Acquisition of Native β-Strand Topology During the Rapid Collapse Phase of Protein Folding†

Martin J. Parker,* Christopher E. Dempsey, Mark Lorch, and Anthony R. Clarke

Department of Biochemistry, University of Bristol, School of Medical Sciences, University Walk, Bristol BS8 1TD, U.K.

Received May 30, 1997; Revised Manuscript Received August 21, 1997⊗

ABSTRACT: The 98 residue C-terminal domain of the cell-surface receptor protein CD2 (CD2.D1) has a β-sandwich fold belonging to the immunoglobulin superfamily but lacking the usual disulfide bridges. Kinetic studies on the folding/unfolding of CD2.D1 reveal that folding proceeds through a rapidly formed intermediate state [Parker, M. J., & Clarke, A. R. (1997) *Biochemistry* 36, 5786−5794]. To characterize the structural properties of this intermediate we have performed a series of amide hydrogen exchange studies using the pH competition method, in which folding and exchange are initiated simultaneously. The complex β-sheet topology of this molecule makes it an ideal object for examining the acquisition of backbone hydrogen bonds made between sequence-local and sequence-distant segments of the chain during folding. The pattern of protected amides in the intermediate reveal that the essential features of the β-sheet topology of CD2.D1 are defined early in the folding pathway, before the development of intimate side chain interactions characteristic of the native state. The results are discussed in light of current issues concerning the mechanistic relevance of kinetic protein folding intermediates.

The native topology of a protein molecule is not secured by a random and exhaustive search of conformational space. Rather, folding is directed along thermodynamically defined pathways where, in most cases, intermediate states transiently accumulate. Such intermediates have been identified in the folding reactions of a number of different proteins. In general, these intermediates are formed rapidly (within the dead time of conventional stopped-flow mixing apparatus; <1 ms), are compact, and contain extensive secondary structure. They do, however, lack the intimate tertiary side chain contacts characteristic of the native state (Kim & Baldwin, 1990; Matthews, 1993; Ptitsyn, 1995; Miranker & Dobson, 1996; Roder & Colón, 1997).

Many hold the view that the formation of such states is an essential step in the folding process, i.e., they are on-pathway species with essentially native-like contacts which direct the chain toward the native fold. Lattice model simulations of heteropolymer organization, on the other hand, have suggested that these globally collapsed states may possess many stable, non-native contacts and that, consequently, the rate-limiting step in folding may involve "undoing" some of these interactions, i.e., it is enthalpic as well as entropic in origin, in terms of intraprotein interactions (Sali *et al*., 1994; Dill *et al*., 1995; Bryngelson *et al*., 1995). The choice between these views depends, to a large extent, upon the exclusivity of interactions required to generate these species. For instance, while many configurations in a lattice model may satisfy a particular number of hydrophobic residue contacts, in a real protein not all states within such an ensemble will be able to fulfill other energetic criteria, such as main chain torsion angle preferences and/or backbone hydrogen bonding.

Resolution of these issues requires detailed knowledge of the structural and energetic properties of protein folding intermediates. The elusive nature of these species, however, poses a demanding problem for the experimentalist. In the main, two experimental approaches offer the greatest potential to resolve the properties of kinetic intermediates. The first takes advantage of the ability to create and interpret the behavior of mutants, providing information on side chain interactions (Fersht *et al*., 1992; Horovitz, 1996). The second combines amide exchange protection studies with NMR (Englander & Kallenbach, 1984; Schmid & Baldwin, 1979; Roder & Wüthrich, 1986; Udgaonkar & Baldwin, 1988; Roder *et al*., 1988), and more recently with electrospray mass spectroscopy (Miranker *et al*., 1993), to provide information on backbone hydrogen bonding and/or burial of the main chain in a protected core. Of particular interest in these studies is the identification of both short- and long-sequence-range interactions in these rapidly collapsed states and the way in which these interactions effectively cooperate to reduce conformational freedom. This is especially important with regard to β-sheet formation, as the interactions responsible for maintaining β-sheets involve residues which are clearly either near or distant in the sequence.

One protein with the potential to serve as a paradigm for the folding of all-β proteins is the 98 residue, C-terminal domain of rat CD2 (denoted CD2.D1;[1] Driscoll *et al*., 1991), which has been shown to fold through a rapidly formed intermediate state (Parker & Clarke, 1997). The topology of CD2.D1 is that of an immunoglobulin variable domain but, unusually, it contains no disulfide bond (Driscoll *et al*.,

† This work was supported by a project grant from the B.B.S.R.C. (U.K.) and equipment funding from the Wellcome Trust. A.R.C. is a Lister Institute research fellow.

* Author to whom correspondence should be addressed.

⊗ Abstract published in *Advance ACS Abstracts,* October 1, 1997.

[1] Abbreviations: CD2.D1, 98-residue C-terminal domain of rat CD2; DQF-COSY, double quantum filter correlated spectroscopy; GuHCl, guanidine hydrochloride; GST, glutathione *S*-transferase; HSQC, heteronuclear single quantum coherence; IPTG, isopropyl β-D-thiogalactoside; NATA, *N*-acetyltryptophanamide; NAYA, *N*-acetyltyrosineamide; NMR, nuclear magnetic resonance; NOESY, nuclear Overhauser spectroscopy; PDLA, poly-D,L-alanine; PDLK, poly-D,L-lysine; pH_m, measured pH; PLE, poly-L-glutamate.

S0006-2960(97)01294-4 CCC: $14.00

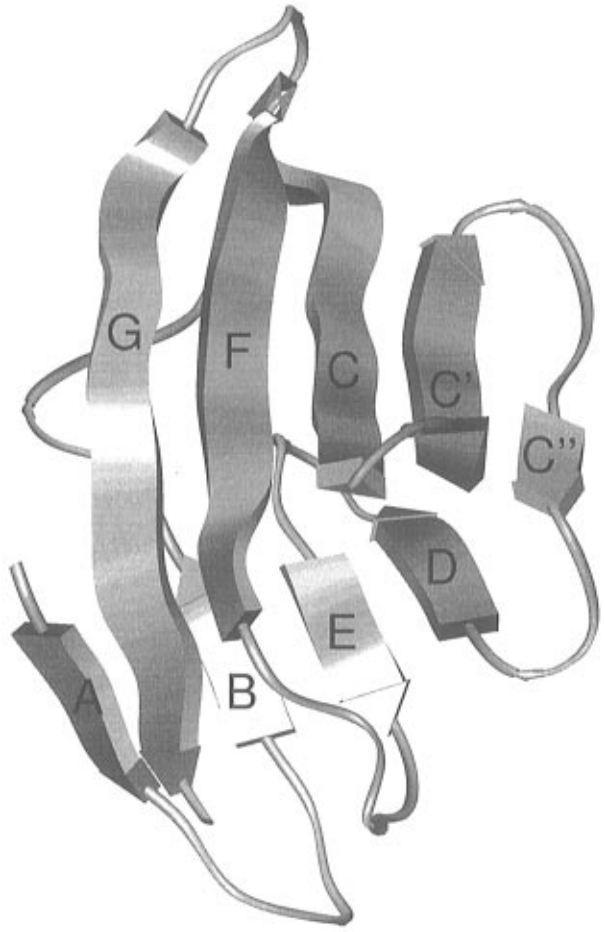

FIGURE 1: Crystal structure of CD2.D1 (Jones *et al*., 1992). The topology of CD2.D1 is that of an immunoglobulin variable (V) domain with two β-sheets of strands conventionally labeled DEB and AGFCC′C″. The positions of the β-strands (large arrows) and the β-turns (small arrows) are based on the DSSP program (Kabsch & Sander, 1983). Despite the absence of the usual disulfide bond, the positions of all the β-strands in the V domain framework are highly conserved (Jones *et al*., 1992).

1991; Jones *et al*., 1992) (Figure 1). This protein is particularly well-suited for the experimental purpose owing to its rather perverse β-strand topology. The molecule has nine strands with four well-defined hairpin structures (CC′, C′C″, DE, and FG) and three strand−strand interactions which are distant in sequence: AG (parallel) and BE and CF (both antiparallel). This pattern of bonding defines a rather complicated and unintuitive fold in which the chain crosses between the sheets three times.

The nature of this β-sheet topology means that we can examine the development of both short-range (hairpin) and long-range (tertiary) interactions as both will be reported by backbone hydrogen bond formation which can, in turn, be detected by the inhibition of amide hydrogen exchange with the solvent. Tertiary interactions between helices or between a helix and a strand cannot be detected in this way owing to the absence of such hydrogen bonding networks between the participating elements. Also, in globular $\alpha\beta$ proteins, attributing the protection of core β-strand amides in folding intermediates to the formation of independently stable β-structure is difficult to address as the side chains of these groups may form interactions with the side chains of α-helices, which generally exhibit a moderate level of independent stability themselves. Another advantage of this β-sandwich structure is that none of the backbone groups is deeply buried in the hydrophobic core in the native state. Hence, any protection measured for amides is likely to reflect the formation of amide-carbonyl hydrogen bonds rather than exclusion from the solvent by core burial.

METHODS

Experimental

Source of Protein. Domain 1 of rat CD2 (residues 1−98; CD2.D1) was expressed in *Escherichia coli* (HB2151 strain; Pharmacia) as a fusion protein with glutathione *S*-transferase (GST) using the pGEX-2T gene fusion vector (Pharmacia) (Driscoll *et al*., 1991). To produce ^{15}N-labeled CD2.D1, cells were grown overnight in M9 minimal media [42 mM Na_2HPO_4, 22 mM KH_2PO_4, 9 mM NaCl, 2 mM $MgSO_4$, 0.2% glucose, 2 μg mL^{-1} thiamine (Sigma)], containing 10 mM $^{15}NH_4Cl$ (Sigma) and 0.5 mg mL^{-1} ampicillin (Beecham Research), at 37 °C. The cells were then diluted 1:10 in fresh media and grown at 37 °C to mid-log phase. Isopropyl β-D-thiogalactoside (IPTG; Sigma) was then added to 1 mM, and the cells were grown for an additional 2−3 h. The protein was then purified as described previously (Parker & Clarke, 1997).

Protein concentrations were estimated by UV absorption of aromatic residues at 280 nm [ϵ = 5500 M^{-1} cm^{-1} for tryptophan (2 residues) and 1100 M^{-1} cm^{-1} for tyrosine (2 residues)].

Equilibrium and Kinetic Folding/Unfolding Measurements. The guanidine hydrochloride (GuHCl)-induced equilibrium unfolding profiles were measured by fluorescence spectroscopy, as described in Parker and Clarke (1997). The GuHCl-dependent folding and unfolding rates were measured by fluorescence stopped-flow, as described in Parker & Clarke (1997). All measurements were made at 25 °C. For both equilibrium and kinetic measurements a mixed buffer solution containing 25 mM sodium borate and 25 mM NaH_2PO_4 (Sigma) was used throughout at the appropriate pH. Sodium sulfate (Na_2SO_4; Sigma) was added to both protein and GuHCl solutions at the appropriate concentration. Ultrapure GuHCl was purchased from Sigma.

Solubility Measurements. Saturated solutions of *N*-acetyltryptophanamide (NATA) and *N*-acetyltyrosineamide (NAYA) (Sigma) were made up in H_2O, containing appropriate concentrations of GuHCl and Na_2SO_4. The solutions were sonicated and then incubated for 48 h at 25 °C with vigorous shaking, to allow complete equilibration. Insoluble material was then removed by centrifugation. The concentration of soluble NATA and NAYA in each sample was then assessed by measuring the absorbance at 280 nm at 25 °C in a Perkin Elmer spectrophotometer.

pH Dependent Competition between Exchange and Folding. A 0.5 mL amount of a 2 mM sample of ^{15}N-labeled CD2.D1 in H_2O, containing 2.8 M GuHCl, was mixed against 8.3 vol of D_2O, containing 25 mM sodium borate and 25 mM NaH_2PO_4, at the appropriate pH_m (measured pH, i.e., not corrected for the glass electrode effect), and 0.448 M Na_2SO_4, at 25 °C in a custom-built, rapid mixing device with a dead time of 2−3 ms (gives final concentrations of Na_2SO_4 and GuHCl of 0.4 and 0.3 M, respectively). Exchange was then quenched by the rapid addition (delay time = 2 s) of 6 mL of ice-cold D_2O containing 100 mM sodium d_3-acetate (Sigma), pH_m 4.0. The sample was then concentrated down to a volume of 0.5 mL at 4 °C, using a 10 mL Amicon Ultrafiltration cell with a 3K molecular

weight cutoff filter. The pH of the breakthrough solution was monitored to ensure that the pH_m of the sample did not exceed 4.0. NMR measurements were performed immediately upon concentration.

Intrinsic Amide Exchange Rates. Changes in the absorbance at 220 nm of poly-D,L-alanine (PDLA), poly-D,L-lysine (PDLK), and poly-L-glutamate (PLE), upon exchange of the amide protons for deuterons, was used to estimate the effect of 0.4 M Na_2SO_4 and 0.3 M GuHCl on the intrinsic rate of exchange of peptide groups, as described by Englander *et al*. (1979). A 10 mg mL^{-1} solution of PDLA, PDLK, or PLE (Sigma) in H_2O was mixed against 10 vol of D_2O, containing 25 mM sodium borate and 25 mM NaH_2PO_4, at the appropriate pH_m, with or without 0.44 M Na_2SO_4 and 0.33 M GuHCl, at 20 °C in an Applied Photophysics stopped-flow apparatus. A wavelength of 220 nm was selected by a single monochromator (slit width 5 nm) from a mercury−xenon light source. Rate constants were determined from an average of 10−20 absorbance transients (average error in rate constants does not exceed 10%). To correct for the buffering potential of these polyamino acids, identical postmix solutions were made up and the pH was measured using a sensitive glass electrode.

All reaction solutions were maintained at the appropriate temperature using thermostated circulating water baths.

NMR Spectroscopy. NMR spectra were obtained using a 500 MHz JEOL α spectrometer operating in the phase-sensitive mode with quadrature detection in both dimensions (States *et al*., 1982). Presaturation of the water signal was achieved using a DANTE pulse train (Morris & Freeman, 1978). $^1H-^{15}N$ HSQC spectra (Bax *et al*., 1990) were acquired with 2048 real points in t_2 (5000 Hz spectral width) with 256 t_1 increments (3000 Hz spectral width) for high-resolution spectra used for spectral assignment, or 64 t_1 increments for cross peak intensity measurements in exchange samples. For exchange spectra the peak intensities were normalized for protein concentration by comparing intensities of resonances in the aliphatic region in high-resolution ^1H-NMR spectra. For the assignments, homonuclear ^1H 2D spectra (2048 real t_2 points over 5000 Hz spectral width; 512 t_1 increments) under conditions of partial exchange were made using NOESY (100 ms mixing time; Kumar *et al*., 1980) and DQF-COSY (Rance *et al*., 1983) at 25 °C on samples of CD2.D1 dissolved directly into the D_2O buffer used for recording spectra of exchange samples (see above). NMR data were zero filled to generate matrices of 2048 × 2048 points (homonuclear 2D spectra) or 2048 × 1024 or 256 points (high-resolution and low-resolution HSQC, respectively) and processed using FELIX 95.

The homonuclear experiments select the set of amides that are sufficiently protected in the native state (average rates of exchange, $k_{ex} < 0.001$ min^{-1} under these conditions) for use as probes of exchange protection during folding. Assignment of these amides was made in conjunction with the chemical shift data (residue-specific NH and CHα ^1H chemical shifts and NH ^{15}N chemical shifts) previously obtained during the NMR structure determination of CD2.D1 (Driscoll *et al*., 1991; Dr. P. C. Driscoll, personal communication). Many of the cross peaks could be assigned directly by chemical shift comparisons since the conditions of spectral acquisition were similar to those used by Driscoll *et al*. (1991). Largely due to the residual Na_2SO_4 and GuHCl in exchange samples, some differences in amide chemical shifts were observed, associated generally with charged residues and their immediate neighbors. Unambiguous assignment of most of the stable amide signals remaining in NOESY and DQF-COSY spectra after dissolving in D_2O buffer was achieved from the retention of networks of NOE interactions associated with β-sheet in the stable core (NH_i-$CH\alpha_i$; $NH_iCH\alpha_{i-1}$), making additional use of the discrete set of interstrand $CH\alpha_i-CH\alpha_j$ NOE's to confirm many CHα chemical shifts under the experimental conditions employed in this study. Of the 45 correlation peaks observed in the $^1H-^{15}N$ HSQC spectrum obtained under conditions of maximal proton retention (folding at pH_m 6.0; see Results), 42 could be unambiguously assigned to specific residues, and these are identified by annotations in the spectra of Figure 4.

Analytical Procedures

Treatment of Solubility Data. The difference in the free energy of solvation of NAYA and NATA, between water and a given concentration of GuHCl and/or Na_2SO_4 (ΔG_s) is given by eq 1:

$$\Delta G_s = -RT \ln(A_{280}/A_{280(w)}) \quad (1)$$

where A_{280} is the absorbance at 280 nm measured at a particular concentration of GuHCl and Na_2SO_4 and $A_{280(w)}$ is that measured in water. For a given, constant concentration of Na_2SO_4, data of ΔG_s *versus* GuHCl concentration ([GuHCl]), for both NAYA and NATA, were fitted to the hyperbolic relationship

$$\Delta G_s = ((\Delta G_{s,max} - \Delta G_{s,0})[\text{GuHCl}]/(K_{den} + [\text{GuHCl}])) + \Delta G_{s,0} \quad (2)$$

where $\Delta G_{s,0}$ is the free energy change of solvation, relative to water, at a specified concentration of Na_2SO_4 and in the absence of GuHCl, and $\Delta G_{s,max}$ is the notional, maximum change in the free energy change of solvation at an infinite GuHCl concentration and a specified concentration of Na_2SO_4, measured relative to water. K_{den} is a denaturation contant, measured in molar terms, and represents the concentration of GuHCl required to reach $(\Delta G_{s,max} + \Delta G_{s,0})/2$.

Denaturant Activity. For the analysis of equilibrium and kinetic data in the absence of Na_2SO_4 [GuHCl] is converted to molar denaturant activity (D), according to the relationship

$$D = C_{0.5}[\text{GuHCl}]/(C_{0.5} + [\text{GuHCl}]) \quad (3)$$

where $C_{0.5}$ is a denaturation constant with the value 7.5 M [see Parker *et al*. (1995)]. For a given, constant concentration of Na_2SO_4, [GuHCl] is converted to molar denaturant activity according to the relationship

$$D = (C_{0.5}[\text{GuHCl}]/(C_{0.5} + [\text{GuHCl}])) + C_{0.5}\Delta G_{s,0}/(\Delta G_{s,max} - \Delta G_{s,0}) \quad (4)$$

In 0−0.4 M Na_2SO_4, the values of $\Delta G_{s,0}/(\Delta G_{s,max} - \Delta G_{s,0})$ calculated for NAYA and NATA are the same, within error (see Results and legend to Figure 3).

Treatment of Equilibrium Data. Equilibrium fluorescence profiles were fitted to the equation

$$I = \alpha_F I_F + \alpha_U I_U \quad (5)$$

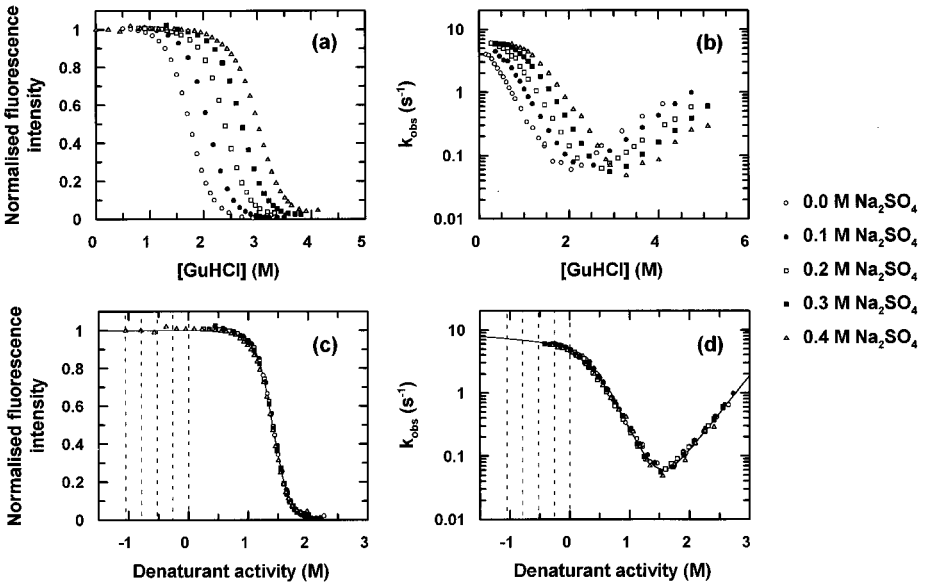

FIGURE 2: Folding dynamics of CD2.D1. (a) GuHCl-induced equilibrium unfolding profiles obtained for CD2.D1 in 0, 0.1, 0.2, 0.3, and 0.4 M Na_2SO_4 (pH 7; 25 °C), measured by tryptophan fluorescence (see Experimental Procedures). For comparison, the data have been normalized for protein concentration. No base line corrections need to be applied to these data. (b) GuHCl dependences of the folding and unfolding rates (k_{obs})—which are both first-order relaxation processes (Parker & Clarke, 1997)—in 0, 0.1, 0.2, 0.3, and 0.4 M Na_2SO_4 (pH 7; 25 °C). (c, d) Equilibrium and kinetic data plotted against denaturant activity, calculated using eq 4 with the values derived from the solubility data in Figure 3 (see Results). As can be seen, when plotted against denaturant activity, the data directly superimpose. The combined equilibrium data set in c has been fitted to eq 5 (fit shown as a continuous curve), which yields values for $K_{F/U(w)}$ and m_U of $(2.7 \pm 0.4) \times 10^4$ and -7.26 ± 0.10 M^{-1}, respectively. The combined kinetic data set in d has been fitted to eq 6 (fit shown as a continuous curve). This yields values for $k_{I-F(w)}$, $k_{F-I(w)}$, and $K_{I/U(w)}$ of 6.0 ± 0.3 s^{-1}, $(5.0 \pm 1.0) \times 10^{-4}$ s^{-1} and 3.6 ± 0.6, respectively, and for m_U, m_I, and m_t of -7.08 ± 0.10 M^{-1}, -2.91 ± 0.17 M^{-1}, and -2.72 ± 0.06 M^{-1}, respectively. The denaturant activities calculated for 0, 0.1, 0.2, 0.3, and 0.4 M Na_2SO_4 are marked as dashed lines in c and d (see legend to Figure 3).

with the following temporary variables:

$$K_{F/U} = K_{F/U(w)} \exp(m_U D)$$

$$\alpha_F = K_{F/U}/(1 + K_{F/U})$$

$$\alpha_U = 1 - \alpha_F$$

where α_F and α_U are the fractional populations of molecules in the folded (F) and unfolded (U) states, respectively; I, I_F, and I_U are the fluorescence intensities (measured, folded, and unfolded), $K_{F/U}$ is the equilibrium constant (F/U) at a given denaturant activity (D), $K_{F/U(w)}$ is this equilibrium constant at $D = 0$, and m_U describes the exponential reduction in $K_{F/U}$ as a function of D: it has units of M^{-1} (Parker *et al*., 1995).

Treatment of Kinetic Data. The analysis of the fluorescence folding and unfolding transients for CD2.D1, which are single, first-order processes, has been described elsewhere (Parker & Clarke 1997). For the intrinsic amide exchange reactions, transients of absorbance (A) *versus* time were fitted to the equation $A = A_a \exp(-k_{ex}t) + A_f$ (where A_a is the absorbance change of the reaction, k_{ex} is the first-order rate constant for exchange, and A_f is the final absorbance). The k_{ex} versus pH_m data were fitted to the equation $\log_{10} k_{ex} = m(pH_m) + c$.

The rate profiles (observed rate constant for folding/unfolding (k_{obs}) *versus* denaturant activity) were fitted to the following equation:

$$k_{obs} = k_{F-I} + k_{I-F}/(1 + 1/K_{I/U}) \quad (6)$$

[see Parker *et al.* (1995) and Parker and Clarke (1997)] where k_{I-F} and k_{F-I} are rate constants describing the forward and reverse transitions, respectively, between the folded and intermediate (I) states and $K_{I/U}$ is the equilibrium constant for the rapid interconversion of the intermediate and unfolded states. In the fitting routine the following temporary variables were used:

$$k_{F-I} = k_{F-I(w)} \exp(-m_t D)$$

$$k_{I-F} = k_{I-F(w)} \exp((m_I - m_t)D)$$

$$K_{I/U} = K_{I/U(w)} \exp((m_U - m_I)D)$$

where the "m" parameters describe the shifts in the stabilities of each state (designated by the subscript) as a function of D and have units of M^{-1} (the subscript "t" denotes the transition state associated with the rate-limiting I to F transition). Note that these values are measured relative to the folded state, i.e., $m_F = 0$.

Treatment of Protection Data. For the calculation of protection factors (P) in the rapidly formed folding intermediate of CD2.D1, the pH dependence of the amide proton intensities (I_{NH}), measured by volume integration of peaks in the $^1H-^{15}N$ HSQC spectra (see above), were fitted to the following equation:

$$I_{NH} = [I_{NHA}(k_{I-F}P/(k_{I-F}P + k_{int})] + I_{NH0} \quad (7)$$

where I_{NHA} is the change in intensity associated with a proton occupancy change of 1 to 0 and I_{NH0} is the baseline intensity (i.e., intensity arising from 0 occupancy). k_{I-F} is the protection rate associated with the rate-limiting I to F folding transition, and k_{int} is the pH-dependent intrinsic rate of exchange, expected for an amide in a random coil polypeptide, taken from Bai *et al.* (1993) (see also Results). Presented data are normalized using the calculated values of I_{NHA} and I_{NH0}.

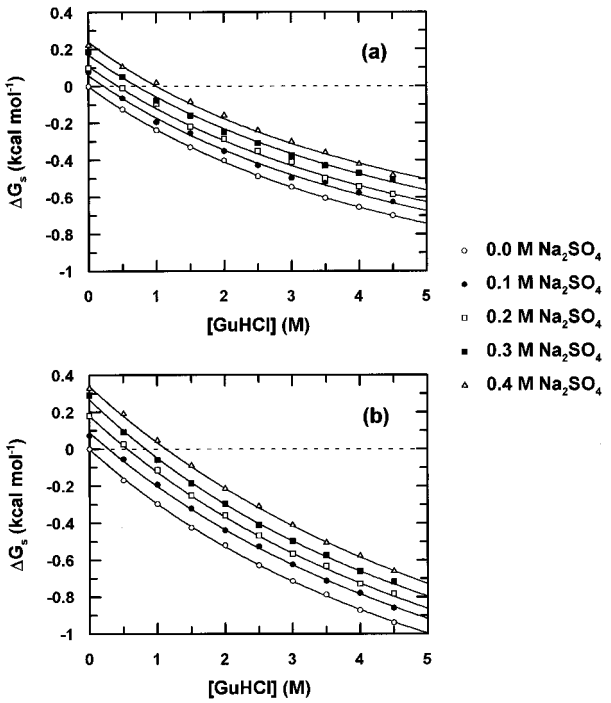

FIGURE 3: GuHCl dependence of hydrocarbon solubility in the presence of Na_2SO_4. The GuHCl dependence of the free energy change of solvation (ΔG_s; see eq 1) at 25 °C in 0, 0.1, 0.2, 0.3, and 0.4 M Na_2SO_4 are shown for NAYA and NATA in a and b, respectively (see Experimental Procedures). These data have been fitted to the hyperbolic relationship given by eq 2 to yield values for $\Delta G_{s,max}$, $\Delta G_{s,0}$, and K_{den} (see Analytical Procedures and Results for explanation of parameters). The fits to the data are shown as continuous curves. The calculated values for $\Delta G_{s,0}$ in 0, 0.1, 0.2, 0.3, and 0.4 M Na_2SO_4 are 0, 0.06, 0.11, 0.17, and 0.24 kcal mol^{-1} for NAYA, respectively, and 0, 0.09, 0.18, 0.27, and 0.34 kcal mol^{-1} for NATA, respectively (average error, 10%). The calculated values for K_{den} are 5.91, 5.85, 6.02, 6.20, and 5.51 M for NAYA, respectively, and 7.19, 7.81, 7.57, 7.01, and 8.25 M for NATA, respectively (average error, 15%). The calculated values for $\Delta G_{s,max} - \Delta G_{s,0}$ are −1.62, −1.59, −1.62, −1.64, and −1.56 kcal mol^{-1} for NAYA, respectively, and −2.43, −2.58, −2.62, −2.56, and −2.83 kcal mol^{-1} for NATA, respectively (average error, 10%). For a given concentration of Na_2SO_4, the values of $\Delta G_{s,0}/(\Delta G_{s,max} - \Delta G_{s,0})$ are the same, within error, for NAYA and NATA: calculated values for NAYA in 0, 0.1, 0.2, 0.3, and 0.4 M Na_2SO_4 are 0, 0.038, 0.068, 0.104, and 0.154, respectively, and those for NATA are 0, 0.035, 0.069, 0.105, and 0.120, respectively (average error, 20%). Using an average of these values in eq 4 yields values for the denaturant activities of 0, 0.1, 0.2, 0.3, and 0.4 M Na_2SO_4 of 0, −0.27, −0.51, −0.78, and −1.03 M, respectively.

All data were fitted using the Grafit analysis software (Erithracus Software, U.K.). When kinetic data were fitted to eq 3 proportional weighting was used, so that the fitted values took account of rate constants equally across the whole range.

RESULTS

Folding Dynamics of CD2.D1

The GuHCl-induced fluorescence equilibrium unfolding profile of CD2.D1 (pH 7.0; 25 °C) is shown in Figure 2a. This profile represents a single structural transition, i.e., the only species populated in the equilibrated system are the fully folded (F) and unfolded (U) states. However, the GuHCl dependence of the folding and unfolding rates (k_{obs}), which are both first-order processes [see Parker and Clarke (1997)], reveals the presence of a rapidly formed intermediate state (I) which is only populated in strong folding conditions (Figure 2b). There is no measurable change in tryptophan fluorescence associated with the U to I transition of CD2.D1. To obtain a measurable signal change for this transition we have labeled the N-terminal amino group of CD2.D1 with dansyl sulfonyl chloride. This modification does not noticeably perturb the energetics of folding/unfolding as the rates obtained by stopped-flow fluorescence are indistinguishable from those obtained for unlabeled CD2.D1. Three distinct tryptophan signals can be resolved for the labeled protein, due to the different degrees by which the dansyl group quenches the trypophan fluorescence in U, I, and F. The burst phase amplitude, associated with the U to I transition, occurs in the dead time of the stopped-flow (<1 ms) and the denaturant dependence of this amplitude produces an equilibrium profile that yields values for $K_{I/U(w)}$ and $m_U - m_I$ which are in accordance with those obtained from the analysis of the rate profile data (M. J. Parker and A. R. Clarke, unpublished data). The mechanism of folding of CD2.D1 is described in Scheme 1 (Parker & Clarke, 1997),

Scheme 1

$$U \underset{}{\overset{K_{I/U}}{\rightleftharpoons}} I \overset{k_{I-F}}{\underset{k_{F-I}}{\rightleftharpoons}} F$$

where $K_{I/U}$ is the equilibrium constant for the rapid U to I transition and k_{I-F} and k_{F-I} are the folding and unfolding rates, respectively, associated with the rate-limiting I to F transition. The denaturant dependence of the relaxation kinetics for folding/unfolding allows deduction of six physical parameters: $K_{I/U(w)}$, $k_{I-F(w)}$, $k_{F-I(w)}$, m_U, m_I, and m_t (Parker *et al*., 1995; see also Analytical Procedures). The first three parameters describe rate and equilibrium constants in water, while the latter three are a measure of how the free energies of U, I, and t (I to F transition state) vary as a function of the solvent conditions, respectively, and are measured relative to the folded state. These *m*-values provide a qualitative measure of the degree to which hydrocarbon is solvated in each state, relative to the folded state (Shortle *et al*., 1988; Staniforth *et al*., 1993; Myers *et al*., 1995; Parker *et al*., 1995; Parker & Clarke, 1997).

It is important to note that the kinetic data in Figure 2 can be equally well modeled by a reaction involving an off-pathway intermediate, i.e., I ⇌ U ⇌ F (Baldwin, 1996). The two solutions for these extreme cases are $k_{obs} = k_{F-I} + k_{I-F}/(1 + 1/K_{I/U})$ for an on-pathway intermediate and $k_{obs} = k_{F-U} + k_{U-F}/(1 + K_{I/U})$ for an off-pathway intermediate; the kinetic data do not distinguish between these mechanisms. Nonetheless, the value of $K_{I/U}$ and the net rate at which the intermediate converts to the folded state are model independent [i.e., k_{I-F} (on-pathway) = $k_{U-F}/K_{I/U}$ (off-pathway)], and the structural and energetic properties of the transiently populated I-state are those which we wish to elucidate.

It is well established that there is a nonlinear relationship between denaturant concentration and both the free energy of solvation of protein hydrocarbon moieties (Tanford, 1970; Pace, 1975; Staniforth *et al*., 1993) and the free energy change of protein unfolding (Johnson & Fersht, 1995). To correct for this nonlinearity, in the analysis of equilibrium and kinetic data, GuHCl concentration ([GuHCl]) is converted to denaturant activity (*D*), according to eq 3 [see Parker *et al.* (1995) and Parker and Clarke (1997)]. This linearized scale allows more reliable extrapolations to conditions where denaturant is absent. For example, the use of this denaturant activity scale for analyzing the kinetic data

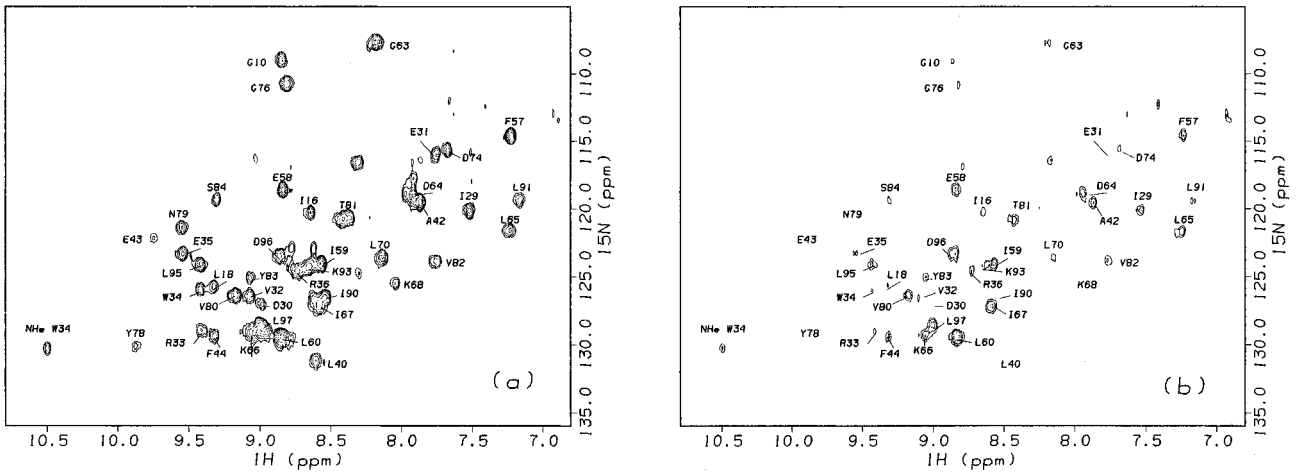

FIGURE 4: ^1H-^{15}N HSQC exchange spectra. (a, b) Representative $^1H-^{15}N$ HSQC spectra acquired for CD2.D1 samples after folding/exchange reactions performed at pH_m 6.0 and 8.0, respectively (see Experimental Procedures). For comparison, the peak intensities in a and b have been normalized for protein concentration. The identities of the amides giving rise to cross peaks in the $^1H-^{15}N$ HSQC spectra are shown. These are based on the original $^1H-^{15}N$ HSQC assignments of Driscoll *et al.* (1991) and also on a series of COSY and NOESY spectra acquired for CD2.D1 in conditions identical to those in the exchange samples (see Experimental Procedures). These spectra were produced using the FELIX 95 software.

in Figure 2 provides a value for $k_{F-I(w)}$ that is in very close agreement to the observed rate of exchange for amide protons in native conditions at the EX1 limit, studied by NMR (M. J. Parker and A. R. Clarke, unpublished results). Here, for amide protons which are fully protected in F, exchange is limited by the global opening rate. If raw concentration is used these rates differ by more than an order of magnitude (data not shown).

Increasing the Stability of the Folding Intermediate

The value of $K_{I/U(w)}$ in these conditions is calculated to be ~4 (see legend to Figure 2), which corresponds to a free energy difference between U and I in water of only ~0.82 kcal mol^{-1}. As, in general, protection factors measured for rapidly formed folding intermediates translate to free energies which are only a fraction of the full free energy change realized between U and I, some method of stabilizing the intermediate must be found, in order that amide protection factors can be confidently measured. A commonly used protein stabilizing agent is sodium sulfate (Na_2SO_4). This compound has been used by others to gain thermodynamic parameters for protein folding intermediates which have been destabilized by mutation to such an extent that they are not populated in water alone [for example see Khorasanizadeh *et al.* (1996)].

Both equilibrium and rate profiles have been obtained for CD2.D1 in the presence of 0.1, 0.2, 0.3, and 0.4 M Na_2SO_4 (Figures 2a and b, respectively). To calculate a denaturant activity scale in the prescence of Na_2SO_4 (up to 0.4 M), we have measured the [GuHCl] dependent solubilities of *N*-acetyltyrosineamide (NAYA) and *N*-acetyltryptophanamide (NATA) in 0, 0.1, 0.2, 0.3, and 0.4 M Na_2SO_4 by UV absorbance (Figure 3; see Experimental Procedures). Here, the solvation free energies in GuHCl and/or Na_2SO_4 are measured relative to water (eq 1). The [GuHCl] dependent data for NAYA and NATA solubility in different concentrations of Na_2SO_4 have been fitted to the hyperbolic relationship given by eq 2 to provide values for $\Delta G_{s,0}$, $\Delta G_{s,max}$, and K_{den}. Here, $\Delta G_{s,0}$ is the free energy change of solvation at a specified concentration of Na_2SO_4, relative to water, $\Delta G_{s,max}$ is the notional, maximum change in the free energy change of solvation at an infinite GuHCl concentration, and K_{den} is the GuHCl concentration required to reach $(\Delta G_{s,m} + \Delta G_{s,0})/2$. These values are given in the legend to Figure 3, for NAYA and NATA in 0, 0.1, 0.2, 0.3, and 0.4 M Na_2SO_4.

The addition of Na_2SO_4 reduces the free energy of solvation of NAYA and NATA (ΔG_s becomes more positive). However, the values of K_{den} for NAYA and NATA are unaffected by the addition of Na_2SO_4, upto 0.4 M (see legend to Figure 3), i.e., the ΔG_s *versus* [GuHCl] curves are parallel but offset on the ΔG_s axis. In other words, Na_2SO_4 does not alter the molar ability of GuHCl to increase the solvation of hydrocarbon. This suggests that the effects of GuHCl and Na_2SO_4 on the solvation of hydrocarbon are additive and independent. In addition, $\Delta G_{s,0}/(\Delta G_{s,max} - \Delta G_{s,0})$ (the free energy of solvation in Na_2SO_4, relative to water, expressed as a fraction of the notional maximum free energy change of solvation, at an infinite concentration of GuHCl) is the same for NAYA and NATA, up to 0.4 M Na_2SO_4 (see legend to Figure 3). As there is a strong correlation between $\Delta G_{s,max}$ (for transfer from water to GuHCl) and the number of carbon atoms for the nonpolar and polar amino acid side chains (Staniforth *et al.*, 1993; Parker *et al.*, 1995; Parker & Clarke, 1997), this suggests that $\Delta G_{s,0}$ will also depend on the number of nonpolar atoms.

For a given concentration of Na_2SO_4, the protein denaturant activity of GuHCl is calculated using the relationship given by eq 4. The last term in eq 4 scales the ability of Na_2SO_4 to decrease hydrocarbon solubility to the molar ability of GuHCl to increase it. This provides negative values for denaturant activity in the presence of Na_2SO_4. The equilibrium and kinetic data measured in 0, 0.1, 0.2, 0.3, and 0.4 M Na_2SO_4 have been plotted against denaturant activity, calculated using eq 4 and the values given in the legend to Figure 3, in Figures 2c and d, respectively. As can be seen, when plotted against the calculated activity scale, both the equilibrium and rate profiles directly superimpose. As well as validating the use of the activity scale, this also demonstrates that the addition of Na_2SO_4 does not noticeably alter the conformational properties of the states on the folding pathway of CD2.D1, i.e., the *m*-values associated with each state are not changed. Hence, Na_2SO_4 increases the stability of each state on the folding pathway by an amount that is

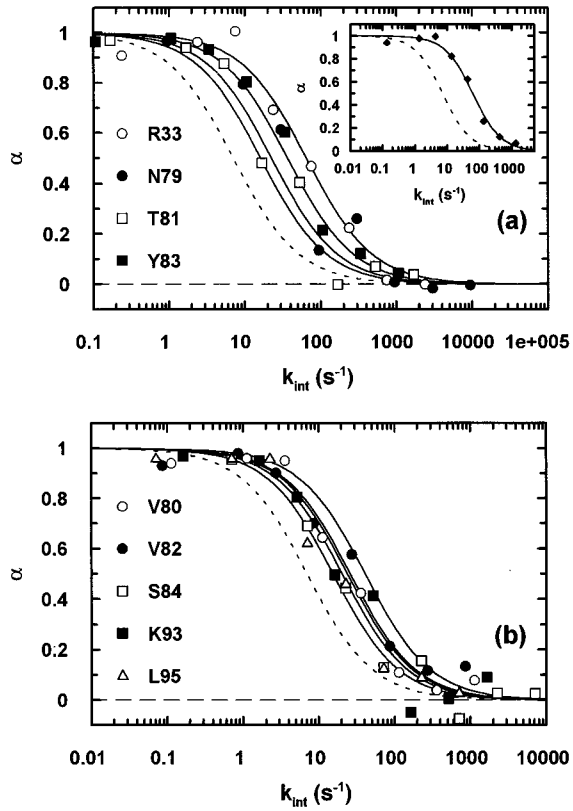

FIGURE 5: pH dependence of amide proton occupancies. To obtain protection factors for the intermediate of CD2.D1 the peak intensities of the amide protons, measured by volume integration of cross peaks in the $^1H-^{15}N$ HSQC exchange spectra (see Experimental Procedures), were plotted against their respective intrinsic exchange rate constants [k_{int}; taken from Bai *et al.* (1993)] and fitted to eq 7. A value for k_{I-F} of 7 s^{-1} was used in the fitting procedure (see Results). Representative data are shown above for residues whose amides make hydrogen bonds in the native state between the FC β-strands (a) and FG β-strands (b) (see Figure 1). Here the proton occupancies (α) were calculated using the values obtained for the amplitude and base line intensities (see Analytical Procedures). The normalized fits to the data are shown as continuous curves. Also shown in a as an inset is the protection data obtained for the indole NH of W34. The rates of H/D exchange for the indole NH group were measured by Bai *et al.* (1993) at 5 °C. To calculate these rates at 25 °C we used values for the activation energies of the base-catalyzed and water-catalyzed exchange rates of 17 and 19 kcal mol^{-1}, respectively [see Bai *et al.* (1993)]. These values provide a conservative estimate for the protection factor of this NH group of ~10. The dashed line in a and b is the α versus k_{int} curve expected for an amide proton which can fully exchange in I, i.e., protection factor, $P = 1$.

proportional to the fraction of buried hydrocarbon. So, Na_2SO_4 influences the free energy of solvation of protein hydrocarbon in an analogous but opposite way to GuHCl. The combined equilibrium and kinetic data in Figures 2c and d have been fitted to eqs 5 and 6, respectively, and the calculated values for $K_{F/U(w)}$, $K_{I/U(w)}$, $k_{I-F(w)}$, $k_{F-I(w)}$, m_U, m_I, and m_t are given in the legend to Figure 2. Marked as dashed lines in Figures 2c and d are the denaturant activities corresponding to 0, 0.1, 0.2, 0.3, and 0.4 M Na_2SO_4.

Measurement of Amide Protection Factors for the Folding Intermediate of CD2.D1

To calculate protection factors for the folding intermediate of CD2.D1 we use the pH competition method (Schmid & Baldwin, 1979; Kim, 1986; Roder, 1989; Englander & Mayne, 1992). In this technique, fully protonated protein, unfolded in denaturant in H_2O, is refolded into deuterated solvent at a specified pH, i.e., folding and exchange are initiated simultaneously. The extent to which a particular amide proton is replaced by a deuteron will depend on the competition between its rate of exchange (k_{ex}) and the rate at which it becomes protected by folding into the native state (k_f). The proton occupancy (fraction of unexchanged amide protons; α) for a particular amide is related to k_{ex} by the following relationship:

$$\alpha = k_f/(k_f + k_{ex}) \quad (8)$$

In conditions where folding proceeds through a rapidly formed intermediate, k_f is given by k_{I-F}, determined from the kinetic analysis described above. The observed rate of exchange for a particular amide proton (k_{ex}) will depend on its intrinsic rate of exchange (i.e., the expected rate of exchange from a random coil polypeptide under these conditions) and any protection afforded to it in the rapidly formed intermediate. The retardation in k_{int} is most usually expressed in the form of a protection factor (P), which is related to k_{ex} by $P = k_{int}/k_{ex}$. CD2.D1 is ideally suited to this type of technique as the dynamics of folding and unfolding are unperturbed over the applied pH range. In fact, the rate and equilibrium profiles obtained for this protein at pH 6, 7, and 10 are directly superimposable (data not shown). The advantage of this pH independence cannot be overstressed, since all kinetic exchange−protection experiments demand exposure of the folding protein to widely differing pH conditions which, in most cases, must perturb the free energies of the system.

As described above, to stabilize the folding intermediate of CD2.D1 the folding/exchange competition reactions were performed in the prescence of Na_2SO_4. The final reaction solution contained 0.4 M Na_2SO_4 and 0.3 M GuHCl (see Experimental Procedures). These conditions correspond to a denaturant activity of approximately −0.76 M (calculated using eq 4 and the values described in the legend to Figure 3). Using eq 6 and the kinetic parameters given in the legend to Figure 2, k_{I-F} and $K_{I/U}$ are calculated to be approximately 7 s^{-1} and 80, respectively, in these conditions. While we have shown previously that D_2O influences the folding dynamics of CD2.D1 relative to H_2O (Parker & Clarke, 1997), the changes in $K_{I/U}$ and k_{I-F} are relatively small and the effect on the determined protection factors is calculated to be within experimental error.

To calculate protection factors for the amide protons of CD2.D1 we use the intrinsic exchange rates determined by Bai *et al.* (1993). To investigate the effect of Na_2SO_4 and GuHCl on these exchange rates we have measured the pH dependence of the rates of amide exchange, in the presence and absence of 0.4 M Na_2SO_4 and 0.3 M GuHCl, for poly-D,L-alanine (PDLA), poly-D,L-lysine (PDLK), and poly-L-glutamate (PLE)—models for random coil polypeptides—using the UV absorption method described by Englander *et al.* (1979) (see Experimental Procedures). The plots of $\log_{10} k_{ex}$ versus pH_m obtained under both sets of conditions produce straight lines with slopes of 1.0 (data not shown). The presence of 0.4 M Na_2SO_4 and 0.3 M GuHCl has only a marginal effect on the exchange rates of PDLA and PDLK (rate of exchange of PDLA is increased by ~1.7 while that of PDLK is decreased by ~0.6). These effects are within the experimental error involved in the determination of intrinsic exchange rate constants. The exchange rate of PLE is affected more significantly, however, and is increased by

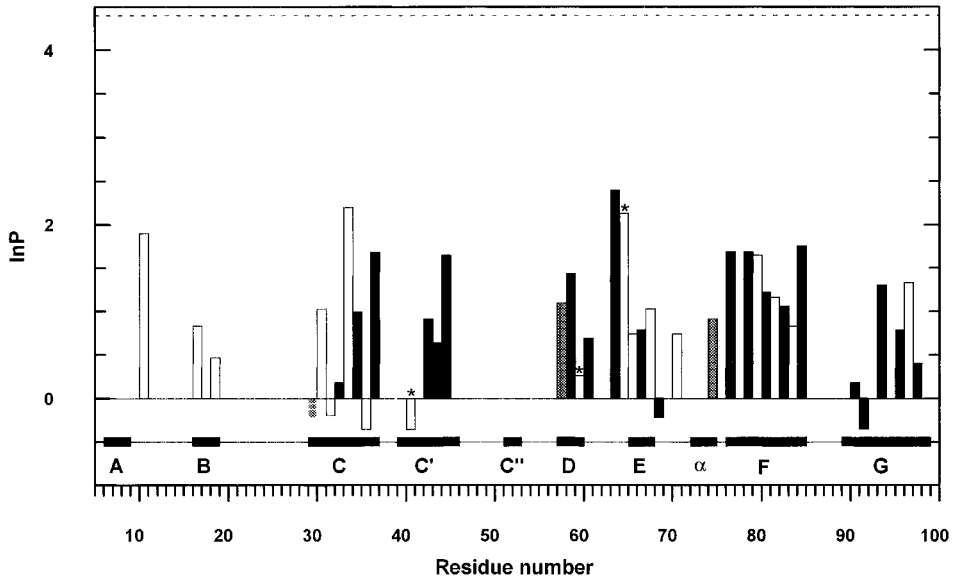

FIGURE 6: Protection factor versus residue number. The natural logarithm of the protection factors (ln *P*) measured for the folding intermediate of CD2.D1 are plotted here against residue number. Also shown is the position of the secondary structural elements in the context of the sequence [based on the DSSP program (Kabsch & Sander, 1983)]. In black are those residues whose amides make hydrogen bonds within β-hairpins of the sheets (sequence-local); these include hydrogen bonds made in the hairpin loops. In white are those residues whose amides make hydrogen bonds between sequence-distant β-strands within the sheets. In grey are those residues whose amides make hydrogen bonds in nonhairpin β-turns. None of the amides followed in these experiments reveal protection factors exactly equal to 1. Hence, in the above plot, a value of zero indicates that a protection factor could not be determined, owing to rapid exchange. L40, I59, and D64 have been marked by an asterisk as the amides of these groups do not make main chain hydrogen bonds in the native state. The amides of L40 and I59 are completely solvent exposed in the native state so their appearance in the exchange spectra is difficult to interpret. The protection measured for D64 could result from the formation of a hydrogen bond with the side chain carbonyl oxygen of asparagine 62 (from inspection of the folded structure). Represented by a dashed line is the protection factor expected for an amide proton that is fully protected in the intermediate, based on the value for $K_{I/U}$ under the experimental conditions ($P = K_{I/U} + 1 = \sim 80$; see Results). The experimental uncertainty in *P* is estimated to be no more than 30%.

~5.7. As discussed by Kim and Baldwin (1982), the addition of salts may influence the rates of base-catalyzed exchange of amides in polyelectrolytes in two ways: (i) by an increased charge screening of the polyelectrolyte itself and of the reactive H_3O^+ and OH^- species; (ii) by affecting the balance of counterion species condensed around the polyelectrolyte, thus influencing the pH in the immediate vicinity of the amides. How these two effects influence the base-catalyzed exchange rates of amides in the context of a protein molecule is extremely difficult to predict, although one might anticipate that the latter effect would not be as significant in a protein molecule due to the much reduced density of charges along the chain compared with an homogenous polyelectrolyte. In addition, even if these salts do increase the exchange rates of glutamate and aspartate residues significantly, this would lead to protection factors for these residues being underestimated.

Representative $^1H-^{15}N$ HSQC spectra collected after folding/exchange reactions carried out at pH_m 6 and 8 are shown in Figure 4a and b, respectively. In total, spectra were collected from folding/exchange reactions carried out at pH_m of 6, 7, 7.5, 8, 8.5, 9, 9.5, and 10. The data for the pH dependences of the amide proton occupancies determined from these exchange spectra are shown in Figure 5. The calculated protection factors for the intermediate state of CD2.D1 are plotted against residue number in Figure 6. The dashed line in Figure 6 is the protection factor expected for an amide proton that is fully protected from exchange in the intermediate, based on the value for $K_{I/U}$ under these conditions (see above). The pattern of protected amides is shown in the context of the native structure in Figure 7.

DISCUSSION

The pattern of protection in the folding intermediate of CD2.D1 is striking. All the β-strand segments show amides which are protected from exchange in the intermediate (i.e., $P > 1$) with the largest protection factors being around 20. Not only is there protection in the central strand positions which have side chains contributing to the hydrophobic core, there is also protection in the turn regions of hairpins where the side chains are exposed to the solvent in the folded structure. Furthermore, both hairpin and long-sequence-range β-strand interactions are formed in this state, demonstrating that the entire, tortuous topology of the molecule is decided within the dead time of a stopped-flow apparatus (< 1 ms). The level of compactness of the structure on this time scale is revealed by the significant protection of the indole NH of W34 (Figure 5). In the native state this group is fully buried in the hydrophobic core and makes a hydrogen bond to the main chain carbonyl oxygen of G63 (CE cross-sheet interaction). The observed protection in the intermediate must result from the formation of a hydrogen bond to the main chain and/or partial exclusion of this side chain from the solvent. Correspondingly, the value of m_I (see legend to Figure 2) also demonstrates a significant degree of chain condensation on this time scale: approximately 60% ($((m_U - m_I)/m_U) \times 100\%$) of the nonpolar groups buried in the folded state are desolvated in the intermediate.

Protection from exchange of amides in β-structure has been observed in the rapidly formed folding intermediates of ribonuclease A (RNase A) (Udgaonkar & Baldwin, 1988), T4 lysozyme (Lu & Dahlquist, 1992), and the N-terminal domain of phosphoglycerate kinase (N-PGK) (Hosszu *et al.*, 1997), but the question of whether these β-structures are

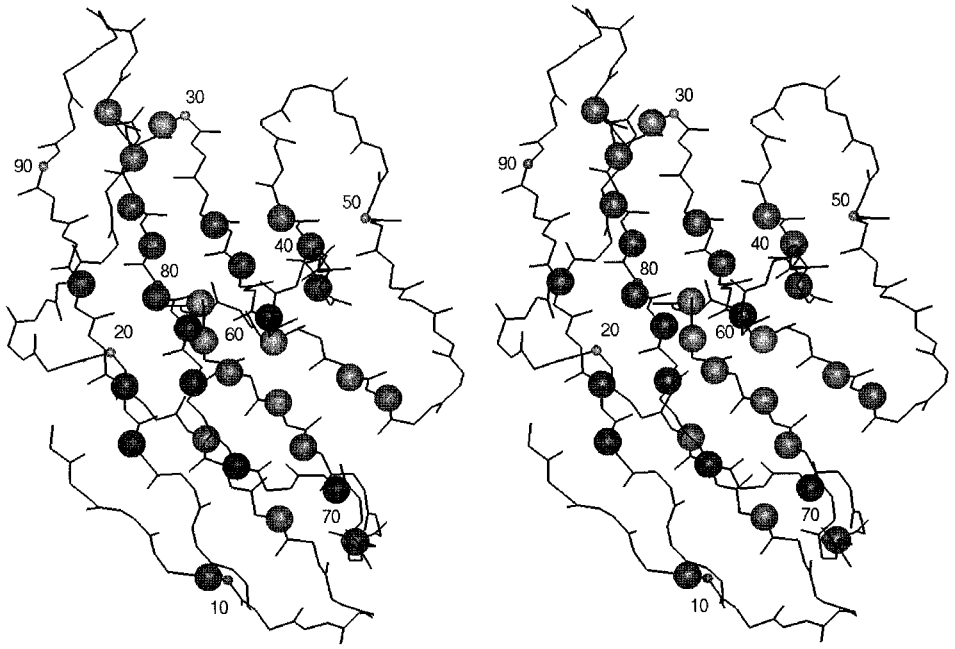

FIGURE 7: Patterns of amide protection in the native topology. The pattern of protected amides is shown here in a stereodiagram of the native structure. The backbone is represented by solid lines, and the amide groups which exhibit ln $P > 0.5$ are represented by dark spheres (produced using INSIGHT 95 molecular graphics software).

independently stable is difficult to address as they are associated with stable α-helices in the native state. In the β-domain of staphylococcal nuclease (SNase), where the β-strands do not form tertiary contacts with α-helices, protection of amide residues is confined to the β-hairpins (Jacobs & Fox, 1994). One example of an all-β protein studied by amide exchange is interleukin-1β (Varley *et al*., 1993). However, while the early folding intermediate of this protein exhibits spectral properties consistent with a high degree of β-structure, it affords no discernable protection of its amide protons.

Owing to the high energetic cost of burying highly polar, hydrogen-bonding groups in hydrophobic environments, the protection factors measured for the intermediate of CD2.D1 are attributed to the formation of intramolecular hydrogen bonds with backbone carbonyl acceptors rather than to solvent protection by closely associated nonpolar side chains. Given this, it is difficult to imagine that the extent of structural organization in the intermediate state arises from a purely random hydrophobic collapse, i.e., one would not expect such configurationally specific bonds between long-sequence-range groups (e.g. AG, FC, and EB strand−strand hydrogen bonds) to be so highly represented in such a degenerate mechanism of structural acquisition.

Recent rapid kinetic data reveal the formation of some level of structure in apomyoglobin on the 10^{-5} s time scale (Ballew *et al*., 1996), and in the case of CD2.D1, the highly ordered structure described here is certainly acquired within a millisecond. These results have a strong bearing on the design of theoretical models intended to predict the relationship between amino acid sequence and protein topology by simulation. They suggest that a polypeptide chain sampling configurations at a rate of 10^9 s^{-1} [as suggested from relaxation times of Cα carbons of the backbone measured by NMR (Glushko *et al*., 1972) and fluorescence anisotropy (Smith *et al*., 1991)] can only investigate a vanishingly small area of configurational space (10 000−1 000 000 structures) in the time taken to adopt a native-like fold, and that precise side chain interactions are not required in this search. If we understood the nature of the seemingly crude physical chemistry which directs the process of rapid structure formation in the kinetic intermediate, then it should be possible to capture this highly directed process in a computational model.

ACKNOWLEDGMENT

We are indebted to Dr. Paul Driscoll and Prof. Ian Campbell for making their original NMR data available to us and to Dr. Richard Sessions for critical comments on the manuscript.

REFERENCES

Bai, Y., Milne, J. S., Mayne, L., & Englander, S. W. (1993) *Proteins: Struct. Funct. Genet. 17*, 75−86.
Baldwin, R. L. (1996) *Folding Des. 1*, R1−R8.
Ballew, R. M., Sabelko, J., & Gruebele, M. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 5795−5764.
Bax, A., Ikura, M., Kay, L. E., Torchia, D. A., & Tschudin, R. (1990) *J. Magn. Reson. 86*, 304−318.
Bryngelson, J. D., Onuchic, J. N., Socci, N. D., & Wolynes, P. G. (1995) *Proteins: Struct. Funct. Genet. 21*, 167−195.
Dill, K. A., Bromberg, S., Yue, K., Fiebig, K. M., Yee, D. P., Thomas, P. D., & Chan, H. S. (1995) *Protein Sci. 4*, 561−602.
Driscoll, P. C., Cyster, J. G., Campbell, I. D., & Williams, A. F. (1991) *Nature 353*, 762−765.
Englander, J. J., Calhoun, D. B., & Englander, S. W. (1979) *Anal. Biochem. 92*, 517−524.
Englander, S. W., & Kallenbach, N. R. (1984) *Q. Rev. Biophys. 16*, 521−655.
Englander, S. W., & Mayne, L. (1992) *Annu. Rev. Biophys. Biomol. Struct. 21*, 243−265.
Fersht, A. R., Serrano, L., & Matouschek, A. (1992) *J. Mol. Biol. 224*, 771−782.
Glushko, V., Lawson, P. J., & Gurd, F. R. N. (1972) *J. Biol. Chem. 247*, 3176−3185.
Horovitz, A. (1996) *Folding Des. 1*, R121−R126.
Jacobs, M. D., & Fox, R. O. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 449−453.
Johnson, C. M., & Fersht, A. R. (1995) *Biochemistry 34*, 6975−6804.
Jones, E. Y., Davis, S. J., Williams, A. F., Harlos, K., & Stuart, D. I. (1992) *Nature 360*, 232−239.

Kabsch, W., & Sander, C. (1983) *Biopolymers 22*, 2577−2637.
Khorasanizadeh, S., Peters, I. D., & Roder, H. (1996) *Nat. Struct. Biol. 3*, 193−205.
Kim, P. S. (1986) *Methods Enzymol. 131*, 136−156.
Kim, P. S., & Baldwin, R. L. (1982) *Biochemistry 21*, 1−5.
Kim, P. S., & Baldwin, R. L. (1990) *Annu. Rev. Biochem. 59*, 631−660.
Kumar, A., Ernst, R. R., & Wüthrich, K. (1980) *Biochim. Biophys. Res. Commun. 95*, 1−6.
Lu, J., & Dahlquist, F. W. (1992) *Biochemistry 31*, 4749−4756.
Matthews, C. R. (1993) *Annu. Rev. Biochem. 62*, 653−683.
Miranker, A. D., & Dobson, C. M. (1996) *Curr. Opin. Struct. Biol. 6*, 31−42.
Miranker, A. D., Robinson, C., Radford, S. E., Aplin, R., & Dobson, C. M. (1991) *Science 262*, 896−900.
Morris, G. A., & Freeman, R. (1978) *J. Magn. Reson. 29*, 433−463.
Murray, A. J., Lewis, S. J., Barclay, A. N., & Brady, R. L. (1995) *Proc. Natl. Acad. Sci. U.S.A. 92*, 7337−7341.
Myers, J. K., Pace, C. N., & Scholtz, J. M. (1995) *Protein Sci. 4*, 2138−2148.
Pace, C. N. (1975) *CRC Crit. Rev. Biochem. 3*, 1−43.
Parker, M. J., & Clarke, A. R. (1997) *Biochemistry 36*, 5786−5794.
Parker, M. J., Spencer, J., & Clarke, A. R. (1995) *J. Mol. Biol. 253*, 771−786.
Ptitsyn, O. B. (1995) *Curr. Opin. Struct. Biol. 6*, 31−42.
Rance, M., Sørenson, O. W., Bodenhausen, G., Wagner, G., Ernst, R. R., & Wüthrich, K. (1983) *Biochim. Biophys. Res. Commun. 117*, 479−485.
Roder, H. (1989) *Methods Enzymol. 176*, 446−473.
Roder, H., & Wüthrich, K. (1986) *Proteins: Struct. Funct. Genet. 1*, 34−42.
Roder, H., & Colón, W. (1997) *Curr. Opin. Struct. Biol. 7*, 15−28.
Roder, H., Elöve, G. A., & Englander, S. W. (1988) *Nature 335*, 700−704.
Sali, A., Shakhnovich E., & Karplus, M. (1994) *J. Mol. Biol. 235*, 1614−1636.
Schmid, F. X., & Baldwin, R. L. (1979) *J. Mol. Biol. 135*, 199−215.
Shortle, D., Meeker, A. K., & Freire, E. (1988) *Biochemistry 27*, 4761−4768.
Smith, C. J., Clarke, A. R., Chia, W. N., Irons, L. I., Atkinson, T., & Holbrook, J. J. (1991) *Biochemistry 30*, 1028−1036.
Staniforth, R. A., Burston, S. G., Smith, C. J., Jackson, G. S., Badcoe, I. G., Atkinson, T., Holbrook, J. J., & Clarke, A. R. (1993) *Biochemistry 32*, 3842−3851.
States, D. J., Haberkorn, R. A., & Ruben, D. J. (1982) *J. Magn. Reson. 48*, 286−292.
Tanford, C. (1970) *Adv. Protein Chem. 24*, 1−95.
Udgaonkar, J. B., & Baldwin, R. L. (1988) *Nature 335*, 694−699.
Varley, P., Gronenborn, A. M., Christensen, H., Wingfield, P. T., Pain, R. H., & Clore, G. M. (1993) *Science 260*, 1110−1113.

BI971294C