Journal of Cellular Biochemistry 29:321–335 (1985)
Protein Structure, Folding, and Design 95–109

Dissecting the Roles of Individual Interactions in Protein Stability: Lessons From a Circularized Protein

David P. Goldenberg

Medical Research Council, Laboratory of Molecular Biology, Cambridge CB2 2QH, England

A circular form of bovine pancreatic trypsin inhibitor (BPTI) has been prepared by introducting a peptide bond between the N- and C-termini, which are in close proximity in the native conformation. The pathway and energetics of the disulphide-coupled folding transition of the circular protein have been studied using methods applied previously to the unmodified protein. The cross-link between the termini was found not to significantly stabilize the native state in spite of the expected reduction in entropy of the unfolded protein. This unexpected result has led to a reexamination of the stabilization expected from a cross-link, considering effects on the native, as well as unfolded, states of the protein. The greatest stabilization is expected when the cross-linked groups are held rigidly in the native protein in the optimum orientation for forming the cross-link. Similar analyses, utilizing thermodynamic cycles, can be applied to other interactions that stabilize native proteins, including disulphide bonds, salt bridges, and hydrogen bonds and to modifications to the protein that remove them. In general, the contribution of an individual interaction to the stability of the native state depends on the extent to which the interaction is favored in the native conformation, which can vary greatly depending on the local environment of the interacting groups.

Key words: salt bridges, disulphide bonds, protein cross-links, protein stability, bovine pancreatic trypsin inhibitor, effects of mutations on protein stability

The demonstration that an active, native protein could be spontaneously regenerated from a polypeptide chain that had been denatured to an unfolded, random-coil state showed that the three-dimensional structures of proteins are determined by their amino acid sequences through intramolecular interactions and interactions with the solvent [1]. The determination of the three-dimensional structures of a large number of proteins, by X-ray crystallography, has illustrated the subtle ways in which the

David Goldenberg's present address is Department of Biology, University of Utah, Salt Lake City, UT 84112.

Received April 17, 1985; revised and accepted August 6, 1985.

© **1985 Alan R. Liss, Inc.**

various residues of a sequence fit together to generate a stable, well packed structure. However, because of the high degree of cooperativity of folded protein structures, surprisingly little is known about the contributions of individual residues or interactions to the stabilities of folded proteins. Still less is known about the roles of particular residues in directing the *process* of folding to the native conformation.

One traditional approach to examining the roles of particular interactions is the use of covalent modifications that either remove or introduce an interaction. Modified proteins have been prepared by both chemical methods and by genetic mutations. The use of genetic manipulations to examine protein structures or functions has been a recurring theme in molecular biology [2–8] but has received a great deal of renewed attention in the past few years as new medthods of *in vitro* DNA manipulation and synthesis have made it possible to alter amino acid sequences with almost unlimited flexibility. By examining the stabilities of proteins that differ only by the presence of a particular interaction, it should be possible to determine the contribution of that interaction to the stability of the protein. An important limitation to this approach is that a measurement of protein stability represents an equilibrium between the native and unfolded states of the protein, and a change in net stability may arise from changes in either the native or unfolded forms or both.

One interesting class of protein modifications, of which several examples have been examined, is that of covalent cross-links between residues that are close together in the native structure but can be quite far apart in the amino acid sequence. Artificial cross-links, introduced by treating native proteins with bifunctional reagents, have been found to stabilize the native conformation by as much as 5 kcal/mol [9–11]. Disulphide bonds between crysteine residues are found in many proteins, and in conditions under which they cannot be reduced or exchanged with other thiols, disulphides are irreversible cross-links. Recently, genetic engineering has been utilized to introduce disulphide bonds into native proteins that do not normally have them [12,13]. This might be a useful method of increasing the stability of a protein to allow it to function under more demanding conditions than might otherwise be possible. The stabilization by a cross-link has been attributed to the reduction in entropy, and the resulting destabilization, of the unfolded state, and various authors have attempted to predict the magnitude of this effect by considering only the unfolded state [9,11,14,15].

A cross-linked form of bovine pancreatic trypsin inhibitor (BPTI) has recently been prepared in which the N- and C-termini were linked together in a peptide bond [16]. The energetics and pathway of the folding transition between the reduced and native forms of this circular protein have been examined using methods previously applied to unmodified BPTI [17]. Surprisingly, the cross-link between the termini was found not to stabilize significantly the native conformation of BPTI.

This unexpected result has prompted a reexamination of the theoretical treatment of protein cross-links, with an emphasis on the relative effects on the unfolded and native states. Thermodynamic cycles are useful conceptual tools for considering this type of problem, and this approach can be extended to treat, at least qualitatively, other types of reversible interactions, and modifications that disrupt them, as discussed here.

CIRCULAR BPTI

Bovine pancreatic trypsin inhibitor has been widely utilized as a model protein for both experimental and theoretical studies of protein conformation and folding [18–

23]. A detailed pathway of folding for BPTI has been experimentally determined by trapping and identifying disulphide-bonded kinetic intermediates in the transition between the reduced, unfolded protein and the native state, which contains three disulphide bonds [19,22]. Because it has been so extensively characterized, BPTI is especially well suited for modification studies, particularly in that the effects of a modification on the folding intermediates, as well as on the native conformation, can be examined.

In the native conformation, the termini of BPTI are in quite close proximity (Fig. 1a) [18,23], and treatment of the native protein with a water-soluble carbodiimide leads to the formation of a peptide bond between the terminal amino and carboxyl groups, generating a circular backbone (Fig. 1B) [16]. This circular form of BPTI (C-BPTI) is fully functional as a trypsin inhibitor and, at moderate pH values, has electrophoretic and chromatographic properties nearly indistinguishable from those of the unmodified protein, indicating that the cross-link does not greatly alter the overall conformation of the protein.

Circular BPTI can be further modified by cleaving the backbone at another position to generate a new, circularly permuted, linear sequence. One permuted form has been prepared by cleaving C-BPTI between Lys-15 and Ala-16 (Fig. 1c), which is the peptide bond that enters the active site of trypsin when the inhibitor binds. Such circularly permuted forms of BPTI might have very interesting folding properties, particularly if the cleavage separates residues that are normally involved in short-range interactions at a crucial stage of folding.

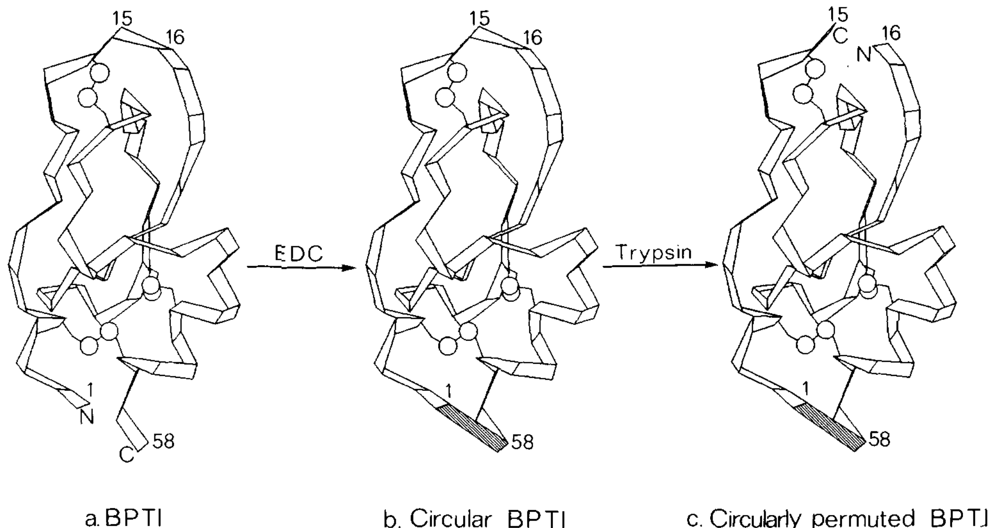


Fig. 1. Backbone representations of BPTI (a), circular BPTI(b) and circularly permuted BPTI (c) drawn from the coordinates of the x-ray crystal structure [18]. The circular form of BPTI was generated by treating the native protein with a water-soluble carbodiimide (1-ethyl-3-(3-dimethylamino-propyl) carbodiimide, EDC) to form a peptide bond between the terminal residues. The circularly permuted form was made by treating the circular form with trypsin (after selectively reducing the 14–38 disulphide bond to allow greater flexibility in the region of the trypsin binding site [37]. In b and c, the connection shown between the termini is only schematic, and no attempt has been made to draw a stereochemically correct model. From Goldenberg and Creighton [16], with permission.

Both the circular and circularly permuted forms of BPTI can refold to the active, native conformation after being unfolded by reduction of their disulphide bonds. The pathway of folding of the circular form has been examined in detail, as was done previously for the unmodified protein. The details of this analysis and the resulting pathway have been published [17] and will be described only briefly here.

The pathway of folding of C-BPTI was studied by trapping kinetic intermediates with various disulphide bonds by chemically blocking any free thiols. The intermediates were fractionated by ion exchange chromatography and the disulphide bonds determined by peptide mapping. The roles of these intermediates were determined by kinetic analysis of both folding and unfolding reactions in the presence of varying concentrations of disulphide and thiol reagents. The concentrations of the intermediates present during the course of these reactions were determined by trapping the various species and resolving them by gel electrophoresis. The data were analyzed by comparing the observed concentrations of the various species with those predicted by particular kinetic models. The predictions were generated by numerical integration of the rate expressions making up the models, using different rate constants, until a single model, made up of a single set of first- and second-order rate constants, was found that accurately simulated all of the experiments.

The outline of the C-BPTI folding pathway, shown schematically in Figure 2, is very similar to that of the unmodified protein. Most of the major intermediates in the folding of BPTI were also identified in the folding of C-BPTI. As was found for the unmodified protein, intramolecular rearrangements of disulphide bonds in intermediates at the 2-disulphide stage of folding are essential for the formation of the native protein.

From the measured forward and reverse rate constants for all of the steps making up the folding pathway, the relative stabilities of the intermediates and the native conformation were calculated and compared with those for the protein without the cross-link, measured in the same way under the same conditions. It was anticipated that the cross-link between the termini would stabilize the native conformation by reducing the entropy of the unfolded chain and would similarly stabilize those intermediates in which the termini tend to be in close proximity. Thus, it would be possible to deduce at what stage of folding of the unmodified protein the termini are brought together. However, it was found that the stability of the native state of the circular protein, relative to the reduced state, was nearly identical to that of the unmodified protein. This result suggests that the native state of the protein is somehow destabilized by the cross-link to compensate for the expected effect on the unfolded state.

THE EFFECTS OF IRREVERSIBLE COVALENT CROSS-LINKS ON PROTEIN STABILITY

Thirty years ago, Schellman [14] and Flory [15] first considered the problem of estimating the effect of a cross-link on the stability of a protein. These and subseqent authors [9,11] utilized a thermodynamic cycle, as illustrated in Figure 3 for the case of circular BPTI, made up of the folded and unfolded states of the protein with and without the cross-link. Linkage between the equilibria of the cycle requires that the ratio of the equilibrium constants for folding the forms with and without the cross-

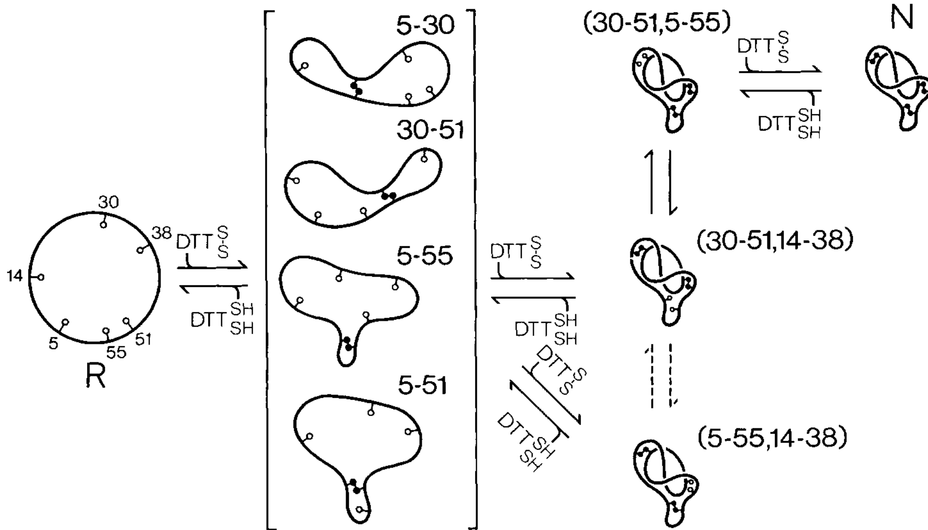


Fig. 2. Pathway of folding and unfolding of circular BPTI determined by trapping and identifying disulphide-bonded intermediates. R is the reduced unfolded protein, N the native protein with three disulphides (30–51, 5–55, and 14–38), and the other species are identified by their disulphide bonds. The one-disulphide intermediates, grouped together in the brackets, appear to be in rapid equilibrium with one another. No attempt has been made to indicate specific conformations in these species. The three major two-disulphide intermediates each contain two of the three disulphides present in the native protein, and their gel electrophoretic mobilities indicate that they have nativelike compact conformations. As for unmodified BPTI, formation of the native protein requires intramolecular rearrangements to produce (30–51, 5–55) before the third disulphide can be incorporated. The thiols of the two-disulphide species (5–55, 14–38) are buried and unreactive with thiol blocking or exchange reagents. It cannot readily form a third disulphide and only slowly rearranges to other two-disulphide species. The details of this pathway and its elucidation are presented in Goldenberg and Creighton [17].

link (K_f^{xl} and K_f) be equal to the ratio of the equilibrium constants for forming the cross-link in the native and unfolded states (K_{xl}^N and K_{xl}^U):

$$K_f^{xl}/K_f = K_{xl}^N/K_{xl}^U \; . \qquad (1)$$

The energetic contribution of the cross-link to the native conformation is then:

$$\Delta G_{xl} = RT\ln\,(K_f^{xl}/K_f) = RT\ln\,(K_{xl}^N/K_{xl}^U) \; . \qquad (2)$$

Thus the degree of stabilization depends on the extent to which the formation of the cross-link is favored in the native conformation relative to the unfolded form.

A convenient device for comparing chemically equivalent intramolecular reactions, such as the formation of the cross-link in the two states, is the "effective concentration" (C_{eff}) [24,25], which is calculated by dividing the intramolecular equilibrium constant (K_{intra}) by that for an equivalent intermolecular reaction (K_{inter}):

$$C_{eff} = K_{intra}/K_{inter} \; . \qquad (3)$$

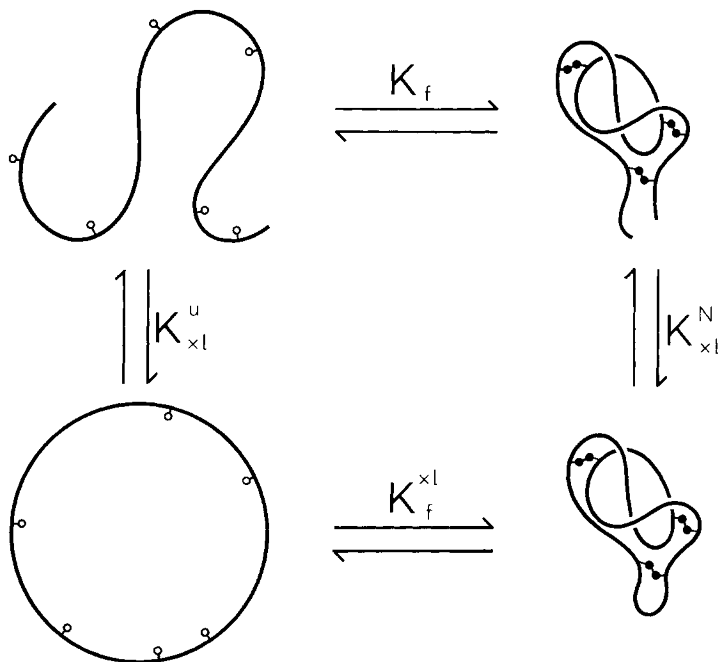


Fig. 3. A hypothetical thermodynamic cycle made up of the unfolded and native conformations of BPTI with and without the peptide bond between the termini. Linkage between the four reactions of the cycle requires that the ratio of equilibrium constants for folding the forms with and without the cross-link (K_f^{xl} and K_f) be equal to the ratio of the equilibrium constants for forming the cross-link in the native and unfolded forms (K_{xl}^N and K_{xl}^U). Thus the stabilization of the native conformation depends on the extent to which the formation of the cross-link is favored in the native state relative to the unfolded state.

Expressed in terms of the effective concentrations of the cross-linked groups in the two states (C_{eff}^N and C_{eff}^U), the stabilization from the cross-link is:

$$\Delta G_{xl} = RT\ln\,(C_{eff}^N/C_{eff}^U)\ . \tag{4}$$

For most cross-links, including the peptide bond of C-BPTI, it is not feasible to measure directly the equilibrium constants, or effective concentrations, for the formation of the cross-link. However, for a disulphide bond, it is quite easy to determine the effective concentration, because the stability of a protein disulphide can be measured by exchange with a similar small molecule intermolecular disulphide:

$$P_{SH}^{SH} + RSSR \leftrightarrow P_{S}^{S} + 2RSH. \tag{5}$$

The effective concentration of the protein disulphide is simply the equilibrium constant for this overall exchange reaction. In this way, the effective concentrations of the disulphides of BPTI have been measured for both the unfolded and native states [22]. The effective concentrations in the unfolded protein are about 10^{-2} M; those in the native protein range from 2×10^2 to 5×10^5 M. These very large values for the

native protein arise because relatively little entropy is lost in forming the protein disulphide as compared with the formation of an intermolecular disulphide [25]. Of equal importance is the great range of values for C^N_{eff} for the same type of interaction within the same folded protein. Presumably, this variation arises because of differences in flexibility, orientation, packing, or possibly other factors in the local environment of the different disulphides. From Equation 4, it is clear that these differences in the native protein will lead to very different stabilization energies for the three different BPTI disulphides, from about 6 to 10 kcal/mol.

In general, it should be anticipated that different cross-links will contribute different net stabilizations and that these differences can arise from differences in effects on the native conformation as well as on the unfolded protein. In that the net stabilities of native proteins are typically only about 15 kcal/mol [26,27], the differences owing to differential effects on the native state can be very significant. The most effective cross-links are expected to be those between groups with high effective concentrations in the native state, *ie,* those that are rigidly held in an optimum orientation for forming the cross-link. This conclusion might appear to be counter-intuitive; it might be thought that the greatest stabilization would come from "tying down" a flexible, weakly bonded region of a molecule. However, a cross-link between groups with a low effective concentration will destabilize the native protein by reducing its entropy or introducing strain.

This relationship between C^N_{eff} and expected stabilization from a cross-link is nicely illustrated by the BPTI disulphides. Under conditions where the disulphides cannot be reduced or undergo exchange with other thiols, the disulphides behave as irreversible cross-links and stabilize the native conformation towards unfolding induced by temperature or denaturants. Forms of BPTI lacking each of the three disulphides can be trapped as intermediates in the disulphide-coupled unfolding or refolding of the protein, and the thermal stabilities of these forms have been measured, it only approximately (Table 1). The disulphide with the highest effective concentration in the native conformation is 5–55, and in its absence the melting temperature is reduced from nearly 100°C to about 35°C. In contrast, removal of the disulphide with the lowest C^N_{eff} , 14–38, also reduces the melting temperature but only by about 20°C.

Several authors have attempted to predict the stabilization from a cross-link by considering only the unfolded chain [9,11,14,15]. These estimates utilize the Jacobson-Stockmayer relationship [28] to calculate the probability (P_U) that the ends of a random coil, made up of N statistical segments of length l, will simultaneously lie within a small volume element (V):

$$P_U = (3/2\pi Nl^2)^{3/2}V . \quad (6)$$

TABLE I. Contributions of the Disulphides of BPTI, as Irreversible Cross-Links, to the Stability of the Native Conformation

Disulphide	Effective concentration in the native state (M) [22]	Melting temperature in the absence of disulphide (°C)
5-55	4.6×10^5	<35 [38]
30-51	1.7×10^3	>37 [39]
14-38	2.3×10^2	75 [40]
Protein with all disulphides intact		95 [27,41]

It is argued that the stabilization of the native protein from the cross-link will be the loss of entropy owing to confining the ends of the unfolded chain in the volume element:

$$\Delta G_{xl} = -RT\ln P_U. \tag{7}$$

This treatment assumes that the cross-link will cause no loss of entropy in the native state and that the probability of the ends lying within the volume element in the native conformation, P_N, is unity. The calculation depends upon choosing a value for the volume element, V. The volume element is related to the effective concentration in the native protein according to:

$$P_N = C_{eff}^N \times V \times N_A = 1, \tag{8}$$

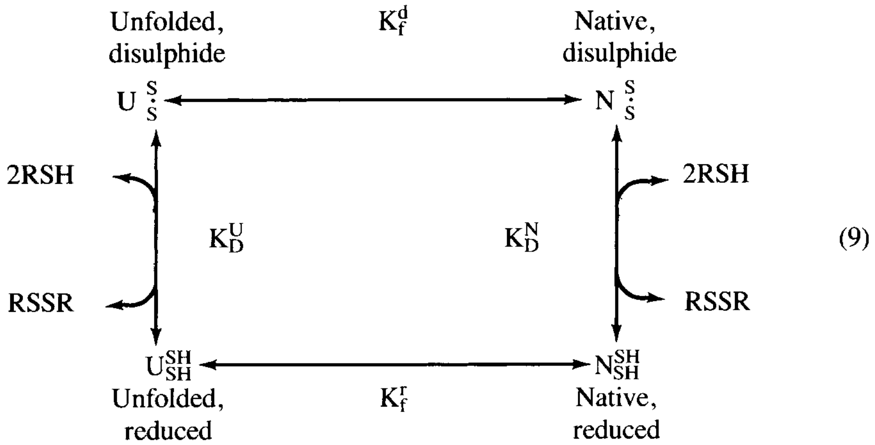
where C_{eff}^N and V have the units of molarity and liter, respectively, and N_A is Avogadro's number. Thus, the chosen volume element represents an assumed value for the effective concentration in the native protein.

Calculations of stabilization from cross-links typically use as volume elements spheres or spherical shells of 1.5–3 Å radius. These volumes correspond to values of 10–10^3 M for C_{eff}^N. Although these values are certainly feasible for folded proteins, measurements of effective concentrations of disulphides show that much higher, and probably lower, values are possible. Thus it should be anticipated that the observed stabilization from a cross-link will be either greater or less than that predicted using an arbitrarily chosen volume element.

An alternative approach is to use Equation 6, or more detailed statistical treatments that consider the geometries of the polypeptide chain [29] or the influence of other cross-links in the unfolded chain [11], to make an estimate of C_{eff}^U, as in the original treatment by Jacobson and Stockmayer [28]. The observed stabilization from a cross-link can then be used to estimate the effective concentration in the native state, according to Eq 4. Such a calculation is limited by the accuracy of the treatment of the unfolded chain but can provide a useful indication of the tendency of the groups to interact in the folded conformation. For the case of the BPTI termini, C_{eff}^U can be estimated to be about 10^{-3} M. A similar value has been measured for the disulphide of an immunoglobulin domain in which the Cys residues are separated by about the same number of residues [30]. Because the stability of the native state of BPTI is not increased by the cross-link, the effective concentration in the native state must also be about 10^{-3} M. This extremely low value, especially as compared with those of the disulphides in the same protein, indicates that the conformation of the native protein disfavors the correct orientation between the terminal amino and carboxyl groups for formation of the peptide bond. As a consequence, the bond strains and destabilizes the native conformation.

THE CONTRIBUTIONS OF DISULPHIDES, AS REVERSIBLE INTERACTIONS, TO PROTEIN STABILITY

In the presence of exchanging thiol and disulphide reagents, protein disulphide bonds are freely reversible, rather than irreversible cross-links as they were treated in the previous section and Table I, and provide a useful model for other interactions in proteins. The simplest case is that of a protein containing two Cys residues that can form a disulphide in either the native or unfolded conformations. The four possible forms are related by a thermodynamic cycle:



(9)

As discussed in the previous section, the two equlibrium constants for forming the protein disulphide:

$$K_d = (P^S_S \times RSH^2)/(P^{SH}_{SH} \times RSSR) , \quad (10)$$

are the effective concentrations of the protein disulphide in the two forms.

Suppose that the various species are at equilibrium in the presence of particular concentrations of the thiol and disulphide reagents, RSH and RSSR, and the overall equilibrium between the native ($N^S_S + N^{SH}_{SH}$) and unfolded ($U^S_S + U^{SH}_{SH}$) forms of the protein is measured. The usual method of making this sort of measurement is to monitor the unfolding transition induced by temperature or denaturants and to extrapolate from the region of the transition, where both native and unfolded forms are detectable, to physiological temperature or zero denaturant [26]. For this particular experiment, the equilibrium would be measured with a method, such as an optical probe, that does not distinguish between the reduced and disulphide-bonded forms of N and U. The observed equilibrium will be:

$$K_f = N/U = \frac{N^{SH}_{SH}}{U^{SH}_{SH}} \times \frac{(1 + C^N_{eff} \times RSSR/RSH^2)}{(1 + C^U_{eff} \times RSSR/RSH^2)} , \quad (11)$$

and the net stability of the native state will be:

$$-\Delta G_f = RT\ln K_f . \quad (12)$$

To determine, experimentally, the contribution of the disulphide bond to this net stability, the formation of the protein disulphide could be prevented by chemically blocking the thiols or, better yet, by genetically converting the Cys residues to Ala or Ser. In the absence of the protein thiols, only the lower half of the cycle of Eq 9 will contribute to the observed equilibrium between N and U. The apparent contribution of the disulphide will be the difference in net stability when the disulphide can and cannot be formed:

$$\Delta G_{s-s} = RT\ln (N/U) - RT\ln (N_{SH}^{SH}/U_{SH}^{SH}) \qquad (13)$$

$$\Delta G_{s-s} = RT\ln \frac{(1 + C_{eff}^{N} \times RSSR/RSH^2)}{(1 + C_{eff}^{U} \times RSSR/RSH^2)}$$

As was seen earlier for the case of an irreversible cross-link, the contribution of a reversible disulphide depends on the extent to which C_{eff}^{N} is greater than C_{eff}^{U} but in a slightly more complex way. In particular, when the disulphide is reversible, its contribution depends on the evironment, in the form of the thiol-disulphide redox potential, as well as the properties of the native and unfolded protein. Provided that C_{eff}^{N} is greater then C_{eff}^{U}, the stabilization provided by the disulphide will be greater the more its formation is favored by the environment.

These relationships are illustrated in Figure 4, in which the expected apparent contribution of a disulphide, calculated according to Eq 13, is plotted over a range of redox potentials, expressed as $RSSR/RSH^2$, for several values of C_{eff}^{N}. For the calculations, C_{eff}^{U} is assumed to be 10^{-2} M. As $RSSR/RSH^2$ is increased, the

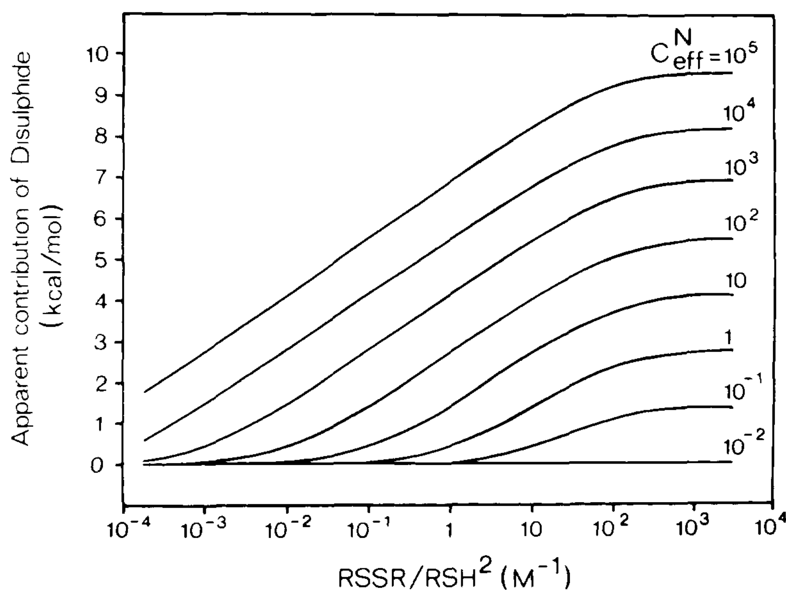


Fig. 4. The predicted contribution of a reversible disulphide bond to the net stability of the folded conformation of a protein as a function of the thio-disulphide redox potential (RSSR/RSH2). The apparent contribution is the reduction in stability when formation of the disulphide is prevented by modifying the protein to remove the thiols. The curves were calculated from Eq 13 assuming that the effective concentration of the disulphide in the unfolded protein, C_{eff}^{U} is, 10^{-2} M and the temperature 25°C for the indicated values of the effective concentration in the native protein (C_{eff}^{N}, M). The contribution is greater the more the formation of the disulphide is favored in the native state relative to the unfolded and the more the disulphide is favored by the redox potential. The abscissa, RSSR/RSH2, represents the equilibrium constant for an intermolecular disulphide under the same conditions, and the graph can be applied to other interactions by treating the abscissa as K_{inter} (M^{-1}) for the appropriate intermolecular reaction.

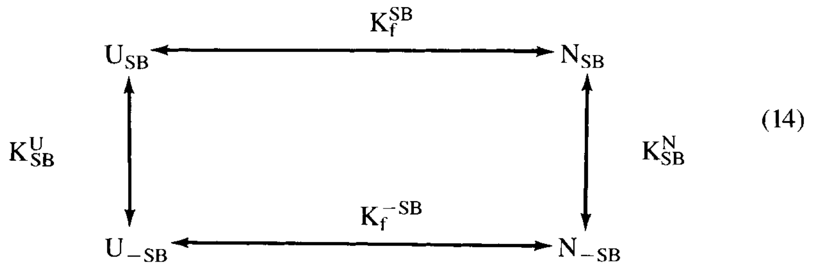
contribution of the disulphide increases asympotically towards the value that would be expected if the disulphide was an irreversible cross-link. Higher values of C^{N}_{eff} have two effects: First, the maximum contribution of the disulphide is greater, and, second, the disulphide can make a significant contribution under more reducing conditions.

Figure 4 illustrates that, because of the high effective concentrations of disulphides in native proteins, they can contribute to the stability of the native state even under conditions that would otherwise appear to be very unfavorable for their formation. For instance, the intracellular concentrations of reduced and oxidized glutathione, 10^{-2} M and 3×10^{-5} M respectively [31,32], are often thought to preclude protein disulphide formation; indeed, the unfolded chain would be expected to form a disulphide only 1% of the time. However, under these conditions the disulphides of *native* BPTI are expected to be formed virtually all of the time and to contribute 2–6 kcal/mol each to the stability of the native state. On the other hand, it is also apparent that these contributions increase significantly, to as much as 4–10 kcal/mol, under more oxidizing conditions. Presumably, this is at least partly the reason that disulphides are found predominately in extracellular proteins.

CONTRIBUTIONS OF OTHER REVERSIBLE INTERACTIONS

Protein disulphide bonds are particularly amenable to the above analysis, because their stabilities can be directly measured by exchange with a small molecule disulphide. However, the same general considerations should apply to other specific interactions that stabilize native proteins, such as hydrogen bonds and salt bridges, even though it is not possible to measure their stabilities directly.

NMR experiments have identified an electrostatic interaction between the terminal amino and carboxyl groups of BPTI, and the contribution of this salt bridge to the stability of the native protein has been estimated by examining a derivative in which the terminal amino group was transaminated [33]. The contribution of this interaction can be usefully analyzed by considering a thermodynamic cycle analogous to that used for the disulphide bond:



(14)

The equilibrium constants for forming the salt bridge, K^U_{SB} and K^N_{SB}, represent the fraction of time that the charged groups are sufficiently close to interact. As for the other intermolecular interactions, these equilibrium constants can be expressed as the products of the equilibrium constant for an equivalent intermolecular salt bridge (K_{inter}) and the effective concentrations of the groups in the two states.

When the terminal amino group is transaminated, the salt bridge can no longer form and only the lower half of the cycle of Eq 14 contributes to the observed equilibrium between the native and unfolded states. Using the arguments applied earlier to the case of a disulphide, the apparent contribution of the salt bridge will be:

$$\Delta G_{SB} = RT\ln\,[(N_{SB} + N_{-SB})/(U_{SB} + U_{-SB})] - RT\ln\,(N_{-SB}/U_{-SB}) \quad (15)$$

$$\Delta G_{SB} = RT\ln\,[(1 + C^{N}_{eff} \times K_{inter})/(1 + C^{U}_{eff} \times K_{inter})]\,. \quad (16)$$

The apparent contribution of this salt bridge to the stability of native BPTI is about 1–1.5 kcal/mol [33]. The intermolecular equilibrium constant for forming a salt bridge in water has been measured to be about 0.5 M^{-1} [34]. If the value of C^{U}_{eff} is taken to be 10^{-3}–10^{-2} M, then the measured stabilization from the salt bridge can be used to calculate the effective concentration of the charged groups in the native protein to be about 14 M, and K^{N}_{SB} to be about 7, significantly greater than in the unfolded state.

This value for C^{N}_{eff} is considerably lower than the values for the disulphides of native BPTI. This is consistent with the C-terminus being quite flexible, as is indicated by the difficulty in identifying the terminal residue in the electron density maps of three crystal forms of BPTI [18,23,35]. Other, more rigidly oriented salt bridges might have higher values of C^{N}_{eff} and be considerably more effective in stabilizing the native state.

Nonetheless, it is striking that the salt bridge provides much more stability than does the peptide bond between the termini in circular BPTI. (The measurements of the effect of the peptide bond were made at a pH greater than the pK_a of the alpha-amino group, so that the salt bridge was not present in the uncross-linked protein for that comparison.) Correspondingly, the effective concentration of the terminal groups in the native protein is much greater for forming a salt bridge than for a peptide bond. Presumably, this difference is owing to the much stricter stereochemical requirements for forming the peptide bond. It should be emphasized that an effective concentraction can apply only to a particular type of interaction. It seems likely that the greater the stereochemical constraints for an interaction the greater is the maximum possible effective concentration, because of the greater loss of entropy in forming the equivalent intermolecular bond, but, also, small deviations from the optimum geometry might cause drastic reductions in C^{N}_{eff}. The salt bridge between the termini of BPTI illustrates how a relatively weak interaction can be quite important in stabilizing a native protein because its formation is significantly favored in the native state.

CONCLUSIONS AND CAVEATS

Using thermodynamic cycles, three types of interactions that can stabilize native proteins have been considered here: irreversible cross-links, reversible disulphide bonds, and a salt bridge. In each case, the extent to which the interaction stabilizes the native state is seen to depend on the extent to which the interaction is favored in the native state relative to the unfolded. The equilibrium constants for these intramolecular interactions can be conveniently expressed and compared as "effective concentrations." Using the limited experimental information available on the stabilities

of disulphide bonds and the contributions of other interactions to protein stability, it is possible to draw some general conclusions.

First, different interactions can have tremendously different effective concentrations even within the same native protein. The disulphides of native BPTI have effective concentrations from 10^2 to 10^5 M, whereas that for forming a peptide bond between the termini is only about 10^{-3} M. As a consequence, different interactions of the same type can vary greatly in their contributions to the stability of the native state.

Second, interactions that are intrinsically quite weak can be important, because their formation is favored in the native state. This is illustrated in Figure 4 for the case of a disulphide bond in which the apparent contribution of a disulphide is plotted as a funciton of the thiol-disulphide redox potential, $RSSR/RSH^2$. In fact, this graph can be used to consider any type of intramolecular interaction, because $RSSR/RSH^2$ simply represents the equilibrium constant for an intermolecular disulphide under the same conditions, and the abscissa can be treated as K_{inter} for other interactions as well. An interaction for which K_{inter} is only 10^{-2} M^{-1} (such as a hydrogen bond in water [37]) will not be siginficant in the unfolded chain, but, if the effective concentration in the native state is sufficiently high that it is usually present there (ie, $> 10^2$ M), then it will contribute significantly to the stability of the native state [24]. At the other extreme, when K_{inter} is sufficiently high that the interaction is always present in both the unfolded and native proteins, ie; for an irreversible cross-link, the stabilization reaches a plateau.

The factors that determine the stabilities of the interactions in native proteins, and have been lumped together in the term C^N_{eff}, are undoubtedly very complex. The tendencies of groups to interact will probably depend on their orientation, flexibility, packing, and electrostatic effects. All these factors depend on the other interactions in the protein; C^N_{eff} represents the stability of an interaction in the presence of all of the other interactions of the native protein and thus is a measure of the cooperativity of the folded state.

Modified proteins should continue to be a rich source of information for dissecting the cooperative roles of different interactions, particularly as the methods of genetic engineering are utilized. For cases when the modifications remove or introduce specific interactions, the arguments presented here might be useful in thinking about the results. An important caveat to this approach is that even the smallest changes, for instance a Cys to Ser change, is likely to have secondary effects, such as in the packing or elecrostatics of the native protein.

Also, it should be expected that many modifications will not be easily interpreted in terms of single specific interactions, as has already been seen for mutations of bacteriophage T4 lysozyme [7] and the alpha subunit of tryptophan synthetase [6]. The important contributions of hydrophobicity and Van der Waals interactions are probably not as easily analyzed as are cross-links and disulphide bonds. However, even for mutations that do not affect easily identified, specific interations, thermodynamic cycles, analogous to that in Figure 1 for a cross-link, could be useful. For any protein modification, the change in net stability of the native state will be the difference in free energies for actually making the modification in the native and unfolded proteins. Thus, a mutation that stabilizes the native state is one for which exchanging the normal amino acid for a mutant one is favored in the native state relative to the unfolded. This argument simply provides an slightly altered viewpoint

for considering the effects of mutations and emphasizes that both the native and unfolded conformations must be considered.

ACKNOWLDEGMENTS

The experiments and ideas presented here evolved out of extensive discussions with Dr. T.E. Creighton, who first suggested the use of effective concentrations to consider protein stability [24] and to whom I am extremely grateful for the privilege of collaboration for the past few years. I have been supported by a postdoctoral fellowship from the United States National Institues of Health (5 F32GM 08010).

REFERENCES

1. Anfinsen CB: Science 181:223, 1973.
2. Brenner S, Streisinger G, Horne RW, Champe SF, Barnett L, Benzer S, Reese MW: J Mol Biol 1:281, 1959.
3. Maling BD, Yanofsky C: Proc Natl Acad Sci USA 47:551, 1961.
4. Edgar RS, Lielausis I: Genetics 52:1187, 1965.
5. Sadler JR, Novick A: J Mol Biol 12:305, 1965.
6. Matthews CR, Crisanti MM, Gepner GL, Velicelbi G, Sturtevant JM: Biochemistry 19:1290, 1980.
7. Hawkes R, Grutter M, Schellman J: J Mol Biol 175:195, 1984.
8. Yutani K, Ogasahara K, Aoki K, Kakuno T, Sugino Y: J Biol Chem 259:14076, 1984.
9. Johnson RE, Adams P, Rupley JA: Biochemistry 17:1479, 1978.
10. Yamada H, Kuroki R, Hirata M, and Imoto T: Biochemistry 22:4551, 1983.
11. Scheraga HA, Denton ME, Konishi Y, Lin SH: Biochemistry 23:5504, 1984.
12. Villafranca JE, Howell EE, Voet DH, Strobel MS, Ogden RC, Abelson JN, Kraut J: Science 222:782, 1983.
13. Perry LJ and Wetzel R: Science 226:555, 1984.
14. Schellman JA: Compt Rend Trav Lab Carlsburg Ser Chim 29:230, 1955.
15. Flory PJ: J Am Chem Soc 78:5222, 1956.
16. Goldenberg DP, Creighton TE: J Mol Biol 165:407, 1983.
17. Goldenberg DP, Creighton TE: J Mol Biol 179:527, 1984.
18. Deisenhofer J, Steigmann W: Acta Cryst B31:238, 1975.
19. Creighton TE: Prog Biophys Mol Biol 22:231, 1978.
20. Wuthrich K, Wider G, Wagner G, Braun W: J Mol Biol 155:311, 1982.
21. Levitt M: J Mol Biol 170:723, 1983.
22. Creighton TE, Goldenberg DP: J Mol Biol 179:497, 1984.
23. Wlodawer A, Walter J, Huber R, Sjolin L: J Mol Biol 180:301, 1984.
24. Creighton TE: Biopolymers 22:49, 1983.
25. Page MI and Jencks WP: Proc Natl Acad Sci USA 68:1678, 1971.
26. Pace CN: Crit Rev Biochem 3:1, 1975.
27. Privalov PL: Adv Prot Chem 33:167, 1979.
28. Jacobson H, Stockmayer WH: J Chem Phys 18:1600, 1950.
29. Mutter M: J Am Chem Soc 99:8307, 1977.
30. Goto Y, Hamaguchi K: J Mol Biol 146:321, 1981.
31. Akerboom TPM, Bilzer M, Sies H: J Biol Chem 257:4248, 1982.
32. Larsson A, Orrenius S, Holmgren A, Mannervik B (eds): "Functions of Glutathione: Biochemical, Physiological, Toxicological and Clinical Aspects." New York: Raven Press, 1983.
33. Brown LR, DeMarco A, Richarz R, Wagner G, Wuthrich K: Eur J Biochem 88:87, 1978.
34. Springs B, Haake P: Biorg Chem 6:181, 1977.
35. Huber R, Kukla D, Bode W, Schwager P, Bartels K, Deisenhofer J, Steigmann W: J Mol Biol 89:73, 1974.

36. Susi H, Timascheff SN, Ard JS: J Biol Chem 239:3051, 1964.
37. Jering H, Tschesche H: Eur J Biochem 61:443, 1976.
38. Creighton TE, Kalef E, Arnon R: J Mol Biol 123:129, 1978.
39. States DJ, Dobson CM, Karplus M, Creighton TE: Nature 286:630, 1980.
40. Wagner G, Kalb(Gilboa) AJ, Wuthrich K: Eur J Biochem 95:249, 1979.
41. Moses E, Hinz H-J: J Mol Biol 170:765, 1983.