Low Energy Pathways and Non-native Interactions
THE INFLUENCE OF ARTIFICIAL DISULFIDE BRIDGES ON THE MECHANISM OF FOLDING*

Jody M. Mason[1]**, Matthew J. Cliff, Richard B. Sessions, and Anthony R. Clarke**
From the Department of Biochemistry, School of Medical Sciences, University of Bristol, Bristol BS8 1TD, United Kingdom

Four versions of a β-sheet protein (CD2.d1) have been made, each with a single artificial disulfide bond inserted into hairpin structures. Folding kinetics of reduced and oxidized forms shows bridge position strongly influences its effect on the folding reaction. Bridging residues 58 and 62 does not affect the rapidly formed intermediate (*I*) or rate-limiting transition (*t*) state, whereas bridging 33 and 38, or 31 and 41, lowers the *t*-state energy, with the latter having the stronger influence. Bridging residues 79 and 90 stabilizes both *I*- and *t*-states. To assess additivity in the energetic effects of these bridges, four double-bridge variants have also been made. All show precise additivity of overall stability, with two showing additivity when ground states and the rate-limiting *t*-state are assessed, *i.e.* no measurable change in the folding mechanism occurs. However, combining 31–41 and 79–90 bridges produces a molecule that folds through a different pathway, with a much more stable intermediate than expected and a much higher *t*-state barrier. This is explained by the artificial introduction of stabilizing, non-native contacts in the *I*-state. More surprisingly, for another double-bridge version (58–62 and 79–90) both *I*- and *t*-states are less stable than expected, showing that conformational constraints introduced by the two bridges prevent formation of non-native contacts that would otherwise stabilize the *I*- and *t*-states, thereby lowering the energy of the folding landscape in the wild-type (unbridged) molecule. We conclude that the lowest energy path for folding has *I*- and *t*-state structures that are stabilized by non-native interactions.

The first domain of the cell surface molecule CD2 is a member of the immunoglobulin family of β-sandwich proteins. It has proved to be a productive experimental model for the study of folding in proteins that comprise only β-sheets, because of the reversibility and high yield of the folding reaction and the large signal change that accompanies the folding process. In addition, the absence of disulfide bonds in the folded structure means that the native state is maintained only by non-covalent interactions and the folding reaction is described by a single exponential process showing that there are no subpopulations that fold by a different route and that there is only one rate-limiting transition state.

Despite the uncomplicated nature of the folding and unfolding kinetics, analysis of the denaturant dependence of the folding rate constant shows that the protein folds through a rapidly formed intermediate state, such that the overall folding process can be accurately described by a three-state mechanism as shown in Equation 1,

$$U \underset{k_u}{\overset{K_{(I/U)}}{\rightleftharpoons}} I \overset{k_f}{\rightleftharpoons} F \qquad \text{(Eq. 1)}$$

where U is the unfolded state, I is a partially folded intermediate state, and F is the fully folded state. The initial rapid collapse, described by the equilibrium constant $K_{(I/U)}$, occurs on the sub-millisecond time scale and leads to the formation of an I-state in which some aspects of overall topology of the molecule are established. Evidence for this is drawn from hydrogen-exchange experiments in which the most extensive I-state protection is detected in a nucleus of four β-strands (see Fig. 1, *B*, *C*, *E*, and *F*). Interestingly, the hydrogen-bonded strand pairings in this nucleus are distant in sequence (B with E and C with F) and form the most topologically complicated substructure within the folded molecule (1, 2). The simpler and more sequence-local substructures, namely the β-hairpins (a β-strand, followed by a turn of any description and a second β-strand that is hydrogen-bonded to the first), are also protected to some extent, particularly the long hairpin created by the hydrogen bonding of the F and G strands.

The rate-limiting transition state for folding of CD2 has also been examined by mutational analysis (3). These studies also emphasize the importance of the B,C,E,F nucleus (see Figs. 1 and 3*b*) in the transition state, because the only residues with transition state φ-values greater than 0.2 are in this region. It is also interesting to note that there are no statistically relevant φ-values greater than 0.5 in the whole protein, implying a rather poorly ordered transition state with respect to the intimacy of side-chain interactions.

In two other proteins of the immunoglobulin superfamily, namely TI 127 and TNfn3 (belonging to the Ig *I*-set and the fn3 set of folds, respectively), it is also the formation of the B,C,E,F nucleus that is the key to folding (4, 5). These findings reinforce the argument that topology and the transition state in the folding pathway are strongly related.

The kinetic influence of disulfide bridges has been used as a tool in dissecting the folding pathway of numerous proteins (6–9). The topological properties of the intermediate and transition states of CD2 have also been probed by examining the effects of preformed disulfide bridges on folding dynamics (10). It was found that, in the majority of cases, well designed cross-links (11, 12) stabilized the folded state by an amount that was predictable from a calculation of the entropic penalty in the unfolded state (13). Surprisingly, however, cross-links that connect β-strands that are distant in sequence were found predominantly to stabilize the rapidly formed intermediate state, suggesting that these strand-strand interactions occur in the initial stages of folding. Cross-links that stabilize local hairpins have their major influence on the second, rate-determining step leading to significant enhancements in the folding rate. The result implies that the slow, rate-limiting step must involve the consolidation of localized, hairpin structures, a process that occurs after the complex elements of overall topology are established. Thus long-range contacts, and therefore molecular order, can be established in early, non-rate-limiting processes.

FIGURE 1. **Ribbon diagram of domain 1 of CD2.** The topology of CD2.d1 is that of an immunoglobulin variable (V) domain with two β-sheets of strands conventionally labeled DEB and AGFCC′C″ (see also Fig. 3*a*). The bridges are given in lowercase letters (*a–d*) and refer to the residue positions in TABLE ONE. Taken from Ref. 17.

To extend the above study of disulfide engineering, in this report we have looked at the additivity of effects, *i.e.* the extent to which the energetic properties of a molecule in which two disulfide bonds have been engineered can be predicted by simply summing the effects of the two single bridges. We undertook this work to address three questions. First, could we combine bridges to produce hyper-stable versions of the protein? This would clearly have implications for engineering extreme robustness, at least into β-sheet-containing proteins. Second, could we combine bridges that individually produced fast-folding proteins to create an ultra-fast folder? Third, and more speculatively, if we found non-additivity in combining the bridges, what information might this reveal about the nature of the folding process?

EXPERIMENTAL PROCEDURES

Mutagenesis—All mutations were made on the rat gene cloned into the pGEX-2T glutathione *S*-transferase fusion vector (14). The proteins were expressed and purified as described previously (2).

Cross-linking Disulfides—To form the disulfide bridges, 10 μM protein was oxidized in an oxygen-purged buffer containing 20 μM zinc chloride, 50 mM triethanolamine hydrochloride at pH 7.5. To break the bridges, 10 μM protein in a helium-purged buffer containing 1 mM EDTA and 50 mM triethanolamine hydrochloride at pH 7.5 was reduced by the addition of 10 mM dithiothreitol. In addition, all other buffers were filtered and purged with the relevant gases. The proteins were then checked for oxidation and reduction using mass spectrometry and Elman's reagent (15).

Kinetic Analysis—The data have been fitted to Equation 2,

$$k_{obs} = k_{F-I} + k_{I-F}/(1 + 1/K_{I/U}) \qquad \text{(Eq. 2)}$$

where k_{F-I} and k_{I-F} are rate constants describing the forward and reverse reactions, respectively, between the folded and intermediate states (*I*). $K_{I/U}$ is the equilibrium constant ($[I]/[U]$) for the rapid interconversion of the intermediate and unfolded states (16). In the fitting routine, the following relationships were used as shown in Equations 3–5,

$$k_{F-I} = k_{F-I(w)}\exp(-m_t D) \qquad \text{(Eq. 3)}$$

$$k_{I-F} = k_{I-F(w)}\exp[(m_I - (m_t)D] \qquad \text{(Eq. 4)}$$

$$K_{I/U} = K_{I/U(w)}\exp[(m_u - m_I)D] \qquad \text{(Eq. 5)}$$

where the subscript *w* describes the rate and equilibrium constants in water, and the *m* parameters describe the shifts in the stabilities of each state (designated by the subscript) as a function of the guanidine hydrochloride (GuHCl) denaturant activity (*D*), where $D = [C_{0.5}\,[\text{GuHCl}]/(C_{0.5} + [\text{GuHCl}])] - 2.6[\text{Na}_2\text{SO}_4]$, where $C_{0.5}$ is a denaturation constant with a value of 7.5 M. This treatment has been explained in detail elsewhere (2, 16, 17), with the coefficient of 2.6 derived from the linear relationship between sodium sulfate (Na_2SO_4) concentration and denaturant activity between 0 and 0.4 M (*i.e.* solutions of 0, 0.1, 0.2, 0.3, and 0.4 M give molar denaturant activities of 0, −0.27, −0.51, −0.78, and −1.03 M, respectively). Na_2SO_4 is a kosmotropic agent and works in a manner analogous but opposite to GuHCl, increasing the free energy of solvation of hydrocarbon and consequently driving folding reaction in favor of more compact and desolvated states. The molar ability of Na_2SO_4 to decrease the extent of hydrocarbon burial has been scaled to the molar ability of GuHCl to increase it, allowing us to calculate the denaturant activity of Na_2SO_4 (16). This results in negative values of denaturant activity.

Folding measurements were initiated by mixing a 10 μM solution of unfolded Domain 1 of Cluster Determinant 2 (CD2.d1)[2] containing 50 mM triethanolamine hydrochloride and 3.27 M GuHCl against 10 volumes of a given concentration of GuHCl at 298 K in an SX.18MV stopped-flow apparatus (Applied Photophysics Ltd.). An excitation wavelength of 295 nm was selected by a single monochromator (bandpass 5 nm) from a mercury-xenon light source. The fluorescence intensity above 340 nm was recorded using an emission cut-off filter. For the unfolding reactions, a 10-mM solution of folded CD2.d1 in 50 mM triethanolamine hydrochloride, pH 7.5, was mixed with 10 volumes of an appropriate concentration of GuHCl at 298 K and the reaction recorded as above. The resulting data points are the result of at least three averages. The data are independent of protein concentration, with temperature control and mixing efficiency remarkably reliable in this apparatus. All reaction solutions were maintained at the appropriate temperature using a thermostated circulating water bath and were monitored continuously with a sensitive thermocouple. From this, the fluctuation in temperature was determined to be no more than ± 0.1 °C. Rate constants determined in the same conditions but in different experimental data sets vary by only 6–7%.

All data were fitted using the Grafit analysis software (Erathracus software). When fitting kinetic data to Equation 2, proportional weighting was used so that the fitted values took account of rate constants equally across the whole range.

Structure—The coordinates for wild-type CD2.d1 were obtained from the Protein Data Bank (1HNG).

RESULTS

Dissection of Folding Pathways—Shown in Fig. 2 are the rate plots (k_{obs} *versus* denaturant activity) for the four single-bridged mutants (*31–41, 33–38, 58–62,* and *79–90*) and the resulting four double-bridged mutants (*31–41 + 58–62, 33–38 + 58–62, 31–41 + 79–90,* and *58–62 + 79–90*) under reducing (no bridge) and oxidizing (intact bridge) conditions. These raw data have all been fitted to Equation 2 to obtain the parameters k_f, k_u, $K_{I/U}$, m_u, m_I, and m_t and are presented in TABLE ONE.

---

[2] The abbreviations used are: CD2.d1, Domain 1 of Cluster Determinant 2; GuHCl, guanidine hydrochloride.
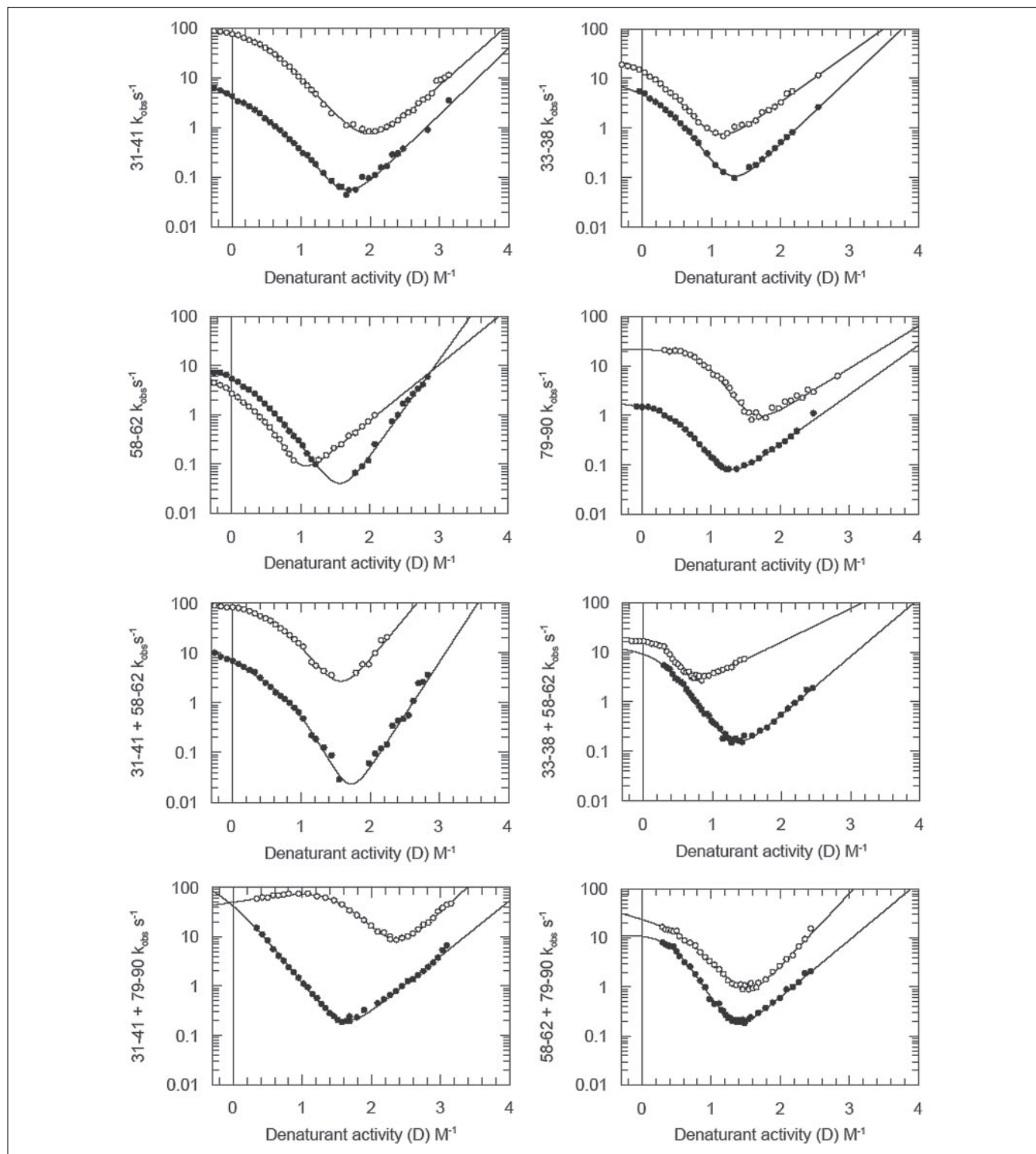
FIGURE 2. **Folding kinetics for the mutant proteins of CD2.d1**. Shown are the observed relaxation rates (k_{obs} *versus* denaturant activity) for the folding/unfolding of CD2.d1 (25 °C and pH 7.5), determined by stopped flow flourimetry (see "Experimental Procedures"). Each mutant is shown in a separate plot labeled upon the *y*-axis. The continuous *curves* represent fits to Equation 2. The calculated values of $k_{I-F(W)}$, $k_{F-I(W)}$, $K_{I/U(W)}$, m_U, m_I, m_t, are given in TABLE ONE. These plots are highly reproducible; repeated experiments with fresh buffers and protein preparations gave the same folding parameters within the measured fitting uncertainties (see TABLE ONE). Each data point represents a fit to at least five averaged kinetic acquisitions. The average error in these kinetic rate constants does not exceed 5%. The *open circles* represent the oxidized version (cross-link intact); the *closed circle* chevron plots represent the reduced version (cross-link broken).

As has been described previously, an analysis of the folding kinetics of the reduced and oxidized forms shows that the position of the bridge strongly influences its effect on the folding reaction (10). The effects of the bridges, with respect to free energy changes, are summarized in

TABLE TWO, which splits the reaction into three stages. The first stage (*U*-to-*I*) is the rapid transition from the unfolded state to the intermediate. The second is the passage from this intermediate state to the rate-limiting transition state (*I*-to-*t*). The third is the descent from this

TABLE ONE

Rate and equilibrium constants and *m*-values obtained from the chevron plots (Fig. 1) are shown for data collected in oxidizing (*ox*) and reducing (*red*) conditions

The bridge position is described by the mutation the single-letter code and by the strands that are joined. Quoted errors are cumulative S.E. calculated from the global least-square fits to the chevron plots (based on a 95% confidence limit) in Fig. 2.

	$k_{I\text{-}F(w)}$	$k_{F\text{-}I(w)}$	$k_{I/U(w)}$	M_u	m_I	m_t	ΔG
	s^{-1}				M^{-1}		*kcal/mol*
(a) R31C-E41C ox	90.0 ± 2.0	1.3e-3 ± 4.1e-4	3.3 ± 0.5	−6.4 ± 0.1	−2.7 ± 0.3	−2.8 ± 0.1	−7.5 ± 0.3
R31C-E41C red	4.6 ± 0.2	2.8e-4 ± 2.7e-5	10.0 ± 3.0	−7.5 ± 0.18	−4.6 ± 0.2	−3.1 ± 0.1	−7.2 ± 0.2
(b) E33C-L38C ox	19.6 ± 2.3	3.4e-2 ± 4.5e-3	2.2 ± 0.7	−6.6 ± 0.2	−2.5 ± 0.2	−2.3 ± 0.1	−4.5 ± 0.3
E33C-L38C red	5.7 ± 0.4	1.1e-3 ± 1.6e-4	4.8 ± 1.4	−7.9 ± 0.2	−3.8 ± 0.2	−3.0 ± 0.1	−6.1 ± 0.3
(c) L58C-D62C ox	2.8 ± 0.1	4.1e-3 ± 4.1e-4	3.6 ± 1.7	−10.2 ± 0.5	−4.8 ± 0.1	−2.6 ± 0.1	−4.8 ± 0.3
L58C-D62C red	6.7 ± 0.8	2.8e-4 ± 1.3e-4	3.8 ± 1.9	−9.1 ± 0.3	−5.4 ± 0.4	−4.5 ± 0.1	−6.1 ± 0.4
(d) T79C-D90C ox	21.4 ± 0.5	2.4e-2 ± 5.0e-4	140 ± 6.8	−7.4 ± 0.0	−2.0 ± 0.0	−1.97 ± 0.0	−7.0 ± 0.3
T79C-D90C red	1.6 ± 0.1	2.7e-3 ± 5.0e-4	3.5 ± 1.0	−7.4 ± 0.0	−2.6 ± 0.0	−2.3 ± 0.0	−4.7 ± 0.1
(a) + (c) E31C-D41C, L58C-D62C ox	85.0 ± 3.5	2.8e-3 ± 7.0e-4	12.1 ± 3.1	−8.1 ± 0.2	−4.1 ± 0.2	−3.9 ± 0.1	−7.7 ± 0.3
E31C-D41C, L58C-D62C red	6.7 ± 0.4	3.2e-5 ± 1.7e-5	61 ± 3.6	−9.4 ± 0.6	−6.6 ± 0.3	−4.8 ± 0.2	−8.6 ± 0.3
(a) + (d) E31C-D41C, T79C-D90C ox	50.3 ± 2.8	3.8e-3 ± 1.1e-3	77 ± 8.0	−6.5 ± 0.1	−2.5 ± 0.1	−3.0 ± 0.1	−8.2 ± 0.4
E31C-D41C, T79C-D90C red	45.9 ± 7.9	3.9e-4 ± 1.0e-4	<1	−6.8 ± 0.1	−3.5 ± 0.0	−2.6 ± 0.1	−6.1 ± 0.2
(b) + (c) E33C-L38C, L58C-D62C ox	16.5 ± 0.5	7.2e-1 ± 5.8e-2	3.5 ± 0.5	−9.1 ± 0.3	−1.4 ± 0.2	−1.6 ± 0.1	−2.8 ± 0.1
E33C-L38C, L58C-D62C red	11.5 ± 1.6	2.4e-3 ± 2.8e-4	3.8 ± 0.6	−7.5 ± 0.1	−2.9 ± 0.8	−2.7 ± 0.1	−6.0 ± 0.1
(c) + (d) L58C-D62C, T79C-D90C ox	25.0 ± 2.8	2.6e-3 ± 3.6e-4	13.7 ± 2.9	−7.8 ± 0.16	−4.2 ± 0.4	−3.4 ± 0.1	−7.1 ± 0.2
L58C-D62C, T79C-D90C red	11.8 ± 0.98	2.8e-3 ± 4.8e-4	9.8 ± 2.7	−7.9 ± 0.2	−2.5 ± 0.46	−2.7 ± 0.08	−6.4 ± 0.3

TABLE TWO

Changes in free energy (ΔΔ*G*) of defined steps in the folding reaction of CD2.d1

All values are derived from the parameters in TABLE ONE and are given in kcal/mol. All ΔΔ*G* values represent $\Delta G_{ox} - \Delta G_{red}$. The value $\Delta\Delta G_{U\text{-to-}I}$ represents the effect upon the intermediate state relative to the unfolded (denatured) state, and $\Delta\Delta G_{I\text{-to-}t}$ and $\Delta\Delta G_{U\text{-to-}F}$ represent energetic changes between the intermediate state and the transition state barrier, and the unfolded state and the folded state. In the last column, absolute stabilities of the oxidized proteins are shown for comparison. Quoted errors are S.E. given to two significant figures and were calculated from the kinetic plots using the standard statistical formula.

	$\Delta\Delta G_{U\text{-to-}I}$	$\Delta\Delta G_{I\text{-to-}t}$	$\Delta\Delta G_{t\text{-to-}F}$	$\Delta\Delta G_{U\text{-to-}F}$
31–41 (CC′)	+0.7 ± 0.3	−1.8 ± 0.1	+0.9 ± 0.1	−0.2 ± 0.4
33–38 (CC′)	+0.5 ± 0.4	−0.7 ± 0.1	+2.0 ± 0.3	+1.8 ± 0.4
58–62 (DE)	0.0 ± 0.5	+0.5 ± 0.1	+1.6 ± 0.4	+2.1 ± 0.6
79–90 (FG)	−2.2 ± 0.2	−1.5 ± 0.1	+1.3 ± 0.1	−2.4 ± 0.2
31–41 58–62 (CC′ + DE)	+1.0 ± 0.1	−1.5 ± 0.1	+2.7 ± 0.5	+2.2 ± 0.4
	(+0.7 ± 0.4)	(−1.3 ± 0.2)	(+2.5 ± 0.6)	(+1.9 ± 0.5)
31–41 79–90 (CC′ + FG)	−3.7	−0.1 ± 0.0	+1.4 ± 0.4	−2.4 ± 0.3
	(−1.5 ± 0.4)	(−3.3 ± 0.2)	(+2.2 ± 0.3)	(−2.6 ± 0.4)
33–38 58–62 (CC′ + DE)	0.0 ± 0.1	−0.2 ± 0.0	+3.4 ± 0.1	+3.2 ± 0.2
	(+0.5 ± 0.7)	(−0.2 ± 0.2)	(+3.6 ± 0.6)	(+3.9 ± 0.6)
58–62 79–90 (DE + FG)	−0.2 ± 0.3	−0.4 ± 0.2	0.0 ± 0.2	−0.6 ± 0.3
	(−2.2 ± 0.6)	(−1.0 ± 0.2)	(+2.9 ± 0.5)	(−0.3 ± 0.5)

high energy state to the folded ground state (*t*-to-*F*). For each stage in the reaction, the difference between the free energy change for the unbridged (reduced) molecule and the bridged (oxidized) molecule is given. The ΔΔ*G* values reported in TABLE TWO for each stage in the reaction are, therefore, calculated by Equation 6.

$$\Delta\Delta G_{(bridge)} = \Delta G_{(ox)} - \Delta G_{(red)} \qquad \text{(Eq. 6)}$$

These ΔΔ*G*(bridge) values should be a true reflection of the topological effect of folding with the bridge constraint compared to the bridge constraint, since the change in volume and hydration potential between two cysteine residues (2 × SH) and a cysteine bridge (S-S) is relatively small. Additionally, by using the reduced state as the reference, rather than the wild-type, small local perturbations caused by replacing the wild-type residue with a cysteine are discounted, and we reveal only the effect of oxidizing the cysteines. For the double-bridge versions the values in brackets represent the expected ΔΔ*G*(bridge) when the appropriate single-bridge values are summed together.

Effects of Single Bridges—The data presented in the upper half of TABLE TWO show the wide spectrum of effects in the single-bridge versions of CD2. For instance, a bridge between residues 58 and 62 has no effect on the rapidly formed intermediate and only slightly increases the energy barrier that defines the rate-limiting transition state for folding. The major effect of the bridge is to destabilize the folded state, presumably by introducing some degree of strain. The insertion of a bridge between residues 33 and 38 leads to a moderate reduction in the size of the transition state barrier (ΔΔ*G*(bridge) = −0.7 kcal/mol) and produces a protein with a 3-fold faster folding rate. The bridge between residues 31 and 41 is in the same hairpin and lowers the energy of the transition state to a greater degree (ΔΔ*G*(bridge) = −1.8 kcal/mol). In this case, we see a 20-fold enhancement in the folding rate. The most striking effect is produced by a bridge between residues 79 and 90, which markedly stabilizes both the intermediate state (ΔΔ*G*(bridge) = −2.2 kcal/mol) and reduces the height of the transition state barrier by 1.5 kcal/mol.
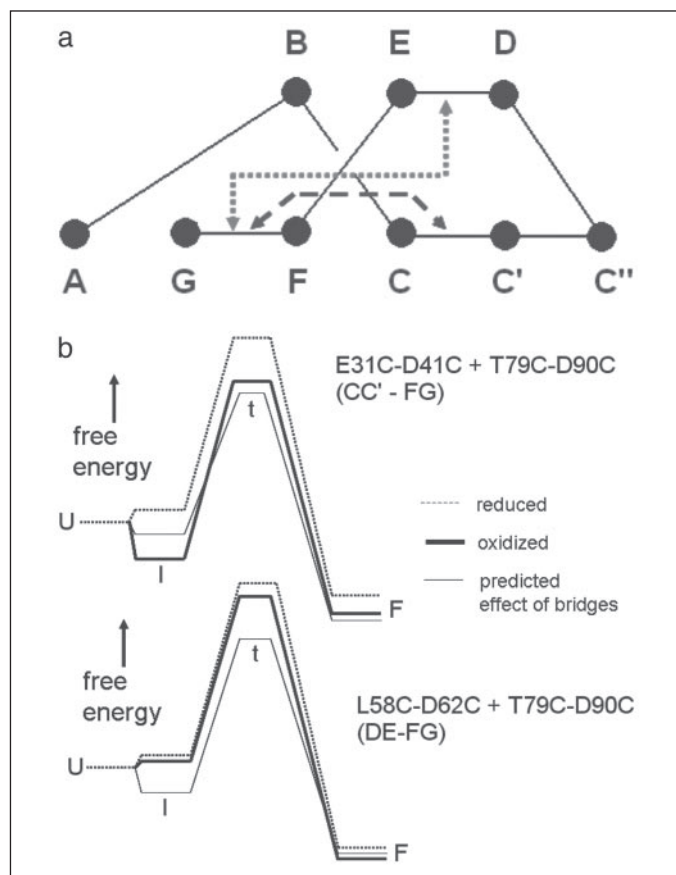
FIGURE 3. **Schematic energy profiles.** *a*, schematic section of CD2 domain 1 showing the hairpin structures. The interaction between CC′ and FG that creates a hyperstable, non-native folding intermediate in the 31–41 + 79–90 double-bridge molecule is indicated by *long-dash arrows*. The non-native interaction between DE and FG that normally leads to a stabilization of the intermediate and transition states, but is abolished in the 58–62 + 79–91 double-bridge version, is indicated by *short-dash arrows*. *b*, energy profiles are shown for the two cases where the effects of the single bridges are not additive. The letters *U*, *I*, *t*, and *F* represent unfolded (denatured), intermediate, transition, and folded states, respectively.

Combination of Bridges; Additivity in Overall Free Energy of Folding—What is more relevant to the study conducted here is the additivity of effects. All the four double-bridge variants show precise additivity with respect to overall stability ($\Delta\Delta G_{(bridge)}$ *U*-to-*F*). That is, the sum of the effects of the single bridges (see values in parentheses) is equal to the effect of the double bridge. This result demonstrates that, at least in principle, engineering high stability by insertion of multiple disulfide bonds into β-sheet regions of proteins is feasible.

Additivity of Energies along the Folding Pathway—The data in TABLE TWO show that the property of additivity of free energy is maintained through every step in the folding mechanism for two of the four double-bridge proteins (31–41 + 58–62 and 33–38 + 58–62). Within error, the measured effect of introducing two bridges into a single molecule is the same as summing the measured effects in the two individual single-bridge versions (see values in brackets in TABLE TWO), showing that the bridges induce no measurable change in the folding mechanism.

However, this property is not shown by the remaining two combinations. Combining the 31–41 and 79–90 bridges produces a molecule that folds through a different pathway, with an intermediate that is much more stable than expected from the additivity principle and a transition state barrier that is much higher in energy (see schematic profile in Fig. 3*a*). More explicitly, the $\Delta\Delta G_{(bridge)}$ for the *U*-to-*I* transition in the double-bridge version is greater than −3.7 kcal/mol (in the

sense of being more stable). The fact that we see no evidence for a measurable population of intermediate in the reduced form means that we do not know the upper limit for this effect. However, from summing the values for the individual bridges we would expect a $\Delta\Delta G_{(bridge)}$ of −1.5 kcal/mol, *i.e.* the measured effect of the double bridge is >2.2 kcal/mol greater than the predicted effect. This can be defined as a synergistic combination of bridges when considering the energy of formation of the folding intermediate, *i.e.* the sum is greater than expected from its parts.

When examining the height of the transition state barrier, it is interesting to note that the reduced and oxidized versions of 31–41 + 79–90 mutant fold at almost the same rate, whereas we would expect a 260-fold enhancement of the rate with its bridges intact. This means that the barrier is 3.2 kcal/mol higher than expected because of the introduction of stabilizing, non-native contacts in the *I*-state that have to be broken before the molecule can achieve the productive transition state.

The hope in combining the CC′ and FG bridges was that we would not alter the folding pathway and so combine the reductions in the height of the folding barrier occasioned by the single bridges. This would have created a variant of CD2 with a folding rate of >10,000 s^{-1} if there were a 260-fold enhancement of the rate constant for the reduced system. However, the distortion in the pathway defeated this aim.

Perhaps more surprisingly, the remaining double-bridge version (58–62 + 79–90) has an intermediate state and a transition state that are less stable than expected. The measured $\Delta\Delta G_{(bridge)}$ for the *U*-to-*I* transition in the double-bridge version (see TABLE TWO) is close to zero. However, from summing the values for the individual bridges we would expect a $\Delta\Delta G_{(bridge)}$ of −2.2 kcal/mol, *i.e.* the measured effect of the double bridge is >2.2 kcal/mol smaller than the predicted effect (see Fig. 3*a*). This can be defined as an anti-synergistic combination of bridges when considering the energy of formation of the folding intermediate, *i.e.* the sum is less than expected from its parts. However, as the term anti-synergistic contains an intrinsic contradiction, this effect will be called a destructive combination.

When examining the height of the transition state barrier, it is interesting to note that this destructive effect persists into the next stage of folding, *i.e.* the transition state barrier. This is 0.6 kcal/mol higher in the double mutant than would be expected from summation of the singles (Fig. 3*b*).

This means that the constraints on conformation introduced by the two bridges in the 58–62 + 79–90 mutant are acting in the opposite way to those described for the 31–41 + 79–90 mutant. That is, they must prevent rather than cause the formation of non-native contacts that would otherwise stabilize the *I*-state and the *t*-state in the folding pathway of the single-bridge versions.

DISCUSSION

Implications of Synergistic Combinations—The principle behind additivity of free energy in a pathway is that when two bridges combine synergistically, *i.e.* their effect on the energy change for a given step is greater than predicted from the sum of the individuals, then the pathway has changed by creating a new conformation of lower free energy than is achieved in either of the single-bridge proteins. This is exemplified by the 31–41 + 79–90 version of CD2, which has an intermediate state that is >2.2 kcal/mol more stable than predicted. This means that there are stabilizing, non-native interactions between the CC′ and FG hairpins in the double-bridge version that are absent in the singles (Fig. 3*b*). However, the fact that additivity is maintained across the whole pathway, from unfolded to fully folded state, means that this advantage

must be lost at a later stage. In fact, the transition state barrier is 3.2 kcal/mol higher than predicted in the doubly bridged molecule. The fact that there is a negative slope in the rate plot for the double mutant at low denaturant concentrations (see Fig. 2) shows that the intermediate is more compact than the transition state. In turn, this demonstrates that the stabilizing non-native contacts formed in the intermediate state have to be broken before the transition state can be attained.

This result can be rationalized in structural terms by the following argument. As shown in the schematic in Fig. 3, the CC′ and FG hairpins are not in contact in the folded state, and analysis of the single-bridge 31–41 version of CD2 shows that the CC′ hairpin is not normally formed in the *I*-state. If the presence of this disulfide bridge forces hairpin formation then there is the possibility of a strong non-native interaction between the two hydrophobic faces of the hairpins. This interaction would then have to be broken to achieve the native topology.

Implications of Destructive Combinations—According with the above paradigm, two bridges can be considered to combine in a destructive manner when their effect is less than predicted from the sum of the individual effects, *i.e.* the pathway is changed in the double mutant by preventing stabilization of non-native interactions that normally occur in the single-bridge proteins. The 58–62 + 79–90 version of CD2 exemplifies this phenomenon, suggesting that such non-native contacts between the DE and FG hairpins occur in the normal folding pathway (see Figs. 1 and 3a). The predicted free energy stabilization ($\Delta\Delta G_{(bridge)}$) in the double mutant for the formation of the *I*-state is −2.2 kcal/mol, whereas the observed is only −0.2 kcal/mol.

In the case of the 79–90 single-bridge version, we see a marked stabilization of the *I*-state (>3.7 kcal/mol). This result shows that the FG hairpin must be at least partially formed in this conformation when the wild-type molecule folds. However, the single 58–62 bridge has, if anything, a destabilizing effect on the *I*-state showing that in the wild-type folding pathway the DE hairpin is not formed in the *I*-state. The destructive nature of the DE-FG combination in the double-bridge molecule can therefore be explained if the FG hairpin in normal folding pathway is stabilized by non-native interactions with the DE segment of the protein that has not, at this stage, formed a hairpin.

It is interesting to note that the expected reduction in the transition state barrier from the summation of the single-bridge effects is 1.0 kcal/mol, whereas the observed stabilization is only 0.4 kcal/mol. This result implies that the non-native contacts that normally stabilize the intermediate state are maintained to a lesser degree in the transition state.

Concluding Remarks—Although one of the objectives of these experiments was to make a fast-folding version of CD2 by combining an artificial bridge across the FG hairpin with one across the CC′ hairpin, we would argue that our failure to do so was instructive. Rather than

making a fast folder, we succeeded only in forming a misfolded overstable intermediate, *i.e.* we distorted the folding pathway by stabilizing a substructure (CC′) that is normally unfolded in the *I*-state, thereby creating non-native interactions between the two hairpins. The lesson here is to engineer more intelligently and, when combining bridges to engineer high folding rates, choose to stabilize hairpins that are in contact in the native state. Hence, non-native interactions that increase the height of the major energy barrier would probably be avoided by pursuing this strategy.

To add to this, we found evidence for the beneficial role of non-native interactions in a folding pathway. When the DE bridge was combined with the FG, we were surprised to find that the double-bridge molecule folded through a less stable intermediate and through a higher energy transition state than would have been predicted from the effects of the two single bridges. This shows that the double bridge has abolished non-native contacts that normally stabilize intermediate and transition states in the folding pathway. The result implies that, perhaps counterintuitively, contacts between regions of the polypeptide chain that are not in contact in the folded native state can assist, rather that inhibit, the kinetics of the folding process.

REFERENCES

1. Hosszu, L. L., Craven, C. J., Parker, M. J., Lorch, M., Spencer, J., Clarke, A. R., and Waltho, J. P. (1997) *Nat. Struct. Biol.* **4,** 801–804
2. Parker, M. J., and Clarke, A. R. (1997) *Biochemistry* **36,** 5786–5794
3. Lorch, M., Mason, J. M., Sessions, R. B., and Clarke, A. R. (2000) *Biochemistry* **39,** 3480–3485
4. Clarke, J., Cota, E., Fowler, S. B., and Hamill, S. J. (1999) *Structure Fold. Des.* **7,** 1145–1153
5. Fowler, S. B., and Clarke, J. (2001) *Structure (Camb.)* **9,** 355–366
6. Clarke, J., and Fersht, A. R. (1993) *Biochemistry* **32,** 4322–4329
7. Yokota, A., Izutani, K., Takai, M., Kubo, Y., Noda, Y., Koumoto, Y., Tachibana, H., and Segawa, S. (2000) *J. Mol. Biol.* **295,** 1275–1288
8. Grantcharova, V. P., Riddle, D. S., and Baker, D. (2000) *Proc. Natl. Acad. Sci. U. S. A.* **97,** 7084–7089
9. Zavodszky, M., Chen, C. W., Huang, J. K., Zolkiewski, M., Wen, L., and Krishnamoorthi, R. (2001) *Protein Sci.* **10,** 149–160
10. Mason, J. M., Gibbs, N., Sessions, R. B., and Clarke, A. R. (2002) *Biochemistry* **41,** 12093–12099
11. Wouters, M. A., and Curmi, P. M. (1995) *Proteins* **22,** 119–131
12. Hutchinson, E. G., Sessions, R. B., Thornton, J. M., and Woolfson, D. N. (1998) *Protein Sci.* **7,** 2287–2300
13. Pace, C. N., Grimsley, G. R., Thomson, J. A., and Barnett, B. J. (1988) *J. Biol. Chem.* **263,** 11820–11825
14. Driscoll, P. C., Cyster, J. G., Campbell, I. D., and Williams, A. F. (1991) *Nature* **353,** 762–765
15. Creighton, T. (ed) (1990) *Protein Structure: A Practical Approach*, pp. 101–133, IRL Press, Oxford
16. Parker, M. J., Spencer, J., and Clarke, A. R. (1995) *J. Mol. Biol.* **253,** 771–786
17. Parker, M. J., Dempsey, C. E., Lorch, M., and Clarke, A. R. (1997) *Biochemistry* **36,** 13396–13405