Protein Science (1998), 7:206–210. Cambridge University Press. Printed in the USA.

FOR THE RECORD

Computation of electrostatic complements to proteins: A case of charge stabilized binding

LILLIAN T. CHONG, SARA E. DEMPSTER, ZACHARY S. HENDSCH, LEE-PENG LEE, and BRUCE TIDOR

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

(RECEIVED August 28, 1997; ACCEPTED October 16, 1997)

Abstract: Recent evidence suggests that the net effect of electrostatics is generally to destabilize protein binding due to large desolvation penalties. A novel method for computing ligand-charge distributions that optimize the tradeoff between ligand desolvation penalty and favorable interactions with a binding site has been applied to a model for barnase. The result is a ligand-charge distribution with a favorable electrostatic contribution to binding due, in part, to ligand point charges whose direct interaction with the binding site is unfavorable, but which make strong intra-molecular interactions that are uncloaked on binding and thus act to lessen the ligand desolvation penalty.

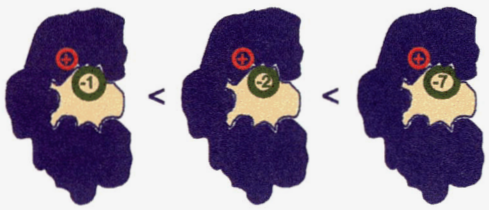
Keywords: continuum electrostatics; electrostatic complement; protein binding; rational ligand design

The relative strengths of interactions involved in protein folding and binding are of fundamental importance for our understanding of these processes and for our ability to design modified or novel proteins and tight-binding ligands. Recent theoretical (Hendsch & Tidor, 1994; Yang & Honig, 1995; Wang et al., 1996) and experimental (Waldburger et al., 1995; Wimley et al., 1996) studies have emphasized the large electrostatic desolvation penalty due to polar and charged groups incurred in protein folding and have led to the realization that the desolvation penalty is not generally recovered in favorable interactions created in the folded state (although exceptions have been noted: Hendsch & Tidor, 1994; Lounnas & Wade, 1997; and, in the context of binding, Xu et al., 1997). For example, Sauer and co-workers replaced a triad of hydrogen-bonded, salt bridging groups with hydrophobic groups (which pay essentially no electrostatic desolvation penalty but recover negligible electrostatic interactions on folding) in Arc repressor and measured enhanced stabilities of 1–2$\frac{1}{2}$ kcal/mol per monomer (Waldburger et al., 1995). Although electrostatic interactions may not generally contribute to stability, they are likely to have a substantial effect on specificity due to the cost of burying but not compensating polar or charged groups (Tanford et al., 1960; Paul, 1982; Hendsch & Tidor, 1994). Similar results are emerging for protein binding as well. Calculations for a number of protein complexes show that the net effect of electrostatic interactions is generally to destabilize the docking of two pre-conformed molecules (Novotny & Sharp, 1992; Misra et al., 1994a, 1994b; Sharp, 1996; Shen & Wendoloski, 1996; Bruccoleri et al., 1997; Novotny et al., 1997). This raises the important question of whether it is possible to design a binding partner for a given site such that the electrostatic contribution to binding is favorable. Here we utilize a recently introduced electrostatic optimization strategy to address this question (Lee & Tidor, 1997).

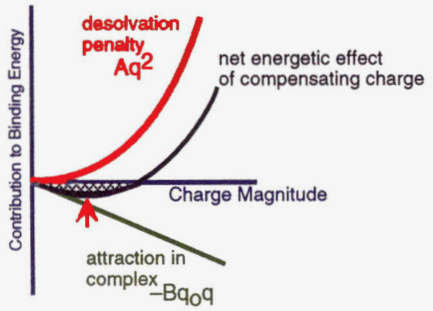
We have developed a scheme (outlined in Fig. 1) to optimize the electrostatic component of the binding free energy within the continuum model (Lee & Tidor, 1997). When a ligand binds to a receptor, there are two competing forces: the desolvation or dehydration penalty (the loss of favorable interaction between each molecule and the solvent) acts against the favorable interaction between the two molecules. Viewing the charge distribution of the receptor as fixed and that of the ligand as variable, the unfavorable ligand dehydration penalty increases with the square of the ligand-charge distribution, whereas the favorable interactions increase only linearly. This results in the existence of an optimal charge distribution that represents the most favorable tradeoff of dehydration and interaction. For the special case in which the ligand and the complex can be considered spherical regions of low dielectric embedded in a higher dielectric solvent, we represent the ligand-charge distribution by a set of variable multipoles whose values are chosen to minimize the electrostatic binding free energy. The monopole component (i.e., the total ligand charge) may either be variable or constrained (e.g., to an integer) to facilitate subsequent molecular construction (Lee & Tidor, 1997).

We have applied this method to the small bacterial ribonuclease from *Bacillus amyloliquefaciens*, barnase. The point-charge distribution from barnase was used for the receptor, and locations for the docking of a series of spherical low-dielectric ligands (of radius 8, 10, and 11 Å) were chosen using molecular graphics. The bound complexes were modeled using the point-charge representation for barnase, the variable multipole distribution for the ligand, and a 31-Å radius spherical dielectric boundary.

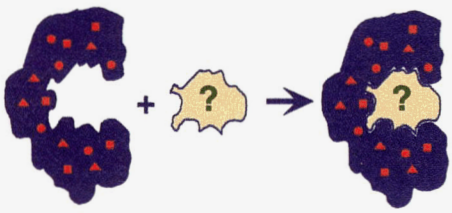
Reprint requests to: Bruce Tidor, Department of Chemistry, Room 6-135, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139-4307; e-mail: tidor@mit.edu.



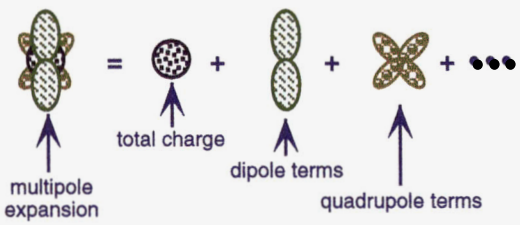
If Coulomb's law interactions between ligand and receptor were the dominant electrostatic effect, then tight-binding ligands would be constructed with very large charges to compensate charges in the active site. For instance, the ligand with charge of –7 would bind better than that with –2, which would bind better than that with –1.



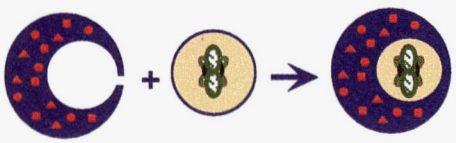
Because ligands pay a large desolvation penalty that increases as the square of the charge (q^2), highly charged ligands are not necessarily tight binders. In fact, because the favorable interaction is only linear in the charge (q), the net contribution to the binding energy as a function of ligand charge is a shifted parabola with a minimum (optimal) binding energy indicated by the arrow.



The primary methodological advance applied here is the development of a technique to solve for the optimal ligand charge distribution — that is, the charge distribution resulting in the most favorable electrostatic contribution to binding — given an arbitrary receptor shape and charge distribution (indicated by the blue region and red symbols, respectively).



For a spherical ligand, this can be achieved by using an arbitrary set of multipoles to represent any ligand charge distribution and solving for that set of multipoles that minimizes the electrostatic contribution to binding. The multipoles describe the total charge, dipolar terms, quadrupolar terms, etc. of the charge distribution.



The theory has been developed and implemented for the special case in which the ligand and the complex are perfect spheres. Work is currently under way to extend the methodology to arbitrarily shaped ligands and complexes.

Fig. 1. Rationale for electrostatic optimization.

Optimal ligand-charge distributions were determined as sets of multipoles. The total charges of the optimal ligands (for the 8, 10, and 11-Å ligands, respectively) were $-1.16e$, $-1.30e$, and $-1.36e$. A separate computation was carried out in which the total charge was constrained to the integer value of $-1e$ for the 8-Å ligand, whose unconstrained total charge was closest to an integer value. The binding energy converged relatively quickly as a function of multipole order and was dominated by the monopole and dipole terms (see Fig. 2A).

To determine the total electrostatic contribution to binding, we computed the ligand dehydration penalty and ligand–receptor interaction energy analytically (because in each case the dielectric boundary was spherical) and computed the receptor dehydration penalty numerically with the DELPHI computer program (Gilson et al., 1988; Gilson & Honig, 1988; Sharp & Honig, 1990). In each of the four cases, the overall electrostatic binding free energy was favorable, including the docking of the ligand with the total charge constrained to $-1e$; this is in contrast to computations on many natural complexes, for which the electrostatic contribution to the binding free energy is unfavorable (Novotny & Sharp, 1992; Misra et al., 1994a, 1994b; Sharp, 1996; Shen & Wendoloski, 1996; Bruccoleri et al., 1997; Novotny et al., 1997). Table 1 shows the

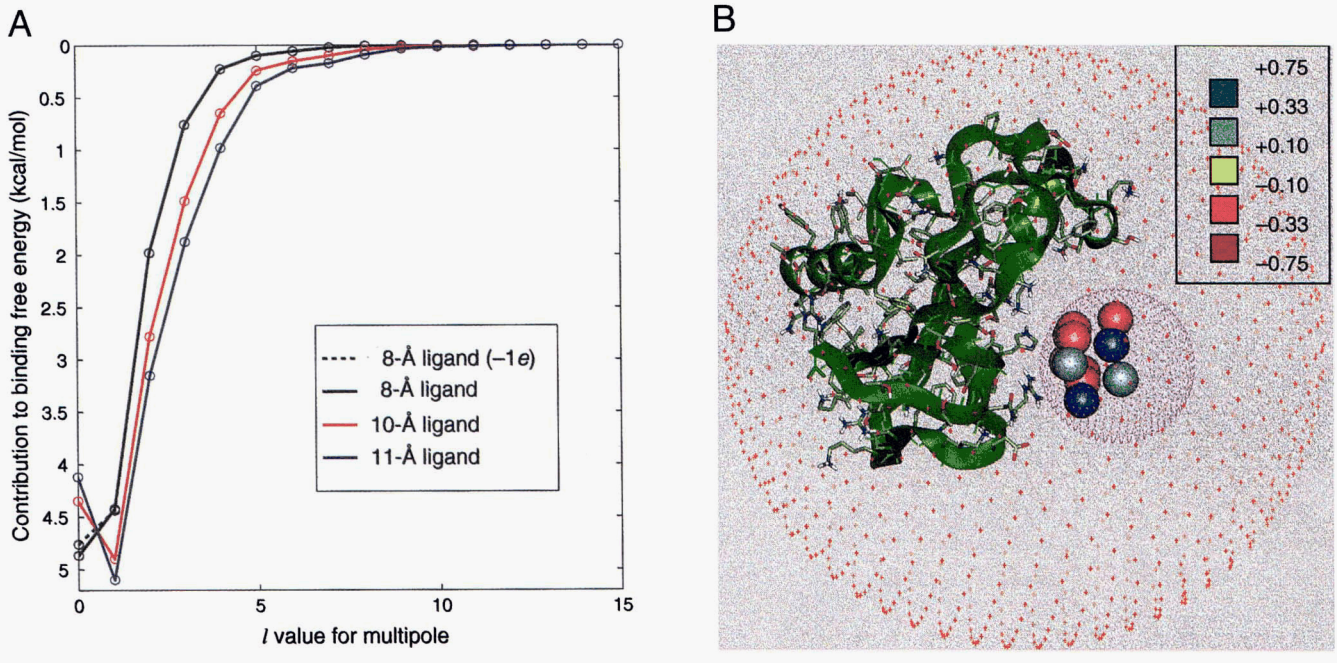


Fig. 2. A: The contribution of the first 16 multipoles to the electrostatic binding free energy is shown for each of the optimized ligand-charge distributions, where $l = 0$ corresponds to the monopole contribution, $l = 1$ corresponds to the dipole contribution, $l = 2$ to the quadrupole contribution, etc. **B:** The structure of barnase and the 11 point-charge distribution fit to the optimal 8-Å ligand (purple dotted sphere) with total charge constrained to $-1e$ is shown. Values of the point charges are indicated by their color (see legend). The dielectric boundary for the complex is indicated by the orange dotted sphere. This figure was prepared with the QUANTA program (Molecular Simulations Inc., San Diego).

contributions to the electrostatic binding free energy for each ligand. The docking of the smallest ligand had the most favorable electrostatic binding free energy (−4.8 kcal/mol), which was diminished by only a tenth of a kcal/mol when the total charge was constrained to $-1e$, due in this case essentially only to changes in the contribution of the monopole term.

The multipole distribution for the 8-Å ligand with the total charge constrained to $-1e$ could be well represented by a set of only 11 point charges (Fig. 2B). These partial charges were fit from a 2-Å cubic grid and ranged from $-0.62e$ to $+0.42e$. The charge locations were restricted to be at least 1 Å from the ligand dielectric boundary because atom-centered point charges must be at least an atomic radius from the molecular surface. The fit between the point charges and the optimal multipole distribution was quite good, as gauged by the small difference (0.2 kcal/mol) between their binding free energies.

The point-charge distribution fit reveals a number of interesting features of ligand design. First, the barnase enzyme carried a net charge of $+1e$, as modeled here, and presented a large number of positive charges facing the ligand binding site (Fig. 2B). The ligand point-charge distribution (which was chosen to have a total charge of $-1e$, but unconstrained optimizations yielded total charges of $-1.16e$ to $-1.36e$) had seven negative partial charges ($-0.13e$ to $-0.62e$) arranged to make strong, favorable electrostatic inter-

Table 1. *Electrostatically optimized ligand-charge distributions*

Ligand radius (Å)	Total charge (e)	$\Delta G_{hyd,L}$[a] (kcal/mol)	$\Delta G_{int,L\text{-}R}$[b] (kcal/mol)	$\Delta G_{hyd,R}$[c] (kcal/mol)	$\Delta G_{binding}$[d] (kcal/mol)
8	−1.16	12.45	−24.91	7.67 (0.02)	−4.78
8	−1.00[e]	11.20	−23.56	7.67 (0.02)	−4.69
10	−1.30	14.74	−29.48	10.23 (0.02)	−4.51
11	−1.36	16.21	−32.42	11.71 (0.02)	−4.50

[a]Ligand desolvation contribution to electrostatic binding free energy.
[b]Screened electrostatic ligand–receptor interaction contribution.
[c]Receptor desolvation contribution. The values are the average of 10 translations on the finite-differences grid. The numerical uncertainty, given in parentheses, is twice the standard deviation of the mean.
[d]Overall electrostatic binding free energy.
[e]The total charge was constrained to $-1e$.

actions (−1.6 to −9.4 kcal/mol) with barnase. Second, the point-charge distribution also contained four positive partial charges. Each of these charges (+0.17*e* to +0.42*e*) made unfavorable interactions with barnase (+1.4 to +3.3 kcal/mol) and had an unfavorable desolvation penalty (+0.3 to +1.3 kcal/mol) upon binding. However, each of the four positive partial charges made improved electrostatic interactions within the ligand (−1.7 to −4.6 kcal/mol) due to reduced solvent screening upon binding. That is, none of these four positive partial charges made favorable interactions with the receptor, yet leaving any one out diminished the binding free energy by up to 0.8 kcal/mol and leaving all four out diminished binding by 4.7 kcal/mol. This binding enhancement can be traced to favorable intra-ligand interactions that are weaker in the unbound state due to screening by solvent and stronger in the bound state due to exclusion of solvent by the receptor. An equally valid description of the role of these four positive partial charges is that they reduce the ligand dehydration penalty by more than the repulsion they introduce with the receptor. Although favorable intramolecular electrostatic interactions are frequently observed within individual binding partners, their role in enhancing molecular association due to a reduced effective dielectric constant on binding has not generally been appreciated.

One limitation of the geometries used in this study is that the ligand binding site is enclosed within the receptor sphere, which may artificially reduce the receptor dehydration penalty. This penalty was recomputed for a number of geometries in which a tunnel of low dielectric was removed from the unbound receptor (tunnel diameters up to 16 Å were computed). This increased the receptor dehydration contribution by up to 1.1 kcal/mol, which resulted in a total electrostatic binding free energy that was still favorable by more than 3.5 kcal/mol.

Another result of the geometries chosen here is that ligand "atoms" could not approach receptor atoms closely enough to make direct hydrogen-bond interactions. This is largely a result of the use of a spherical ligand, although we purposely chose to maintain a substantial distance of closest approach between receptor point charges and the receptor surface (2.3 Å) so that the finite-difference computation of the receptor dehydration penalty could be done especially accurately (Gilson et al., 1988). The full consequences of using a spherical ligand and a spherical bound complex are unknown at this point. It seems likely that the general picture of electrostatic complementarity will be similar for flat, curved, and bumpy interfaces, but the details might very well be different in important ways. As the algorithm is extended to treat actual molecular shapes, it will be important to see whether electrostatically favorable binding energies can still be achieved.

In summary, we have applied a novel charge optimization scheme to compute electrostatic complements for the protein barnase and found that these charge distributions (1) produce favorable electrostatic binding free energies, and (2) can be fit to a relatively small number of partial charges that are of small enough magnitude (under 0.7*e* here) and of sufficient spacing (at least 2 Å apart) that molecules may be conceived with similar point-charge distributions. By using these electrostatic complements as a guide, it is possible that this scheme will prove useful in the design of tight-binding ligands, where significant improvements in electrostatic interactions may be realizable. The covalent and non-covalent constraints of chemistry may make it impossible to construct molecules as effective as the complements described here, but it may be possible to approach this optimum. The algorithm presented here is rapid (multipole distributions can be determined in about an hour on a laboratory workstation) because the ligand and complex were chosen to have spherical geometries, allowing analytic solutions to the electrostatic equations to be used (Lee & Tidor, 1997). Extensions to realistic molecular shapes are in progress; such computations are substantially slower due to the need to repeatedly solve the linearized Poisson–Boltzmann equation numerically for all members of a representative charge basis set, but we anticipate that they may be even more useful in molecular design studies. A further result of this study is the observation that strong intramolecular electrostatic interactions that are screened by solvent in the unbound state can be uncloaked by the exclusion of solvent in the bound state. The consequent enhancement can contribute substantially to binding free energy and is likely to be effective for arbitrarily shaped ligands and complexes.

Methods: The coordinates for the barnase receptor were taken from the crystal structure of the barnase–barstar complex (chain A of 1BRS in the Brookhaven Protein Data Bank; Bernstein et al., 1977; Buckle et al., 1994). Polar-hydrogen positions were built using the HBUILD facility of CHARMM (Brooks et al., 1983; Brünger & Karplus, 1988). Partial atomic charges for barnase were taken from the CHARMM PARAM19 parameter set (Brooks et al., 1983). In other work (ZSH, CV Sindelar, & BT, unpubl. obs.) we have found this parameter set to give results similar to the PARSE parameter set (Sitkoff et al., 1994). The spherical boundary defining the complex was centered at (−6.781, 66.603, −38.750), enclosing all point charges and radii in the barnase–barstar complex. Ligand spheres were chosen such that they fit wholly within the position of the barstar ligand. Ligand spheres of radius 8, 10, and 11 Å were centered at (−13.788, 72.885, −41.659), (−15.317, 74.256, −42.294), and (−16.0288, 74.8939, −42.5893), respectively. The optimal multipole expansion for each ligand was computed with the algorithm of Lee and Tidor (1997). Multipole expansions were evaluated to 58 poles ($l_{max} = l_{cut} = 57$). The ligand and complex interior dielectric constant (ϵ_{int}) and the solvent exterior dielectric constant (ϵ_{ext}) were 4 and 80, respectively.

The receptor dehydration penalty was computed numerically with the DELPHI computer program (Gilson et al., 1988; Gilson & Honig, 1988; Sharp & Honig, 1990) as the difference in the total electrostatic energy for the barnase charge distribution embedded in a low-dielectric region representing the bound and unbound state. The Poisson equation was solved (corresponding to zero ionic strength) with the low-dielectric ($\epsilon_{int} = 4$) region for the receptor defined as the spherical region for the complex with the spherical ligand region (and, for some computations, a tunnel leading from the spherical ligand to the exterior) subtracted. The remaining volume had $\epsilon_{ext} = 80$. Each calculation was run using a finite-difference grid consisting of 131 points in each of the three Cartesian directions. A dual-level focusing approach filling 23% and then 92% of the grid was used. The results presented are the average of ten translations at the fine grid spacing (corresponding to 0.52 Å per grid unit).

Point charges were fit to the multipole distribution (up to $l = 27$) on a 2-Å Cartesian grid of point-charge centers using singular-value decomposition (Press et al., 1992; Strang, 1993). Iterative cycles were used in which point-charge locations were successively removed from the basis set to achieve a good fit with a small number of point charges. In each iteration singular-value decomposition was used to perform a least-squares fit of the multipoles to the current basis set, point charges with very small magnitudes

were removed from the basis, and the multipoles were re-fit to the new basis.

Acknowledgments: The authors thank Barry Honig for making the DELPHI program package available. We also thank Barry Honig, Wayne L. Hubbell, Peter S. Kim, Robert T. Sauer, Lawrence J. Stern, Gilbert Strang, and James R. Williamson for helpful discussions, comments, and suggestions. This work was supported by the National Institutes of Health (GM47678 and GM56552) and the MIT Science Partnership Fund.

References

Bernstein FC, Koetzle TF, Williams GJB, Meyer EF Jr, Brice MD, Rodgers JR, Kennard O, Shimanouchi T, Tasumi M. 1977. The Protein Data Bank: A computer-based archival file for macromolecular structures. *J Mol Biol 112*:535–542.

Brooks BR, Bruccoleri RE, Olafson BD, States DJ, Swaminathan S, Karplus M. 1983. CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. *J Comput Chem 4*:187–217.

Bruccoleri RE, Novotny J, Davis ME, Sharp KA. 1997. Finite difference Poisson-Boltzmann electrostatic calculations: Increased accuracy achieved by harmonic dielectric smoothing and charge antialiasing. *J Comput Chem 18*:268–276.

Brünger AT, Karplus M. 1988. Polar hydrogen positions in proteins: Empirical energy placement and neutron diffraction comparison. *Proteins Struct Funct Genet 4*:148–156.

Buckle AM, Schreiber G, Fersht AR. 1994. Protein–protein recognition: Crystal structure analysis of a barnase–barstar complex at 2.0-Å resolution. *Biochemistry 33*:8878–8889.

Gilson MK, Honig B. 1988. Calculation of the total electrostatic energy of a macromolecular system: Solvation energies, binding energies, and conformational analysis. *Proteins Struct Funct Genet 4*:7–18.

Gilson MK, Sharp KA, Honig BH. 1988. Calculating the electrostatic potential of molecules in solution: Method and error assessment. *J Comput Chem 9*:327–335.

Hendsch ZS, Tidor B. 1994. Do salt bridges stabilize proteins? A continuum electrostatic analysis. *Protein Sci 3*:211–226.

Lee L-P, Tidor B. 1997. Optimization of electrostatic binding free energy. *J Chem Phys 106*:8681–8690.

Lounnas V, Wade, RC. 1997. Exceptionally stable salt bridges in cytochrome P450cam have functional roles. *Biochemistry 36*:5402–5417.

Misra VK, Hecht JL, Sharp KA, Friedman RA, Honig B. 1994a. Salt effects on protein-DNA interactions: The λcI repressor and EcoRI endonuclease. *J Mol Biol 238*:264–280.

Misra VK, Sharp KA, Friedman RA, Honig B. 1994b. Salt effects on ligand–DNA binding: Minor groove binding antibiotics. *J Mol Biol 238*: 245–263.

Novotny J, Bruccoleri RE, Davis M, Sharp KA. 1997. Empirical free energy calculations: A blind test and further improvements to the method. *J Mol Biol 268*:401–411.

Novotny J, Sharp K. 1992. Electrostatic fields in antibodies and antibody/antigen complexes. *Prog Biophys Mol Biol 58*:203–224.

Paul CH. 1982. Building models of globular protein molecules from their amino acid sequences. I. Theory. *J Mol Biol 155*:53–62.

Press WH, Teukolsky SA, Vetterling WT, Flannery BP. 1992. *Numerical recipes in C: The art of scientific computing*, 2nd ed. Cambridge: Cambridge University Press.

Sharp KA. 1996. Electrostatic interactions in hirudin–thrombin binding. *Biophys Chem 61*:37–49.

Sharp KA, Honig B. 1990. Electrostatic interactions in macromolecules: Theory and applications. *Annu Rev Biophys Biophys Chem 19*:301–332.

Shen J, 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.

Sitkoff D, Sharp KA, 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, Massachusetts: Wellesley-Cambridge Press.

Tanford C, De PK, Taggart VG. 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.

Waldburger CD, Schildbach JF, Sauer RT. 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, Hermans J. 1996. Energetic decomposition of the α-helix–coil equilibrium of a dynamic model system. *Biopolymers 39*:479–489.

Wimley WC, Gawrisch K, Creamer TP, White SH. 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.

Xu D, Lin SL, Nussinov R. 1997. Protein binding versus protein folding: The role of hydrophilic bridges in protein associations. *J Mol Biol 265*: 68–84.

Yang A-S, Honig B. 1995. Free energy determinants of secondary structure formation: I. α-helices. *J Mol Biol 252*:351–365.