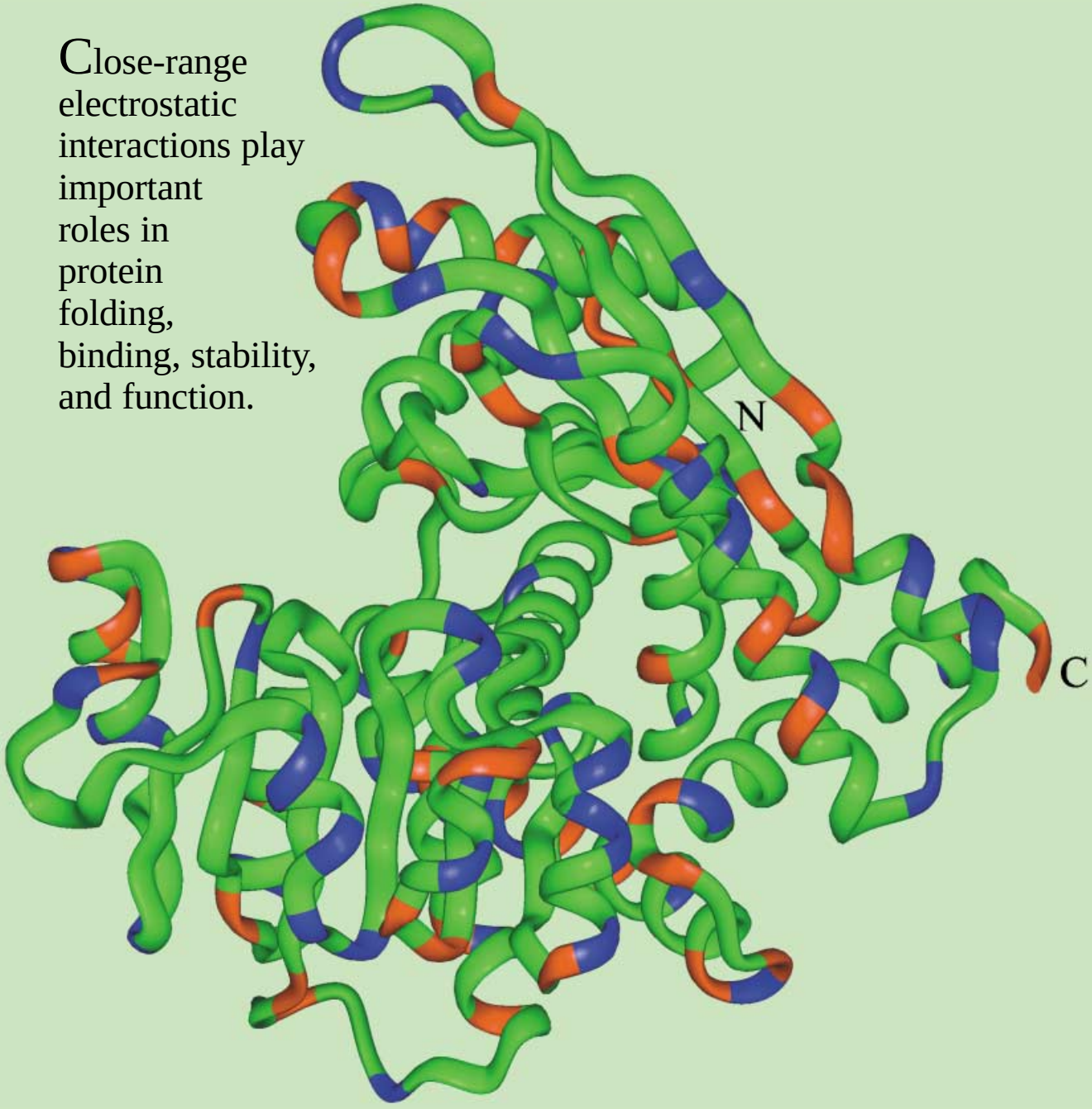
Close Range Electrostatic Interactions in Proteins

Close-range electrostatic interactions play important roles in protein folding, binding, stability, and function.



As an example, a ribbon diagram of a monomer of the thermophilic enzyme glutamate dehydrogenase from *P. furiosus* is shown. Red indicates residues with positively charged side chains. The residues with negatively charged side chains are shown in blue. Other residues are shown in green.

1439-4227/02/03/07 $ 20.00+.50/0 ChemBioChem **2002**, 3, 604 – 617

Close-Range Electrostatic Interactions in Proteins

Sandeep Kumar[c] and Ruth Nussinov*[a, b]

Two types of noncovalent bonding interactions are present in protein structures, specific and nonspecific. Nonspecific interactions are mostly hydrophobic and van der Waals. Specific interactions are largely electrostatic. While the hydrophobic effect is the major driving force in protein folding, electrostatic interactions are important in protein folding, stability, flexibility, and function. Here we review the role of close-range electrostatic interactions (salt bridges) and their networks in proteins. Salt bridges are formed by spatially proximal pairs of oppositely charged residues in native protein structures. Often salt-bridging residues are also close in the protein sequence and fall in the same secondary structural element, building block, autonomous folding unit, domain, or subunit, consistent with the hierarchical model for protein folding. Recent evidence also suggests that charged and polar residues in largely hydrophobic interfaces may act as hot spots for binding. Salt bridges are rarely found across protein parts which are joined by flexible hinges, a fact suggesting that salt bridges constrain flexibility and motion. While conventional chemical intuition expects that salt bridges contribute favorably to protein stability, recent computational and experimental evidence shows that salt bridges can be stabilizing or destabilizing. Due to systemic protein flexibility, reflected in small-scale side-chain and backbone atom motions, salt bridges and their stabilities fluctuate in proteins. At the same time, genome-wide, amino acid sequence composition, structural, and thermodynamic comparisons of thermophilic and mesophilic proteins indicate that specific interactions, such as salt bridges, may contribute significantly towards the thermophilic – mesophilic protein stability differential.

KEYWORDS:

electrostatic interactions · flexibility · protein folding · protein structures · salt bridges

1. Introduction

Specific interactions play important roles in protein folding, binding, flexibility, stability, and function. Here we focus on close-range electrostatic interactions. Throughout this review, we refer to a pair of oppositely charged residues (Asp or Glu with Arg, Lys, or His) as an ion pair. An ion pair is defined as a salt bridge if the centroids of the side-chain charged-group atoms in the residues lie within 4.0 Å of each other and at least one pair of Asp or Glu side-chain carbonyl oxygen and side-chain nitrogen atoms of Arg, Lys, or His are also within this distance. Such a definition ensures that the oppositely charged residues in a salt bridge are spatially close and interact.[1, 2]

2. The Role of Specific Interactions in Protein Folding/Binding

The initial fast step in protein folding is the hydrophobic collapse, which leads to the molten globule (MG) state. In the next step, specific interactions are optimized; this shifts the equilibrium toward the native state. Specific electrostatic interactions play a critical role in reaching the native fold. A good example is the α-lactalbumin (α-LA). If the Ca^{2+} ion is removed, α-LA is observed to be in its molten globule state. At low pH values (around 2.0) α-LA also exists in the molten globule state, referred to as the acid state (the A state).[3a] Other examples include nucleic acid binding proteins, which on their own are frequently disordered. However, upon binding to the DNA (or RNA) they become stabilized.[3b] Alternatively, in some cases, depending on the location and charge, local disorder may be observed. A nice example is that of the adenine binding domain in dihydrofolate reductase. By itself, this domain is unstable. However, when bound to other domains, or to its NADPH nucleotide cofactor, it is stabilized.[3c]

The hierarchical model is among the several models proposed for protein folding. In this model folding initiates locally. Local folded elements associate in a step-wise fashion to yield the native structure.[4] Formation of salt bridges is consistent with hierarchical protein folding. Recently, we have analyzed a database of 222 salt bridges from 36 nonhomologous monomeric protein crystal structures solved to high resolution (1.6 Å or better). For approximately half of these salt bridges, the number of intervening residues was ≤ 10. Hence, the oppositely charged residues that form salt bridges are often near each other in the amino acid sequences of the proteins.[2a] Many charged residues that form salt bridges have helical conformations. In α

[a] *Prof. R. Nussinov*
Intramural Research Support Program—SAIC
NCI-Frederick, Building 469, Room 151
Frederick, MD 21702 (USA)
Fax: (+1) 301-846-5598
E-mail: ruthn@ncifcrf.gov

[b] *Prof. R. Nussinov*
Sackler Institute of Molecular Medicine
Department of Human Genetics and Molecular Medicine
Sackler School of Medicine
Tel Aviv University
Tel Aviv, 69978 (Israel)

[c] *Dr. S. Kumar*
Laboratory of Experimental and Computational Biology
NCI-Frederick, Building 469, Room 151
Frederick, MD 21702 (USA)

helices, negatively charged residues, Asp and Glu, are favored at the N terminus, and at positions in the first turn. Positively charged residues, Lys and His, are favored at or near the C terminus. These residues have an important role in helix termination and in the formation of helix-capping motifs. In the middle of α helices, oppositely charged residue pairs tend to have greater propensities to occur at adjacent positions (i, $i \pm 1$) or on the same face (i, $i \pm 3,4$). However, the middle positions of α helices are neutral (neither favored nor avoided) for most charged residues and the average propensities for oppositely and like charged residue pairs to occur at (i, $i \pm 1,2,3,4$) positions are similar.[5] Nevertheless, in surveys of α-helical structures,[6](i, $i \pm 3,4$) salt-bridge pairs are observed. In α-helical peptides, salt bridges and their networks stabilize the helical structure to varying extents.[7] There have been relatively fewer studies on salt-bridge formation in β sheets. However, these studies also show that salt bridges stabilize the β sheets to similar extents as the α helices.[8] These studies also indicate that salt bridges are often formed in protein secondary structural elements.

A study on the conservation of salt bridges in different protein families has shown that buried salt bridges are more likely to be conserved than the surface exposed ones.[9] Electrostatic interactions are optimized locally and appear to evolve in the direction of avoiding electrostatic repulsions. Nevertheless, buried unsatisfied charges exist, and their destabilizing effect has been controversial. While some studies have suggested that such charges may considerably destabilize protein structures, others have suggested that their apparent destabilizing effect is alleviated by local unfolding, reorientation of backbone charged groups, and penetration of water. Interestingly, with regard to optimization of electrostatic interactions, formation of additional salt bridges is secondary.[10] In addition to pair-wise salt bridges, more complex associations of the charged residues in proteins are also observed. Musafia et al.[11] have carried out a statistical analysis of complex salt bridges (involving at least three charged residues) in 94 proteins. They find that complex salt bridges are often formed, a fact indicating a tendency of the charged residues to form cooperative networks. These salt bridge networks are more often found at subunit – subunit interfaces. Arginine, which contains guanidium group in its side chain, acts as a connector in such networks.

Oliveberg and Fersht[12] have developed a method to study transition-state structures in the folding pathway, by using the proton-titration behavior of charged protein residues. They find that a partially buried salt bridge Arg 69 – Asp 93 in Barnase is formed early in the folding process. Electrostatic interactions may also be important kinetically. A triple mutant of the Arc repressor dimer which replaces a triad of charged residues with hydrophobic amino acids has been shown to fold faster than the wild type. Hence, formation of buried polar interactions may be a slow step in protein folding.[13] Theoretical calculations indicate that electrostatic interactions affect unfolding rates of thermophilic and mesophilic rubredoxins.[14] Torshin and Harrison[15] have suggested that centroids of positive and negative charges may match protein folding cores detected by hydrogen-exchange experiments.

Recently, Nussinov and co-workers have proposed a hierarchical model for protein folding. In this model, folding initiates locally. First, building blocks consisting of 15 or more residues are formed. A building block consists of a single secondary structure or a set of contiguous secondary structures (super-secondary structures). The conformation of a building block seen in the native protein structure may (or may not) be same as the most

***Ruth Nussinov**, born in 1943, is a professor in the Department of Human Genetics, School of Medicine, Tel Aviv University, Israel, and a senior scientist at the National Cancer Institute—Frederick, Maryland, USA. She received her BSc degree in Microbiology at Washington University, Seattle, USA, in 1967, and her MSc in Biochemistry in 1968 from Rutgers University. She received her PhD in Biochemistry from Rutgers University in 1977. Dr. Nussinov was a fellow at the Weizmann Institute and a Visiting Scientist in the Chemistry Department at the University of California, Berkeley, and in the Biochemistry Department at Harvard University. She joined the Medical School at Tel Aviv in 1985 as an associate professor, and in 1990 became a full professor. Her association with National Institutes of Health began in 1983, first with the National Institute of Child Health and Human Development, and, since 1985, with the National Cancer Institute. She has authored or coauthored more than 200 scientific papers. Until 1990, her papers addressed nucleic acid sequence and structure and protein – nucleic acid interactions. In 1990 she switched to proteins; her research currently focuses on protein folding and binding.*

***Sandeep Kumar**, born in 1968, has been a postdoctoral visiting fellow at Laboratory of Experimental and Computational Biology, the National Cancer Institute—Frederick, Maryland, USA, since January 1998. He obtained his BSc(Hons) Physics degree from the University of Delhi, India, in 1989. But he then became more inclined towards biology and, hence, obtained an MSc in Molecular Biology and Biotechnology from the G.B. Pant University of Agriculture and Technology, Pant Nagar, India, in 1992. In the same year, he joined the Molecular Biophysics Unit at the Indian Institute of Science, Bangalore, India, for his doctoral studies. In 1998, he was awarded a PhD degree in recognition of his work on sequence and structural relationships in α helices. In his postdoctoral work, he has focused on issues related with protein stability and the role of close-range electrostatic interactions in proteins. His research interests focus on understanding sequence – structure – function relationships in proteins. So far, he has been an author or coauthor of 25 technical papers.*

populated conformation of the corresponding isolated peptide fragment in solution. Mutual conformational selection leads to a combinatorial assembly of the building blocks into hydrophobic folding units. A hydrophobic folding unit contains a sufficiently large buried hydrophobic core and is capable of independent thermodynamic existence. One or more hydrophobic folding units associate to form domains. Domains then associate into subunits and subunits associate to yield the protein quaternary structure. Dissection of proteins into their anatomical parts yields information on their most likely protein folding pathways.[16] Analysis of salt bridges and hydrogen bonds shows that most of these interactions are formed within building blocks, hydrophobic folding units, domains, or subunits, rather than across these,[17] consistent with the model.

Protein folding and binding are similar processes. Their difference is in the absence of chain connectivity.[18a] The energy landscape of the bound molecule can be described by fusing the folding funnels of the constituent unbound molecules.[18d] Just like in protein folding, close-range electrostatic interactions play important roles in protein – protein binding. Tsai et al.[16a,b, 19] have created a nonredundant data set of protein – protein interfaces. Analyses performed on 362 structurally unrelated protein – protein interfaces and 57 symmetry-related oligomeric interfaces indicate that a higher proportion of charged and polar residues are buried at protein – protein interfaces than in the protein core. However, protein – protein interfaces are poorer in charged residues than the protein surface. The hydrophobic effect measured in terms of the buried nonpolar surface area is smaller at the interfaces than in the protein cores. Although variable, nevertheless, the hydrophobic effect is dominant in the majority of protein – protein interfaces, as in protein cores. Analysis of protein – protein interfaces shows more hydrogen bonds and salt bridges across the interfaces. However, the geometries of these interactions are less optimal at the interfaces than in the cores, and water mediates such interactions in the interfaces to a greater extent. Salt bridges formed across protein – protein interfaces are mostly stabilizing and the number of such interactions is correlated with the binding free energy.[20]

Protein – protein interfaces often contain "hot spots" for binding. The residues forming hot spots contribute more towards the free energy of binding than the residues outside these hot spots. The role of the surrounding hydrophobic residues is to occlude bulk water. Analysis of alanine-scanning mutants has shown that hot spots are enriched in tryptophan, tyrosine, and arginine. An analysis of 11 families of interfaces showed that although overall the binding sites are hydrophobic, they contain conserved polar residues hot spots.[21] Electrostatic complementarity between the individual molecules further optimizes binding.[22] Inclusion of electrostatic terms in the binding free energy function of the molecular docking programs results in a better performance. If electrostatics-based filters are used in screening the docking results for protein – protein complexes, the chances of finding the native or near-native solutions are improved.[23] This however is the case only if the initial solutions are already in near-native positions. On the other hand, if solutions submitted to electrostatic calculations are far from the binding sites, such calculations do not provide efficient filtering. Recently, electrostatic interactions have also been implicated in precipitation of soluble proteins upon aggregation induced by amyloids.[24]

3. Specific Interactions and Protein Flexibility

Proteins and protein – protein complexes show a continuous spectrum of flexibility. Formation of specific interactions, such as close-range electrostatic interactions, appears to shift the equilibrium toward the native state[25] and to constrain backbone flexibility.[26] A molecular dynamics study on cytochrome b_5 (cyt b_5)[27] has indicated a periodic dynamic behavior of the cyt b_5 surface. A cleft is formed, which enables access to the prosthetic group heme through a hydrophobic channel. A salt bridge and a disulfide bond introduced into mutants prevented the opening of this cleft.

Proteins exhibit two types of flexibilities, systemic and segmental.[17a] Systemic flexibility refers to small-scale fluctuations in side-chain and main-chain atoms of the proteins in their native states. Systemic flexibility is distributed throughout the protein. The time-scale of systemic protein flexibility is fast. On the other hand, segmental flexibility refers to the motion of one part of the protein molecule with respect to the other in response to a molecular event related to the protein function. The motion is mostly restricted to a small segment of the protein, such as a hinge. Segmental protein flexibility has slower time scales. Protein movements due to segmental mobility are much larger than the movements due to systemic flexibility.

Systemic protein flexibility can be studied by comparing conformational isomers of the proteins. The ensembles of protein conformations can be obtained either by computer simulations or by study of the nuclear magnetic resonance (NMR). Both have advantages and disadvantages. In simulations, the computational resources and time required to sufficiently sample the conformational space of the proteins are still quite expensive. A force field to accurately simulate the behavior of the protein in solution is essential. In NMR spectroscopy, it is difficult to separate artifacts due to the structure calculation protocol from genuine protein motion. However, the sampling of protein conformational space obtained by simulations and by NMR measurements usually shows a good qualitative agreement. Protein flexibility can also be judged from atomic B factors in the protein crystal structures. The availability of multiple protein structures for the same protein is valuable in studying protein flexibility. In our studies,[2b,c] we have used NMR conformer ensembles and protein crystal structures to study systemic protein flexibility. Figure 1 provides an example of protein flexibility and illustrates individual conformers in the NMR conformer ensemble of the *Escherichia coli* chemotaxis protein CheY as well as its crystal structures. The fluctuations in atomic coordinates of the charged residues due to systemic protein flexibility lead to fluctuations in their locations in the protein and in the geometries of the close-range interactions formed by the charged residues and consequently to fluctuations in their stabilities. Our studies show that, due to systemic protein flexibilities, the salt bridges seen in the protein crystal

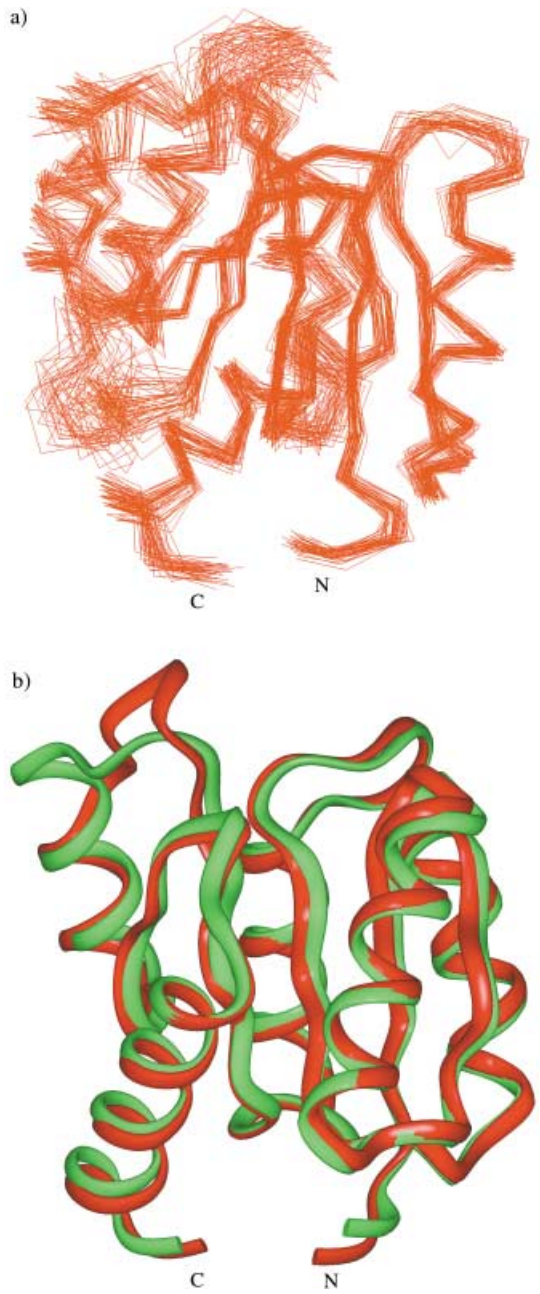


Figure 1. *An example of protein flexibility. a) 46 individual conformers in the NMR conformer ensemble of* E. coli *chemotaxis protein, CheY. The atomic coordinates of the individual conformers were obtained from the Protein Data Bank (PDB)[47] entry 1cey. b) Ribbon diagrams showing the superposition of two crystal structures of CheY. The Mg^{2+}-bound form of CheY (PDB entry 1chn) is shown in red and the Mg^{2+}-deficient form (PDB entry 3chy) is shown in green. Close-range electrostatic interactions relate to protein flexibility.*

structures may break and reform easily in different conformers of the protein. The identities of the charged residue pairs forming the salt bridges may also fluctuate across different conformers.

Segmental protein flexibility can be studied by using "open" and "closed" (active or inactive) conformations of proteins. Here, using crystal structures containing open and closed conformations, we have studied the interactions between the moving parts that are joined by hinges. The moving part can be a fragment (building block), a domain, or a subunit. Our analysis shows that electrostatic interactions are limited between moving protein parts. However, substantial nonpolar, buried surface area could still be present between the two parts both in open and in closed conformations.[17b]

Protein flexibility is observed in protein – protein, protein – ligand, enzyme – substrate, and antigen – antibody binding. These are frequently assigned into two binding modes, lock-and-key and induced-fit.[28] Complexes of hen egg white lysozyme (HEL) with anti-HEL antibodies have been studied for the role of electrostatic interactions across the antigen – antibody interface. Formation of salt bridges by Lys 97 (HEL) with Asp 32 and Asp 96 of (HyHEL-10 VH) contribute to the specificity of the antigen – antibody association and entropically stabilize the complex.[29] Residues in the catalytic triad of *Rhizomucor miehei* lipase are involved in a larger electrostatic network around the active site. This network stabilizes the active site geometry and is conserved in the lipase family.[30] The presence of electrostatic interactions across the protein – protein interface results in specificity in complex formation. However, the structural plasticity facilitated by their limited presence also serves useful purposes. The recognition of several ligands by a single molecule has important immunological consequences. For example, the Fc fragment of IgG binds to many ligands, including protein A, protein G, rheumatoid factor, and neonatal Fc receptor.[31]

Another issue in protein flexibility (rigidity) relates to thermal adaptation of proteins.[32] We found that salt bridges and their networks increase in thermophilic proteins as compared to their mesophilic homologues. In one family, the homologous thermophilic and mesophilic glutamate dehydrogenases, there is a greater formation of salt bridges and their networks around the active site of the thermophilic glutamate dehydrogenase. This observation appears reasonable. The thermophilic protein has a greater need to protect its active site from the larger disorder at high temperatures.[32] Consistently, we have also compared the locations of salt-bridge-forming residues in the crystal structures of citrate synthase from thermophilic, mesophilic, and psychrophilic organisms. Thermophilic and psychrophilic citrate synthases are more similar to each other in a sequence- and structure-wise manner and contain a larger number salt bridges than their mesophilic homologue. However, in the thermophilic citrate synthase the salt bridges and their networks are located closer to the active site, while in the psychrophilic citrate synthase they are located further from the active site.[32d]

4. Free Energy Contribution of Electrostatic Interactions towards Protein Stability

The electrostatic description of proteins is considerably more than a list of close-range electrostatic interactions. Long-range electrostatic interactions also play an important role in the stability of proteins and in protein – protein complexes. Furthermore, the total electrostatic energy calculations also include terms for self-energy and local polarity. Protein relaxation and reorganization also affects the charge – charge interactions and the dielectric constants. Several methodologies for computing

the overall free energy contribution due to electrostatics have been developed.[33] Such calculations are important for relating protein structure with function. The focus of these is to calculate pK_a shifts in the ionizable side chains in catalytically important residues and to compute redox potentials and the electrostatic contribution to binding free energies for protein – protein, protein – ligand, enzyme – substrate, and antigen – antibody interactions. For example, see the excellent papers by Warshel and co-workers.[33a,d, 34] Work in this direction is also the focus in many other groups.[35]

During protein folding/binding, the charged groups in the proteins desolvate as their environments change from aqueous (water) to largely nonpolar solvent. This desolvation process is energetically unfavorable and the charged residues pay desolvation energy penalties as the protein folds. Consistently, burial of polar groups in proteins results in a decrease in the heat capacity change between native and denatured states.[36] Even though biochemical intuition suggests that electrostatics is stabilizing towards folded proteins or bound complexes, failure to pay the desolvation energy penalty may result in a net destabilizing effect.[22a,e, 37] Optimization of charge – charge interactions leads to substantial improvement in binding electrostatics.[22b] In the GCN4 leucine zipper, binding of the two helices improves the intrahelical electrostatic interactions, although their overall contribution remains destabilizing.[22a] Estimates of the desolvation penalty paid by the charged residues and screening of charge – charge interactions in the protein medium depend upon the value of the dielectric constant used for the protein. In classical electrostatics, the media are assumed to be homogeneous and thus have a single dielectric constant. However, the proteins are nonhomogeneous and different regions of the proteins have different polarizabilities. For example, the charges at or near the protein surface may experience a different dielectric constant than the charges buried in the protein core. Estimates of the effective dielectric constant experienced by a salt bridge and the energetic contributions by the salt bridge need to take into account the effect of protein relaxation and reorganization of the polar groups.[33d, 34j] Hence, the statement that close-range electrostatic interactions such as salt bridges are stabilizing towards proteins is controversial. There is considerable theoretical and experimental evidence both in favor and against this statement.[1b, 2, 17, 20b, 22, 33d, 34j–n, 37, 38] In many instances, these interactions contribute only marginally towards protein stability (or destability).[6a, 39] Warshel and co-workers[34k–n] have observed that ion pairs may not be stabilizing in a low dielectric environment and reorganization of the polar environment in the protein may help to stabilize them. Similar observations have also been made by others.[37, 38h]

To understand the electrostatic properties of a protein in aqueous solution, one needs an accurate description of the protein (solute), the water (solvent), and the interaction between the protein and the water. There are different methods that can be used to provide this description. One of the most commonly used is the continuum electrostatics approach. It is based on classical electrostatics. In the continuum electrostatics approach, the protein is described in atomic detail, but water is only described in terms of its bulk properties.[40] Essentially, we follow the method described by Hendsch and Tidor in an excellent paper.[37] This method calculates the free energy of a salt bridge relative to a computer mutation of the salt-bridging residues to their hydrophobic isosteres. A hydrophobic isostere of a charged residue is the charged residue with its side-chain functional-group atomic charges set to zero. This method has been widely used in the literature.[1b, 2, 17, 32b, 37, 39a]

The total electrostatic free energy of a salt bridge, $\Delta\Delta G_{tot}$, can be partitioned into three terms. $\Delta\Delta G_{dslv}$ is the sum of the unfavorable desolvation penalties incurred by the individual salt bridging residues due to the change in their environment from water to the protein interior. $\Delta\Delta G_{brd}$ is the favorable bridge energy due to the interaction of charged side-chain functional groups with each other. $\Delta\Delta G_{prt}$ represents the interaction of the salt-bridging side chains with the charges in the rest of the protein. The stability of a salt bridge can also be measured by the association energy, $\Delta\Delta G_{assoc}$. It refers to the desolvation of the whole salt bridge and the interaction between the salt-bridging side chains, but it does not consider the interaction of the salt bridge with the rest of the protein. Though not part of $\Delta\Delta G_{tot}$, $\Delta\Delta G_{assoc}$ is useful, since it measures the free energy change associated with bringing two charged residues from a solvent of high dielectric medium into a low dielectric protein medium without regard to the other changes in the protein. Hence, the association between the charged residues forming the salt bridge would be stabilizing if the interaction between the charged residues is strong enough to overcome the desolvation energy penalty paid by the salt bridge.

Electrostatic calculations involve a numerical solution of the Poisson – Boltzmann equation with a finite difference approximation. These calculations can be performed by using the DELPHI package developed by Honig and co-workers.[41] The use of the PARSE3 set of partial atomic charges and radii allows the experimental data to be reproduced for a wide range of small organic molecules and ions representing amino acid side chains.[42] In each case, the protein molecule is mapped on to a three-dimensional grid. A rough calculation with the molecule occupying a smaller volume of the grid provides boundary conditions for the more focused calculations.[43] In order to improve the accuracy of these calculations, it is recommended that the protein structure is appropriately optimized.[44]

We have identified 222 nonequivalent salt bridges in 36 nonhomologous monomeric protein crystal structures.[2a] This database captures the salt bridges in all structural contexts. Approximately one third of the salt bridges are buried in the protein core, that is, both salt-bridging charged residues are $\leq 20\%$ solvent exposed. The remaining two-thirds are classified as solvent exposed. Less than one-tenth of the salt bridges are networked, to form four triads and three tetrads. The remaining are isolated salt bridges. Continuum electrostatic calculations show that most ($\approx 86\%$) of the salt bridges are stabilizing with respect to their hydrophobic isosteres. These include most of the buried salt bridges. Most stabilizing salt bridges contain at least one hydrogen bond between the atoms in their side-chain charged groups.

The electrostatic strengths (electrostatic free energy contribution towards protein stability) of the 222 salt bridges show a wide variation. On average, the electrostatic strength of the salt bridges appears to be determined by two terms with opposite signs, namely, $\Delta\Delta G_{dslv}$ and $\Delta\Delta G_{brd}$. The term $\Delta\Delta G_{prt}$ has a secondary, but not minor, contribution. Hence, the interaction between side-chain charged groups of the salt-bridging residues is primarily responsible for overcoming the unfavorable desolvation energy paid by the salt-bridging residues. However, in the case of the networked residues, both $\Delta\Delta G_{brd}$ and $\Delta\Delta G_{prt}$ have similar magnitudes. All networked salt bridges are stabilizing. The stabilizing salt bridges pay smaller desolvation energy penalties and have stronger $\Delta\Delta G_{brd}$ and $\Delta\Delta G_{prt}$ terms than the destabilizing ones. The hydrogen-bonded salt bridges are stronger than the non-hydrogen-bonded ones.

The buried salt bridges in our database are more stabilizing than the exposed ones, although they pay higher desolvation energy penalties. This observation runs counter to the currently accepted view in the literature that burial of charged residues destabilizes the protein.[13, 38e–h, 45] However, recent evidence suggests that buried charged residues in the proteins occur more frequently than previously thought. Furthermore, a single buried charged residue may also be stabilizing.[46] The rationale for these observations is as follows. Even though a charged residue pays a large desolvation energy penalty due to burial in the low dielectric protein core, its interaction with the other charged residues and (or) with the protein-backbone charged atoms is also much stronger due to lesser solvent screening. This is due to the difference in the dielectric constants of water and the protein core. If we take the value of dielectric constant to be 80 for water and 4 for the protein core, the electrostatic interaction between two oppositely charged residues would be twenty times stronger in the protein core than on the protein surface, provided that otherwise nothing has changed.

Geometrical orientation of the salt-bridging side-chain charged groups (salt-bridge geometry) plays a critical role in determining salt-bridge stability. Salt bridges with favorable geometries are likely to be stabilizing anywhere in the protein structure. Salt-bridge geometry is characterized by (1) the distance (r) between the side-chain charged-group centroids and (2) the angular orientation (θ) of the side-chain charged groups in the two salt-bridging residues. This is the angle between two unit vectors. Each unit vector joins a C^{α} atom and a side-chain charged-group centroid in a charged residue. Figure 2 shows a polar plot of salt-bridge geometries in our database. Since all the salt bridges have good geometries, a majority of them is stabilizing.

A consequence of the fluctuations in atomic coordinates of the protein, due to flexibility, is that the charged residues within the pairs move with respect to each other in different conformers of the protein. The location of the charged residues in the protein is also affected. Hence, it can be expected that systemic protein flexibility will affect the electrostatic strength of the interacting charged residues. Using the NMR conformer ensembles, we have analyzed the electrostatic strengths of intra- and interhelical ion pairs and a five-residue ion-pair network (IPN-5) in individual conformers in the NMR ensemble and in the average energy-minimized structure of the c-Myc-Max leucine zipper (Protein Data Bank (PDB)[47] entries 1a93 and 2a93). All the ion pairs and the ion-pair network show extensive conformer-dependent fluctuations in their electrostatic strengths and geometries, as well as the location of the charged residues. However, the most surprising observation was that each ion pair, as well as the ion pair network (IPN-5), interconverted between being stabilizing and being destabilizing.[2b] This indicates that the overall electrostatic contribution of the ion pairs toward proteins is conformer-population dependent.

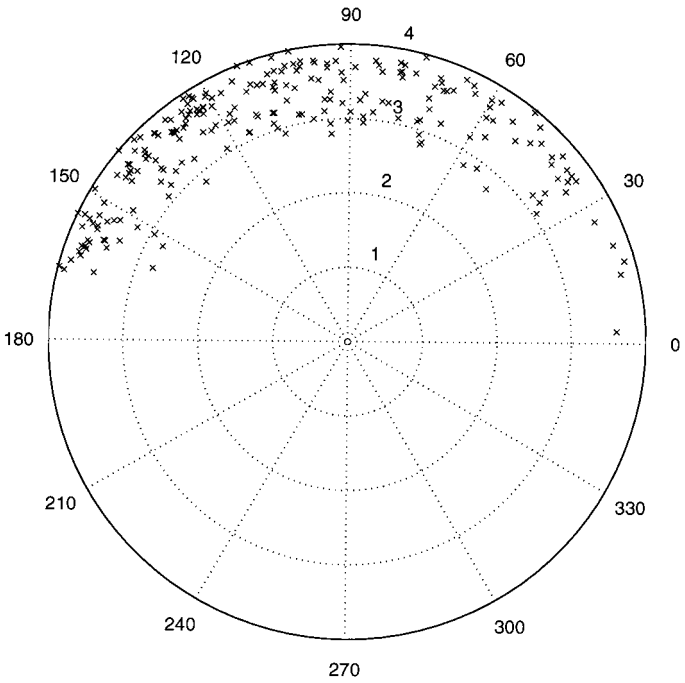


Figure 2. *Polar plot showing the orientation of salt-bridging side chains in 222 salt bridges. The distance between charged-group centroids in the salt bridges is plotted along the radius of the plot. The angle is between the two unit vectors joining C^{α} atoms with their respective charged-group centroids for the salt-bridging residues. The observed data points are indicated by the symbols 'x'. The salt bridges in our database have good geometry. Hence, most of them are stabilizing.*

To probe this issue further, we have performed an extensive analysis of electrostatic strengths of 22 ion pairs in NMR conformer ensembles of 11 different proteins.[2c] These ion pairs form salt bridges in the crystal structures, in the average energy-minimized structures, or in the "most representative" conformer of the proteins. We again found conformer-dependent fluctuation in the electrostatic contributions of these ion pairs. Most of the ion pairs interconverted between being stabilizing and destabilizing in different protein conformers. The observed fluctuations reflected the variabilities in the ion-pair geometries as well as the location of the charged residues in different conformers of the proteins. We also found that the salt bridges seen in the crystal structures could easily break and reform in different conformers. Ion pairs which do not form salt bridges in the crystal structures were often seen to form salt bridges in conformers of the ensembles. Hence, both the identity of the charged residues that form close-range electrostatic interactions

and the electrostatic strengths of these interactions are conformer-population dependent. These observations appear reasonable if we realize that the energy landscapes of the proteins are dynamic and shift in response to the changes in the protein's environments.[18] These fluctuations are seen not only in NMR conformers but also in different crystal structures of the same protein.

The fluctuations observed in ion-pair geometries and in their electrostatic strengths are interrelated. The availability of extensive data on ion-pair geometries and their electrostatic strengths provides an opportunity to study this relationship. We have divided the database of ion pairs according to their geometries into salt bridges, N–O bridges, and longer range ion pairs. Salt bridges are those ion pairs which follow both geometrical criteria in our definition of a salt bridge.[2] N–O bridges follow only one, that is, they have at least a pair of side-chain functional-group nitrogen and oxygen atoms within 4 Å distance, but the side-chain functional-group centroids are >4.0 Å apart. The longer-range ion pairs are those which violate both criteria of the salt bridge definition. Figure 3 shows an example of these three types of ion pairs. The geometrical orientation of the side-chain charged groups in the ion-pairing residues is most favorable in salt bridges, less favorable in the N–O bridges, and most unfavorable in the longer-range ion pairs. The electrostatic strengths of the ion pairs show the same trend. The electrostatic interaction between the ion-pair residues is the strongest when they form salt bridges, weaker when they form N–O bridges, and the weakest when they form longer-range ion pairs. Most of the salt bridges are stabilizing with respect to their hydrophobic isosteres. A considerably reduced majority of N–O bridges are also stabilizing. A majority of the longer-range ion pairs are destabilizing.[48]

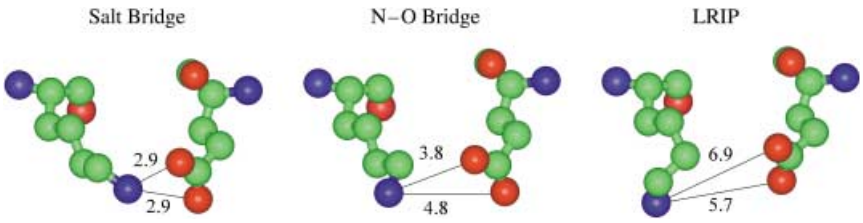


Figure 3. *Examples of different geometries of ion pairs in proteins described in the text. a) Salt bridge, b) N–O bridge, and c) longer-range ion pair (LRIP) are shown for Lys 4 – Glu 15 in the NMR conformer ensemble (⟨1GB1⟩) of protein G. Oxygen atoms are in shown red, nitrogen atoms are in blue, and carbon atoms are in green. The number along the line joining two atoms shows the distance (in Å) between the atoms. Ion-pair geometries are most optimal when the charged residues form a salt bridge, less optimal when they form an N–O bridge, and least optimal when they form an LRIP. Most of the salt bridges are stabilizing towards proteins. The majority of N–O bridges are stabilizing. The majority of the longer-range ion pairs are destabilizing towards the proteins.*

In summary, the free energy contributions of the electrostatic interactions depend upon the protein conformer population distribution and hence upon the experimental conditions. Geometrical optimization of electrostatic interactions is an important factor in an ion pair's free energy contribution towards the protein stability. These observations further validate our criteria for identifying salt bridges in protein structures.

5. Protein Thermostability and Close-Range Electrostatic Interactions

Proteins from organisms that live at high temperatures show greater thermal stability as compared to proteins from mesophilic organisms. An understanding of the thermodynamic stability difference between thermophilic and mesophilic proteins is important not just for theoretical reasons but also for practical industrial applications.[32, 49] Thermodynamic stability of proteins can be studied by means of differential scanning calorimetry (DSC) and spectroscopy (circular dichroism, fluorescence, UV/Vis). For a reversible two state ($N \Leftrightarrow D$) folding protein, a protein-stability curve describing the variation of its thermodynamic stability ($\Delta G(T)$) with temperature (T) can be obtained by using the Gibbs – Helmholtz equation.[50]

$$\Delta G(T) = \Delta H_G(1 - T/T_G) - \Delta C_p[(T_G - T) + T\ln(T/T_G)]$$

To plot the stability curve, three experimentally determined thermodynamic quantities are needed, ΔH_G, the enthalpy change between the native (N) and denatured (D) states of the protein, ΔC_p, the heat capacity change, and T_G, the melting temperature of the protein.[50] The shape of the protein stability curve is a skewed parabola. Using the protein-stability curve, one can compute the free energy change between the native (N) and denatured (D) states of the protein ($\Delta G(T)$) at a given temperature, T. One parameter that is of interest is the free energy change at the temperature of maximal protein stability, $\Delta G(T_S)$. Recently, we have compared the protein-stability curves and the thermodynamic parameters in five families containing homologous thermophilic and mesophilic proteins. Interpreting our results in terms of the protein structure, we see that greater protein stability is obtained by a greater formation of specific interactions, such as salt bridges and their networks, in the thermophilic proteins.[51]

These observations are consistent with our earlier analysis of sequence and structural differences in 18 nonredundant families of homologous thermophilic and mesophilic proteins.[32a] We analyzed various factors, hydrophobicity, compactness, proline content, disulfide bridges, residue composition, secondary structure content, surface areas, insertions/deletions, oligomerization, salt bridges, and hydrogen bonds. The thermophilic and mesophilic proteins have similar percentages of nonpolar surface buried in the core (hydrophobicities) and compactness (atomic packing), which indicates a similar extent of hydrophobic contribution towards the stability of these proteins. The distribution of buried and exposed polar and nonpolar surface areas is quite uniform in proteins and does not change among thermophiles and mesophiles. The occurrence of main chain – main chain hydrogen bonds that define protein secondary structure is also similar between thermophiles and mesophiles. Hence, the factors that may determine the overall protein fold

have similar values for thermophilic and mesophilic proteins. Proline content, insertions/deletions, and oligomerization do not show consistent trends between thermophiles and mesophiles. On the other hand, the amino acid distributions in thermophiles and mesophiles are significantly different, despite the high sequence identities among thermophiles and mesophiles in our database.[32a] Thermophilic proteins appear to favor residues with larger side chains and to avoid thermolabile residues. An increase in electrostatic interactions (salt bridges and side chain – side chain hydrogen bonds) in thermophiles as compared to their homologous mesophiles is the most consistent trend. These sequence and structural features may simultaneously raise ΔH_G and lower ΔC_p of a thermophilic protein as compared to its mesophilic homologue, thereby resulting in the greater thermodynamic stability of the thermophilic protein.

We have compared the electrostatic strengths of salt bridges in glutamate dehydrogenase from a hyperthermophile (*Pyrococcus furiosus*) and mesophile (*Clostridium symbiosum*).[32b] *Pyrococcus furiosus* glutamate dehydrogenase (PfGDH) is extremely thermostable, with its melting temperature being 113 °C. The mesophilic *Clostridium symbiosum* glutamate dehydrogenase (CsGDH) shares 34% sequence identity with PfGDH and the monomers of the two proteins superimpose with a root mean square deviation (RMSD) of 1.38 Å. In both organisms, the biochemically active GDH is a homohexamer. However, CsGDH has a half life of only 20 minutes at 52 °C and its melting temperature is 55 °C. Previously[32a] we found an increase in salt-bridge formation in PfGDH. Continuum electrostatic calculations performed on monomers of PfGDH and CsGDH show that the salt bridges in PfGDH are highly stabilizing. The salt bridges in CsGDH are marginally stabilizing. Salt bridges in PfGDH form extensive salt-bridge networks. Due to this, the interactions of charged side chains in the salt-bridge-forming residues with the rest of protein are almost as significant as the interaction of these side chains with each other. Hence, the cooperative nature of electrostatic interactions may lead to increased stability of PfGDH.

Analyses based on protein sequences from complete genomes of thermophilic and hyperthermophilic organisms and comparisons of sequence/structural properties of homologous thermophilic and mesophilic proteins have also consistently indicated a significant increase in the proportion of charged residues for the thermophiles.[52] The improvement in electrostatics for thermophiles is reflected in alleviation of electrostatic repulsions, increased occurrence of ion pairs and their networks, and geometrical optimization of charged residue positions to yield a favorable energetic contribution towards protein stability.[53] Different protein families may optimize these factors differently. In the following section we present a detailed case study of the molecular basis of greater thermostability of *Thermus thermophilus* ribonuclease H.

6. Molecular Analysis of Thermostability of *Thermus thermophilus* Ribonuclease H

Ribonuclease H is a single domain protein. Its function is to cleave DNA – RNA hybrids. *T. thermophilus* ribonuclease H (TtRnaseH) and *E. coli* ribonuclease H (EcRnaseH) are very similar. The amino acid sequences of the two proteins are ≈55% identical. The X-ray crystal structures[54] for the mesophilic and thermophilic Rnase H superimpose with an overall C^{α} RMSD of 1.23 Å with 140 matching residues, as determined by a sequence-order-independent structural superposition technique.[55] Hollien and Marqusee[56] have compared the thermodynamic stabilities of cysteine free mutants of TtRnaseH and EcRnaseH and performed hydrogen – deuterium exchange to identify residues contributing towards the stability of TtRnaseH. Figure 4 presents the protein stability curves of EcRnaseH and

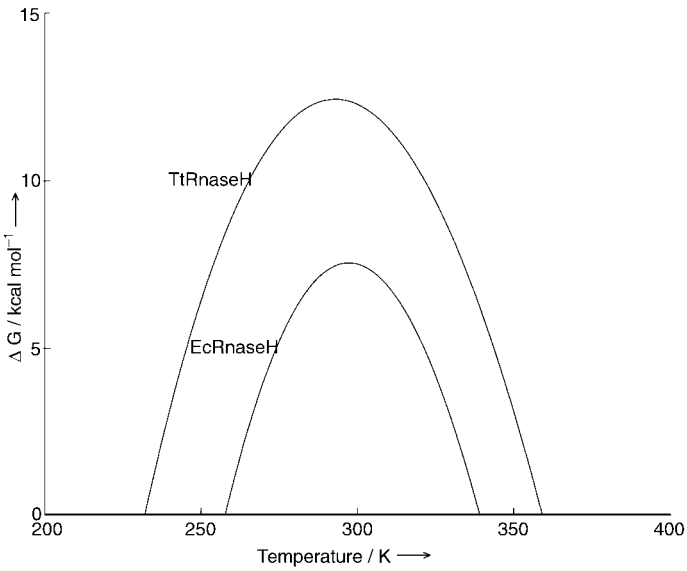


Figure 4. *Comparison of protein stability curves for* E. coli *and* T. thermophilus *ribonuclease H cysteine-free mutants. The stability curve for the thermophilic ribonuclease H (TtRnaseH) is broader and up-shifted compared to the stability curves for the mesophilic ribonuclease H (EcRnaseH). The y-axis represents the free energy difference (ΔG) between denatured (D) and native (N) states of the proteins and the x-axis represents the temperature. The protein stability curves were plotted by using the Gibbs-Helmholtz equation (see text).[50] Data on* ΔH_G, ΔC_p, *and* T_G *for TtRnaseH and EcRnaseH were taken from ref. [56b].*

TtRnaseH cysteine free mutants plotted by using their data. The protein stability curve of TtRnaseH is broader and up-shifted as compared to that of EcRnaseH, indicating greater thermodynamic stability for TtRnaseH. The living temperature of *T. thermophilus* is 68.5 °C. At 68.5 °C, the unfolded state of a cysteine-free variant of *E. coli* Rnase H (EcRnaseH) would be favorable by 0.91 kcal mol^{-1}. However, for the cysteine-free variant of *T. thermophilus* Rnase H (TtRnaseH), the folded state is favorable by 5.6 kcal mol^{-1} at this temperature. Here, we attempt to rationalize this stability difference between TtRnaseH and EcRnaseH in terms of the differences in the sequence and structural properties of TtRnaseH and EcRnaseH.

Using hydrogen/deuterium exchange experiments, Hollien and Marqusee have measured the free energy contributions for 39 out of 52 residues which contain slow-exchanging amide protons (Table 1 in ref. [56a]). These residues stabilize the folded state of TtRnaseH by 5.5 – 16.5 kcal mol^{-1}. They have also reported that these residues are distributed throughout the

protein structure and that thermostability is achieved in a delocalized manner. The sequence alignment of EcRnaseH and TtRnaseH shows that 19 out of the 39 residues are conserved between the two proteins, while the remaining 20 residues are mutated. Ishikawa et al.[54b] have suggested that replacement of Lys 95 in EcRnaseH by Gly 100 in TtRnaseH contributes to protein thermostability by relieving steric hindrance. Lys 95 in EcRnaseH is in a left-handed helical (α_L) conformation. The formation of an intramolecular disulfide bond also contributes to the stability of TtRnaseH.[57]

Table 1 compares the microscopic sequence and structural parameters between EcRnaseH and TtRnaseH. Both have similar secondary structural (α and β) content, fraction of surface accessibility, hydrophobicity, compactness and average occluded surface parameters,[58] and main chain – main chain hydrogen bonds. Figure 5 plots contour maps $C^\alpha - C^\alpha$ distances for

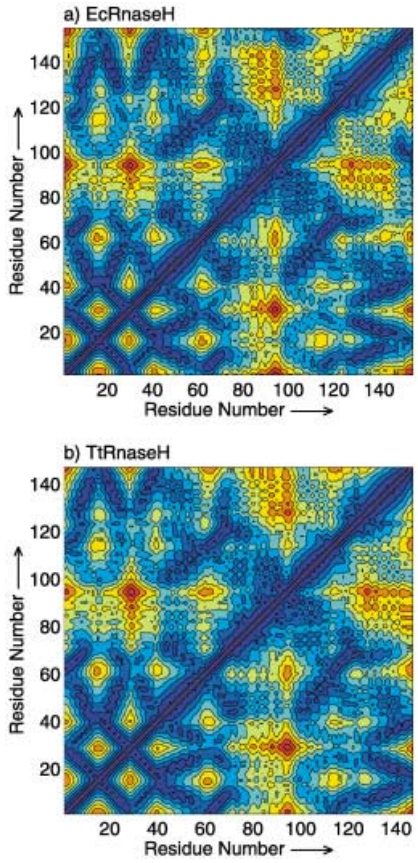


Figure 5. *Contour plots of C^α-C^α distances in a) EcRnaseH and b) TtRnaseH. Contours are filled with different colors (violet to red). The C^α-C^α distances increase from violet to red. Note that the two contour plots are very similar.*

EcRnaseH and TtRnaseH. The maps are very similar. These results indicate that both proteins have similar atomic packing and similar extents of nonspecific interactions. These results are similar to those obtained previously for 18 families of homologous thermophilic and mesophilic proteins.[32a]

Despite the 55% sequence identity, the amino acid distributions of EcRnaseH and TtRnaseH are significantly different at the 95% level of confidence (χ^2 value = 36.1; Table 1). For 19 parameter systems such as amino acid distributions in EcRnaseH and TtRnaseH, the χ^2 value should be greater than 30.14 to reject the null hypothesis (H_o: Two distributions are similar) at the 95% level of confidence (the probability of accepting the null hypothesis, $P \leq 0.05$). This observation is further supported by a large Hamming distance[5] between the two protein sequences in 20-dimensional percent amino acid composition space. The

Table 1. *Microscopic properties of thermophilic and mesophilic ribonuclease H.*

Property	EcRnaseH	TtRnaseH
general comparison		
number of residues	155	166
PDB file	2RN2	1RIL
resolution	1.48 Å	2.8 Å
number of residues in crystal structure	155	147
sequence identity	55%	
C^α RMSD[a]	1.39 Å	
number of matching residues[a]	140	
sequence and structural comparison		
α content[b]	34.8%	36.1%
β content[b]	28.4%	21.1%
hydrophobicity[c]	0.80	0.79
compactness[c]	1.71	1.63
average occluded surface parameter[d]	0.37 ± 0.15	0.37 ± 0.14
fractional ASA[e]	53.7%	47.6%
MC – MC hydrogen bonds[c]	68	60
MC – SC hydrogen bonds[c]	34	19
SC – SC hydrogen bonds[c]	15	7
N–O bridges[f]	6	14
salt bridges[f]	4	3
χ^2 value[g]	36.10	
Hamming distance[g]	9.3	
charged residues (D, E, H, K, R)	45 (29.0%)	55 (33.1%)
polar residues (N, Q, S, T)	29 (18.7%)	22 (13.3%)
aromatic residues (F, Y, W)	13 (8.4%)	13 (7.8%)
apolar residues (G, A, V, L, I)	56 (36.1%)	57 (34.3%)
β-branched residues (I, V, T)	26 (16.8%)	16 (9.6%)
number of proline residues	5 (3.2%)	12 (7.2%)
thermolabile residues (C, M, N, Q)	22 (14.2%)	16 (9.6%)

[a] The C^α root mean square deviation (RMSD) and the number of matching residues were obtained by superimposing the two structures with a computer-vision-based sequence-order-independent structure comparison method.[55] [b] α and β content indicate the fraction of residues in α-helical and β-strand conformations, respectively. These values were taken from promotif summary pages of PDB files 1ril and 2rn2, available at the PDBSUM website: http://www.biochem.ucl.ac.uk/bsm/pdbsum/. [c] Hydrophobicity, compactness, and numbers of main chain – main chain (MC – MC), main chain – side chain (MC – SC), and side chain – side chain (SC – SC) hydrogen bonds were calculated according to the procedures described by Kumar et al.[32a] from the crystal structure of EcRnaseH and TtRnaseH. TtRnaseH contains atomic coordinates for 147 (out of 166) residues. [d] Average occluded surface parameter (OSP)[58] values were calculated by using the OS71 program available at http://www.csb.yale.edu. Along with compactness values, these parameters qualitatively measure packing in EcRnaseH and TtRnaseH. [e] Fraction of protein surface area exposed to water, calculated by using the accesssurf routine in the ProStat program in the Homology module of the molecular modeling package INSIGHTII (98.0) from MSI. [f] Salt bridges in EcRnaseH and TtRnaseH were identified by using the method of Kumar and Nussinov.[2a] N–O bridges were inferred if at least one pair of oxygen and nitrogen atoms in side-chain functional groups of two oppositely charged residues are within 4.0 Å distance. [g] χ^2 value and Hamming distance computed according to Kumar and Bansal.[5a] Hamming distance indicates the distance between EcRnaseH and TtRnaseH sequences in 20-dimensional amino acid composition space.

proportion of the charged residues (Asp, Glu, His, Lys, and Arg) increases in TtRnaseH (33.1%) as compared to the proportion of these residues in EcRnaseH (29%) by 4.1%. The proportion of polar uncharged residues (Asn, Gln, Ser, and Thr) decreases in TtRnaseH as does the proportion of the thermolabile residues (Cys, Met, Asn, and Gln). The apolar residues (Gly, Ala, Val, Leu, and Ile) occur with similar proportions in TtRnaseH (34.3%) and EcRnaseH (36.1%). The proportion of proline residues increases in TtRnaseH. In general, it appears that TtRnaseH favors residues with longer side chains, such as Glu, Leu, and Arg.

Consistent with the increase in the occurrence of charged residues is the formation of a larger number of close-range electrostatic interactions. The crystal structure of TtRnaseH (PDB entry 1RIL, 2.8 Å resolution with coordinates for 147 out of 166 residues[54b])has 17 ion pairs. Fourteen of these are N–O bridges and the remaining three are salt bridges. It also contains two ion-pair networks, a hexad (six-residue network) and a tetrad (four-residue network). The crystal structure of EcRnaseH (PDB entry 2RN2, 1.48 Å resolution with coordinates for all 155 residues[54a]) has 10 ion pairs (6 N–O bridges and 4 salt bridges) and two ion-pair networks, a pentad and a triad. Figure 6 highlights the location of charged residues in TtRnaseH and EcRnaseH.

The sequence alignment of EcRnaseH and TtRnaseH[54b] indicates eight positions where apolar residues have been replaced by charged residues. These substitutions are L2R, M50K, V54E, I66D, V101R, L136E, A139R, and L146K. We have computed the electrostatic free energy contribution ($\Delta\Delta G_{elec}$) for six of these eight charged residues towards the stability of TtRnaseH by using continuum electrostatic calculations based on the method described by Tidor and co-workers.[53a] In this procedure, $\Delta\Delta G_{elec}$ for a charged residue consists of two terms, $\Delta\Delta G_{dslv}$ and $\Delta\Delta G_{int}$. $\Delta\Delta G_{dslv}$ is the desolvation energy penalty paid by the charged residue and $\Delta\Delta G_{int}$ is the free energy of the electrostatic interaction between the charged residue and all the charges in the rest of the proteins. Hence,

$$\Delta\Delta G_{elec} = \Delta\Delta G_{dslv} + \Delta\Delta G_{int}$$

The remaining two substitutions L2R and L146K fall at the N and C termini of TtRnaseH. Since the atomic coordinates for the adjoining residues are missing in the crystal structure of TtRnaseH, we did not calculate the $\Delta\Delta G_{elec}$ for these substitutions.

Five out of the six charged residue substitutions in TtRnaseH have stabilizing electrostatic free energy contributions with respect to their hydrophobic isosteres (Figure 7). Out of the five, three residues, D66 in strand D and E136 and R139 in helix αV, appear to have large electrostatic stabilization. D66 is also part of a six-residue ion-pair network (hexad; formed by residues R2, R4, E64, D66, R115, and R117) in TtRnaseH. Glu 136 and Arg 139 form a salt bridge. This salt bridge is stabilizing towards TtRnaseH by -5.6 kcal mol^{-1}. Of the remaining two residues with stabilizing $\Delta\Delta G_{elec}$ values, K50 is part of a four-residue ion-pair network (tetrad formed by E39, R46, K50, and D102) and E54 forms a salt bridge with K57 in TtRnaseH. The salt bridge E54 – K57 stabilizes TtRnaseH by -1.1 Kcal mol^{-1}. Both ion-pair networks, the hexad and the tetrad, have stabilizing electrostatic contributions. The

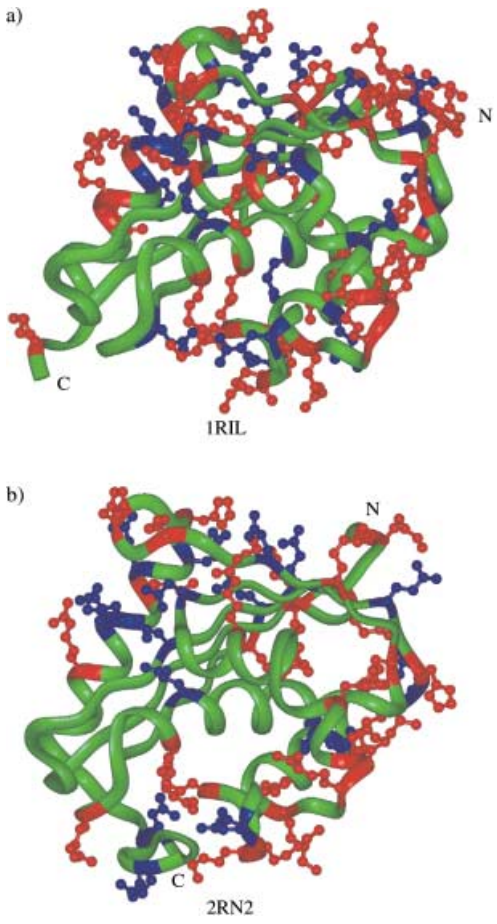


Figure 6. *Ribbon diagrams showing the distribution of charged residues in a)* T. thermophilus *ribonuclease H (TtRnaseH) and b)* E. coli *ribonuclease H (EcRnaseH) crystal structures (PDB codes: 1RIL and 2RN2, respectively). All charged residues are shown in ball-and-stick representation. The positively charged residues are shown in red and the negatively charged residues are shown in blue. The ribbon for all other residues is shown in green.*

electrostatic free energy contribution by the hexad (R2, R4, E64, D66, R115, and R117) in TtRnaseH is -7.8 kcal mol^{-1} and that for the tetrad (E39, R46, K50, and D102) is -2.0 kcal mol^{-1}. Hydrogen-exchange experiments by Hollien and Marqusee[56a] also indicate stabilizing roles for four of these residues. Arg 101 has destabilizing electrostatic free energy contribution ($\Delta\Delta G_{elec} = +2.04$ kcal mol^{-1}) towards TtRnaseH. Arg 101 lies at the N terminus of helix α^{IV} [54b] and does not form a salt bridge or ion pair. Arginine residues are avoided at positions near the α-helix N terminus.[5a] Hence, a mutation of R101 may further enhance the stability of TtRnaseH.

The electrostatic free energy values reported here for single charged residues, salt bridges, and ion-pair networks correspond to room temperature, pH 7.0, and zero ionic strength. At the living temperature of *T. thermophilus*, these values are expected to further decrease by approximately 1 kcal mol^{-1} due to the reduced hydration free energy changes for the residues, reduced dielectric constant for water, and reduced solvent screening of the electrostatic interactions.[32b, 35c] This indicates that $\Delta\Delta G_{elec}$ values would be less destabilizing for R101 and more stabilizing for K50, E54, D66, E136, and R139 in TtRnaseH. It is not clear how

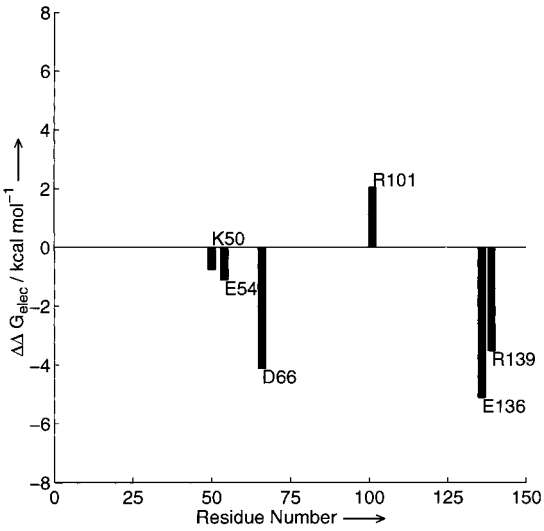


Figure 7. *A bar diagram showing the electrostatic free energy contribution towards the stability of TtRnaseH for six charged residues. EcRnaseH contains apolar residues at these positions. A negative value for $\Delta\Delta G_{elec}$ (y-axis) indicates that the charged residue stabilizes TtRnaseH. The procedure for computing $\Delta\Delta G_{elec}$ is described briefly in the text. Amino acids are shown in their single letter codes. Five of the six charged residues have stabilizing electrostatic contributions. Of these, three residues D66, E136 and R139 have large contributions. R101 has a destabilizing electrostatic free energy contribution. D66 is part of a highly stabilizing six-residue ion-pair network. E136 and R139 form a highly stabilizing salt bridge.*

the electrostatic free energy contributions of these charged residues, salt bridges, and ion-pair networks would be affected if the atomic coordinates of the missing 19 residues at the N and C termini of TtRnaseH were known.

7. Summary and Outlook

Here, we have discussed the role of specific interactions in protein folding, structure, and function. Though not the driving force in protein folding and binding, these interactions play important roles in fine-tuning the protein structure for optimum function. Salt bridges are among the best studied noncovalent interactions in proteins. The location of salt bridges is consistent with the hierarchical model for protein folding. While some experiments suggest formation of buried salt bridges and their networks may constitute a slow step in protein folding, others regard these as being involved in early events. Possibly these could be related to the location and sequence separation of the salt-bridging residues. Formation of salt bridges and their networks appears to constrain protein flexibility. Yet, both the identities of the charged residues forming the salt bridges and the electrostatic strengths of the salt bridges fluctuate. Salt bridges are avoided in protein parts that show segmental flexibility, such as hinge-bending motion.

It can be debated whether salt bridges stabilize or destabilize proteins. There is considerable theoretical and experimental evidence in favor of both. In theory, the effects due to internal dielectric constants of the protein experienced by the salt bridges as well as protein relaxation and reorganization of the polar groups upon salt-bridge formation are not completely understood. In our work, we find that geometrical orientation of the salt-bridging-residue side-chain charged groups with respect to each other is a crucial factor in determining the electrostatic strength of a salt bridge. Salt bridges with favorable geometrical orientation of the side-chain charge groups may be stabilizing *anywhere* in the protein structure.

While the stabilizing effect of salt bridges remains a controversial issue, it is becoming increasingly clear that they contribute significantly towards the homologous thermophilic–mesophilic protein stability differential. That is, among the homologous thermophilic and mesophilic proteins, the thermodynamic stability is modulated by optimizing the close-range electrostatic interactions. These optimizations include relieving the electrostatic stress due to repulsions, geometrical optimization of the salt bridges and their networks, and formation of additional salt bridges and their networks, particularly around the active sites of the thermophilic proteins. On the one hand, we note that salt bridges and their networks constrain protein flexibility. On the other hand, an increase in the number of charged residues and greater formation of salt bridges and their networks are the most consistent trends observed in the thermophilic proteins. This raises the possibility that thermophilic proteins are more rigid. However, at the living temperatures of their source organisms, the thermophilic proteins are flexible enough to perform their functions optimally.[59]

The overall free energy contributions of electrostatics may also be stabilizing or destabilizing towards proteins and protein–protein complexes. It is possible that the overall contribution of electrostatics may be destabilizing even if the free energy contribution of the close-range electrostatic interactions is stabilizing.[22e] Considerable work is needed to further improve current electrostatic models.[33, 34j]

We thank Dr. Buyong Ba, Dr. Chung-jung Tsai, Dr. Neeti Sinha, Dr. Gunasekaran Kannan, Dr. David Zanuy, and in particular Dr. Jacob V. Maizel for numerous helpful discussions. We are indebted to Dr. Yuk Yin Sham for critical reading of our manuscript. We also thank Prof. Arieh Warshel for sending us reprints of several of his papers on protein electrostatics. The research of R.N. in Israel has been supported in part by grant number 95-00208 from BSF, Israel, by a grant from the Israel Science Foundation administered by the Israel Academy of Sciences, by the Magnet grant, by the Ministry of Science, by the Tel Aviv University Basic Research grants, and by the Center of Excellence, administered by the Israel Academy of Sciences. This project has been funded in whole or in part with federal funds from the National Cancer Institute, National Institutes of Health, under contract number NO1-CO-56000. The content of this publication does not necessarily reflect the view or policies of the Department of Health and Human Services, nor does the mention of trade names, commercial products, or organization imply endorsement by the U.S. Government.

[1] a) D. J. Barlow, J. M. Thornton, *J. Mol. Biol.* **1983**, *168*, 867–885; b) V. Lounnas, R. C. Wade, *Biochemistry* **1997**, *36*, 5402–5417.
[2] a) S. Kumar, R. Nussinov, *J. Mol. Biol.* **1999**, *293*, 1241–1255; b) S. Kumar, R. Nussinov, *Proteins: Struct. Funct. Genet.* **2000**, *41*, 485–497; c) S. Kumar, R. Nussinov, *Proteins: Struct. Funct. Genet.* **2001**, *43*, 433–454.

[3] a) E. A. Permyakov, L. J. Berliner, *FEBS Lett.* **2000**, *473*, 269 – 274; b) P. E. Wright, H. J. Dyson, *J. Mol. Biol.* **1999**, *293*, 321 – 331; c) Y. Y. Sham, B. Ma, C. J. Tsai, R. Nussinov, *Protein Sci.* **2001**, *10*, 135 – 148.

[4] a) R. L. Baldwin, G. D. Rose, *Trends Biochem. Sci.* **1999**, *24*, 26 – 33; b) R. L. Baldwin, G. D. Rose, *Trends Biochem. Sci.* **1999**, *24*, 77 – 84.

[5] a) S. Kumar, M. Bansal, *Proteins: Struct. Funct. Genet.* **1998**, *31*, 460 – 76; b) "Geometry and sequence correlation studies on α-helices in globular proteins": S. Kumar, PhD thesis, Indian Institute of Science, Bangalore, India, **1997**; c) D. Gandini, L. Gogioso, M. Bolognesi, D. Bordo, *Proteins: Struct. Funct. Genet.* **1996**, *24*, 439 – 449.

[6] a) M. Sundaralingam, Y. C. Sekharudu, N. Yathindra, V. Ravichandran, *Proteins: Struct. Funct. Genet.* **1987**, *2*, 64 – 71; b) D. Walther, P. Argos, *Protein Eng.* **1996**, *9*, 471 – 478.

[7] a) P. C. Lyu, P. J. Gans, N. R. Kallenbach, *J. Mol. Biol.* **1992**, *223*, 343 – 350; b) M. J. Bodkin, J. M. Goodfellow, *Protein Sci.* **1995**, *4*, 603 – 612; c) S. Marqusee, R. L. Baldwin, *Proc. Natl. Acad. Sci. USA* **1987**, *84*, 8898 – 8902; d) C. A. Olson, E. J. Spek, Z. Shi, A. Vologodskii, N. R. Kallenbach, *Proteins: Struct. Funct. Genet.* **2001**, *44*, 123 – 132.

[8] a) C. A. Blaise, J. M. Berg, *Biochemistry* **1997**, *36*, 6218 – 6222; b) J. S. Merkel, J. M. Sturtevant, L. Regan, *Structure* **1999**, *7*, 1333 – 1343.

[9] O. Schueler, H. Margalit, *J. Mol. Biol.* **1995**, *248*, 125 – 135.

[10] a) M. T. Oliva, J. Moult, *Protein Eng.* **1999**, *12*, 727 – 735; b) V. Z. Spassov, B. P. Atanasov, *Proteins: Struct. Funct. Genet.* **1994**, *19*, 222 – 229.

[11] B. Musafia, V. Buchner, D. Arad, *J. Mol. Biol.* **1995**, *254*, 761 – 770.

[12] a) M. Oliveberg, A. R. Fersht, *Biochemistry* **1996**, *35*, 2726 – 2737; b) M. Oliveberg, A. R. Fersht, *Biochemistry* **1996**, *35*, 6795 – 6805.

[13] C. D. Waldburger, T. Jonsson, R. T. Sauer, *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 2629 – 2634.

[14] S. Cavagnero, D. A. Debe, Z. H. Zhou, M. W. W. Adams, S. I. Chan, *Biochemistry* **1998**, *37*, 3369 – 3376.

[15] I. Y. Torshin, R. W. Harrison, *Proteins: Struct. Funct. Genet.* **2001**, *43*, 353 – 364.

[16] a) C. J. Tsai, R. Nussinov, *Protein Sci.* **1997**, *6*, 24 – 42; b) C. J. Tsai, R. Nussinov, *Protein Sci.* **1997**, *6*, 1426 – 1437; c) C. J. Tsai, J. V. Maizel, R. Nussinov, *Protein Sci.* **1999**, *8*, 1591 – 1604; d) C. J. Tsai, J. V. Maizel, R. Nussinov, *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 12038 – 12043; e) S. Kumar, Y. Y. Sham, C. J. Tsai, R. Nussinov, *Biophys. J.* **2001**, *80*, 2439 – 2454; f) C. J. Tsai, B. Ma, S. Kumar, H. Wolfson, R. Nussinov, *Crit. Rev. Biochem. Mol. Biol.* **2001**, *36*, 399 – 433.

[17] a) S. Kumar, H. Wolfson, R. Nussinov, *IBM J. Res. Dev.* **2001**, *45*(3/4), 513 – 523; b) N. Sinha, S. Kumar, R. Nussinov, *Structure* **2001**, *9*, 1165 – 1181.

[18] a) D. Xu, C. J. Tsai, R. Nussinov, *Folding Des.* **1998**, *3*, R71-R80; b) C. J. Tsai, S. Kumar, B. Ma, R. Nussinov, *Protein Sci.* **1999**, *8*, 1181 – 1190; c) B. Ma, S. Kumar, C. J. Tsai, R. Nussinov, *Protein Eng.* **1999**, *12*, 713 – 720; d) S. Kumar, B. Ma, C. J. Tsai, H. Wolfson, R. Nussinov, *Cell Biochem. Biophys.* **1999**, *31*, 141 – 164; e) C. J. Tsai, B. Ma, R. Nussinov, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 9970 – 9972; f) S. Kumar, B. Ma, C. J. Tsai, N. Sinha, R. Nussinov, *Protein Sci.* **2000**, *9*, 10 – 19; g) B. Ma, S. Kumar, C. J. Tsai, Z. Hu, R. Nussinov, *J. Theor. Biol.* **2000**, *203*, 383 – 387; h) N. Sinha, R. Nussinov, *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 3139 – 3144.

[19] a) C. J. Tsai, S. L. Lin, H. J. Wolfson, R. Nussinov, *Protein Sci.* **1997**, *6*, 53 – 64; b) C. J. Tsai, S. L. Lin, H. J. Wolfson, R. Nussinov, *Crit. Rev. Biochem. Mol. Biol.* **1996**, *31*, 127 – 152; c) C. J. Tsai, S. L. Lin, H. J. Wolfson, R. Nussinov, *J. Mol. Biol.* **1996**, *260*, 604 – 620.

[20] a) D. Xu, C. J. Tsai, R. Nussinov, *Protein Eng.* **1997**, *10*, 999 – 1012; b) D. Xu, S. L. Lin, R. Nussinov, *J. Mol. Biol.* **1997**, *265*, 68 – 84.

[21] a) T. Clackson, J. A. Wells, *Science* **1995**, *267*, 383 – 386; b) A. A. Bogan, K. S. Thorn, *J. Mol. Biol.* **1998**, *280*, 1 – 9; c) Z. Hu, B. Ma, H. J. Wolfson, R. Nussinov, *Proteins: Struct. Funct. Genet.* **2000**, *39*, 331 – 342; d) B. Ma, H. J. Wolfson, R. Nussinov, *Curr. Opin. Struct. Biol.* **2001**, *11*, 364 – 369.

[22] a) Z. S. Hendsch, B. Tidor, *Protein Sci.* **1999**, *8*, 1381 – 1392; b) L.-P. Lee, B. Tidor, *Protein Sci.* **2001**, *10*, 362 – 377; c) A. J. McCoy, V. C. Epa, P. M. Colman, *J. Mol. Biol.* **1997**, *268*, 570 – 584; d) M. Vijaykumar, K.-Y. Wong, G. Schreiber, A. R. Fersht, A. Szabo, H.-X. Zhou, *J. Mol. Biol.* **1998**, *278*, 1015 – 1024; e) F. B. Sheinerman, R. Norel, B. Honig, *Curr. Opin. Struct. Biol.* **2000**, *10*, 153 – 159.

[23] a) J. G. Mandell, V. A. Roberts, M. E. Pique, V. Kotlovyi, J. C. Mitchell, E. Nelson, I. Tsigelny, L. F. Ten Eyck, *Protein Eng.* **2001**, *14*, 105 – 113; b) R. Norel, F. Sheinerman, D. Petrey, B. Honig, *Protein Sci.* **2001**, *10*, 2147 – 2161.

[24] T. Konno, *Biochemistry* **2001**, *40*, 2148 – 2154.

[25] C. J. Tsai, B. Ma, Y. Y. Sham, S. Kumar, R. Nussinov, *Proteins: Struct. Funct. Genet.* **2001**, *44*, 418 – 427.

[26] D. Sahal, P. Balaram, *Biochemistry* **1986**, *25*, 6004 – 6013.

[27] a) E. M. Storch, V. Daggett, *Biochemistry* **1995**, *34*, 9682 – 9693; b) E. M. Storch, V. Daggett, *Biochemistry* **1996**, *35*, 11596 – 11604; c) E. M. Storch, V. Daggett, W. M. Atkins, *Biochemistry* **1999**, *38*, 5054 – 5064; d) E. M. Storch, J. S. Grinstead, A. P. Campbell, V. Daggett, W. M. Atkins, *Biochemistry* **1999**, *38*, 5065 – 5075.

[28] a) E. Fischer, *Ber. Dtsch. Chem. Ges.* **1894**, *27*, 2985 – 2991; b) D. E. Koshland, Jr., *Proc. Natl. Acad. Sci. USA* **1958**, *44*, 98 – 123.

[29] a) M. Shiroishi, A. Yokata, K. Tsumoto, H. Kondo, Y. Nishimiya, K. Horii, M. Matsushima, K. Ogasahara, K. Yutani, I. Kumagai, *J. Biol. Chem.* **2001**, *276*, 23042 – 23050; b) S. J. Smith-Gill, T. B. Lavoie, C. R. Mainhart, *J. Immunol.* **1984**, *133*, 384 – 393; c) L. N. W. Kam-Morgan, S. J. Smith-Gill, M. G. Taylor, L. Zhang, A. C. Wilson, J. F. Kirsch, *Proc. Natl. Acad. Sci. USA* **1992**, *90*, 3958 – 3962; d) E. A. Padlan, E. W. Silverton, S. Sheriff, G. H. Cohen, S. J. Smith-Gill, D. R. Davies, *Proc. Natl. Acad. Sci. USA* **1989**, *86*, 5938 – 5942.

[30] S. Herrgard, C. J. Gibas, S. Subramaniam, *Biochemistry* **2000**, *39*, 2921 – 2930.

[31] E. J. Sundberg, R. A. Mariuzza, *Structure* **2000**, *8*, R137-R142.

[32] a) S. Kumar, C. J. Tsai, R. Nussinov, *Protein Eng.* **2000**, *13*, 3, 179 – 191; b) S. Kumar, B. Ma, C. J. Tsai, R. Nussinov, *Proteins: Struct. Funct. Genet.* **2000**, *38*, 4, 368 – 383; c) S. Kumar, R. Nussinov, *Cell. Mol. Life Sci.* **2001**, *58*, 9, 1216 – 1233; d) S. Kumar, R. Nussinov, unpublished results.

[33] a) A. Warshel, A. Papazyan, *Curr. Opin. Struct. Biol.* **1998**, *8*, 211 – 217; b) D. J. Tobias, *Curr. Opin. Struct. Biol.* **2001**, *11*, 253 – 261; c) T. Simonson, *Curr. Opin. Struct. Biol.* **2001**, *11*, 243 – 252; d) Y. Y. Sham, I. Muegge, A. Warshel, *Biophys. J.* **1998**, *74*, 1744 – 1753.

[34] a) A. Warshel, J. Aqvist, *Chem. Scr.* **1989**, *29A*, 75 – 83; b) A. Warshel, J. Aqvist, *Annu. Rev. Biophys. Biophys. Chem.* **1991**, *20*, 267 – 298; c) J. D. Madura, Y. Nakajima, R. M. Hamilton, A. Wierzbicki, A. Warshel, *Struct. Chem.* **1996**, *7*, 131 – 138; d) F. S. Lee, Z.-T. Chu, M. B. Bolger, A. Warshel, *Protein Eng.* **1992**, *5*, 215 – 228; e) Y. Y. Sham, Z. T. Chu, H. Tao, A. Warshel, *Proteins: Struct. Funct. Genet.* **2000**, *39*, 393 – 407; f) F. S. Lee, Z. T. Chu, A. Warshel, *J. Comp. Chem.* **1993**, *14*, 161 – 185; g) I. Muegge, H. Tao, A. Warshel, *Protein Eng.* **1997**, *10*, 1363 – 1372; h) I. Muegge, T. Schweins, A. Warshel, *Proteins: Struct. Funct. Genet.* **1998**, *30*, 407 – 423; i) Z. Z. Fan, J. K. Hwang, A. Warshel, *Theor. Chem. Acc.* **1999**, *103*, 77 – 80; j) C. N. Schutz, A. Warshel, *Proteins: Struct. Funct. Genet.* **2001**, 400 – 417; k) A. Warshel, S. T. Russell, *Q. Rev. Biophys.* **1984**, *17*, 283 – 421; l) A. Warshel, *Biochemistry* **1981**, *20*, 3167 – 3177; m) A. Warshel, S. T. Russell, A. K. Churg, *Proc. Natl. Acad. Sci. USA* **1984**, *81*, 4785 – 4789; n) J. K. Hwang, A. Warshel, *Nature* **1988**, *334*, 270 – 272.

[35] a) R. Luo, M. S. Head, J. Moult, M. K. Gilson, *J. Am. Chem. Soc.* **1998**, *120*, 6138 – 6146; b) J. Antosiewicz, J. A. McCammon, M. K. Gilson, *Biochemistry* **1996**, *35*, 7819 – 7833; c) A. H. Elcock, J. A. McCammon, *J. Mol. Biol.* **1998**, *280*, 731 – 748; d) P. H. Hunenberger, V. Helms, N. Narayana, S. S. Taylor, J. A. McCammon, *Biochemistry* **1999**, *38*, 2358 – 2366; e) J. Lamotte-Brasseur, A. Dubus, R. C. Wade, *Proteins* **2000**, *40*, 23 – 28; f) N. Froloff, A. Windemuth, B. Honig, *Protein Sci.* **1997**, *6*, 1293 – 1301; g) C. J. Gibas, P. Jambeck, S. Subramaniam, *Methods (San Diego)* **2000**, *20*, 292 – 309; h) E. G. Alexov, M. R. Gunner, *Biophys. J.* **1997**, *74*, 2075 – 2093; i) E. G. Alexov, M. R. Gunner, *Biochemistry* **1999**, 38, 8253 – 8270.

[36] V. V. Loladze, D. N. Ermolenko, G. I. Makhatadze, *Protein Sci.* **2001**, *10*, 1343 – 1352.

[37] Z. S. Hendsch, B. Tidor, *Protein Sci.* **1994**, *3*, 211 – 226.

[38] a) A. Horovitz, A. R. Fersht, *J. Mol. Biol.* **1992**, *224*, 733 – 740; b) S. Marqusee, R. T. Sauer, *Protein Sci.* **1994**, *3*, 2217 – 2225; c) K. Pervushin, M. Billeter, G. Siegal, K. Wüthrich, *J. Mol. Biol.* **1996**, *264*, 1002 – 1012; d) L. Xiao, B. Honig, *J. Mol. Biol.* **1999**, *289*, 1435 – 1444; e) S. Dao-Pin, U. Sauer, H. Nicholson, B. W. Matthews, *Biochemistry* **1991**, *30*, 7142 – 7153; f) S. Dao-Pin, D. E. Anderson, W. A. Baase, F. W. Dahlquist, B. W. Matthews, *Biochemistry* **1991**, *30*, 11521 – 11529; g) C. D. Waldburger, J. F. Schildbach, R. T. Sauer, *Nat. Struct. Biol.* **1995**, *2*, 122 – 128; h) B. H. Honig, W. L. Hubell, *Proc. Natl. Acad. Sci. USA* **1984**, *81*, 5412 – 5416.

[39] a) X. Barril, C. Aleman, M. Orozco, F. J. Luque, *Proteins: Struct. Funct. Genet.* **1998**, *32*, 67 – 79; b) U. C. Singh, *Proc. Natl. Acad. Sci. USA* **1988**, *85*, 4280 – 4284; c) L. Serrano, A. Horovitz, B. Avron, M. Bycroft, A. R. Fersht, *Biochemistry* **1990**, *29*, 9343 – 9352; d) E. J. Spek, A. H. Bui, M. Lu, N. R. Kallenbach, *Protein Sci.* **1998**, *7*, 2431 – 2437.

[40] B. H. Honig, A. Nicholls, *Science* **1995**, *268*, 1144 – 1149.

[41] a) M. K. Gilson, A. Rashin, R. Fine, B. H. Honig, *J. Mol. Biol.* **1985**, *183*, 503 – 516; b) M. K. Gilson, K. A. Sharp, B. H. Honig, *J. Comput. Chem.* **1988**, *9*, 327 – 335; c) M. K. Gilson, B. H. Honig, *Nature* **1987**, *330*, 84 – 86; d) M. K. Gilson, B. H. Honig, *Proteins: Struct. Funct. Genet.* **1988**, *4*, 7 – 18; e) K. A. Sharp, B. H. Honig, *Annu. Rev. Biophys. Biophys. Chem.* **1990**, *19*, 301 – 332; f) B. H. Honig, K. A. Sharp, A. Yang, *J. Phys. Chem.* **1993**, *97*, 1101 – 1109.

[42] a) D. Sitkoff, K. A. Sharp, B. H. Honig, *J. Phys. Chem.* **1994**, *98*, 1978 – 1988; b) A. Radzicka, R. Wolfenden, *Biochemistry* **1988**, *27*, 1664 – 1670.

[43] I. Klapper, R. Hagstrom, R. Fine, K. A. Sharp, B. H. Honig, *Proteins: Struct. Funct. Genet.* **1986**, *1*, 47 – 59.

[44] J. E. Nielsen, K. V. Anderson, B. Honig, R. W. W. Hooft, G. Klebe, G. Vriend, R. C. Wade, *Protein Eng.* **1999**, *12*, 657 – 662.

[45] a) W. C. Wimley, K. Gawrisch, T. P. Creamer, S. H. White, *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 2985 – 2990; b) S. Miller, J. Janin, A. M. Lesk, C. Chothia, *J. Mol. Biol.* **1987**, *196*, 641 – 656; c) D. Eisenberg, A. D. McLachlan, *Nature* **1986**, *319*, 199 – 203.

[46] a) T. Kajander, P. C. Kahn, S. H. Passila, D. C. Cohen, L. Lehtio, W. Adolfsen, J. Warwicker, U. Schell, A. Goldman, *Structure* **2000**, *8*, 1203 – 1214; b) A. Warshel, *Proc. Natl. Acad. Sci. USA* **1978**, *75*, 5250 – 5254; c) A. Giletto, C. N. Pace, *Biochemistry* **1999**, *38*, 13 379 – 13 384.

[47] a) F. C. Bernstein, T. F. Koetzle, G. J. Williams, E. E. Shimanouchi, M. Tasumi, *J. Mol. Biol.* **1977**, *112*, 535 – 542; b) H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, P. E. Bourne, *Nucleic Acids Res.* **2000**, *28*, 235 – 242.

[48] S. Kumar, R. Nussinov, *Biophys. J.* **2002**, in press.

[49] R. Sterner, W. Liebl, *Crit. Rev. Biochem. Mol. Biol.* **2001**, *36*, 39 – 106.

[50] a) W. Becktel, J. A. Schellman, *Biopolymers* **1987**, *26*, 1859 – 1877; b) P. L. Privalov, *Crit. Rev. Biochem. Mol. Biol.* **1990**, *25*, 281 – 305.

[51] S. Kumar, C. J. Tsai, R. Nussinov, *Biochemistry* **2001**, *40*, 14 152 – 14 165

[52] a) C. Cambillau, J.-M. Claverie, *J. Biol. Chem.* **2000**, *275*, 32 383 – 32 386; b) S. Chakravarty, R. Varadarajan, *FEBS Lett.* **2000**, *470*, 65 – 69; c) P. J. Haney, J. H. Badger, G. L. Buldak, C. I. Reich, C. R. Woese, G. J. Olsen, *Proc. Natl. Acad. Sci. USA* **1996**, 3578 – 3583.

[53] a) S. Spector, M. Wang, S. A. Carp, J. Robblee, Z. S. Hendsch, R. Fairman, B. Tidor, D. P. Raleigh, *Biochemistry* **2000**, *39*, 872 – 879; b) G. R. Grimsley, K. L. Shaw, L. R. Fee, R. W. Alston, B. M. Huyghues-Despointes, R. L. Thurlkil, J. M. Scholtz, C. N. Pace, *Protein Sci.* **1999**, *8*, 1843 – 1849; c) V. V. Loladze, B. Ibarra-olero, J. M. Sanchez-Ruiz, G. I. Makhatadze, *Biochemistry* **1999**, *38*, 16 419 – 16 423; d) J. M. Sanchez-Ruiz, G. I. I. Makhatadze, *Trends Biotechnol.* **2001**, *19*, 132 – 135.

[54] a) K. Katayanagi, M. Miyagawa, M. Matsushima, M. Ishikawa, S. Kanaya, M. Ikehara, T. Matsuzaki, K. Morikawa, *Nature* **1990**, *347*, 306 – 309; b) K. Ishikawa, M. Okumura, K. Katayanagi, S. Kimura, S. Kanaya, H. Nakamura, K. Morikawa, *J. Mol. Biol.* **1993**, *230*, 529 – 542.

[55] C. J. Tsai, S. L. Lin, H. Wolfson, R. Nussinov, *J. Mol. Biol.* **1996**, *260*, 604 – 620.

[56] a) J. Hollien, S. Marqusee, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 13 674 – 13 678; b) J. Hollien, S. Marqusee, *Biochemistry* **1999**, *38*, 3831 – 3836.

[57] N. Hirano, M. Haruki, M. Morikawa, S. Kanaya, *Biochemistry* **1998**, *37*, 12 640 – 12 648.

[58] P. J. Fleming, F. M. Richards, *J. Mol. Biol.* **2000**, *299*, 487 – 498.

[59] R. Jaenicke, *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 2962 – 2964.

Received: September 7, 2001 [A 293]