Protein Science (1993), *2*, 1551-1558. Cambridge University Press. Printed in the USA.

REVIEW

Disulfide bonds and the stability of globular proteins

STEPHEN F. BETZ

The DuPont Merck Pharmaceutical Company, Wilmington, Delaware 19880-0328

(RECEIVED May 28, 1993; ACCEPTED July 21, 1993)

Abstract

An understanding of the forces that contribute to stability is pivotal in solving the protein-folding problem. Classical theory suggests that disulfide bonds stabilize proteins by reducing the entropy of the denatured state. More recent theories have attempted to expand this idea, suggesting that in addition to configurational entropic effects, enthalpic and native-state effects occur and cannot be neglected. Experimental thermodynamic evidence is examined from two sources: (1) the disruption of naturally occurring disulfides, and (2) the insertion of novel disulfides. The data confirm that enthalpic and native-state effects are often significant. The experimental changes in free energy are compared to those predicted by different theories. The differences between theory and experiment are large near 300 K and do not lend support to any of the current theories regarding the stabilization of proteins by disulfide bonds. This observation is a result of not only deficiencies in the theoretical models but also from difficulties in determining the effects of disulfide bonds on protein stability against the backdrop of numerous subtle stabilizing factors (in both the native and denatured states), which they may also affect.

Keywords: disulfide bonds; protein stability; thermodynamics

The determination of the forces that govern protein stability is of fundamental importance for our ability to understand and control the interactions of complex biological molecules. It has been known since the 1960s that the primary structure of a protein dictates its three-dimensional fold (see Anfinsen, 1973), yet a comprehensive understanding of the factors that impart thermodynamic stability to proteins is elusive. This is because the tertiary folds of native proteins are defined by a large number of weak interactions: hydrogen bonding (Stickle et al., 1992), hydrophobic interactions (Privalov & Gill, 1988), salt bridges, and weakly polar interactions (Burley & Petsko, 1988). In addition to these noncovalent forces, certain proteins are also stabilized covalently by disulfide bonds. This review is concerned with the effects of disulfides on monomeric, globular proteins (oligomeric systems have been excluded). Particularly, it examines the disruption of native disulfides and the incorporation of novel disulfides and their effects on thermodynamic stability.

The overwhelming majority of data concerning the stability of proteins comes from examination of the reversible denaturation of small, globular proteins. For many monomeric, single-domain proteins, denaturation by heat or chemicals (such as GdmCl or urea) is often consistent with a single, cooperative transition from the folded, N-state to a (mostly) disordered form, the D-state. This "two-state model" can be written as:

$$\mathrm{N} \rightleftarrows \mathrm{D}. \qquad (1)$$

The difference in free energy between the D- and N-states is defined as ΔG_d, which is equivalent to $-RT$ ln[D]/[N]. Therefore, the more stable the protein, the more positive ΔG_d will be. The advent of molecular bi-

Reprint requests to: Stephen F. Betz, The DuPont Merck Pharmaceutical Company, Wilmington, Delaware 19880-0328.

Abbreviations: BPTI, bovine pancreatic trypsin inhibitor; C_m, the concentration of GdmCl at which half the protein is denatured; D-state, the denatured state of a protein; GdmCl, guanidinium chloride; hew, hen egg white; m_{den}, the slope of the line of $-\Delta G_d$ versus denaturant concentration; N-state, the native state of a protein; RNase, ribonuclease; T_m, temperature at which half the protein is denatured; $t_{1/2}$, time required for enzymatic activity to decrease 50%; ΔC_p, the difference in heat capacity between the denatured and native states; ΔG_d, ΔG for protein denaturation; $\Delta G_{d,\mathrm{H_2O}}$, ΔG_d determined from chemical denaturation; $\Delta G_{d,T}$, ΔG_d determined from thermal denaturation; ΔH_{conf}, the change in conformational enthalpy between the native and denatured states; ΔH_d, ΔH for protein denaturation; ΔH_{hyd}, the change in solvational enthalpy between the native and denatured states; ΔH_m, ΔH for protein denaturation at T_m; ΔS_{conf}, the change in conformational entropy between the native and denatured states; ΔS_d, ΔS for protein denaturation; ΔS_{hyd}, the change in solvational entropy between the native and denatured states.

ology and the ability to generate variant proteins without excessive difficulty has spawned an intense reinvestigation of the factors that affect ΔG_d (Shortle, 1989).

The change in free energy for protein denaturation as a function of temperature ($\Delta G_{d,T}$) can be calculated using the Gibbs–Helmholtz equation:

$$\Delta G_{d,T} = \Delta H_m(1 - T/T_m) - \Delta C_p[(T_m - T) + T\ln(T/T_m)] \quad (2)$$

where ΔH_m is the enthalpy change upon denaturation at T_m, T_m is the temperature at which half the protein is denatured, and ΔC_p is the difference in heat capacity between the D- and N-states. In addition, ΔG_d at one particular temperature can be calculated from chemical denaturation data using the equation:

$$\Delta G_d = -m_{den} \times [\text{den}] + \Delta G_{d,H_2O} \quad (3)$$

where $\Delta G_{d,H_2O}$ is the free energy of denaturation in the absence of denaturant, and m_{den} is the slope of the line of $-\Delta G_d$ versus denaturant concentration (Pace et al., 1990) and has been related to the difference in solvent-accessible, nonpolar surface area between the D- and N-states (Schellman, 1978; Shortle, 1989). The denaturant concentration at which half the protein is denatured is defined as C_m. Equations 2 and 3 are valid provided denaturation is reversible and the concentration of intermediates is small.

Theoretical and computational insights

Treating the D-state of proteins as a random coil, Flory (1956) predicted that disulfide bonds increase the free energy of the D-state by decreasing its configurational entropy, thus driving the two-state equilibrium (Equation 1) to the left. The decrease in entropy is estimated by calculating the probability that the ends of the chain will occupy the same volume element, v_s. The mathematical description of this for a polymer containing one crosslink becomes:

$$\Delta S = -R\ln[3/(2\pi l^2 N)^{3/2}]v_s \quad (4)$$

where R is the gas constant, N is the number of statistical segments in the loop, and l is the average length of a statistical segment, 3.8 Å for an amino acid residue within a protein (Pace et al., 1988). The choice of the value for v_s has been a source of discussion (Poland & Scheraga, 1965; Pace et al., 1988). Pace et al. (1988) determined 57.9 Å^3 for v_s based on the closest possible approach of two thiols. Using the above values, the effect upon the configurational entropy reduces to:

$$\Delta S = -2.1 - 3/2(\ln n) \quad (5)$$

where n in the number of residues encompassed by the disulfide. Pace et al. (1988) determined that predicted values of $\Delta\Delta G_d$ correlate well with those obtained experimentally for the removal of disulfide bonds from RNase T1, lysozyme, and other proteins. Regardless of the choice for v_s, the conclusions from this sort of analysis are clear: the stabilization of proteins by disulfide bonds is driven entropically, and this effect is enhanced with increasing loop size.

The thermodynamic details of this increased stability however, appear to be more complex and involve both enthalpic and entropic effects from not only the protein chain, but also the solvent. Alternatively, Doig and Williams (1991) suggest that the dominant effect of disulfide bonds on stability is enthalpic.

Consider that for ΔG_d to increase, either the free energy of the N-state must decrease or the free energy of the D-state must increase. Additionally, ΔG_d will become more positive if ΔH_d becomes more positive or ΔS_d (the entropy change upon denaturation) becomes less positive. Rewriting the two-state hypothesis to include solvent:

$$n\text{H}_2\text{O} + \text{N} \rightleftarrows \text{D}(\text{H}_2\text{O})_n. \quad (6)$$

Partition ΔS_d into two components: ΔS_{conf} for the protein chain, and ΔS_{hyd} for the solvent (i.e., $\Delta S_d = \Delta S_{conf} + \Delta S_{hyd}$). Both ΔS_d and ΔS_{conf} are positive at and above physiological temperatures (Doig & Williams, 1991 and references therein). Equation 5 shows that, for a single-chain, non-crosslinked protein, ΔS_{hyd} is negative, and therefore, $|\Delta S_{conf}| > |\Delta S_{hyd}|$.

Upon introducing a disulfide, the conformational entropy of the D-state ($S_{conf,D}$) will decrease while the conformational entropy of the N-state ($S_{conf,N}$) should change little. This will result in ΔS_{conf} decreasing compared to the non-crosslinked protein, in short, the Flory argument. A crosslink has a compensatory effect on ΔS_{hyd}. The constrained D-state makes $S_{hyd,D}$ more positive while $S_{hyd,N}$ is unaltered. This results in ΔS_{hyd} being less negative for the crosslinked protein compared to the non-crosslinked protein. Both effects (ΔS_{conf} and ΔS_{hyd}) are diminished upon the introduction of the crosslink. However, Doig and Williams (1991) have calculated from model compound data that $|\Delta(\Delta S_{hyd})|$ is larger in magnitude than $|\Delta(\Delta S_{conf})|$, and thus, ΔS_d actually becomes more positive (i.e., destabilizing) for a disulfide-bonded protein compared to its non-crosslinked counterpart.

With this conclusion, the analysis must turn to an examination of enthalpic factors. Partition ΔH_d into two components: ΔH_{conf} and ΔH_{hyd}. It is known that ΔH_{hyd} near 300 K is essentially zero (Creighton, 1993), and so any enthalpic effect upon incorporation of a disulfide bond must come from ΔH_{conf}. Because no covalent bonds are broken upon denaturation, ΔH_{conf} must derive from noncovalent interactions.

Consider Equation 6 in terms of hydrogen bonding. The native state can be imagined as having essentially all its potential hydrogen bonds satisfied in protein–protein

interactions (Privalov & Khechinashvili, 1974). In this sense, the N-state of a disulfide-bonded and non-crosslinked protein should be equivalent. The D-state of the non-crosslinked protein can be thought of as having no intraprotein hydrogen bonds, but again being fully satisfied by protein–solvent hydrogen bonds. (The assumption that the D-state has no intraprotein interactions is very likely a gross oversimplification [vide infra; Dill & Shortle (1991)] but does not affect the conclusions of this argument, provided the N-state possesses a greater number of intramolecular hydrogen bonds than the D-state.) For a disulfide-bonded protein, the D-state will also possess a vast majority of protein–solvent hydrogen bonds. However, due to the steric interactions forced by the disulfide linkage, certain hydrogen-bonding groups near the crosslink will be unable (or poorly able) to interact with the protein *or* the solvent and as such be left as unsatisfied donor–acceptor pairs. This enthalpic effect destabilizes the D-state of disulfide-bonded proteins and dominates the compensatory entropic effect discussed above.

Molecular dynamics and normal mode analysis on the N-state configurational entropy of BPTI with either two or three disulfide bonds suggests that the presence of disulfides affects the N-state of proteins and not only their D-states (Tidor & Karplus, 1993). Favorable entropic contributions to free energy of 2–5 kcal mol^{-1} were found for the *reduction* of either of the three naturally occurring disulfides in BPTI. (Note that the magnitude of these effects is on the order of *stabilizing* effects of disulfides measured experimentally [vide infra], as well as those calculated using Equation 5.) This effect correlates with the structure of the molecule near the crosslink. Constraining flexible regions decreases the N-state entropy more than crosslinks between rigid regions. These entropic effects are potentially countered by similar but opposite enthalpic effects: crosslinks between rigid regions of proteins may cause torsional strain and local repacking (decreasing ΔH_d), whereas crosslinks between flexible regions are viewed as relatively free to repack with little loss of favorable enthalpic interactions. The underlying conclusion from this analysis is that N-state entropic and enthalpic effects must be considered as well as D-state entropic ones. However, crystallographic data on novel disulfides introduced into subtilisin BPN′ (Katz & Kossiakoff, 1986, 1990), indicate that crosslinks are more readily accommodated by flexible regions rather than rigid ones, suggesting that enthalpic N-state effects may dominate negative entropic considerations.

Experimental evidence

Removal of natural disulfides

Recent examinations of the destabilizing effect of removing one naturally occurring disulfide bond from different lysozymes (Cooper et al., 1992; Kuroki et al., 1992) permit examination of the Doig and Williams (1991) hypothesis. Selective disruption and modification of the disulfide between residues 6 and 127 of hen egg white (hew) lysozyme causes the T_m of the protein to drop approx. 25 °C at low pH (Cooper et al., 1992). Although ΔG_d values are not tabulated, the parameters needed to do so using Equation 2 are presented. The modified protein is 7.5 kcal mol^{-1} less stable than the protein containing all three disulfide bonds at pH 3.8 and 300 K. The theory of Doig and Williams (1991) predicts that the modified protein should have a smaller ΔH_d and ΔS_d, whereas ΔC_p should increase. Cooper et al. (1992) determined that ΔH_d (at 300 K) is essentially unchanged, ΔS_d increases, and ΔC_p either does not change or actually decreases. These results are at odds with Doig and Williams at every turn and much more strongly support a Flory-type argument.

The thermodynamics of changing cysteines 77 and 95 of human lysozyme to alanine have been examined (Kuroki et al., 1992). The disulfide-containing protein is approx. 2 kcal mol^{-1} more stable than the non-crosslinked variant at 300 K. The native (i.e., crosslinked) protein has a higher T_m (>14 °C) and a greater ΔH_d compared to the non-crosslinked variant. These findings are in line with the predictions of Doig and Williams (1991). However, the native and non-crosslinked lysozymes have the same ΔC_p. To complicate these results further, the crystal structure of the disulfide-lacking protein indicates that *B*-factors for atoms near positions 77 and 95 are larger than those of the native protein, reminiscent of the effects discussed previously for BPTI (Tidor & Karplus, 1993). Additionally, the replacement of both cysteines with alanine creates a potentially destabilizing pocket within the native structure of the protein (Inaka et al., 1991). It is clear in this case that removal of this disulfide bond affects both the N- and D-states.

A similar study has been performed on RNase T1 (Pace et al., 1988), where one or two naturally occurring disulfide bonds were selectively reduced and chemically modified. The difference in ΔG_d between the native protein and the variant with one disulfide bond is 3.3 kcal mol^{-1} at 300 K. The difference between the one disulfide-containing protein and the variant in which both disulfides are reduced is 3.8 kcal mol^{-1}. The latter value is somewhat suspect because of uncertainty in the reversibility of the denaturation of the disulfide-free protein. That work was done primarily by chemical denaturation, and so it is difficult to compare those results with the lysozyme experiments described above (Cooper et al., 1992; Kuroki et al., 1992) or with the theory of Doig and Williams (1991), which are based on thermal denaturation and parameterization.

The reduction and modification of the disulfide between positions 14 and 38 of BPTI were studied by thermal denaturation (Schwartz et al., 1987). The ΔH_d (at 298 K) for the crosslinked species was more than 20 kcal mol^{-1} greater than the modified variant. For the disul-

fide-containing protein, ΔG_d was approx. 8 kcal mol^{-1} greater at 300 K than that of the most stable modified derivative at either pH 2 or pH 5. These data are in accord with the work of Doig and Williams (1991). Normal mode analysis of BPTI (Tidor & Karplus, 1993) determined that reduction of this disulfide bond resulted in the greatest increase in entropy in the N-state. Apparently, D-state entropic and/or overall enthalpic effects overcome this destabilizing effect. There are difficulties in that the reversibility of the system is not complete. Additionally, the relatively high number (3) of disulfides in BPTI for such a small protein (58 residues) may make its behavior atypical.

Ikeguchi et al. (1992) have examined the effect of the reduction and modification of a disulfide of α-lactalbumin on the equilibrium thermodynamics between the N-, D-, and A- (molten globule) states. Chemical denaturation was the only thermodynamic experiment performed. For the N to D transition, the disulfide-linked protein is calculated to be 3 kcal mol^{-1} more stable at 278 K than the non-crosslinked variant in the presence of Ca^{2+}. The A-state of the crosslinked protein is also stabilized relative to the non-crosslinked protein, albeit to a lesser extent. The authors suggest that the crosslink may stabilize secondary structure within the A-state. Similar effects are seen in the D-state of cytochrome *c* containing a novel disulfide (Betz & Pielak, 1992; vide infra).

An examination of crystal structures of disulfide-containing proteins demonstrates that disulfide bonds are generally solvent inaccessible (Srinivasan et al., 1990). Yet, after disruption of a disulfide bond, the effects of burying two thiols or blocked thiols are rarely taken into account. The measurement of the free energy of transfer of model compounds from cyclohexane to water suggests that burying cystine is 0.5 kcal mol^{-1} more favorable than burying two cysteines (Saunders et al., 1993). Because of the reactivity of free thiols, reduced disulfides are often chemically modified with large, fairly polar groups (e.g., carboxymethyl or carboxamidomethyl). The relative polarities of these "residues" should not be overlooked. It is a testament to the robust nature of proteins that disruption and blocking of completely buried disulfides can sometimes result in at least partially active proteins (Pace et al., 1988 and references therein).

Novel disulfides

In addition to examining the destabilizing effects of disrupting disulfide bonds in natural proteins, there has been considerable interest in increasing the stability of proteins by incorporating novel disulfide bonds (Wetzel, 1987). The value of these studies in understanding the mechanism of the stabilizing effects of disulfide bonds varies greatly, but in general, most of these studies suffer from incomplete thermodynamic analysis. In some studies, ΔG_d cannot be determined because denaturation is irreversible. Instead, increased stability relative to a wild-type protein is usually defined as an increase in either C_m or T_m, or in the time required for enzymatic activity to decrease 50% ($t_{1/2}$). Even if denaturation is reversible, and the concentration of intermediates is small, determination of T_m (Equation 2), C_m (Equation 3), or $t_{1/2}$ (Lumry et al., 1966) alone is insufficient to determine ΔG_d.

Attempts to increase the stability of monomeric proteins by introduction of novel disulfides have met with mixed success. Only proteins in which both the crosslinked and native proteins are reversibly denatured are considered. Of these, the most extensively studied is T4 lysozyme, which has no disulfides in its native form. Five different novel disulfide bonds have been introduced into this protein (Perry & Wetzel, 1984; Wetzel et al., 1988; Matsumura et al., 1989a,b). Two have been found to lower T_m by 20 °C at neutral pH. In addition, ΔH_d for these variants is greatly reduced, indicating that the presence of these crosslinks destabilizes the protein (Wetzel et al., 1988). Three other novel disulfides increase T_m by 5–10 °C (Matsumura et al., 1989a,b) at low pH, and combinations of these crosslinks are nearly additive in their ability to increase T_m (Matsumura et al., 1989b), although the tendency of these crosslinks to increase resistance to denaturation at neutral pH (Wetzel et al., 1988) or under low-temperature denaturing conditions (Anderson et al., 1990) is diminished. However, each of these "stabilizing" crosslinks raises ΔC_p and lowers ΔH_d (Matsumura et al., 1989a). These data do not support the hypothesis of Doig and Williams (1991). Remarkably, independent calculation of ΔG_d at 300 K (using Equation 2) for each of these proteins using the values presented in Figure 2 and Table 3 of Matsumura et al. (1989a) indicates that the disulfide-containing variants are 0.4–1.6 kcal mol^{-1} more stable than their reduced counterparts, but are actually 0.6–1.1 kcal mol^{-1} *less stable* than wild-type T4 lysozyme.

One of the more successful, well-characterized novel disulfide bonds is one introduced into RNase H by Kanaya et al. (1991). The presence of the disulfide increases C_m and raises T_m by 12 °C. In this case, ΔG_d was raised approx. 2.8 kcal mol^{-1} at 300 K, but comparisons of ΔH_d and ΔC_p are not possible because chemical denaturation was the only method used to estimate free energy changes. There are further complications because residues in the native protein other than those that formed the disulfide bond were changed, and controls that consider the effects of those substituted wild-type residues were not reported.

Two different novel disulfides were introduced into the RNase barnase as probes for folding kinetics (Clarke & Fersht, 1993). These two disulfides were found to increase C_m significantly in urea or GdmCl, depending upon the crosslink. One disulfide was found to increase ΔG_d by 1.2 kcal mol^{-1} while slightly decreasing m_{urea}. The other crosslink increased ΔG_d by 4.1 kcal mol^{-1}, but left m_{GdmCl} unaltered (both measurements at 298 K). Interestingly, the more stabilizing crosslink encompasses

fewer residues than the other (17 versus 37 residues), which is contradictory to the decreased chain entropy theory of Flory (1956). Equation 5 predicts that the 18-residue crosslink should stabilize the protein by 3.2 kcal mol^{-1} at 298 K and is in reasonable accord with what is observed, whereas the 38-residue crosslink should be worth 3.8 kcal mol^{-1}. The more-stabilizing disulfide is in a flexible region of the molecule and links a strand of β-sheet to a loop, whereas the less-stabilizing crosslink connects a helix to a sheet. Katz and Kossiakoff (1990) suggest that helices may be less able to dissipate strain energy from disulfides compared to loops and sheets, and these results support that idea. Also, these results argue that N-state entropic effects (Tidor & Karplus, 1993) are small compared to D-state or enthalpic effects.

A crosslink introduced into dihydrofolate reductase increases C_m slightly but decreases T_m (Villafranca et al., 1987). The authors conclude that the crosslinked form of the protein is more stable than the native protein. However, their analysis is flawed (see Shortle, 1989), and ΔG_d for the disulfide-containing protein is considerably decreased. Additionally, uncertainty concerning the two-state behavior of the crosslinked variant makes analysis of this system difficult.

A novel disulfide, modeled on that of bullfrog cytochrome *c* (Brems et al., 1982), has been introduced into yeast iso-1-cytochrome *c* between positions 20 and 102 (Betz & Pielak, 1992). Spectroscopic analysis suggests that the N-states of the wild-type and crosslinked proteins are nearly identical. The disulfide-bonded variant was found to be 2–4 kcal mol^{-1} less stable under varying conditions, despite having a greater T_m than the non-crosslinked protein. Additionally ΔH_d, ΔS_d ΔC_p, and m_{GdmCl} were also found to decrease. Spectroscopic analysis of the D-states (chemically or thermally induced) of the two proteins demonstrates that the effect of the crosslink is to stabilize the D-state into a more structured and compact form. These data are in complete opposition to the arguments of Flory (1956). This effect may arise from the cooperation of the novel crosslink with the covalently bound heme of cytochrome *c*, which has been postulated to modulate structure in the D-state of the native protein (Muthukrishnan & Nall, 1991).

A series of novel disulfides was introduced into a designed, parallel coiled-coil (Zhou et al., 1993). The repetitive heptad repeat of the coiled-coil was used as a scaffold to determine the preference of certain core positions to accept a disulfide bond as measured by resistance to chemical denaturation. The results show that one subset of core residues formed disulfides favorably, whereas another set destabilized the coiled-coil structure. These effects are more pronounced toward the center of the molecule than near its termini. Inspection of putative models suggested that the favorable crosslinks could adopt a typical disulfide geometry (vide infra), whereas the other position introduced highly strained conformers. No attempt to measure the ΔG_d of these constructs was reported.

Eder and Wilmanns (1992) have reported the insertion of a novel disulfide into the β/α barrel enzyme *N*-(5′-phosphoribosyl)anthranilate isomerase from yeast. The loop includes 83% of the primary sequence, yet ΔG_d is only increased 1 kcal mol^{-1} at pH 6.5 for the crosslinked variant, and there is no difference in ΔG_d between the crosslinked and native forms at pH 7.8 as measured by chemical denaturation. The weak stabilization of this crosslink may derive from a poor model of the N-state (the *Escherichia coli* enzyme crystal structure was used) or from a nonadherence to a two-state equilibrium.

Novel disulfide bonds have been introduced into subtilisin BPN′ (Wells & Powers, 1986; Pantoliano et al., 1987; Mitchinson & Wells, 1989), subtilisin E (Takagi et al., 1990), and troponin C (Gusev et al., 1991), but irreversible denaturation is observed for either the crosslinked or non-crosslinked forms or both, obviating thermodynamic treatment based on Equation 1.

Conclusions

As crystal structures and experimental evidence suggest, an important consideration when introducing novel disulfides is their effect(s) on the native state. Substantial information is available concerning spatial requirements for disulfide bonds (Richardson, 1981; Thornton, 1981; Srinivasan et al., 1990). In an analysis of high-resolution crystal structures of disulfide-containing proteins, Richardson (1981) determined that the torsional angles, χ_1 and χ_3, and the C_α-C_α atomic separation were the most highly conserved elements in disulfide bond structure. This has been confirmed and extended by the analysis of Srinivasan et al. (1990), who analyzed a much larger data set.

Clearly, the underlying message of these data is that the original hypothesis of Flory (1956) although sound, is incomplete given the subtleties of protein structure. Figure 1 displays a crosslink's entropic contribution to free energy as a function of loop size (Equation 5) compared to the data reviewed in this article. Near 300 K, the correlation between theory and experiment is poor and does not support a hypothesis based solely on configurational entropy. Indeed, as many of the crosslinks are destabilizing as are favorable. A comparison of the same data extrapolated to 335 K again shows scatter, but in this case all of the disulfides are at least marginally stabilizing. The enthalpic argument of Doig and Williams (1991) has received a great deal of attention, but experimental data supporting this idea are inconclusive. Equation 18 in the paper by Doig and Williams (1991) calculates $\Delta G_d(T)$ as a function of the number of residues and number of disulfides within the protein. At 300 K, $\Delta\Delta G_d$ for a crosslinked protein versus its non-crosslinked counterpart is calculated to be 4.4 kcal mol^{-1}. This value is essentially

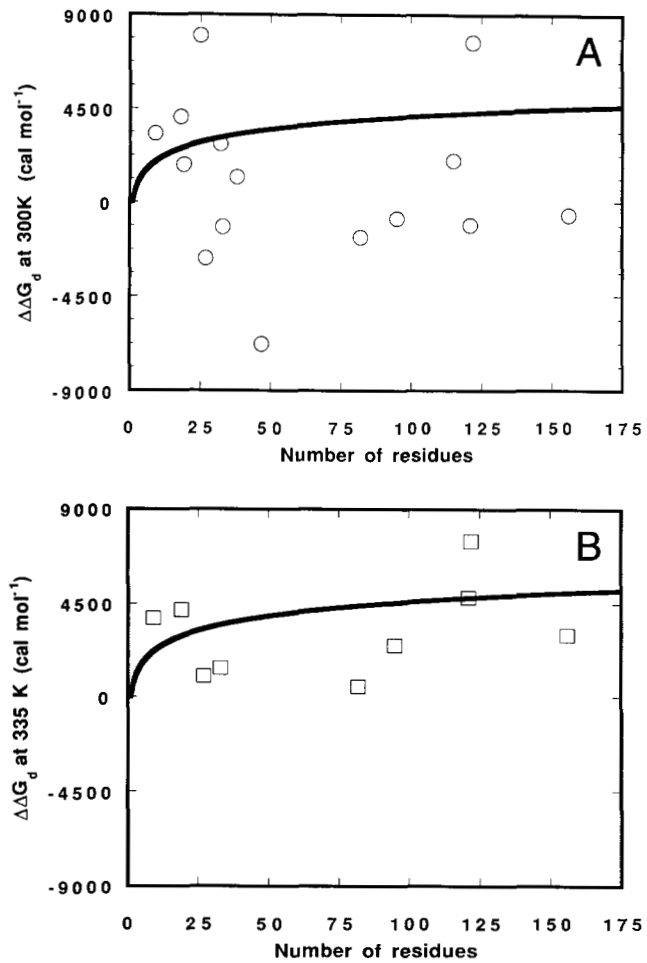

Fig. 1. Comparison of theoretical and experimental $\Delta\Delta G_d$ for disulfide-bonded proteins at 300 K **(A)** and 335 K **(B)**, as a function of loop size. $\Delta\Delta G_d$ is defined as ΔG_d (disulfide-bonded) − ΔG_d (disulfide-free). The solid line represents the theoretical increase based upon Equation 5. Experimental data from this article are individual points. There are fewer points in B because chemical denaturation data at 300 K cannot be extrapolated to other temperatures.

independent of protein length and loop size. Coincidentally, this value is very similar to the stabilization predicted by Equation 5 for crosslinks greater than 130 residues at 300 K. There is clearly no better correlation between the data presented here and the Doig and Williams (1991) hypothesis than that from the configurational entropy argument displayed in Figure 1.

At 335 K, however, the method of Doig and Williams (1991) predicts that $\Delta\Delta G_d$ should be 3.2 kcal mol^{-1}, whereas Equation 5 predicts that $\Delta\Delta G_d$ should increase slightly. Examination of Figure 1B suggests that the Doig and Williams (1991) hypothesis may more accurately predict $\Delta\Delta G_d$ than Equation 5, although the limited data, combined with its scatter, make this conclusion speculative.

As outlined in a discussion of the rationale for the design of disulfides in T4 lysozyme (Matsumura & Matthews, 1991), two residues close in space are not sufficient for a stabilizing disulfide to be formed. Effects on the packing and flexibility of the N-state must be considered (Tidor & Karplus, 1993). In the case of novel disulfides, this can often be addressed by examination of a high-resolution structure of the native protein, aided by computer programs developed to find optimal potential pairs (Pabo & Suchanek, 1986; Hazes & Dijkstra, 1988). Unfortunately, this is only half the picture because effects in the D-state can also determine the overall behavior of the crosslinked protein. To assume that the D-state of a protein is a random coil is certainly an oversimplification, and a number of recent reports have confirmed this idea (Dill & Shortle, 1991; Betz & Pielak, 1992; Green et al., 1992; Shortle et al., 1992; Sosnick & Trewella, 1992).

The role of disulfide bonds in protein structure and function is not limited to increasing thermodynamic stability of globular proteins. Indeed, if thermodynamic stability is the predominant effect of disulfides, and loop size is proportional to the stabilization imparted (see Equation 5 and Fig. 1), it seems odd that almost half of disulfides encompass less than 24 residues, and that the average separation for cysteines within a disulfide bond is 15 residues (Thornton, 1981). Examination of Figure 1 suggests that disulfides may stabilize proteins more effectively at higher temperatures. Yet, for extreme stability versus thermal denaturation, disulfides may be poor choices because cystine may oxidize at high temperatures (Hocking & Harris, 1980). Evidence from thermodynamic analysis of proteins from thermophilic organisms suggests that electrostatic interactions, and not disulfide bonds, contribute more significantly to the stability of these proteins (Straume et al., 1992; Kelly et al., 1993).

Disulfides perform diverse functions, from constraining small peptides into bioactive conformations, to linking the chains of immunoglobins. Also, the effect of disulfides on the kinetics of folding has not been considered here. None of the current theories adequately predict the effects of disulfide bonds on protein stability, and this may reflect the variety of purposes that disulfides serve. The other side of this coin is that none of the experiments described here can truly gauge the effects of the presence of a disulfide. Disulfide removal experiments are problematic because disruption and blocking of cysteines involve groups that may cause steric and polarity clashes. Mutagenesis of cysteine to other residues is beset with similar difficulties in that volume and polarity characteristics are difficult to replace equally. As with other forces that dictate protein structure and function, the effects of disulfide bonds are subtle, multifaceted, and challenging to extract from the interdependent whole.

Acknowledgments

I thank David Shortle and Gary Pielak for comments and suggestions on early versions of the manuscript and Dan Raleigh for helpful comments and discussions.

References

Anderson, W.D., Fink, A.L., Perry, J., & Wetzel, R. (1990). Effect of an engineered disulfide bond on the folding of T4 lysozyme at low temperatures. *Biochemistry 29*, 3331–3337.

Anfinsen, C.B. (1973). Principles that govern the folding of protein chains. *Science 181*, 223–230.

Betz, S.F. & Pielak, G.J. (1992). Introduction of a disulfide bond into cytochrome *c* stabilizes a compact denatured state. *Biochemistry 31*, 12337–12344.

Brems, D.N., Cass, R., & Stellewagen, E. (1982). Conformational transitions of frog heart cytochrome *c*. *Biochemistry 21*, 1488–1493.

Burley, S.K. & Petsko, G.A. (1988). Weakly polar interactions in proteins. *Adv. Protein Chem. 39*, 125–189.

Clarke, J. & Fersht, A.R. (1993). Engineered disulfide bonds as probes of the folding pathway of barnase: Increasing the stability of proteins against the rate of denaturation. *Biochemistry 32*, 4322–4329.

Cooper, A., Eyles, S.J., Radford, S.E., & Dobson, C.M. (1992). Thermodynamic consequences of the removal of a disulphide from hen lysozyme. *J. Mol. Biol. 225*, 939–943.

Creighton, T.E. (1993). In *Proteins: Structures and Molecular Properties*. Freeman, New York.

Dill, K.A. & Shortle, D. (1991). Denatured states of proteins. *Annu. Rev. Biochem. 60*, 795–825.

Doig, A.J. & Williams, D.H. (1991). Is the hydrophobic effect stabilizing or destabilizing in proteins: The contribution of disulphide bonds to protein stability. *J. Mol. Biol. 217*, 389–398.

Eder, J. & Wilmanns, M. (1992). Protein engineering of a disulfide in a β/α barrel protein. *Biochemistry 31*, 4437–4444.

Flory, P.J. (1956). Theory of elastic mechanisms in fibrous proteins. *J. Am. Chem. Soc. 78*, 5222–5235.

Green, S.M., Meeker, A.K., & Shortle, D. (1992). Contributions of the polar, uncharged amino acids to the stability of staphylococcal nuclease: Evidence for mutational effects on the free energy of the denatured state. *Biochemistry 31*, 5717–5728.

Gusev, N.B., Grabarek, Z., & Gergely, J. (1991). Stabilization by a disulfide bond of the N-terminal domain of a mutant troponin C (TnC48/2). *J. Biol. Chem. 266*, 16622–16626.

Hazes, B. & Dijkstra, B.W. (1988). Model building of disulfide bonds in proteins with known three dimensional structure. *Protein Eng. 2*, 119–125.

Hocking, D. & Harris, J.I. (1980). D-Glyceraldehyde-3-phosphate dehydrogenase: Amino acid sequence of the enzyme from the extreme thermophile *Thermus aquaticus*. *Eur. J. Biochem. 208*, 567–579.

Igekuchi, M., Sugai, S., Fujino, M., Sugawara, T., & Kuwajima, K. (1992). Contribution of the 6–120 disulfide bond of α-lactalbumin to the stabilities of its native and molten globule states. *Biochemistry 31*, 12695–12700.

Inaka, K., Taniyama, Y., Kikuchi, M., Morikawa, K., & Matsushima, M. (1991). The crystal structure of a mutant human lysozyme C77/95A with increased secretion efficiency in yeast. *J. Biol. Chem. 266*, 12599–12603.

Kanaya, S., Katsuda, C., Kimura, S., Nakai, T., Kitakuni, E., Takamura, H., Katayanagi, K., Morikawa, K., & Ikehara, M. (1991). Stabilization of *Escherichia coli* ribonuclease H by introduction of an artificial disulfide bond. *J. Biol. Chem. 266*, 6038–6044.

Katz, B. & Kossiakoff, A. (1986). The crystallographically determined structures of atypical strained disulfides engineered into subtilisin. *J. Biol. Chem. 261*, 15480–15485.

Katz, B. & Kossiakoff, A. (1990). Crystal structures of subtilisin BPN′ variants containing disulfide bonds and cavities: Concerted structural rearrangements induced by mutagenesis. *Proteins Struct. Funct. Genet. 7*, 343–357.

Kelly, C.A., Nishiyama, M., Ohnishi, Y., Beppu, T., & Birktoft, J.J. (1993). Determinants of protein thermostability observed in the 1.9 Å crystal structure of malate dehydrogenase from the thermophilic bacterium *Thermus flavus*. *Biochemistry 32*, 3913–3922.

Kuroki, R., Inaka, K., Taniyama, Y., Kidokoro, S., Matsushima, M., Kikuchi, M., & Yutani, K. (1992). Enthalpic destabilization of a mutant human lysozyme lacking a disulfide bridge between cysteine-77 and cysteine-95. *Biochemistry 31*, 8323–8328.

Lumry, R., Biltonen, R., & Brandts, J.F. (1966). Validity of the "two state" hypothesis for conformational transitions of proteins. *Biopolymers 4*, 917–944.

Matsumura, M., Becktel, W.J., Levitt, M., & Matthews, B.W. (1989a). Stabilization of phage T4 lysozyme by engineered disulfide bonds. *Proc. Natl. Acad. Sci. USA 86*, 6562–6566.

Matsumura, M. & Matthews, B.W. (1991). Stabilization of functional proteins by introduction of multiple disulfide bonds. *Methods Enzymol. 202*, 336–356.

Matsumura, M., Signor, G., & Matthews, B.W. (1989b). Substantial increase of protein stability by multiple disulfide bonds. *Nature 342*, 291–293.

Mitchinson, C. & Wells, J.A. (1989). Protein engineering of disulfide bonds in subtilisin BPN′. *Biochemistry 28*, 4807–4815.

Muthukrishnan, K. & Nall, B.T. (1991). Effective concentration of amino acid side chains in an unfolded protein. *Biochemistry 30*, 4706–4710.

Pabo, C.O. & Suchanek, E.G. (1986). Computer-aided model-building strategies for protein design. *Biochemistry 25*, 5987–5991.

Pace, C.N., Grimsley, G.R., Thomson, J.A., & Barnett, B.J. (1988). Conformational stability and activity of ribonuclease T1 with zero, one, and two intact disulfide bonds. *J. Biol. Chem. 263*, 11820–11825.

Pace, C.N., Shirley, B.A., & Thomson, J.A. (1990). Measuring the conformational stability of a protein. In *Protein Structure: A Practical Approach* (Creighton, T.E., Ed.), pp. 287–310. IRL Press, Oxford.

Pantoliano, M.W., Ladner, R.C., Bryan, P.N., Rollence, M.L., Wood, J.F., & Poulos, T.L. (1987). Protein engineering of subtilisin BPN′: Enhanced stabilization through the introduction of two cysteines to form a disulfide bond. *Biochemistry 26*, 2077–2082.

Perry, L.J. & Wetzel, R. (1984). Disulfide bond engineered into T4 lysozyme: Stabilization of the protein towards thermal inactivation. *Science 226*, 555–557.

Poland, D.C. & Scheraga, H.A. (1965). Statistical mechanics of noncovalent bonds in polyamino acids: VIII. Covalent loops in proteins. *Biopolymers 3*, 379–399.

Privalov, P.L. & Gill, S.J. (1988). Stability of protein structure and hydrophobic interaction. *Adv. Protein Chem. 39*, 191–234.

Privalov, P.L. & Khechinashvili, N.N. (1974). A thermodynamic approach to the problem of stabilization of globular protein structure: A calorimetric study. *J. Mol. Biol. 86*, 665–684.

Richardson, J. (1981). The anatomy and taxonomy of protein structure. *Adv. Protein Chem. 34*, 167–339.

Saunders, A.J., Young, G.B., & Pielak, G.J. (1993). The polarity of disulfide bonds. *Protein Sci. 2*, 1183–1184.

Schellman, J.A. (1978). Solvent denaturation. *Biopolymers 17*, 1305–1322.

Schwartz, H., Hinz, H., Mehlich, A., Tschesche, H., & Wenzel, H.R. (1987). Stability studies on derivatives of the bovine pancreatic trypsin inhibitor. *Biochemistry 26*, 3544–3551.

Shortle, D. (1989). Probing the determinants of protein folding and stability with amino acid substitutions. *J. Biol. Chem. 264*, 5315–5318.

Shortle, D., Chan, H.S., & Dill, K.A. (1992). Modelling the effects of mutations on the denatured states of proteins. *Protein Sci. 1*, 201–215.

Sosnick, T.R. & Trewella, J. (1992). Denatured states of ribonuclease A have compact dimensions and residual secondary structure. *Biochemistry 31*, 8329–8335.

Srinivasan, N., Sowdhamini, R., Ramakrishnan, C., & Balaram, P. (1990). Conformations of disulfide bridges in proteins. *Int. J. Pept. Protein Res. 36*, 147–155.

Stickle, J.F., Presta, L.G., Dill, K.A., & Rose, G.D. (1992). Hydrogen bonding in globular proteins. *J. Mol. Biol. 226*, 1143–1159.

Straume, M., Murphy, K.P., & Friere, E. (1992). Thermodynamic strategies for protein design: Increased temperature stability. In *Biocatalysis Near and Above 100 °C*. ACS Symposium Series, ACS Books, Washington, D.C.

Takagi, H., Takahashi, T., Momose, H., Inouye, M., Maeda, Y., Matsuzawa, H., & Ohta, T. (1990). Enhancement of the thermostability of subtilisin E by introduction of a disulfide bond engineered on the basis of structural comparison with a thermophilic serine protease. *J. Biol. Chem. 265*, 6874–6878.

Thornton, J.A. (1981). Disulphide bridges in globular proteins. *J. Mol. Biol. 151*, 261–287.

Tidor, B. & Karplus, M. (1993). The contribution of cross-links to protein stability: A normal mode analysis of the configurational entropy of the native state. *Proteins Struct. Funct. Genet. 15*, 71–79.

Villafranca, J.E., Howell, E.E., Oatley, S.J., Xuhong, N., & Kraut, J.

(1987). An engineered disulfide bond in dihydrofolate reductase. *Biochemistry 26*, 2182-2189.

Wells, J.A. & Powers, D.B. (1986). In vivo formation and stability of engineered disulfide bonds in subtilisin. *J. Biol. Chem. 261*, 6564-6570.

Wetzel, R. (1987). Harnessing disulfide bonds using protein engineering. *Trends Biochem. Sci. 12*, 478-482.

Wetzel, R., Perry, L.J., Baase, W.A., & Becktel, W.J. (1988). Disulfide bonds and thermal stability in T4 lysozyme. *Proc. Natl. Acad. Sci. USA 85*, 401-405.

Zhou, N.E., Kay, C.E., & Hodges, R.S. (1993). Disulfide bond contribution to protein stability: Positional effects of substitution in the hydrophobic core of the two-stranded α-helical coiled-coil. *Biochemistry 32*, 3178-3187.

Forthcoming Papers

Knowledge-based model building of proteins: Concepts and examples
J. Bajorath, R. Stenkamp, and A. Aruffo

Mechanism of phage P22 tailspike protein folding mutations
M. Danner and R. Seckler

New insights into protein-tyrosine kinase receptor signaling complexes
M.J. Fry, G. Panayotou, G.W. Booker, and M.D. Waterfield

Import of cytochrome b_2 to the mitochondrial intermembrane space: The tightly folded heme-binding domain makes import dependent upon matrix ATP
B.S. Glick, C. Wachter, G.A. Reid, and G. Schatz

Design of a functional calcium channel protein: Inferences about an ion channel-forming motif derived from the primary structure of voltage-gated calcium channels
A. Grove, J.M. Tomich, T. Iwamoto, and M. Montal

Macromolecular solvation energies derived from small molecule crystal morphology
D.C. Rees and G.M. Wolfe

Resolution of the fluorescence equilibrium unfolding profile of *trp* aporepressor using single tryptophan mutants
C.A. Royer, C.J. Mann, and C.R. Matthews