellesley-Cambridge Press, Wellesley, Massachusetts.
- Tanford, C., De P.K., and Taggart, V.G. 1960. The role of the α -helix in the structure of proteins. Optical rotatory dispersion of β -lactoglobulin. *J. Am. Chem. Soc.* **82**: 6028–6034.
- Tidor, B. and Karplus, M. 1994. The contribution of vibrational entropy to molecular association. The dimerization of insulin. *J. Mol. Biol.* **238**: 405–414.
- Vanderbei, R.J. 1997a. *Linear Programming: Foundations and Extensions*. Kluwer Academic Publishers, Boston, Massachusetts.
- . 1997b. *LOQO v3.10* Princeton University, Princeton.
- . 1998. *LOQO v4.01*. Princeton University, Princeton.
- Waldburger, C.D., Schildbach, J.F., and Sauer, R.T. 1995. Are buried salt bridges important for protein stability and conformational specificity? *Nat. Struct. Biol.* **2**: 122–128.
- Wang, L., O'Connell, T., Tropsha, A., and Hermans, J. 1996. Energetic decomposition of the α -helix-coil equilibrium of a dynamic model system. *Biopolymers* **39**: 479–489.
- Wimley, W.C., Gawrisch, K., Creamer, T.P., and White, S.H. 1996. Direct measurement of salt-bridge solvation energies using a peptide model system: Implications for protein stability. *Proc. Natl. Acad. Sci. USA* **93**: 2985–2990.
- Yang, A.S. and Honig, B. 1995. Free energy determinants of secondary structure formation: I. α -helices. *J. Mol. Biol.* **252**: 351–365.