Protein Fragments as Models for Events in Protein Folding Pathways: Protein Engineering Analysis of the Association of Two Complementary Fragments of the Barley Chymotrypsin Inhibitor 2 (CI-2)

Javier Ruiz-Sanz,[‡] Gonzalo de Prat Gay, Daniel E. Otzen,[§] and Alan R. Fersht*

MRC Unit for Protein Function and Design and Cambridge Centre for Protein Engineering, Department of Chemistry, Cambridge University, Lensfield Road, Cambridge CB2 1EW, United Kingdom

Received September 19, 1994; Revised Manuscript Received November 28, 1994[⊗]

ABSTRACT: Two fragments of chymotrypsin inhibitor-2, CI-2(20−59) and CI-2(60−83), derived from cyanogen bromide cleavage at Met-59, associate to give a native-like structure. We analyze the kinetics and equilibria of association of mutant fragments derived from cleaving mutant proteins at the same methionine residue. The changes in free energy of association have been measured both from isothermal studies of the binding of fragments and from thermal denaturation of the complexes. In general, there is a good correlation between the changes on mutation of the free energy of association of fragments and the changes in free energy of folding of the uncleaved parent protein. The notable exceptions are for residues in regions of the fragments that form nonnative hydrophobic clusters in the isolated fragments; mutation of the hydrophobic residues involved in these clusters decreases the equilibrium constant for formation of the noncovalent complex less than it does the equilibrium constant for folding of intact protein. The dissociated fragments must be destabilized by mutation of those hydrophobic residues, but to a lesser extent than is the complex itself. These clusters are thus less important energetically in the denatured state of the intact protein. The second-order rate constants for the major phase of association change with mutation, similar results being obtained from fluorescence measurements of the regain of tertiary structure and from circular dichroism measurements of the regain of secondary structure. The rate constants for association correlate well, in general, with the rate constants of refolding of the respective uncleaved proteins. Fragments that have mutations in the regions of nonnative hydrophobic clusters associate faster than expected from the correlation. Thus, breaking up the clusters facilitates the rate of folding. It is remarkable that the two fragments associate via a transition state that is very similar to that for the folding of the intact protein.

Early events in protein folding may include either the formation of secondary structural elements and/or a hydrophobic collapse (Ptitsyn, 1973; Karplus & Weaver, 1976; Sali & Karplus, 1994; Kim & Baldwin, 1990; Moult & Under, 1991; Dill & Shortle, 1991). Protein fragments are being used as models for studying these early stages of the folding process (Dyson *et al.*, 1992; Sancho *et al.*, 1992; Dyson & Wright, 1993, and references therein; Waltho *et al.*, 1993; Shortle & Abeygunawardana, 1993; Prat Gay & Fersht, 1994, and references therein) because they are more amenable than the unfolded states of intact proteins for structural characterization and structure in fragments may reflect local structural tendencies. Structural studies alone are not sufficient to characterize folding pathways: they must be linked to kinetic studies to see what features are relevant and which are on side paths (Fersht, 1993). Further, the response of changes in structure to changes in sequence are informative. We have recently characterized the association of the two complementary fragments of chymotrypsin inhibitor-2 (CI-2), CI-2(20−59) and CI-2(60−83), generated by cleaving with cyanogen bromide the native wild-type protein at the reactive site bond Met-59−Glu-60, which is in a loop (Prat Gay & Fersht, 1994; Prat Gay *et al.*, 1994a). The association process generates a native-like structure starting from an "unfolded state", i.e., the fragments in water. The scissile bond is present in most CI-2 mutants generated in this laboratory.

The tertiary structure of CI-2 consists of a four-stranded mixed parallel and antiparallel β-sheet packed against an α-helix to form the hydrophobic core (McPhalen & James, 1987). Both the kinetics and equilibrium of folding fit a simple two-state mechanism (Jackson & Fersht, 1991). The transition state of CI-2 in the folding pathway has been characterized by protein engineering (Otzen *et al.*, 1994). A molecular dynamics simulation of the unfolding of CI-2 gives a description of the transition state that is in excellent agreement with the experiments (Li & Daggett, 1994). We have shown previously that the kinetics of association of the two fragments consists of a second-order step combined with a rearrangement of structure, akin to a protein folding process (Prat Gay *et. al.*, 1994a). The family of mutant fragments generated by cleavage at Met-59 allows systematic complementation studies in the same way as a protein engineering procedure that was used to analyze the folding of barnase (Matouschek *et al.*, 1989; Fersht *et al.*, 1992) and intact CI-2. This has enabled a description of the transition state for the association (Prat Gay *et al.*, 1994b). It is remarkably similar to that for the folding of the intact protein, except that the N-terminal region of the single helix is completely

* To whom correspondence should be addressed.
[‡] Recipient of an EC fellowship.
[§] Recipient of a fellowship from The Danish Natural Science Council.
[⊗] Abstract published in *Advance ACS Abstracts,* January 15, 1995.

0006-2960/95/0434-1695$09.00/0

formed in the association reaction, whereas the structure is present in a weakened form in the intact protein.

There is evidence that there is residual structure in the cleaved fragments, manifested in the formation of nonnative hydrophobic clusters (Prat Gay *et al.*, 1994b). In the present study, we analyze the association equilibrium and kinetics of the mutant fragments of CI-2(20−59) and CI-2(60−83) to investigate the mechanism of their association. We also characterise the thermal dissociation/unfolding of the complexes by circular dichroism. We selected mutants in the hydrophobic core (LA27, VA38, IV39, IV48, VA70, and IV76), the α-helix (SA31, EN33, EA33EA34, SA31EA33EA34, KA36, KA37, KG37 and KA43), at positions in β-sheets (TA22, IV49, IA49, VT53, and VA53), and a small hydrophobic "minicore" (LA51, LV51, and FL69).

EXPERIMENTAL PROCEDURES

Materials. Cyanogen bromide (CNBr) and trifluoroacetic acid (TFA) were purchased from Fluka, Switzerland. Sodium dihydrogen orthophosphate and disodium hydrogen orthophosphate were analytical grade and acetonitrile was HPLC grade, all of them from Fisons, England. All other reagents were of highest purity available. Water was deionized and purified with an Elgastat system. Expression, purification, and description of the wild-type and mutant CI-2 proteins used in this work have been described (Jackson *et al.*, 1993; Otzen *et al.*, 1994).

Cleavage and Purification of the Fragments. Wild-type and mutant CI-2 proteins were dissolved in water and the pH taken to 2.3 with HCl. The cleavage and purification procedures were as described (Prat Gay & Fersht, 1994). Two fragments were obtained from each mutant, CI-2(20−59) and CI-2(60−83), described by the residue numbering used previously (Prat Gay & Fersht, 1994), the first residue being number 20. The molecular mass of each fragment was measured by electron spray ionization mass spectrometry and found to be in excellent agreement with those calculated from the primary sequence.

Fluorescence. All the methods for fluorescence experiments, equilibria and kinetics, carried out in this work have been described (Prat Gay & Fersht, 1994; Prat Gay *et al.*, 1994a).

Circular Dichroism. CD spectra and kinetic experiments were performed as described (Prat Gay & Fersht, 1994; Prat Gay *et al.*, 1994a). The thermal scan speed was 50 °C/h, and data were collected every 0.2 °C. The data were analyzed using Kaleidagraph (Abelbeck software) and fitted to a two-state transition, in which the dissociation and unfolding processes are occurring simultaneously: Complex [(20−59)•(60−83)] (folded state)$_R$ CI-2(20−59) + CI-2(60−83) (unfolded state).

RESULTS

Equilibrium Association of Mutant Fragments Measured by Fluorescence and Circular Dichroism. CI-2 contains only one tryptophan residue (Trp 24), which is responsible for the fluorescence emission increase on unfolding the protein with a maximum at 356 nm (Jackson & Fersht, 1991). The fluorescence of wild-type CI-2(20−59) decreases on mixing it with increasing concentrations of CI-2(60−83); the fluorescence spectrum of the complex at equimolar concentrations of the fragments is identical to that of native uncleaved CI-2 (Prat Gay & Fersht, 1994). The far-UV circular dichroism spectrum of the [CI-2(20−59)•(60−83)] complex is also identical to that of native uncleaved CI-2. This indicates that the secondary and tertiary structures of the fragments judged by their spectra are very similar to those of the uncleaved protein (Prat Gay & Fersht, 1994).

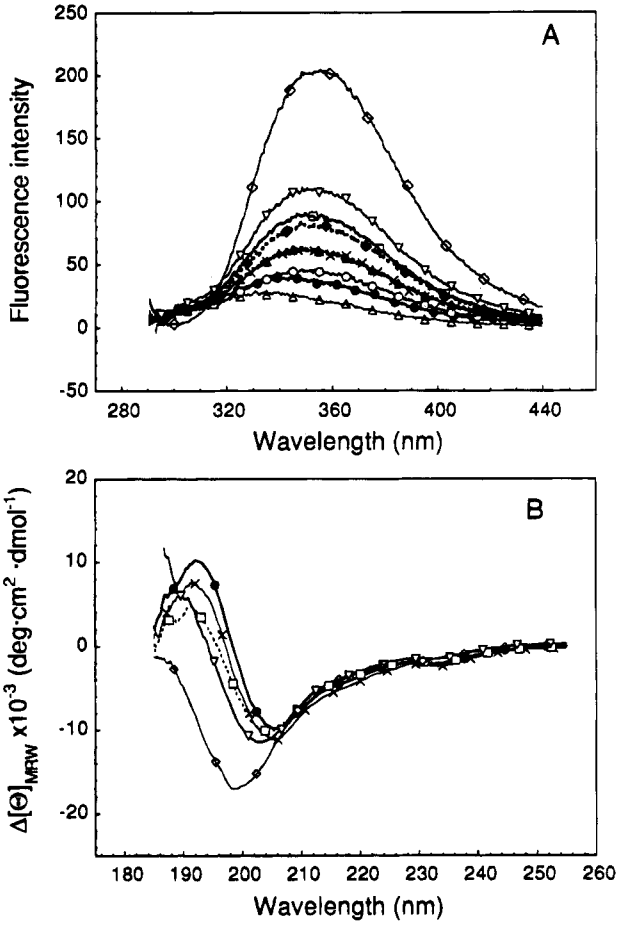

FIGURE 1: (A) Fluorescence spectra of CI-2 complexes and of wild-type CI-2 fragment (20−59) in 10 mM phosphate, pH 6.3. (△) IV76 complex; (●) wild-type complex; (○) KA43 complex; (×) KG37 complex; (s) IV39 complex; (◆) VA70 complex; (o) SA31/EA33/EA34 complex; (▽) LA27 complex; and (◇) wild-type CI-2 fragment (20−59). The scans were recorded at 50 nm/min with a step resolution of 0.2 nm, using an excitation wavelength of 280 nm. (B) Circular dichroism spectra of CI-2 complexes and wild-type CI-2 fragment (20−59) in 10 mM phosphate, pH 6.3. (●) Wild-type complex; (×) KG37 complex; (o) SA31/EA33/EA34 complex; (▽) LA27 complex; and (◇) wild-type CI-2 fragment (20−59). The scans were recorded at 50 nm/min with a step resolution of 0.2 nm, using a 0.5 cm path length cuvette. The concentration of each sample in both CD and fluorescence was 5 μM. The experiments were carried out at 25 °C.

Fluorescence spectra of the mutant fragments of CI-2(20−59) are identical to those of wild-type CI-2(20−59) (not shown). The fluorescence spectra for several different mutant complexes and for the wild-type CI-2(20−59) are shown in Figure 1A; both the fluorescence intensity and the wavelength maximum decrease on association. The differences between the fluorescence spectra and those of wild-type complex result only from the incomplete formation of the complexes. For this reason, the spectra of the complexes that have high dissociation constants are affected strongly by the fluorescence of the free fragment CI-2(20−59), being

Table 1: Equilibrium Dissociation Constants and Thermal Stability Data of Complexes

CI-2 mutant	K_D[a] (μM)	T_m (K)	ΔH_{T_m} (kcal mol^{-1})	$\Delta\Delta G_D$[b] (kcal mol^{-1})	$\Delta\Delta G_{312}$[c] (kcal mol^{-1})	$\Delta\Delta G_{F\text{-}U}$[d] (kcal mol^{-1})
wt[e]	0.041 ± 0.002	319.6 ± 0.4	34.1 ± 1.0		0.00 ± 0.04	
SA31	0.270 ± 0.006	311.2 ± 3.3	35.4 ± 2.7	1.12 ± 0.03	0.84 ± 0.39	0.89 ± 0.05
SA31EA33EA34	0.173 ± 0.010	305.5 ± 1.5	53.0 ± 4.8	0.86 ± 0.05	1.92 ± 0.31	1.67 ± 0.05
EA33EA34	0.242 ± 0.001	311.2 ± 1.2	39.0 ± 2.6	1.06 ± 0.03	0.85 ± 0.16	0.76 ± 0.06
EN33	0.099 ± 0.017	315.8 ± 0.4	36.1 ± 4.7	0.53 ± 0.11	0.33 ± 0.08	0.70 ± 0.05
KA36	0.095 ± 0.021	307.0 ± 1.9	27.2 ± 4.5	0.50 ± 0.13	1.22 ± 0.21	0.49 ± 0.03
KA37	0.107 ± 0.026	314.4 ± 1.1	34.8 ± 4.0	0.57 ± 0.15	0.49 ± 0.13	−0.23 ± 0.13
KG37	0.673 ± 0.190	308.8 ± 1.0	37.6 ± 4.0	1.66 ± 0.17	1.15 ± 0.14	0.97 ± 0.08
KA43	0.104 ± 0.013	313.8 ± 1.5	24.6 ± 3.0	0.55 ± 0.08	0.61 ± 0.12	0.65 ± 0.06
LA27	ND	297.1 ± 3.6	34.6 ± 8.2	ND	2.75 ± 0.71	2.68 ± 0.15
VA38	0.096 ± 0.008	314.9 ± 0.2	30.0 ± 1.5	0.51 ± 0.06	0.48 ± 0.05	0.48 ± 0.06
IV39	0.346 ± 0.040	304.8 ± 0.4	26.3 ± 0.4	1.27 ± 0.09	1.43 ± 0.06	1.30 ± 0.05
IV48	0.064 ± 0.020	312.4 ± 1.3	27.8 ± 3.5	0.27 ± 0.19	0.71 ± 0.12	1.11 ± 0.05
IA49	0.269 ± 0.018	ND	ND	1.12 ± 0.05	ND	2.12 ± 0.06
IV49	0.269 ± 0.018	ND	ND	1.12 ± 0.05	ND	2.12 ± 0.06
IV49	0.044 ± 0.002	321.0 ± 0.9	50.9 ± 6.2	0.05 ± 0.04	−0.59 ± 0.22	−0.08 ± 0.08
VA70	0.336 ± 0.068	302.8 ± 1.6	33.5 ± 5.1	1.25 ± 0.12	1.86 ± 0.27	1.95 ± 0.07
IV76	0.017 ± 0.004	320.7 ± 0.8	34.6 ± 2.9	−0.52 ± 0.14	−0.11 ± 0.11	−0.19 ± 0.10
TA22	0.071 ± 0.004	316.7 ± 1.1	ND	0.33 ± 0.04	ND	0.85 ± 0.05
VA53	0.029 ± 0.019	316.8 ± 1.1	40.5 ± 5.0	−0.20 ± 0.39	0.16 ± 0.15	0.63 ± 0.11
VT53	0.163 ± 0.053	311.3 ± 1.4	34.0 ± 4.8	0.82 ± 0.19	0.82 ± 0.16	1.03 ± 0.05
LA51	0.364 ± 0.008	310.5 ± 1.3	42.5 ± 6.2	1.30 ± 0.03	0.95 ± 0.19	2.37 ± 0.05
LV51	0.250 ± 0.046	319.1 ± 1.1	37.3 ± 6.3	1.07 ± 0.11	−0.03 ± 0.18	0.50 ± 0.06
FL69	0.209 ± 0.027	305.5 ± 3.7	27.5 ± 8.5	0.97 ± 0.08	1.38 ± 0.44	2.11 ± 0.06

[a] The values are averages of at least two experiments. [b] $\Delta\Delta G_D$ is the change in free energy of dissociation, calculated from eq 1. [c] $\Delta\Delta G_{312}$ is the change in free energy of association obtained from eq 5 for thermal dissociation, calculated for a mean T_m value of 312 K. [d] $\Delta\Delta G_{F\text{-}U}$ is the change in free energy of folding of the protein on mutation. Data from Jackson *et al.* (1993) and Otzen *et al.* (1994) calculated from guanidinium chloride-induced unfolding at equilibrium. These values are within experimental error of those determined by differential scanning calorimetry in the absence of guanidinium chloride. [e] Previously obtained value is 0.042 ± 0.02 (Prat Gay & Fersht, 1994).

the sum of the fluorescence of the free fragment and of the complex in a native-like structure.

Far-UV circular dichroism (CD) spectra recorded between 255 and 185 nm for the CI-2(20−59) mutant fragments and their complexes are identical to their equivalents of wild-type CI-2(20−59). The fragments have a minimum at 200 nm, characteristic of a disordered structure, and a shoulder at 212 nm which could be indicative of a nonrandom coil structure, as described for the wild-type CI-2(20−59) fragment (Prat Gay & Fersht, 1994). The far-UV CD spectra of the complexes show differences from those of the fragments (Figure 1B). The minimum at 200 nm is shifted to wavelengths between 203−205 nm, and a strong positive band appears at 190 nm, which is present in native uncleaved CI-2. However, these spectra vary between the mutant complexes, the differences being related to the extent of their formation (as calculated from their measured dissociation constants) and not to different structures, as is the case of the fluorescence spectra.

Determination of Dissociation Constants. The change in the fluorescence intensity upon formation of the complex (Figure 1A) was used to determine the equilibrium dissociation constant, K_D, using a fixed concentration of 0.65 μM of CI-2(20−59) (the fragment containing the Trp) and varying concentrations of CI-2(60−83) (Table 1).

The change in the free energy of association on mutation ($\Delta\Delta G_D$), is calculated from equation 1

$$\Delta\Delta G_D = -RT \ln (K_D^{WT}/K_D^{Mut}) \quad (1)$$

The values of $\Delta\Delta G_D$ parallel $\Delta\Delta G_{F\text{-}U}$, the change in free energy of folding of the intact protein on mutation (Table 1). This suggests that the association process is directly linked to the folding process of the uncleaved protein.

Thermal Stability of Wild-Type and Mutant [CI-2(20−59)•CI-2(60−83)] Complexes. The wild-type complex dissociates and denatures with increasing temperature (Prat Gay & Fersht, 1994). A detailed thermodynamic analysis using differential scanning calorimetry (DSC) is not possible since the wild-type CI-2 complex has a nonreversible broad transition above pH 4 (not shown) and the complex dissociates at low pH (Prat Gay & Fersht, 1994). However, the thermal denaturation of the complexes at lower concentrations could be monitored by far-UV CD. The thermal denaturation of the wild-type complex is reversible at 5 μM of complex in 10 mM phosphate buffer pH 6.3 at a temperature scan speed of 50 °C/h (Figure 2A). The change in ellipticity observed with temperature results from the unfolding/dissociation of the complex and not the unfolding of residual structure in the free fragments since the thermal denaturation of the fragments CI-2(20−59) and CI-2(60−83) does not show cooperative transitions (Figure 2B).

The unfolding data were fitted to theoretical curves for simultaneous dissociation and unfolding processes. The quantity α is defined as the fraction of the complex that is denatured at any temperature for the equilibrium between associated complex and free fragments ($WY_RW + Y$). The equilibrium constant, $K(T)$, is defined by

$$K(T) = \frac{[W][Y]}{[WY]} \quad (2)$$

We define α = the fraction of dissociation of the complex = $(0.5[W])/C_t$. Then

$$\alpha = \frac{-K(T) + \sqrt{K(T)^2 + 16C_tK(T)}}{8C} \quad (3)$$

Using the procedure described by Marky and Breslauer

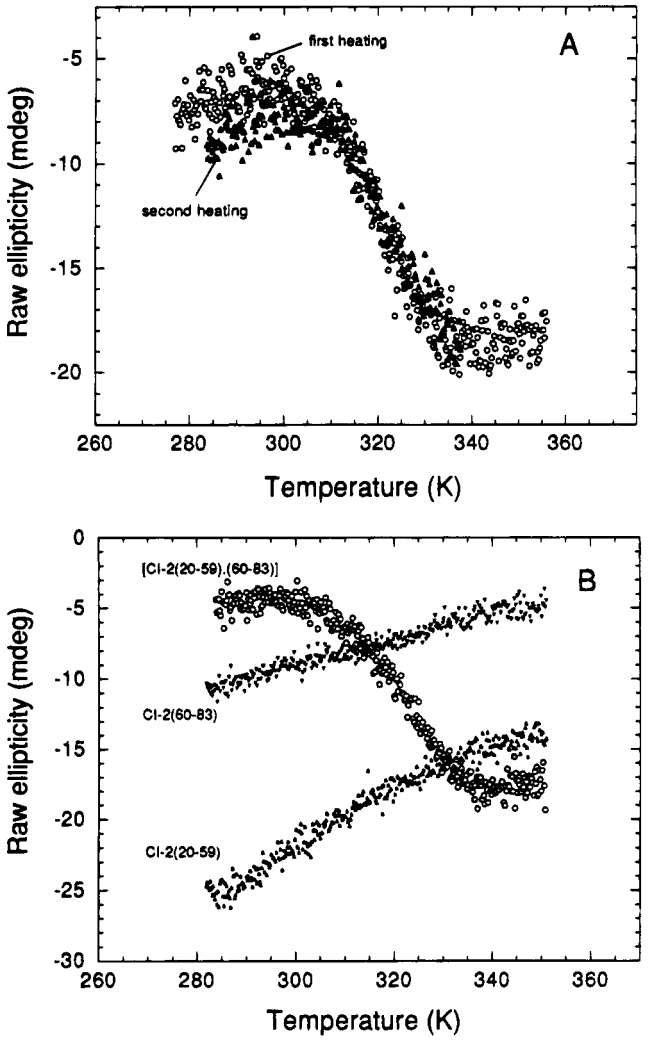

FIGURE 2: Thermal denaturation experiments followed by measurement of the change in ellipticity at 200 nm. (A) Reversibility of the thermal transition of the wild type CI-2 complex; reheating of the same sample was carried out after cooling it from 80 °C down to 10 °C. (B) Temperature scans of wild-type complex, CI-2(20−59) and CI-2(60−83) fragments. All the scans were performed at 50 °C/h and at a fixed concentration of 5 μM complex.

(1987) and Steif *et al.* (1993), we obtain

$$K(T) = \exp[[-\Delta H_m(1 - T/T_m) + \Delta C_p(T - T_m - \ln(T/T_m)) - RT\ln(2C_t)]/RT] \quad (4)$$

where C_t corresponds to the initial concentration of the complex (5 μM in these experiments), ΔC_p is the difference in the molar heat capacity between the native and denatured states, and T_m is defined as the temperature at which $\alpha = 0.5$. We have assumed that the ΔC_p value for each of the complexes studied is similar to that of the uncleaved mutant since it cannot be calculated from DSC experiments; this value was found to be 720 cal mol^{-1} K^{-1} for the wild-type CI-2 and assumed to be constant for all the mutants (Jackson *et al.*, 1993). If we use the theoretical approximation of multiplying the number of residues in the protein by 12 cal mol^{-1} K^{-1} (Edelhoch & Osborne, 1976), the ΔC_p is 768 cal mol^{-1} K^{-1}, and this translates to only 0.4% difference in T_m. For the purpose of this study, it is the change in T_m, rather than the absolute value of T_m, which is important; T_m is negligibly altered by small errors in ΔC_p.

All experiments were performed at a fixed initial concentration of complex (5 μM) because of the concentration dependence of the equilibrium (eqs 3 and 4). A T_m value for each complex at this concentration is obtained by fitting the CD data of the thermal scans to eqs 3 and 4 (Table 1).

Calculations of $\Delta\Delta G_D$ from Thermal Denaturation Experiments. The absolute value of ΔG_D for a dissociation reaction depends on the choice of standard states. We can simplify the problem of calculating $\Delta\Delta G_D$ by choosing a constant set of reference concentrations and measuring all the mutants relative to those. For simplicity, we use the values of 2.5 μM complex in equilibrium with 2.5 μM of each of the fragments, i.e., the concentrations at the T_m in our experiments. We can extrapolate the value of ΔG_D for each mutant to a common temperature T by inserting the values of ΔH at the $T_m(\Delta H(T_m))$ and the corresponding value of ΔS into the standard equation ($\Delta G = \Delta H - T\Delta S$) to give

$$\Delta G(T) = \Delta H(T_m)(1 - T/T_m) + \Delta C_p[(T - T_m) - T\ln(T/T_m)] \quad (5)$$

As we have pointed out previously, the errors in the quantities $\Delta H(T_m)$ and ΔC_p are attenuated by the use of this equation (Matouschek *et al.*, 1994). It is easily shown by taking partial derivatives that the error in $\Delta\Delta G_D$ due to an error of $\delta\Delta H(T_m)$ is given by

$$\delta\Delta\Delta G(T)_{er\Delta H} = (1 - T/T_m)\delta\Delta H(T_m) \quad (6)$$

that due to an error of $\delta\Delta C_p$ in ΔC_p by

$$\delta\Delta\Delta G(T)_{er\Delta C_p} = [(T - T_m) - T\ln(T/T_m)]\delta\Delta C_p \quad (7)$$

and that due to an error of δT_m in T_m by

$$\delta\Delta\Delta G(T)_{er\Delta T_m} = (\Delta H(T_m)T/(T_m)^2 - \Delta C_p(1 - T/T_m))\delta T_m \quad (8)$$

The best temperature to which to extrapolate is the mean value observed for T_m, here 312 K. By far the largest contribution to the observed error in general is $\delta\Delta\Delta G(T)_{er\Delta T_m}$. $\delta\Delta\Delta G(T)_{er\Delta C_p}$ for a ±20% error in ΔC_p is negligible, being generally <0.01 kcal mol^{-2}, and $\delta\Delta\Delta G(T)_{er\Delta H}$ is <0.05 kcal mol^{-2} for half the examples. The combined total errors are listed in Table 1.

Kinetics of Association of Mutant Fragments. Association rate constants were measured first under pseudo-first order conditions where [CI-2(60−83)] ≫ [CI-2(20−59)]. [CI-2(60−83)] is used in excess because it does not contain a tryptophan residue. The reaction has two distinguishable exponential phases; a major one which changes with mutation due to the association of fragments and a minor phase that is caused by a small proportion of fragment 20−59 being in a different conformation, probably because of *cis−trans* isomerization about a peptidyl−prolyl bond (Prat Gay *et al.*, 1994a). The kinetics of the major phase fits the binding process

$$\text{CI-2(20−59)} \underset{k_{off}}{\overset{k_{on}\text{CI-2(60−83)}}{\rightleftharpoons}} \text{[CI-2(20−59)·(60−83)]} \quad (9)$$

The standard rate law for eq 9 is

$$k_1 = k_{off} + k_{on}[\text{CI-2(60}-\text{83)}] \quad (10)$$

where k_1 is the observed first-order rate constant for formation of complex, k_{on} is the second-order rate constant, derived from the linear plot of k_1 versus [CI-2(60−83)], and k_{off} is the dissociation rate constant. The value of k_{off} is very low compared with the measured values of k_1 under accessible conditions and so is derived poor accuracy from the long extraoplation. The value of k_{on} is readily derived, however. It clearly changes with mutation (Table 2) and reflects, therefore, the main folding event. The slow phases of the mutant complexes are first-order as described for the wild type and do not change with mutation confirming the proposed mechanism of association/folding, i.e., the slow phase is not a main folding event, presumably a proline isomerization (Prat Gay *et al.*, 1994a). An average value of 0.010 ± 0.001 s^{-1} is obtained for the first-order rate constant, k_1 (Table 2). A few control experiments were performed to show that the fragments associate with the same rate constants when mixed together from solutions in guanidinium chloride.

The association rate constants may also be determined under second-order conditions (Prat Gay *et al.*, 1994a). The values of k_{on} obtained from mixing 5 μM of each fragment are identical to those determined from pseudo-first-order experiments (Table 2). Both pseudo-first- and second-order experiments give identical values.

The rate constants vary considerably. The significantly high rate constant for KA43 from both fluorescence and CD (see below) is also observed for refolding of the uncleaved mutant (Otzen *et al.*, 1994). It may result from Ala43 making a favorable interaction in the transition state, since mutation to glycine at the same position slows down the refolding of the intact protein (Otzen *et al.*, 1994).

Fluorescence versus Ellipticity Change in Fragment Association/Folding Kinetics. A protein engineering analysis of the folding of intact CI-2 (i.e., a kinetic analysis of the folding of mutants) showed a highly cooperative global folding and a single rate-determining transition state (Otzen *et al.*, 1994). The secondary and tertiary structures of wild-type [CI-2(20−59)·(60−83)] complex are also formed in parallel as shown by CD kinetics (Prat Gay *et al.*, 1994a). Similarly, the regain of secondary structure on formation of the mutant complexes measured under second-order conditions from the change of the ellipticity at 200 nm under the same conditions as used for the fluorescence equimolar experiments. Within experimental error, the rate constants from ellipticity change are identical to the regain of tertiary structure measured from fluorescence experiments (Table 2): there is a linear correlation with a slope of 0.97 ± 0.06. Secondary and tertiary structure are formed in parallel for all 22 mutants.

Kinetics of Association of the Complex and Refolding of the Intact Protein. All the parent mutants of the fragments presented in this work have been characterized kinetically and thermodynamically (Jackson *et al.*, 1993; elMasry & Fersht, 1994; Otzen *et al.*, 1994). A plot of ln k_{on} versus ln k_f (where k_f is the first-order rate constant for folding) is linear with slope of 1.1 ± 0.2 (Figure 3). Although the correlation is generally good, there are a few notable deviations: the fragments of SA31EA3EA34 and SA31 associate more slowly than expected from the rate constants for folding of their uncleaved parents, and the fragments of FL69, LA51, and VT53 associate more rapidly.

DISCUSSION

The association of fragments was studied both by isothermal equilibrium procedures and thermal denaturation. Although the standard errors calculated for $\Delta\Delta G_D$ from the measurements of K_D are the lower, there are probably hidden systematic errors when dealing with the low concentrations of reagents that are required for these experiments. The

Table 2: Rate Constants of the Association of Mutant Fragments Obtained from Fluorescence and Circular Dichroism Experiments

	pseudo-first-order		second order	
	fluorescence		fluorescence	circular dichroism
mutant	k_{on} (s^{-1} M^{-1})	k_1 (s^{-1})[a]	k_{on} (s^{-1} M^{-1})	k_{on} (s^{-1} M^{-1})
wt[b]	3900 ± 480	0.011 ± 0.001	3980 ± 180	3600 ± 270
SA31	910 ± 40	0.010 ± 0.001	1140 ± 10	1150 ± 100
SA31EA33EA34	510 ± 70	0.008 ± 0.001	700 ± 10	630 ± 80
EA33EA34	1390 ± 80	0.009 ± 0.001	1110 ± 10	1010 ± 70
EN33	1800 ± 210	0.009 ± 0.001	1680 ± 20	1430 ± 10
KA36	3550 ± 230	0.010 ± 0.001	5040 ± 110	3890 ± 390
KA37	3400 ± 90	0.010 ± 0.001	3720 ± 60	3900 ± 200
KG37	1320 ± 160	0.009 ± 0.001	1240 ± 10	1320 ± 90
KA43	7300 ± 350	0.012 ± 0.001	12010 ± 310	7750 ± 1940
LA27	ND	ND	2320 ± 100	2620 ± 470
VA38	4400 ± 730	0.012 ± 0.002	4170 ± 80	3870 ± 260
IV39	1590 ± 100	0.011 ± 0.001	1460 ± 30	1390 ± 100
IV48	2770 ± 180	0.011 ± 0.002	2870 ± 60	3290 ± 450
IA49	1610 ± 200	0.010 ± 0.001	3100 ± 60	ND
IV49	3320 ± 500	0.011 ± 0.001	3580 ± 40	2400 ± 290
VA70	1640 ± 220	0.006 ± 0.001	2490 ± 50	2330 ± 220
IV76	3590 ± 280	0.009 ± 0.001	2790 ± 40	2720 ± 170
TA22	3790 ± 530	0.009 ± 0.001	3670 ± 110	ND
VA53	3820 ± 480	0.010 ± 0.001	4610 ± 70	3940 ± 390
VT53	4130 ± 580	0.011 ± 0.001	5870 ± 90	3130 ± 380
LA51	2820 ± 470	0.012 ± 0.001	3880 ± 80	3130 ± 420
LV51	4060 ± 740	0.011 ± 0.002	5360 ± 110	4050 ± 540
FL69	2430 ± 340	0.009 ± 0.001	2900 ± 60	2850 ± 400

[a] Average value: 0.010 ± 0.001. [b] Previous values obtained were 3700 ± 300 and 0.011 ± 0.001 for k_{on} and k_1, respectively (Prat Gay *et al.*, 1994).

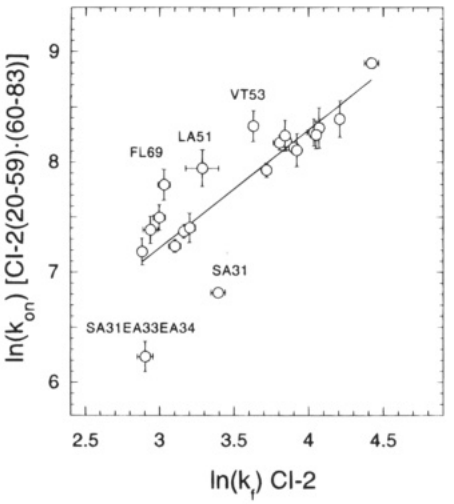

FIGURE 3: Plot of ln k_{on}, (rate constant for association of fragments) against ln k_f (rate constant for folding of the uncleaved parent CI-2).

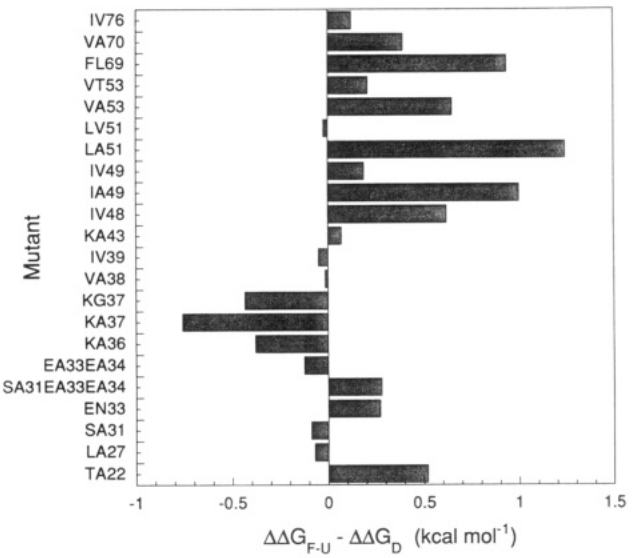

FIGURE 4: Bar chart of the differences between the change on mutation of the free energy of folding of the parent mutant ($\Delta\Delta G_{F-U}$) and the free energy of association of the mutant fragments ($\Delta\Delta G_D$).

somewhat higher errors for $\Delta\Delta G_{T_m}$ from T_m determinations are likely to be random, and so these values are less precise but possibly more accurate. Nevertheless, there is good agreement between the values of $\Delta\Delta G_D$ at 25 °C and $\Delta\Delta G_{T_m}$. Mutations that affect the stability of the intact protein also affect the dissociation constant of the complex of fragments, clearly seen on comparing the values of $\Delta\Delta G_D$, calculated from K_D, with those of $\Delta\Delta G_{F-U}$, the changes in free energy of folding of uncleaved parent on the same mutation. Most mutations are destabilizing (higher K_D), a few show little or no effect, and one, IV76, is significantly stabilizing. This increase in stability is also observed in the uncleaved parent mutant (Jackson *et al.*, 1993a). The residue Ile76 is located in a loosely packed region of the hydrophobic core; the crystal structure of the uncleaved mutant IV76 exhibits small structural rearrangements in this region (Jackson *et al.*, 1993a).

Differential Effects of Mutation on the Stabilities of Intact CI-2 and the Noncovalent Complex. Free energies of folding or association are the differences in energies between the folded and unfolded states or the associated and dissociated states, respectively. The nominally unfolded states formed by dissociation of the complexes and by the guanidinium chloride denaturation of the parent uncleaved mutants are not strictly identical, being free fragments in water at 25° C for the determination of K_D and possibly less structured, more unfolded, uncleaved protein in denaturant. There could also be differences between the dissociated fragments at 25 °C and those at higher temperatures produced in the analysis of the thermal denaturation experiments. An excellent agreement has been found, however, between $\Delta\Delta G_{F-U}$ for mutations in the hydrophobic core (Jackson et al., 1993a) and for uncharged groups in the helix (elMasry & Fersht, 1994) of intact CI-2 measured by guanidinium chloride denaturation and differential scanning calorimetry in the absence of denaturant so that the denatured state of the protein does not change radically with conditions. There are some discrepancies between the value of $\Delta\Delta G_{T_m}$ (or $\Delta\Delta G_D$) and $\Delta\Delta G_{F-U}$, and they are probably real and informative.

A chart of $\Delta\Delta G_{F-U} - \Delta\Delta G_D$ (Figure 4, where $\Delta\Delta G_D$ is the average value of the thermal and isothermal data) reveals some systematic differences. It is not known a priori whether the differences result from effects on the folded or unfolded (or dissociated) states. The structure of the folded complex has now been solved in this laboratory by NMR (B. Davis, unpublished) and X-ray crystallography (A. Buckle, unpublished) and is fully native in the regions mutated here so the differences presumably result from effects on the dissociated fragments. Further, we have presented evidence from NMR studies on the isolated fragments that these contain clusters of hydrophobic residues (Prat Gay *et al.*, 1994b; B. Davis, G. de Prat Gay, and A. R. Fersht, unpublished). Fragment (20−59) has nonnative hydrophobic clusters around residues 22−24 and 48−57 (clusters 1 and 2) and fragment (60−83), and a cluster between residues 61 and 70, around the burial of Phe69 (cluster 3). The most significant deviations where mutations destabilize the complex of the fragments less than the intact protein (greater than 0.3−0.4 kcal mol^{-1}) occur in clusters 2 and 1 in CI-2(20−59) and cluster 3 in CI-2(60−83). These correspond to residues Thr22 (β-strand 1), Leu51(β-strand 3), and Phe69 (β-strand 4). The most reasonable explanation is that these clusters are present in the fragments in water but not as much in the denatured state of the intact protein, and mutation destabilizes the fragments as well as the complex, but to a lesser extent.

Mutation of residues Lys36 and Lys37, which are part of the α-helix, destabilizes the complexes more than it does the uncleaved protein. Charged side chains are expected, however, to be affected by the ionic strength of guanidinium chloride, and so it is not expected that the energies should be the same in the denaturant and in water.

Differential Effects of Mutation on the Kinetics of Folding of Intact CI-2 and the Association of Fragments. A comparison of the structures of the transition state for folding and association by Φ-value analysis has been made elsewhere (Prat Gay et al., 1994b). In general, they are very similar except that the N-terminal region of the helix (residues 31−34) is completely formed in the transition state for association whereas it is largely formed in that for folding. For example, the triple mutation SA31EA33EA34 destabilizes the transition state of association by the same energy as the folded complex. Here, we are concerned with some of the differ-

ences. The structures of transition states can be studied only by kinetics. The kinetics depends on the energy difference between the transition and ground states. A plot of ln k_{on} (rate constant for association) against ln k_f (rate constant for folding of uncleaved parent mutant) shows a strong correlation (Figure 3) with slope 1.1 ± 0.2, but with a few deviations. The most notable deviations are for the mutants involving Ser31, whose rate of association is below the regression line. Ser31 is the N-cap of the helix and its -OH group stabilizes the N-terminal region. The equilibrium free energies of destabilization on mutation of Ser31 → Ala are the same, within error, for the noncovalent complexes and the intact proteins. The slower rate of refolding reflects the complete formation of the helix in the transition state for association compared with the incomplete formation in that for folding. The transition state for association is thus more sensitive to destabilization. Mutants of Leu51, Val53, and Phe69 have the largest positive deviations from the regression line; that is, the kinetics of association is less affected by mutation than is that for folding. These residues are ones that have been implicated in forming non-native structures (hydrophobic clusters) in the fragments and show large deviations between $\Delta\Delta G_{F\text{-}U}$ and $\Delta\Delta G_D$ (Figure 4). The most likely explanation is that as the side chains of Leu51, Val53, and Phe69 stabilize the ground states of the fragments, mutation of them destabilizes the ground states and so lowers the activation energy for association. Another possible explanation, for which there is no supporting evidence, is that these mutations cause more favorable interactions in the transition state of the complex because of its looser structure compared with the transition state for folding of the intact protein.

The differences between the folding of intact protein and the association of fragments do not result from the effects of guanidinium chloride. The rate constants for refolding of intact protein are the same when measured in the absence of guanidinium chloride (for refolding from the the acid-denatured state) and for the guanidinium chloride denatured state when extrapolated to the absence of denaturant. Further, the same rate rate constants are found for the association of fragments when diluted from guanidinium chloride solution and when mixed together in simple aqueous solution. The strong correlation between ln k_{on} and ln k_f for most of the mutants shows that the structures of the transition states of association and folding are in general very similar. The effects of mutation on the residues that form hydrophobic clusters, both on the equilibria and kinetics, suggest that these clusters are less important in the denatured state in guanidinium chloride or at higher temperatures in equilibrium studies or in the denatured state of the intact protein from which folding is initiated in kinetic studies in water.

It is noteworthy that two quite different starting points lead to formation of native (or native-like) structure through a similar transition state; either starting from two separate fragments of the protein that have nonnative structure or from intact denatured protein.

ACKNOWLEDGMENT

We thank Laura Itzhaki, Nadia elMasry, and Sophie Jackson for CI-2 mutant proteins. This work was partly supported by a grant of the European Community from its Biotechnology Programme.

REFERENCES

Becktel, W. J., & Schellman, J. A. (1987) *Biopolymers 26*, 1859−1877.
Dill, K. A., & Shortle, D. (1991) *Annu. Rev. Biochem. 60*, 795−825.
Dyson, H. J., & Wright, P. E. (1993) *Curr. Opin. Struct. Biol. 3*, 60−65.
Dyson, H. J., Merutka, G., Waltho, J. P., Lerner, R. A., & Wright, P. E. (1992) *J. Mol. Biol. 226*, 795−817.
Edelhoch, H., & Osborne, J. C., Jr. (1976) *Adv. Protein Chem. 30*, 183−250.
elMasry, N. F., & Fersht, A. R. (1994) *Protein Eng. 7*, 777−782.
Fersht, A. R. (1993) *FEBS Lett. 325*, 5−16.
Fersht, A. R., Matouschek, A., & Serrano, L. (1992) *J. Mol. Biol. 224*, 771−782.
Fersht, A. R., Itzhaki, L. S., elMasry, N., Matthews, J. M., & Otzen, D. E. (1994) *Proc. Natl. Acad. Sci. U.S.A.* 10426−10429.
Jackson, S. E., & Fersht, A. R. (1991) *Biochemistry 30*, 10428−10435.
Jackson, S. E., Moracci, M., elMasry, N., Johnson, C. M., & Fersht, A. R. (1993) *Biochemistry 32*, 11259−11269.
Karplus, M., & Weaver, D. L. (1976) *Nature 260*, 404.
Kim, P. S., & Baldwin, R. L. (1990) *Annu. Rev. Biochem. 59*, 631−660.
Li, S., & Daggett, V. (1994) *Proc. Natl. Acad. Sci. U.S.A.* (in press).
Ludvigsen, S, Shen, H., Kjær, M., Madsen, J. C., & Pouslen, F. M. (1991) *J. Mol. Biol. 222*, 621−635.
Marky, L. A., & Breslauer, K. J. (1987) *Biopolymers 26*, 1601−1620.
Matouschek, A., Kellis, J. T., Jr., Serrano, L., & Fersht, A. R. (1989) *Nature 342*, 122−126.
Matouschek, A., Matthews, J. M., Johnson, C. M., & Fersht, A. R. (1994) *Protein Eng. 7*, 1089−1095.
McPhalen, C. A., & James, M. N. G. (1987) *Biochemistry 26*, 261−269.
Moult, J., & Unger, R. (1991) *Biochemistry 30*, 1591−1594.
Otzen, D., Itzhaki, L., elMasry, N., Jackson, S. E., & Fersht, A. R. (1994) *Proc. Natl. Acad. Sci. U.S.A.* 10422−10425.
Prat Gay, G. de, & Fersht, A. R. (1994) *Biochemistry 33*, 7957−7963.
Prat Gay, G. de, Ruiz-Sanz, J., & Fersht, A. R. (1994a) *Biochemistry 33*, 7964−7970.
Prat Gay, G. de, Ruiz-Sanz, J., Davis, B., & Fersht, A. R. (1994b) *Proc. Natl. Acad. Sci. U.S.A.* 10943−10946.
Ptitsyn, O. B. (1973) *Dokl. Akad. Nauk 210*, 1213−1215.
Sâli, A., Shakhnovich, E., & Karplus, M. (1994) *Nature* 369, 248−251.
Sancho, J., Neira, J. L., & Fersht, A. R. (1992) *J. Mol. Biol. 224*, 749−758.
Shortle, D., & Abeygunawardana, C. (1993) *Structure 1*, 121−134.
Steif, C., Weber, P., & Hinz, H.-J. (1993) *Biochemistry 32*, 3867−3876.
Waltho, J. P., Feher, V. A., Merutka, G., Dyson, H. J., & Wright, P. E. (1993) *Biochemistry 32*, 6337−6347.

BI942206G