

# Electrostatic effects on protein stability

An-Suei Yang and Barry Honig

Columbia University, New York, USA

Recent experimental and theoretical studies have led to new insights into the contribution of ionizable amino acids to protein stability. The role of polar groups is less clear, in part because their interactions are difficult to control experimentally.

Current Opinion in Structural Biology 1992, 2:40–45

## Introduction

The past few years have witnessed major advances in our understanding of electrostatic interactions in proteins (summarized in a number of recent reviews [1,2]). Electrostatic interactions, as defined here, arise both from ionizable amino acids and from polar groups that contain permanent dipoles. The major focus of this review will be recent developments in elucidating the contributions of ionizable amino acids to protein stability. Progress has been facilitated by a combination of modern experimental techniques, such as NMR and site-directed mutagenesis, and more traditional methods such as the measurement of pH and salt effects on denaturation free energies. In parallel, new theoretical methods now make it possible to interpret experimental results in terms of specific molecular interactions. The interplay between increasingly reliable theoretical methods and the improved ability of experimental methods to probe well defined interactions will be highlighted in this review.

Only a short section of this review is devoted to the discussion of the contribution of dipolar groups to protein stability. In our view, there has been only limited progress in this area; indeed, the fundamental question as to whether or not hydrogen bonds stabilize proteins has not yet been definitively resolved. In addition, little attention has been devoted to the destabilizing effects that result from the burial of permanent dipoles in the protein interior. We will offer our opinion on these questions in a brief section near the end of the review.

## Basic principles

In the first place, it is worth considering what is meant by the statement that a particular interaction stabilizes or destabilizes a protein. If the focus is on the effect of a single charge, the clearest experimental measure of its contribution to protein stability can be obtained by neutralizing it through changes in pH. Unless one is dealing with small, designed peptides, however, changes in

pH affect more than one charged group, requiring that an entire titration curve be deconvoluted through some combination of experimental and theoretical techniques.

Site-directed mutagenesis offers another way of neutralizing a charge but the method substitutes a new interaction for an old one, which introduces additional complexity to the process of separating variables. For example, the effects on protein stability of pairwise Coulombic interactions have been identified through the use of double and triple mutant cycles in an important study by Fersht and coworkers [3]. Electrostatic interactions among three ionizable groups were found to have a net stabilizing effect but, nevertheless, the effect of replacing all three with alanines was an increase in protein stability. That alanines provide a greater net stabilization of the native state than ionizable groups of course says nothing as to whether the charges are themselves stabilizing or destabilizing.

The change in the pKa of a group between the native and denatured states offers a direct measure of the effect of that group on protein stability. One way of measuring pKa's is to monitor proton release as a function of pH for all the ionizable residues in a protein using two-dimensional NMR [4]. Measurement of the titration curves of proteins is another possibility [5] but, as mentioned above, the pKa's of individual ionizable groups are not easy to resolve. If the pKa's of all ionizable residues are known for the native and denatured states of a protein, the denaturation energy of the protein as a function of pH can be obtained from an equation based on an expression given by Tanford [6]:

$$\Delta G_{\text{den}}^0(\text{pH}) = 2.3kT \int_{\text{pH}^{1/2}}^{\text{pH}} \Delta Z(\text{pH})d\text{pH} \quad (1)$$

where  $\Delta G_{\text{den}}^0(\text{pH})$  is the denaturation energy at a given pH with respect to that at some reference pH,  $\Delta Z(\text{pH})$  is the difference in net charge between the denatured and native protein, and  $\text{pH}_{1/2}$  is the midpoint of the acid denaturation curve. It should be emphasized that Equation 1 is an expression for a relative free energy, and does

not give the total electrostatic free energy of a system. Two proteins may have the same set of pKa's and, consequently, the same relative free energies as functions of pH, but may have different total electrostatic free energies. Thus, knowledge of pKa's is necessary but not sufficient for a full understanding of electrostatic interactions in proteins.

## Experimental observations

### Salt bridges (hydrogen-bonded ion pairs)

A number of laboratories have used site-directed mutagenesis or chemical synthesis to study the stabilizing effects of acidic and basic residues which have been placed in locations (such as the  $i$  and  $i + 4$  positions on  $\alpha$ -helices) where they are able to form salt bridges [7•,8,9•,10,11]. In general, it appears that the engineered salt bridges have only marginal effects, contributing  $\sim 0$  to  $-0.5$  kcal mol $^{-1}$  to stabilization of the folded state. These findings appear to contradict the earlier work of Marqusee and Baldwin [12], who found that lysines and glutamic acids placed in positions  $i$  and  $i + 4$  in an alanine-based  $\alpha$ -helix stabilized the helical conformation. The actual stabilization energy was not determined in that study, however, and, based on the small pKa shifts of the residues involved, the effects were not large.

The factors that account for the small stabilizing effects of salt bridges have been clearly elucidated in the above studies. Firstly, coulombic interactions between charged residues on the surface of proteins are largely screened by the high dielectric solvent. Secondly, there is a free energy penalty associated with desolvating both charges when a salt bridge forms. Finally, there is an entropic cost associated with fixing the mobile side chains in a well defined salt bridge. Indeed, in many cases, the residues involved in the salt bridge are not seen in the crystal structure, apparently preferring to remain mobile and well solvated. Nonetheless, if some other factor (such as hydrophobic interactions between the hydrocarbon side chains of lysine and glutamate and the intervening alanines in a helix) also favors fixing the side chains in the favored salt-bridge conformation, the apparent stabilizing effect of the ion pair can be high. Some combination of such salt-bridge-favoring interactions is probably responsible for the large stabilizing effect of the Asp70-His31 ion pair in T4 lysozyme [13•] and may have contributed to the effects observed by Marqusee and Baldwin [12] mentioned above.

### Total charge effects

Pace *et al.* [14•] have studied the urea denaturation of RNase A and RNase T1 at pH 2 to 10. Over this range, the total charge of RNase A varies from 18 to  $-2$  while the denaturation energy varies from 2 to 10 kcal mol $^{-1}$ . For RNase T1, the total charge varies from 6 to  $-12$ , and the denaturation energy is 3–9 kcal mol $^{-1}$ . These observations, in keeping with earlier studies of Hollecker and Creighton [15], indicate that the magnitude of electro-

static interactions involving charged residues are of the same order as the denaturation energies of proteins. It should be emphasized that these effects can arise only from the few groups that have anomalous pKa's, as nicely demonstrated by the recent study of McNutt *et al.* [16].

### Buried charges

Whereas charges on surfaces may do very little to stabilize proteins, buried charges are likely to be extremely destabilizing unless the protein succeeds in compensating for the loss of aqueous solvation. For example, the active-site Asp26 in *Escherichia coli* thioredoxin is buried in a hydrophobic region and has an anomalous pKa of 7.5, suggesting that it destabilizes the native protein by 5 kcal mol $^{-1}$  [17•]. Similarly, a lysine that replaces Val66 of staphylococcal nuclease was found to be buried in the hydrophobic core, to have a pKa  $< 6.4$  and to destabilize the protein [18•]. In contrast, McGrath *et al.* (M McGrath *et al.*, unpublished data) have replaced Ser214 in trypsin with a lysine residue which, though buried, retains its charge as a result of favorable interactions with hydrogen-bonding groups in the vicinity.

### Acid denaturation

Hughson *et al.* [19•] have detected a partially unfolded intermediate in apomyoglobin which exists at pH 4–5. On the basis of two-dimensional NMR experiments, which were used to identify slowly and rapidly exchanging amide protons, a model for the intermediate was proposed in which a compact subdomain consisting of helices A, G and H retains native-like structure. The remaining helices appear to unfold at the same pH. The partial unfolding is likely to be caused by the presence of histidines with anomalous pKa's, such as His113, His119 and His24, which titrate in this pH range. We have found that the protonation of a small number of these residues significantly decreases the stability of the native protein, partly as a result of their burial in the interior and partly because of their repulsive interactions with other charges (A-S Yang and B Honig, unpublished data). As these groups are located near the interface between the stable subdomain and the remaining helices, this destabilizing electrostatic force can be relieved by partial unfolding.

The classical model of Linderstrom-Lang assumes that the electrostatic free energy of a protein results from the net repulsions between charged groups in the native state [6]. Stigter *et al.* [20] have introduced a model that can account for net charge effects in the unfolded state as well. The recent experimental evidence summarized above, however, suggests that, in some cases, the contribution of ionizable amino acids to protein stability arises from a small number of residues with anomalous pKa's rather than from the effects of net charge. Indeed, the fact that increasing salt concentration destabilizes apomyoglobin near neutral pH, and stabilizes a partially denatured intermediate at low pH [21], suggests that coulombic interactions are attractive in the former case and only repulsive at low pH.

## Calculating electrostatic properties

### pKa's

The intrinsic pKa of a single ionizable group in a protein is usually defined as:

$$\text{pKa}_i^{\text{int}} = \text{pKa}_i^0 - \frac{\gamma(i)\Delta\Delta G_i^{\text{env}}}{2.3kT} \quad (2)$$

where  $\text{pKa}_i^0$  is the pKa of the  $i^{\text{th}}$  group in an isolated model compound and  $\gamma(i)$  is +1 when the group is basic and -1 when acidic.  $\Delta\Delta G_i^{\text{env}}$  is the change in electrostatic free energy associated with charging the group in the protein environment relative to the same process in the model compound [1,22•].  $\Delta\Delta G_i^{\text{env}}$  depends on both the extent to which the group is desolvated in the protein and on any compensating interactions with permanent dipoles.

When a protein has more than one titratable group, the proton affinities at titrating sites become mutually dependent. In this case,  $\text{pKa}_i$  can only be determined from a calculated titration curve, which will yield the pH at which the group is 50% protonated. When more than one ionizable group is being considered, it is convenient to define the intrinsic  $\text{pKa}_i$  as the pKa that group  $i$  would have if all other groups were in their neutral form. The pKa of group  $i$  can now be written in the form:

$$\text{pKa}_i = \text{pKa}_i^{\text{int}} + \Delta\text{pKa}_i^{\text{tr}} = \text{pKa}_i^0 - \frac{\gamma(i)\Delta\Delta G_i^{\text{env}}}{2.3kT} + \Delta\text{pKa}_i^{\text{tr}} \quad (3)$$

where  $\Delta\text{pKa}_i^{\text{tr}}$  is simply the difference between the intrinsic pKa and the one obtained from the titration curve.

The titration curve of each ionizable group in a protein can be obtained from a statistical mechanical average over the  $2^N$  states that arise from  $N$  ionizable residues. The average charge,  $\langle\rho_i\rangle$  of each group is defined by the expression:

$$\langle\rho_i\rangle = \frac{\sum_{n=1}^{2^N} \delta_n(i)\gamma(i)e^{-\frac{\Delta G^n}{kT}}}{Q} \quad (4)$$

where  $\delta_n(i) = 0$  when group  $i$  is neutral in state  $n$ , and  $\delta_n(i) = 1$  when group  $i$  is charged.  $Q$  is the partition function of the system of  $2^N$  states. A reference state of zero free energy is defined to correspond to all ionizable groups in their neutral forms.  $\Delta G^n$  is the free energy of the  $n^{\text{th}}$  state and is given by:

$$\Delta G^n = \sum_{i=1}^N \delta_n(i) \left\{ \gamma(i)2.3kT(\text{pH} - \text{pKa}_i^{\text{int}}) + \sum_{1 \leq j < i} \delta_n(j)\Delta G^{ij} \right\} \quad (5)$$

where  $\Delta G^{ij}$  is the electrostatic interaction energy between the charged forms of groups  $i$  and  $j$ . Thus, if both groups  $i$  and  $j$  are charged in the  $n^{\text{th}}$  state, this term will account for the additional free energy resulting from their interaction relative to the reference state where both are neutral.

Equation 4 and 5 provide a prescription for calculating the titration curves of  $N$  ionizable residues. It is only necessary to determine the average protonation states, or the fractional charges, at each pH of interest. The large number of possible states, however, makes this approach impractical as  $N$  reaches several tens of titratable groups. The first attempt to circumvent the problem was introduced by Tanford and Roxby [23], who assumed that each residue could be defined in terms of an average protonation state, which was then used to calculate inter-residue interactions. This approximation is generally valid but breaks down when the two groups have similar pKa's and the interaction between them is large [22•].

Bashford and Karplus [22•] have calculated electrostatic free energies with the finite difference Poisson-Boltzmann method (see review by Sharp and Honig [1]) and obtained pKa's from the statistical mechanical expression. They limited the number of states that had to be considered by introducing a reduced site approximation in which residues whose pKa's were far from the pH of interest were assumed to be in their appropriate protonation states. More recently, Beroza *et al.* [24] used a Monte Carlo approach to deal with the problem of sampling a large number of possible states. We have developed a hybrid method which exploits the Tanford-Roxby treatment for weakly interacting groups and uses the statistical mechanical expression for strong interactions (A-S Yang *et al.*, unpublished data).

Perhaps the greatest uncertainty in such calculations arises from the uncertainty in the three-dimensional structure. Problems can arise from: possible errors in the X-ray coordinates; conformational changes that may accompany protonation and deprotonation; uncertainty in the position of polar hydrogens; and, possible differences between the crystal and solution conformations. Nevertheless, good agreement with experiment has been obtained for those proteins that have been studied to date. Calculations are generally within about 1 pKa unit from experimental values, although larger errors are obtained in a number of cases. At this stage, the calculations provide an important tool with which to interpret experimental results but they have not yet reached the stage where they have true predictive value.

### Free energies of unfolding

The contribution of ionizable groups to the electrostatic free energy of a folded protein can be obtained by a statistical mechanical sum over the  $2^N$  possible protonation states, thus:

$$\Delta G^{\text{elec}}(\text{protein}) = \frac{\sum_{n=1}^{2^N} \Delta G^n e^{-\frac{\Delta G^n}{kT}}}{Q} \quad (6)$$

where  $\Delta G^{\text{elec}}(\text{protein})$  is the average electrostatic free energy of the native protein with respect to the neutral reference state.

It should be clear from Equations 4-6 that the ability to calculate a titration curve of both the native and dena-

tured states implies the ability to predict the electrostatic contributions to denaturation. We have found that the summation over  $2^N$  states contained in Equation 6 can be approximated accurately by a sum over the  $N$  ionizable groups, with the expression:

$$\Delta G^{\text{elec}}(\text{folded}) = \sum_{i=1}^N \left\{ \langle \rho_i \rangle 2.3kT(\text{pH} - \text{pK}a_i^{\text{int}}) + \sum_{1 \leq j < i} |\langle \rho_i \rangle| |\langle \rho_j \rangle| |\Delta G^{\text{ij}}| \right\} \quad (7)$$

where  $\langle \rho_i \rangle$  is the average charge on ionizable group  $i$  (A-S Yang and B Honig, unpublished data). In the denatured state, it is generally assumed that the intrinsic  $\text{pK}a$ 's are those of the individual amino acids, and that screening by high dielectric solvent and mobile ions reduces the interaction between ionizable groups to zero. Thus:

$$\Delta G^{\text{elec}}(\text{unfolded}) = \sum_{i=1}^N \langle \rho_i' \rangle 2.3kT(\text{pH} - \text{pK}a_i^0) \quad (8)$$

where  $\langle \rho_i' \rangle$  is the average charge on ionizable group  $i$  of the unfolded protein. Note that the terms  $\langle \rho_i \rangle$ ,  $\langle \rho_i' \rangle$ ,  $\Delta G^{\text{elec}}(\text{folded})$ , and  $\Delta G^{\text{elec}}(\text{unfolded})$  in Equations 7 and 8 are all functions of  $\text{pH}$ . The evaluation of Equation 7 for the folded state involves the same terms that must be evaluated in the calculation of titration curves and can be carried out quite rapidly. The difference,  $\Delta G^{\text{elec}}(\text{unfolded}) - \Delta G^{\text{elec}}(\text{folded})$ , is the  $\text{pH}$ -dependent electrostatic contribution to protein unfolding.

Equations 1, 7 and 8 imply that if the titration curves of both the native and denatured protein are known, the relative denaturation energy of the protein can be determined as a function of  $\text{pH}$ . The titration curves of native proteins can be determined experimentally, or from calculations. Thus, the tools are now in hand to combine theory and experiment so as to gain a full understanding of  $\text{pH}$  and salt effects on protein stability.

### Dipolar groups

There has been considerable uncertainty for many years as to whether hydrogen bonds contribute to protein stability. The most widely held point of view has been that they do not, as every group that forms an intramolecular hydrogen bond in the native state will form a hydrogen bond with water in the denatured state. The accumulating evidence (see above) that  $\alpha$ -helices are marginally stable in water, however, suggests that peptide  $\text{NH}\dots\text{CO}$  hydrogen bonds do in fact contribute to protein stability. Scholtz *et al.* [25•] have found that each hydrogen bond makes an enthalpic contribution of  $\sim 1 \text{ kcal mol}^{-1}$ , which may compensate in part for the loss of conformational entropy associated with fixing the peptide backbone into an  $\alpha$ -helix.

The location of hydrogen bonds in a putative preformed  $\alpha$ -helix is very different from their location in a final three-dimensional structure. The process of tertiary-struc-

ture formation requires that peptide dipoles, as well as many dipolar groups in amino acid side chains, be removed from proximity to water and buried in the protein interior. We have recently found that the free energy penalty associated with this process is quite large; indeed, it appears to be roughly equivalent in magnitude (although opposite in sign) to the hydrophobic free energy resulting from tertiary interactions between non-polar side chains (A-S Yang, K Sharp and B Honig, unpublished data). The existence of these two 'compensating' interactions may account in part for the fact that, despite large individual contributions to the free energy balance of proteins (conformational entropy, hydrophobic effect, etc.), proteins of very different size and average hydrophobicities are all only marginally stable [26].

### Conclusions

A considerable body of experimental evidence now indicates that electrostatic interactions involving ionizable amino acids make only a small contribution to the total free energy balance of protein stability. Nevertheless, the magnitude of the interactions is of the same order as the denaturation energies of proteins and, consequently, charged groups can make important contributions to protein stability. Moreover, electrostatic interactions, and hence protein stabilities, are relatively easy to control through changes in  $\text{pH}$  and salt concentration.

Perhaps the most surprising insight to emerge recently is that most ionizable amino acids in proteins have unshifted  $\text{pK}a$ 's (relative to the isolated amino acid) and hence make essentially no contribution to protein stability. The  $\text{pH}$  and salt effects that have been observed arise from a relatively small number of amino acids with abnormal  $\text{pK}a$ 's rather than from net charge effects which are predicted from the classical Linderstrom-Lang model. Abnormal  $\text{pK}a$ 's can arise from salt bridges between groups whose locations have been fixed in part by non-electrostatic interactions, or from individual residues buried in a non-polar environment.

Recent theoretical advances now make it possible to calculate  $\text{pK}a$ 's in proteins and hence to account for  $\text{pH}$  and salt effects on protein stability. Although the reliability of the calculations is reduced by uncertainties in three-dimensional structure and in the potential functions that are used, they are accurate enough to provide an important tool in the interpretation of experimental results. The existence of marginally stable elements of secondary structure suggests that peptide hydrogen bonds make a significant contribution to protein stability. Theoretical calculations, however, indicate that the burial in the protein interior of these groups and polar groups on amino acid side chains has a large destabilizing effect.

### Acknowledgements

We thank M Gunner and K Sharp for stimulating discussions. This work was supported by NIH grant GM30518.

## References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

- of special interest
- of outstanding interest

1. SHARP KA, HONIG B: Electrostatic Interactions in Macromolecules: Theory and Application. *Annu Rev Biophys Chem* 1990, 19:301–332.
  2. WARSHEL A, AQVIST J: Electrostatic Energy and Macromolecular Function. *Annu Rev Biophys Chem* 1991, 20:267–298.
  3. HOROVITZ A, SERRANO L, AVRON B, BYCROFT M, FERSHT AR: Strength and Co-operativity of Contributions of Surface Salt Bridges to Protein Stability. *J Mol Biol* 1990, 216:1031–1044.
  4. KOHDA D, SAWADA T, INAGAKI F: Characterization of pH Titration Shifts for All the Nonlabile Proton Resonances in a Protein by Two-dimensional NMR: the Case of Mouse Epidermal Growth Factor. *Biochemistry* 1991, 30:4896–4900.
  5. BARKER PD, MAUK MR, MAUK AG: Proton Titration Curve of Yeast Iso-1-ferricytochrome c. Electrostatic and Conformational Effects of Point Mutations. *Biochemistry* 1991, 30:2377–2383.
  6. TANFORD C: Protein Denaturation, Part C. *Adv Protein Chem* 1970, 25:1–95.
  7. DAO-PIN S, SAUER U, NICHOLSON H, MATTHEWS BW: Contribution of Engineered Surface Salt Bridges to the Stability of T4 Lysozyme Determined by Directed Mutagenesis. *Biochemistry* 1991, 30:7142–7153.
- Several mutants of T4 lysozyme are designed to test the free energy contributions of the engineered salt bridges at different locations in the protein. The salt bridges have little effect on stability, which is consistent with the idea that they are not located in a fixed conformation. The crystal structures of the mutant proteins are consistent with this interpretation.
8. BRADLEY EK, THOMASON JF, COHEN FE, KOSEN PA, KUNTZ ID: Studies of Synthetic Helical Peptides Using Circular Dichroism and Nuclear Magnetic Resonance. *J Mol Biol* 1990, 215:607–622.
  9. SALI D, BYCROFT M, FERSHT AR: Surface Electrostatic Interactions Contribute Little to Stability of Barnase. *J Mol Biol* 1991, 220:779–788.
- Site-directed mutagenesis is used to engineer an ion pair (Glu28...Lys32) on the surface of barnase. The side chains are separated by one turn of an  $\alpha$ -helix such that they are able to form a salt bridge. A double mutant cycle is used to measure the electrostatic interaction between these groups, which is found to be only 0.2 kcal mol<sup>-1</sup>.
10. SERRANO L, HOROVITZ A, AVRON B, BYCROFT M, FERSHT AR: Estimating the Contribution of Engineered Surface Electrostatic Interactions to Protein Stability by Using Double-mutant Cycles. *Biochemistry* 1990, 29:9343–9352.
  11. MERUTKA G, SHALONGO W, STELLWAGEN E: A Model Peptide with Enhanced Helicity. *Biochemistry* 1991, 30:4245–4248.
  12. MARQUESE S, BALDWIN RL: Helix Stabilization by Glu<sup>-</sup>...Lys<sup>+</sup> Salt Bridges in Short Peptides of *de novo* Design. *Proc Natl Acad Sci USA* 1987, 84:8898–8902.
  13. ANDERSON DE, BECKTEL WJ, DAHLQUIST FW: pH-induced Denaturation of Proteins: a Single Salt Bridge Contributes

3–5 kcal/mol to the Free energy of Folding of T4-lysozyme. *Biochemistry* 1990, 29:2403–2408.

The free energy contribution of a salt bridge (Asp70...His31) in T4 lysozyme is studied using both NMR and heat-denaturation methods. The results indicate that the salt bridge has an unusually large electrostatic effect, contributing 3–5 kcal mol<sup>-1</sup> to the stability of the protein.

14. PACE CN, LAURENTS DV, THOMSON JA: pH Dependence of the Urea and Guanidine Hydrochloride Denaturation of Ribonuclease A and Ribonuclease T1. *Biochemistry* 1990, 29:2564–2572.

The denaturation energies of ribonuclease A and T1 as functions of pH are reported. The proteins are most stable at pHs which are close, but not identical, to the isoelectric points. The pH dependence suggests that the electrostatic contributions to the stability of these ribonucleases amount to only a few kilocalories per mole.

15. HOLLECKER M, CREIGHTON TE: Effect on Protein Stability of Reversing the Charge on Amino Groups. *Biochim Biophys Acta* 1982, 701:395–404.

16. MCNUTT M, MULLINS LS, RAUSHEL FM, PACE CN: Contribution of Histidine Residues to the Conformational Stability of Ribonuclease T1 and Mutant Glu-58→Ala. *Biochemistry* 1990, 29:7572–7576.

17. LANGSETMO K, FUCHS JA, WOODWARD C, SHARP KA: Linkage of Thioredoxin Stability to Titration of Ionizable Groups with Perturbed pKa. *Biochemistry* 1991, 30:7609–7614.

The pKa of a buried aspartic acid in thioredoxin is determined. A thorough theoretical analysis generates possible interpretations of the unusual properties of this group.

18. STITES WE, GITTIS AG, LAITMAN E, SHORTLE D: In a Staphylococcal Nuclease Mutant the Side-chain of a Lysine Replacing Valine 66 is Fully Buried in the Hydrophobic Core. *J Mol Biol* 1991, 221:7–14.

A buried lysine in a hydrophobic environment is found to have a pKa <6.4 and to significantly destabilize the protein. The paper contains a thoughtful analysis of the free energy changes that accompany the mutation.

19. HUGHSON FM, WRIGHT PE, BALDWIN RL: Structural Characterization of a Partly Folded Apomyoglobin Intermediate. *Science* 1990, 249:1544–1548.

Proton-exchange NMR is used to characterize the acid-denatured intermediate of apomyoglobin at pH 4.2. The partly unfolded intermediate appears to consist of a subdomain involving folded helices A, G and H, whereas helices B and E appear to remain unfolded.

20. STIGTER D, ALONSO DOV, DILL KA: Protein Stability: Electrostatic and Compact Denatured States. *Proc Natl Acad Sci USA* 1991, 88:4176–4180.

21. GOTO Y, FINK AL: Phase Diagram for Acidic Conformational States of Apomyoglobin. *J Mol Biol* 1990, 214:803–805.

22. BASHFORD D, KARPLUS M: pKa's of Ionizable Groups in Proteins: Atomic Detail from a Continuum Electrostatic Model. *Biochemistry* 1990, 29:10219–10225.

The finite difference Poisson–Boltzmann method is used to calculate the pKas of the ionizable groups in lysozyme. An important innovation, in addition to the introduction of the theoretical methodology, is the use of a reduced site approximation to limit the sum over 2<sup>N</sup> states required if all interactions are considered. Most of the calculated pKas are consistent with experimental data, although a number of noticeable deviations are encountered.

23. TANFORD C, ROXBY R: Interpretation of Protein Titration Curves. Application to Lysozyme. *Biochemistry* 1972, 11:2192–2198.

24. BEROZA P, FREDKIN DR, OKAMURA MY, FEHER G: **Protonation of Interacting Residues in a Protein by a Monte Carlo Method: Application to Lysozyme and the Photosynthetic Reaction Center of *Rhodobacter sphaeroides*. *Proc Natl Acad Sci USA* 1991, 88:5804-5808.**
25. SCHOLTZ JM, MARQUESE S, BALDWIN R, YORK EJ, STEWART JM, SANTORO M, BOLEN DW: **Calorimetric Determination of the Enthalpy Change for the Alpha-helix to Coil Transition of an Alanine Peptide in Water. *Proc Natl Acad Sci USA* 1991, 88:2854-2858.**

A calorimetric method is used to estimate the enthalpy of  $\alpha$ -helix to random-coil transition. A range of  $\Delta H$  [0.9-1.3 kcal (mol residue)<sup>-1</sup>] is

determined for the unfolding of the  $\alpha$ -helix. The enthalpy of unfolding results largely from the disruption of peptide hydrogen bonds.

26. PRIVALOV P, GILL SJ: **Stability of Protein Structure and Hydrophobic Interaction. *Adv Protein Chem* 1988, 39:191-234.**

A-S Yang and B Honig, Department of Biochemistry and Molecular Biophysics, Columbia University, 630 West 168 St, New York, New York 10032, USA.

Introduction

Understanding protein stability, which is determined by the difference between the free energies of the native and denatured conformations, is an important part of understanding how the sequence of a protein determines its three-dimensional structure. Moreover, issues of stability affect how we think about protein evolution, design, and folding pathways. Comparing the stability and structure of proteins differing by one or a few amino acids has been one of the major ways of testing and refining our knowledge of the determinants of protein stability.

Tolerance to amino acid substitutions

Genetic studies can be used to determine the tolerance of positions in a protein to amino acid substitutions and thereby to determine the relative importance of different residues to protein structure and function [1,2,3-5]. The most complete work in this area is that of Fersht and his colleagues [2\*], who have used mutant suppression to determine the effects of substituting Asp, Ala, Ser, Cys, Phe, Leu, Glu, His, Phe, Tyr, Arg, and Lys at 105 of the 104 residue positions in phage T4 lysozyme. Using a genetic screen in which 5% or more of the wild-type level of activity was required, only 328 of the 3025 substitutions were found to be deleterious. At more than half of the positions, all of the substitutions were acceptable. At the remaining positions, an average of two-thirds of the substitutions were acceptable.

Positions in the hydrophobic core of T4 lysozyme were generally most sensitive to substitution: the residues were more with proline and the polar amino acids Glu, Lys, and Arg being notably excluded from most interior positions [2\*]. Similar trends have been observed in other systems and in general, buried residues are far less tolerant of substitution than surface residues [3,6]. Some surface residues, however, including those involved in salt bridges, hydrogen bonds and aromatic-aromatic interactions, can also have restricted substitution patterns which suggest a role for them in protein stability [2\*].

Two somewhat surprising results of the T4 lysozyme study were that glutamine could be accommodated at many buried positions and that proline could be accommodated at many surface positions in a helix [2\*]. In contrast to buried residues in  $\lambda$  repressor, strongly polar residues like glutamine were never allowed at buried positions and proline was never allowed in a helix [3,7]. These differences probably reflect the fact that such substitutions are destabilizing (see below), with some positions being more stable and thus less sensitive to destabilization. Alternatively, such differences could reflect the sensitivity of the genetic tests of protein activity used in the two systems.

Tolerance to insertions and deletions

If proteins can accommodate most side-chain substitutions without loss of function, then what about more dramatic changes such as insertions or deletions? Several studies [8\*] concentrated and examined the effects of single glycine or alanine insertions at 30 sites within  $\alpha$ -helices,  $\beta$ -strands, turns, loops or extended regions of staphylococcal nuclease. Although some of the mutants had extremely low enzymatic activity, indicating changes in structure, the remarkable finding was that mutants with insertions at 14 of the sites were fully active. Clearly, although there must be some structural changes in the mutant proteins, such changes are not propagated to the active site. As might be expected, the insertion mutants were less stable than wild-type nuclease, with an average destabilization of roughly 0.5 kcal mol<sup>-1</sup> for insertions in  $\alpha$ -helices and  $\beta$ -strands, and 1.0 kcal mol<sup>-1</sup> for those in loops, turns, etc. Although these changes may seem large, at many positions the insertions were no more destabilizing than substitution mutants in which one or the other of the flanking residues were replaced with glycine or alanine. Deletion of a side chain to the C $\alpha$  or C $\beta$  position will simply remove atoms from the structure whereas an insertion of an entire residue will move the flanking residues apart by at least 3 Å and will introduce a number of additional atoms. That the protein is able to accommodate such modifications indicates that there must be