Interdomain Communication of T-Cell CD4 Studied by Absorbance and Fluorescence Difference Spectroscopy Measurements of Urea-Induced Unfolding[†,||]

Susan W. Tendian,[‡] David G. Myszka,[§] Raymond W. Sweet,[§] Irwin M. Chaiken,[§] and Christie G. Brouillette*[,‡]

Southern Research Institute, Birmingham, Alabama 35205, and SmithKline Beecham Pharmaceuticals, King of Prussia, Pennsylvania 19406

Received December 20, 1994; Revised Manuscript Received March 1, 1995[⊗]

ABSTRACT: CD4 is a transmembrane glycoprotein expressed on T-lymphocytes. It is a receptor for class II major histocompatibility complex (MHC) molecules and for the HIV envelope glycoprotein gp120. The extracellular portion of CD4 (sCD4) is a rod-shaped molecule consisting of four domains designated D1 through D4. Denaturant-induced unfolding of sCD4 and of isolated CD4 domains, D1D2 and D3D4, was measured using both ultraviolet absorbance and fluorescence difference spectroscopy. Though both absorbance and fluorescence changes arise from changes in the solvent exposure of the intrinsic tryptophan chromophores, the unfolding curves obtained with the two techniques looked very different for sCD4 and D3D4. This dissimilarity is indicative of a greater than two-state unfolding mechanism. The global three-state fit for sCD4, which simultaneously fit both absorbance and emission data to a model with one thermodynamically stable unfolding intermediate, was significantly better than the best two-state fit, suggesting that there are two unfolding regions. Unfolding of isolated D3D4 also fit a three-state model while unfolding of isolated D1D2 fit satisfactorily to a two-state model. The unfolding of these two isolated fragments could not be summed to yield the unfolding profile of sCD4, implying that an interaction between D2 and D3 is lost by splitting sCD4 between these domains. The unfolding data of isolated D1D2 and D3D4 were used with the solvent-accessible surface area of tryptophans calculated from atomic crystal structure coordinates of human D1D2 and rat D3D4 to assign the unfolding steps. The data are consistent with a model for sCD4 unfolding wherein the one stable intermediate appears to contain only the D4 domain unfolded. The remaining three domains apparently unfold as a unit. Furthermore, interactions between domains D1, D2, and D3 appear to stabilize D4, suggesting that stabilizing interactions exist between D3 and D4 even though the unfolding of the D3D4 fragment is best fit by a three-state model. This report is the first to describe a thermodynamic basis for a wide range of biological properties implicated for CD4.

CD4 is a cell surface receptor expressed on T-lymphocytes that plays an important role in the cellular immune response by its recognition of class II major histocompatibility complex (MHC)[1] molecules. In humans, CD4 serves as the receptor for the envelope glycoprotein gp120 of the human immunodeficiency virus (HIV),[1] thereby initiating viral entry into the cell. CD4 consists of a 372 amino acid extracellular portion, a 23 amino acid transmembrane segment, and a 38 amino acid cytoplasmic tail (Maddon et al., 1985, 1987). The 44 000 MW extracellular portion, soluble CD4 (sCD4),[1] consists of four contiguous folding domains (D1−D4).[1] Ultracentrifugation measurements give evidence that sCD4 is an extended rodlike molecule that is not bent between domains (Kwong et al., 1990).

Before the crystal structures of D1D2 and D3D4 were available, these domains were predicted on the basis of sequence homology to belong to the immunoglobulin superfamily (Clark et al., 1987; Maddon et al., 1985, 1987; Williams, 1987). The N-terminal domain, D1, shares the greatest sequence homology with immunoglobulin variable domains (particularly V_κ). Two groups independently solved the crystal structure of the D1D2 fragment of human CD4 (Ryu et al., 1990, 1994; Wang et al., 1990; Garrett et al., 1993), confirming that these domains possess an immunoglobulin fold. Each domain is an antiparallel β barrel, with nine β strands in domain 1 and seven β strands in domain 2 (Figure 1). The last strand of D1 continues straight into the first strand of D2 with residue 98 forming a hydrogen bond in a sheet of D1 and residue 99 forming a hydrogen bond in a sheet of D2. A consequence of this unique connection between the D1 and D2 domains is that the domains pack closely against one another with a large hydrophobic interface.

Although there is no published crystal structure for sCD4 or human D3D4, the crystal structure of the D3D4 fragment of rat CD4 has been solved (Brady et al., 1993; Lange et al., 1994). Rat and human CD4s share 55% overall sequence identity and 70% sequence similarity. Like human D1 and D2, both D3 and D4 are antiparallel β barrels, and rat D3 and D4 have the same number of β strands as human D1 and D2, respectively (Figure 1). Rat D3 and D4 are also connected with a continuous β strand, and they form a closely associated rodlike structure. Unlike the other three domains,

[†] This work was supported by NIH Grant AI32687.

[||] Coordinates for D1D2 were obtained from Brookhaven Protein Databank files 1CD4 and 2CD4, and coordinates for D3D4 were obtained from Brookhaven Protein Databank file 1C1D.

* To whom correspondence should be addressed.

[‡] Southern Research Institute.

[§] SmithKline Beecham Pharmaceuticals.

[⊗] Abstract published in *Advance ACS Abstracts,* May 1, 1995.

[1] Abbreviations: MHC, major histocompatibility complex; HIV, human immunodeficiency virus; sCD4, soluble CD4; D1, domain 1; D2, domain 2; D3, domain 3; D4, domain 4.

0006-2960/95/0434-6464$09.00/0

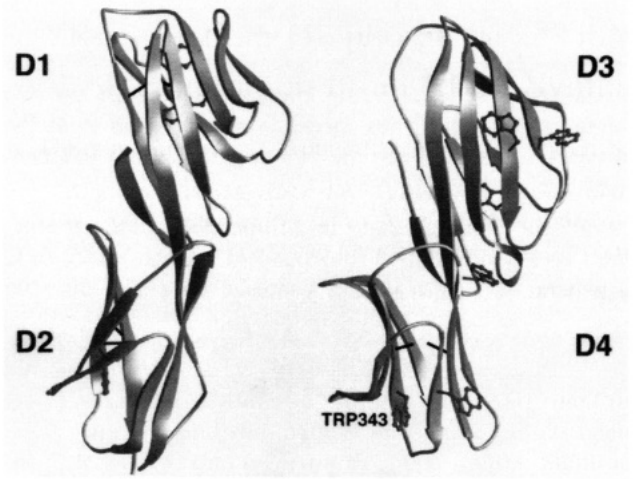

FIGURE 1: Backbone structure of the human D1D2 (left) and rat D3D4 (right) fragments of sCD4 (from Brookhaven PDB entries 3CD4 and 1CID, respectively). Disulfide bonds are shown as solid lines. Tryptophan residues are shown. Four of the tryptophans shown are not in rat D3D4, but the residues were changed for this figure to show the location of tryptophans in the human sequence. This figure was drawn with the program RIBBONS version 2.5 by M. Carson (1987).

rat and human D3 lack a disulfide bond. This difference results in an increased separation of the two β sheets in D3 compared to D1, D2, and D4. While rat D3 is expanded relative to human D1, rat D4 is more compact than D2 of human CD4.

Regions of functional significance have been identified throughout the CD4 molecule. Domains 1 and 2 and possibly 3 play a role in the recognition of MHC II molecules (Clayton et al., 1989; Fleury et al., 1991; Lamarre et al., 1989; Moebius et al., 1992, 1993). The interaction of CD4 with its coreceptor, the T-cell receptor (TCR), is not as well understood but is thought to involve the membrane-proximal D3 and D4 domains of CD4 (Chuck et al., 1990; Dianzani et al., 1992). Domain 1 contains the binding site for HIV gp120 on an exposed loop as summarized in Garrett et al. (1993), Moore and Sweet (1993), and Ryu et al. (1994). Antibody binding to regions outside of the exposed D1 loop and mutations of residues in D2, D3, and certain buried residues in D1 have also been found to interfere with gp120 binding to CD4 (Clayton et al., 1988; Mizukami et al., 1988; Fleury et al., 1991) and/or other steps in the infection process (Healey et al., 1990; Burkly et al., 1992). These results suggest the existence of functionally significant intradomain structural effects in D1 and interdomain communication between D1 and D2D3. The cytoplasmic portion of CD4 is noncovalently associated with the protein tyrosine kinase $p56^{lck}$ (Rudd et al., 1988; Veillette et al., 1988).

We have studied interdomain linkage in human sCD4 by making thermodynamic measurements of denaturant-induced unfolding of sCD4 as well as isolated fragments of sCD4. To monitor the unfolding, we used fluorescence and absorption difference spectroscopy which rely on the sensitivity of intrinsic protein chromophores to changes in solvent exposure resulting from changes in protein conformation. The number of states in sCD4 unfolding was determined, thus allowing the detection of a stable unfolding intermediate. By combining the results from the unfolding of sCD4 fragments with crystal structures of D1D2 and D3D4, a structure for the sCD4 unfolding intermediate is suggested.

MATERIALS AND METHODS

Materials. Ultrapure urea was obtained from Gibco BRL (Grand Island, NY) and was not further purified. Chymotrypsin−agarose was purchased from Sigma Chemical Co. (St. Louis, MO). Concanavalin A−Sepharose (Con A−Sepharose) was purchased from Pharmacia LKB (Piscataway, NJ). Affi-prep 10 was purchased from BioRad (Hercules, CA).

Purification of Soluble CD4 and D1D2. Recombinant human sCD4 (residues 1−369) and D1D2 (residues 1−183) were expressed in CHO cells and purified as described previously (Deen et al., 1988; Arthos et al., 1989). sCD4 and D1D2 were purified by a combination of ion-exchange and size-exclusion chromatography. sCD4 and D1D2 concentrations were determined from the absorbance at 280 nm using the extinction coefficients of 1.45 $mL \cdot mg^{-1} \cdot cm^{-1}$ and 0.9 $mL \cdot mg^{-1} \cdot cm^{-1}$ and molecular weights of 44 000 and 20 262, respectively.

D3D4 Generation and Isolation. D3D4 was produced by proteolytic cleavage of sCD4 with chymotrypsin coupled to agarose (Healey et al., 1990). A 1.5 mg/mL sCD4 solution was incubated in 150 mM sodium chloride, 50 mM methionine, 50 mM HEPES [*N*-(2-hydroxyethyl)piperazine-*N*′-2-ethanesulfonic acid], 1 mM calcium chloride, and 1 mM manganese chloride (pH 7.9) with mixing at room temperature for 2 h with N^{α}-tosyllysine chloromethyl ketone-treated chymotrypsin−agarose (5 units per 1.5 mg of sCD4). After the 2 h proteolysis, the chymotrypsin−agarose was spun down, and the supernatant was removed and treated with phenylmethanesulfonyl fluoride (PMSF) to irreversibly inhibit any contaminating chymotrypsin beads. D3D4 was the predominant proteolysis fragment; N-terminal sequencing (following separation by SDS−PAGE and transfer onto PVDF) identified the first residue as Ala-178.

D3D4 was purified by affinity chromatography by taking advantage of the presence of oligosaccharides in D3D4 and their absence in the rest of sCD4. Proteolyzed sCD4 was loaded on a Con A-Sepharose column, and then the column was washed to remove unbound small molecular weight fragments (<10 000). D3D4 and unproteolyzed sCD4 were then eluted with 0.5 M mannopyranoside. D3D4 preparations with more than ~5% contamination from sCD4 were further purified with a column containing HIV gp120 immobilized on Affi-prep 10. D3D4 concentrations were determined from the absorbance at 280 nm using an extinction coefficient of 1.75 $mL \cdot mg^{-1} \cdot cm^{-1}$ and a molecular weight of 24 500.

Measurements of Equilibrium Unfolding by Difference Spectroscopy. Denaturant-induced unfolding of sCD4, D1D2, and D3D4 was monitored by ultraviolet absorbance using a Cary 3E UV/visible spectrophotometer (Varian Australia Pty. Ltd., Victoria, Australia) and by fluorescence using an Aminco-Bowman Series 2 fluorometer (SLM-Aminco, Rochester, NY). Urea stock solutions (10 M) prepared daily in 50 mM sodium acetate were adjusted to pH 5.5 with acetic acid. The concentrations of stock solutions of urea were confirmed by refractive index measurements (Warren & Gordon, 1966) using an Abbe Refractometer (American Optical Corp., Buffalo, NY). The kinetics of unfolding at several urea concentrations were determined to establish the time necessary to reach equilibrium. Protein samples were incubated with various concentrations of urea at 25 °C in

50 mM sodium acetate buffer at pH 5.5 for 36−50 h to allow samples to reach equilibrium before measuring absorbance or fluorescence. Because of the lengthy time period required to reach equilibrium, measurements were made on individual samples instead of titrating a single sample with urea.

Absorption spectra were obtained for 0.15−0.50 mg/mL protein samples (A_{280} = 0.25−0.45) from 250 to 330 nm at 0.1 nm intervals with a signal averaging time of 0.73 s at each data point and a spectral band width of 1.5 nm. Because the extinction coefficient at ~270 nm is independent or nearly independent of experimental conditions (Demchenko, 1986), a 270 nm correction was used to correct absorption spectra for slight unavoidable differences in protein concentrations arising from the preparation of multiple protein samples. The wavelength of the isosbestic point near 270 nm was determined, and spectra were normalized for the average intensity at this wavelength. This correction decreased the scatter in the Δabs values, but it did not affect the overall data trends. Difference spectra were then generated by subtracting spectra of protein samples containing urea from the spectrum of protein in the absence of urea. The change in absorbance at ~294 nm (ΔA_{294}), the position of the larger of two peaks in the difference spectra, was determined and divided by the molar protein concentration and the cell path length (Schmid, 1989).

Fluorescence emission spectra of 0.02−0.03 mg/mL samples were measured between 300 and 450 nm using an excitation wavelength of 294 or 271 nm. Photomultiplier high voltage was kept constant during all the measurements for a single unfolding profile. (Emission spectra from one sample prepared at the same concentration used in the absorbance measurements confirmed that the results were independent of concentration.) Emission spectra were corrected by subtracting the appropriate buffer or urea blanks from the protein spectra. Difference spectra were then generated by subtracting spectra of protein samples containing urea from the spectrum of protein in the absence of urea. The fluorescence difference was monitored at its maximum which was located between 320 and 330 nm.

Curve Fitting of Unfolding Data. SigmaPlot (Jandel Scientific Corp., San Rafael, CA), which applies the Marquardt−Levenberg algorithm, was used to fit the unfolding data to the nonlinear equations described below. Data were fit to multiple unfolding models including two-state (Santoro & Bolen, 1988), three-state (Matthews & Crisanti, 1981; Pace, 1986), four-state, and modified four-state (Rowe & Tanford, 1973) as described by eqs 1−4, respectively. The

$$\mathrm{F(folded)} \overset{K_D}{\rightleftharpoons} \mathrm{U(unfolded)} \quad (1)$$

$$\mathrm{F} \overset{K_1}{\rightleftharpoons} \mathrm{I(intermediate)} \overset{K_2}{\rightleftharpoons} \mathrm{U} \quad (2)$$

$$\mathrm{F} \overset{K_1}{\rightleftharpoons} \mathrm{I_1} \overset{K_2}{\rightleftharpoons} \mathrm{I_2} \overset{K_3}{\rightleftharpoons} \mathrm{U} \quad (3)$$

$$\mathrm{F} \overset{K_1}{\rightleftharpoons} \mathrm{I_1} \overset{K_2}{\rightleftharpoons} \mathrm{U}, \mathrm{F} \overset{K_2}{\rightleftharpoons} \mathrm{I_2} \overset{K_1}{\rightleftharpoons} \mathrm{U} \quad (4)$$

modified four-state model differs from a true four-state in that only two equilibrium constants define the unfolding equilibria. In structural terms, the modified four-state can be thought of in the following way. The folded region in intermediate 1 is the unfolded region in intermediate 2 and vice versa. The equilibrium constants named K_1 for $F \rightleftarrows I_1$ and $I_2 \rightleftarrows U$ are equivalent as are the two equilibrium constants labeled K_2, because this model assumes there are no stabilizing or destabilizing interactions between the two folding regions; thus, this model is also referred to as the independent model since the two regions unfold independently.

Raw data plotted as Δabs or Δfluor (dependent variable) versus urea concentration (independent variable) were fit to the general eq 5. Equation 6 was used for global fits after

$$\Delta y = \frac{\Delta y_f + \Delta y_u K_{app}}{1 + K_{app}} \quad (5)$$

$$f_{app} = \frac{K_{app}}{1 + K_{app}} \quad (6)$$

$$f_{app} = \frac{\Delta y - \Delta y_f}{\Delta y_u - \Delta y_f} \quad (7)$$

converting Δabs or Δfluor to f_{app} using eq 7, where Δy is the experimental difference spectroscopy measurement, Δy_f is the preunfolding value, and Δy_u is the postunfolding value. Δy_f and Δy_u were not generally constants, but instead were variables linearly dependent upon the urea concentration. They were determined by linear regression of the first few low urea points or the last few high urea points in the unfolding curve ($\Delta y_f = \Delta y_f^\circ + m_f[\mathrm{urea}]$ and $\Delta y_u = \Delta y_u^\circ + m_u[\mathrm{urea}]$) (Pace et al., 1989). Occasionally when there were not enough initial points to determine a folded base line, Δy_f was set equal to zero. Simulations of unfolding data have shown that this assumption normally does not introduce unacceptable errors in the derived thermodynamic data (Eftink, 1994).

For each of the thermodynamic models described in eqs 1−4, the overall K_{app}, the urea-dependent equilibrium constant for the unfolding process, is defined respectively by eqs 8−11:

$$\text{two-state: } K_{app} = K_d = \exp\left(\frac{\Delta G_{H2O} - m_D[D]}{-RT}\right) \quad (8)$$

$$\text{three-state: } K_{app} = \frac{K_1K_2 + ZK_1}{1 + (1 - Z)K_1} \quad (9)$$

$$\text{four-state: } K_{app} = \frac{K_1K_2K_3 + Z_1K_1 + Z_2K_2}{1 + (1 - Z_1)K_1 + (1 - Z_2)K_2} \quad (10)$$

$$\text{modified four-state: } K_{app} = \frac{K_1K_2 + ZK_1 + (1 - Z)K_2}{1 + (1 - Z)K_1 + ZK_2} \quad (11)$$

where

$$K_1 = \exp\left(\frac{\Delta G_{1H2O} - m_{1D}[D]}{-RT}\right), K_2 = \exp\left(\frac{\Delta G_{2H2O} - m_{2D}[D]}{-RT}\right), K_3 = \exp\left(\frac{\Delta G_{3H2O} - m_{3D}[D]}{-RT}\right)$$

[D] is the experimental denaturant concentration. The parameters determined by the fit, Z, ΔG_{H2O}, and m_D, are,

respectively, the fractional change in the measured optical value for the transition from the native state to the intermediate, the intercept, and the negative slope of the linear extrapolation to zero urea concentration of unfolding free energy changes which were measured as a function of urea concentration. In the two-state model, therefore, ΔG_{H2O} is the free energy difference between the native and the unfolded protein in the absence of urea. $\Delta G_{H2O}/m_D$ equals $[urea]_{1/2}$, which is the urea concentration at the midpoint of the unfolding transition. In the two-state model, $[urea]_{1/2}$ is the urea concentration where the concentration of native protein equals the concentration of unfolded protein. In the three- and four-state models, ΔG_{1H2O}, m_{1D}, ΔG_{2H2O}, m_{2D}, ΔG_{3H2O}, and m_{3D} have the same meaning as in the two-state model, except that parameters with subscript "1" apply to the unfolding step defined by K_1 (i.e., $F \rightleftarrows I_1$), subscript "2" is used for parameters that apply to the unfolding step defined by K_2, and so on (see eqs 1−4).

In global fits, which were used for the final analyses, multiple data sets including both absorption and emission data for a given protein were used to determine a single set of thermodynamic parameters while allowing, in three-state and four-state fits, different fractional changes of the fluorescence and absorbance during each unfolding step. The number of data points per fitted parameter ranged from 27 to 41 for global two-state fits and from 9 to 14 for global three-state and modified four-state fits. Global fits were not affected by the use of weighting to compensate for an unequal number of fluorescence and absorbance data sets. Weighting based on differences in the size of errors associated with data collected by the two techniques was not necessary since fluorescence and absorbance measurements exhibited approximately equal errors as determined by comparison of residuals. Essentially equivalent fits were obtained whether Z was fit as a constant or urea dependence was included.

Summation of Unfolding of Fragments. Two types of sums of D1D2 and D3D4 unfolding were generated. One type of sum is the mathematical addition of the f_{app}s for D1D2 and D3D4 unfolding after multiplying each f_{app} by a fractional contribution term determined from observed absorbance and fluorescence changes and tryptophan solvent accessibilities in the crystal structures. For the absorbance sum, f_{app} values for both D1D2 and D3D4 were multiplied by 0.5. For the fluorescence sum, D1D2 contributed 60% while D3D4 contributed 40%. This sum simulates independent unfolding of the two fragments. The other type of sum is a fit of experimental sCD4 unfolding data to the four-state unfolding equation (eq 10) where K_2 is drawn from D1D2 unfolding and K_1 and K_3 come from D3D4 unfolding. The four-state fits were obtained after setting the free energies constant and equal to values from D1D2 and D3D4 fits (G_{1H2O} = 2.47 kcal/mol, G_{2H2O} = 2.51 kcal/mol, G_{3H2O} = 2.62 kcal/mol), letting m_D's only vary by a factor of 2 from the D1D2 and D3D4 values ($0.9 < m_{1D} < 4$, $0.5 < m_{2D} < 2$, $0.5 < m_{3D} < 2$), and letting Z's vary within tight ranges determined from (a) the Z values of D3D4, (b) the solvent accessibilities of tryptophans determined from crystal structures, and (c) the relative absorbance differences observed with D1D2 and D3D4 unfolding (from D3D4: $0.38 < Z_{1A} < 0.48$, $0 < Z_{1F} < 0.01$, $Z_{1A} + Z_{2A} < 0.87$, and $Z_{1F} + Z_{2F} < 0.73$; and from D1D2: $0.36 < Z_{2A} < 0.50$ and $0.57 < Z_{2F} < 0.73$). The data were also fit to the four-state equation with free energy values that were scaled down so the sum of the free energies more closely matched the sum of the free energies measured for sCD4 (G_{1H2O} = 1.78 kcal/mol, G_{2H2O} = 1.89 kcal/mol, G_{3H2O} = 1.81 kcal/mol). The fit, however, was not dramatically improved.

Refolding Measurements. Difference spectroscopy was also used to measure refolding of sCD4, D1D2, and D3D4. Protein samples were incubated in urea, at a concentration which corresponded to ~50% or >90% fraction unfolded on the unfolding curve, for a sufficient time to achieve at least 75% of the change observed for that urea concentration as determined from the unfolding kinetic measurements. The samples were then diluted with buffer to lower the urea to a concentration where <10% of the protein should be unfolded, if the unfolding were reversible. Repetitive absorption or fluorescence emission spectral measurements, using the same spectrophotometric settings as for the equilibrium unfolding measurements, were performed until a steady state was reached. Also, spectra of protein samples, incubated with urea at concentrations equivalent to those before and after dilution, were measured to obtain the starting point and the theoretical infinity point for completely reversible unfolding. Data analysis, except for curve fitting, was the same as for the equilibrium unfolding measurements.

Molecular Graphics. To aid in the interpretation of the absorbance and fluorescence measurements, the molecular graphics program SYBYL (version 6.03; Tripos Associates Inc., St. Louis, MO) was used to examine the location of tryptophans in the crystal structures of human D1D2 and rat D3D4. Their positions relative to possible quenchers such as disulfide bonds, histidines, charged residues, tyrosines, and putative hydrogen bonding partners of the tryptophans also were identified. Coordinates were obtained from Brookhaven Protein Databank files 1CD4 (Ryu et al., 1990) and 2CD4 (Wang et al., 1990) for D1D2, and 1CID (Brady et al., 1993) for D3D4. A Connolly solvent-accessible surface program [Quantum Chemistry Program Exchange (QCPE), Bloomington, IN] interfaced to SYBYL was used with a 1.4 Å radius sphere to determine the solvent-accessible surface area of each tryptophan. To determine the buried surface area at domain interfaces, the sum of the surface area of the appropriate individual domains (D1, D2, D3, and/or D4) was compared to the surface area of the domain pair—D1D2 or D3D4. Four residues in rat D3D4, Arg31, Met122, Gln150, and Gly181, were changed to tryptophans using SYBYL's "biopolymer/change residue" procedure. These residues correspond to tryptophans 214, 306, 334, and 365 in human CD4. The aforementioned measurements/observations were also made for these "mutated" tryptophans.

Differential Scanning Calorimetry. Calorimetry of sCD4 was performed on a MC-2 high-sensitivity differential scanning calorimeter (Microcal, Inc., Northampton, MA). sCD4 samples were dialyzed into 50 mM histidine, pH 7, before being loaded into the calorimeter for heating scans. After calorimetric data were converted to heat capacity, a base line was generated and subtracted; the resulting heat capacity peaks were integrated to determine calorimetric enthalpy. Overlapping peaks were resolved using the Origin software provided by Microcal, Inc., according to three basic models: (a) two-state transitions; (b) independent, non-two-state transitions; and (c) linked or sequential, non-two-state transitions.

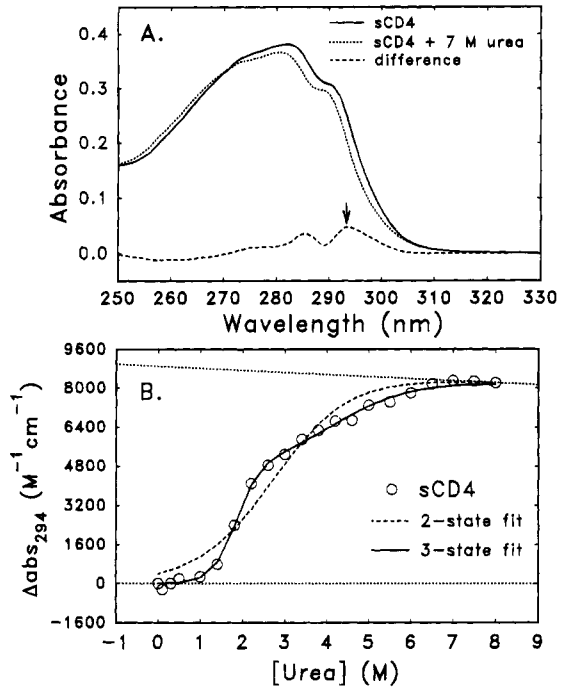

FIGURE 2: (A) Absorption and difference spectra of sCD4. sCD4 in 50 mM sodium acetate, pH 5.5, at 0.255 mg/mL (5.8 μM) was scanned after 39 h incubation at 25 °C with either 7 M urea (dotted line) or in the absence of urea (solid line). The UV difference spectrum (dashed line) was obtained by subtracting the spectrum of sCD4 with urea from the spectrum of sCD4 without urea. The arrow indicates the peak of the difference spectrum that was used to monitor unfolding (≈294 nm). (B) Urea dependence of the absorbance difference. Δabs_{294} values, the peak heights from the sCD4 difference spectrum divided by the molar protein concentration and path length, are plotted versus urea concentration for one set of sCD4 unfolding data. The folded and unfolded base lines shown (dotted lines) were used in fitting the data to the two-state and three-state equations. These two fits are indicated with dashed and solid lines, respectively.

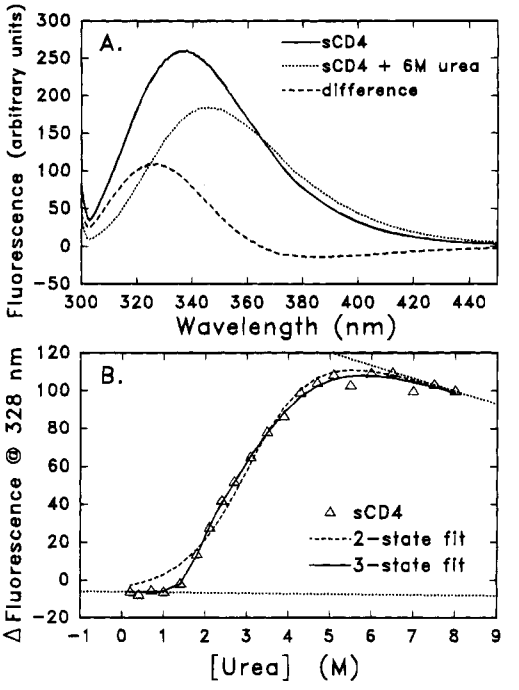

FIGURE 3: (A) Fluorescence emission and difference spectra of sCD4. Sample emission spectra of sCD4 (pH 5.5) at 0.03 mg/mL (0.7 μM) are shown for protein excited at 294 nm and incubated with either 6 M urea (dotted line) or in the absence of urea (solid line). The fluorescence emission difference spectrum (dashed line) was obtained by subtracting the spectrum of sCD4 with urea from the spectrum of sCD4 without urea. (B) Urea dependence of the fluorescence difference. The fluorescence change at 328 nm is plotted versus urea concentration for one set of sCD4 unfolding data. The folded and unfolded base lines shown (dotted lines) were used in fitting the data to the two-state and three-state equations. These two fits are indicated with dashed and solid lines, respectively.

RESULTS

Figure 2A shows typical absorption spectra of human sCD4 in the presence and absence of urea, as well as the difference spectrum obtained by subtraction of these spectra. Similar spectra (not shown) were obtained for the urea-induced unfolding of D1D2 and D3D4. As expected, urea-induced unfolding is accompanied by a decrease in the maximum absorbance and a blue-shift of the absorption spectrum. There are 10 tryptophans and only 3 tyrosines in sCD4 (3 Trp and 1 Tyr in D1D2 and 7 Trp and 2 Tyr in D3D4). Because of the disproportionately high number of tryptophans in sCD4 and the lower absorption coefficient of tyrosines, the tryptophans in sCD4 are expected to contribute ~93% of the absorbance. Therefore, the peaks in the difference spectra were expected to occur at the wavelengths characteristic of tryptophans, 292−295 and 284−286 nm, rather than at 286−289 nm as expected for tyrosines (Demchenko, 1986). This was indeed the case, and unfolding for sCD4, D1D2, and D3D4 was monitored through the Δabs of the peak that occurred at ~294 nm (marked by the arrow in Figure 2A) since this peak was larger than the ~285 nm peak. In Figure 2B, one set of sCD4 unfolding difference data, Δabs_{294} (M^{-1} cm^{-1}), is plotted versus urea concentration.

Figure 3A shows typical fluorescence emission spectra of sCD4 in the presence and absence of urea. The unfolded protein has a lower fluorescence intensity than folded protein, and the peak wavelength shifts from ~337 to ~346 nm, indicating an increase in the polarity of the average environment of the tryptophans upon unfolding. Similar spectra (not shown) were obtained for the urea-induced unfolding of D1D2 and D3D4. The difference spectrum obtained by subtraction is also shown in Figure 3A. The maximum difference in emission intensities for sCD4, D1D2, and D3D4 was between 320 and 330 nm. The results from one sCD4 unfolding experiment measured by fluorescence are shown in Figure 3B.

Two-state and three-state fits using eqs 8 and 9 are shown for the sCD4 unfolding measurements in Figures 2B and 3B. For both the sCD4 absorbance and the fluorescence data, the three-state fits were significantly better than the two-state fits at ≥97.5% confidence levels based on F tests.

In order to put absorbance and fluorescence measurements on the same scale for global fitting and in order to make the comparison of the unfolding of sCD4 and the isolated domains D1D2 and D3D4 easier, the experimental data were converted to apparent fraction unfolded, f_{app}, using eq 7. Fractions calculated from multiple absorbance and fluorescence measurements of sCD4, D1D2, and D3D4 unfolding are shown in Figure 4. Refolding measurements of sCD4, D1D2, and D3D4 (filled symbols in Figure 4) showed that unfolding was largely reversible. Also shown in Figure 4 are global fits, wherein multiple data sets including both absorption and emission data for a given protein were used

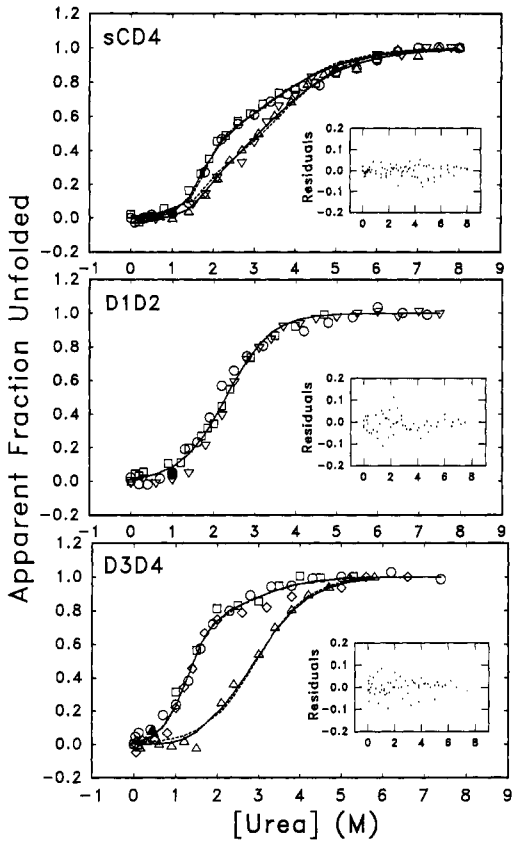

FIGURE 4: Urea unfolding of sCD4, D1D2, and D3D4. Dependence of the fractional changes in Δabs (○, □, and ◇) and Δfluor (△ and ▽) on urea concentration which are equivalent to apparent fractions of non-native material (f_{app}) is shown for multiple data sets obtained on different days. Also shown for sCD4 and D3D4 are the global three-state fits (solid lines) and global modified four-state fits (dashed lines) and the global two-state fit for D1D2 (solid line). Residuals from the global three-state fits (sCD4 and D3D4) and global two-state fit (D1D2) are shown in the graph insets. Filled symbols are fractional changes measured after refolding of sCD4, D1D2, or D3D4 from 6 M urea, 5.5 M urea, or 4.4 M urea, respectively (all >75% unfolded before refolding began).

to determine a single set of thermodynamic parameters while allowing fluorescence and absorbance measurements to have different Z values. The thermodynamic parameters that were determined from these global fits are shown in Table 1. Residuals from the two-state and three-state fits are shown in the graph insets, and they do not show any substantial systematic deviations from zero, indicating that the data are consistent with the fitting equations. A two-state fit is shown for D1D2, since the three-state or modified four-state equations did not significantly improve the fit. If an intermediate is present during D1D2 unfolding, then it is not clearly detectable by absorbance and fluorescence. After comparing two-state, three-state, four-state, and modified four-state fits, the best fits obtained for sCD4 and D3D4 were the three-state fits (solid lines) shown in Figure 4. Statistical analysis indicates that we can be 95% confident that sCD4 unfolding data are more accurately described by the three-state model than by the modified four-state model (dashed lines in Figure 4) while the confidence level for the improved fit of D3D4 data with the three-state equation is only 70%. The improvement of the fit gained for both sCD4 and D3D4 from the use of the three-state compared to the modified four-state equations is, however, subtle. This is probably due to the presence of only one significantly populated intermediate in the four-state model as shown in Figure 5. Four-state unfolding will reduce to three-state unfolding if the two equilibrium constants, K_1 and K_2, are significantly different. In denaturant-induced unfolding, the latter occurs because of significant differences in stability (ΔG_{1H2O} and ΔG_{2H2O}) of the two domains or in the urea dependence of the free energies (m_{1D} and m_{2D}). Three-state unfolding with a single intermediate also arises when an interdomain interaction energy exists since this will cause the stability of one domain to depend on the folded state of the other (Brandts et al., 1989).

In addition to urea unfolding, we also measured **thermal** unfolding of sCD4 at pH 7 by differential scanning calorimetry (DSC) at scan rates of 60 °C/h (Figure 6) and 25 °C/h. At both scan rates, the calorimetric enthalpy was greater than the van't Hoff enthalpy, indicating non-two-state thermal unfolding of sCD4 that proceeds through one or more intermediate(s). This calorimetric measurement of thermal unfolding supports the conclusion drawn from spectroscopic measurements of denaturant-induced unfolding, namely, that sCD4 unfolding is non-two-state. Because the T_m varied with scan rate, we were not able to use the DSC data to determine the thermodynamic parameters of unfolding.

DISCUSSION

Interdomain communication may be important in the function of CD4 since association of extracellular CD4 domains with MHC II and the T-cell receptor (TCR) results in an intracellular response which is believed to be at least partially due to $p56^{lck}$, the tyrosine kinase which is noncovalently associated with the cytoplasmic domain of CD4. Since different CD4 domains have been proposed to interact with MHC II and TCR, cooperativity can best be achieved if the CD4 domains communicate their structural interactions with each other.

Multidomain proteins have domain interactions that fall into three general classes: (1) minimal or nonexistent interactions where domains fold independently [e.g., ovomucoid and papain (Privalov, 1982), immunoglobulin light chain (Tsunenaga et al., 1987), Taka-amylase A (Fukuda et al., 1987)]; (2) very strong interactions where multiple domains unfold in a single cooperative unit and no equilibrium intermediates are detected [e.g., calcium-free parvalbumin (Filimonov et al., 1978) and yeast phosphoglycerate kinase under most conditions (Hu & Sturtevant, 1987; Szpikowska et al., 1994)]; and (3) the intermediate case where domains are linked so that their unfolding characteristics are altered relative to the isolated domains, yet unfolding of the separate domains is distinguishable from inspection of the unfolding curve or upon deconvolution [e.g., diptheria toxin (Ramsay & Friere, 1990), prothrombin (Lentz et al., 1991), B-subunit of cholera toxin (Bhakuni et al., 1991), and molecular chaperone DnaK (Montgomery et al., 1993)]. Thermodynamic measurements of the unfolding of multidomain proteins are a useful means of identifying interdomain interactions and of distinguishing between the three interaction classes. We have chosen to measure the thermodynamics of denaturant-induced unfolding of sCD4 by absorbance and fluorescence spectroscopy and to make limited measurements of thermal unfolding by differential scanning calorimetry. Our thermodynamic measurements of

Table 1: Thermodynamic Parameters for Urea Unfolding[a]

protein	Z	ΔG_{1H2O} (kcal/mol)	m_{1D} [kcal/(mol·M)]	$[urea]_{1/2,1}$ (M)	ΔG_{2H2O} (kcal/mol)	m_{2D} [kcal/(mol·M)]	$urea_{1/2,2}$ (M)
			3-State Fits				
sCD4							
abs & fluor	$Z_A = 0.28 \pm 0.16$ $Z_F = 0 \pm 0.20$	3.74 ± 0.96	2.26 ± 0.64	1.65	1.74 ± 0.48	0.56 ± 0.08	3.11
D3D4							
abs & fluor	$Z_A = 0.75 \pm 0.06$ $Z_F = 0 \pm 0.11$	2.47 ± 0.34	1.93 ± 0.31	1.28	2.62 ± 0.57	0.90 ± 0.15	2.91
			2-State Fit				
D1D2							
abs & fluor		2.51 ± 0.20	1.09 ± 0.08	2.30			

[a] ΔG_{H2O} is the free energy of unfolding in the absence of denaturant, m_D is the urea dependency of the unfolding free energy, and Z is the fractional change in the measured quantity occurring during the native to intermediate unfolding step (Z_A for absorbance and Z_F for fluorescence). The parameters obtained from curve fitting are shown as 95% confidence intervals (assuming a normal distribution of the errors).

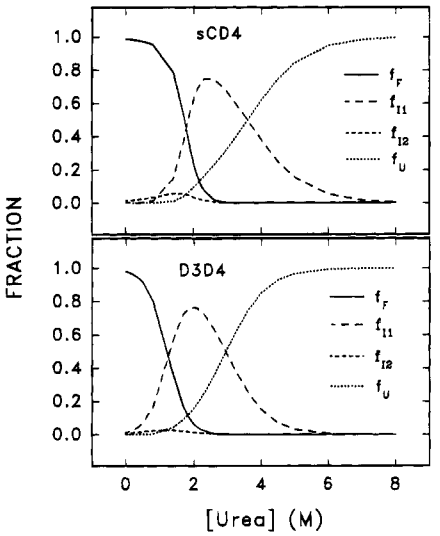

FIGURE 5: Urea dependence of the population of sCD4 and D3D4 states. The upper panel shows fractions of folded (f_F), intermediate 1 (f_{I1}), intermediate 2 (f_{I2}), and unfolded (f_u) sCD4, and the lower panel shows the fractions of these states during D3D4 unfolding. Fractions were calculated using the parameters obtained from the modified four-state fits. In this model, the folded regions in intermediate 1 are unfolded in intermediate 2 and vice versa.

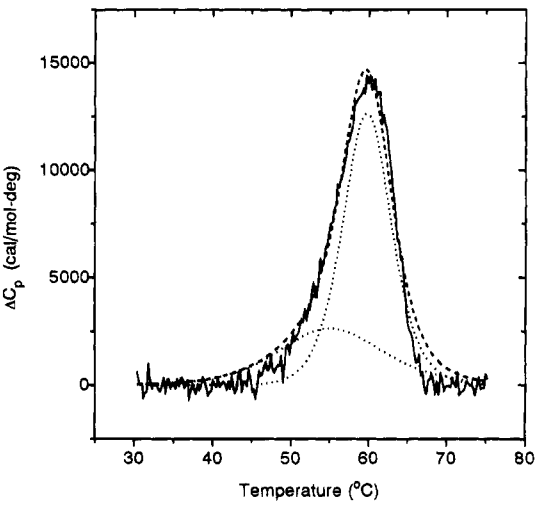

FIGURE 6: Thermal unfolding of sCD4 measured by differential scanning calorimetry. The solid line is the experimental heat capacity profile after base line subtraction and concentration normalization. Protein concentration in this scan is 0.02 mM. The dashed line is the fit of sCD4 data to the non-two-state transition model. The dotted peaks are the two constituent transitions.

sCD4 unfolding show that its domain interactions fall into the third class listed above. This report will add to the few currently in the literature where spectroscopic measurements of denaturant unfolding have been used to detect the unfolding of individual domains of a multidomain protein (e.g., Matthews & Crisanti, 1981; Tsunenaga et al., 1987; Consler & Lee, 1988; Morjana et al., 1993).

One should note that the combination of absorption and emission measurements was necessary to definitively determine the number of unfolding steps and to accurately determine the unfolding parameters of sCD4 and of its isolated domains. For example, D3D4 unfolding measured by fluorescence alone is described fairly well by the two-state equation because of the very low fluorescence change in the formation of the intermediate ($Z_f \approx 0$). Absorbance measurements, on the other hand, indicated that D3D4 unfolding was probably three-state, but, without a global fit that included fluorescence data, the parameters for the second step were poorly defined. When absorbance and fluorescence data were combined for a global three-state fit, the errors associated with the fitted parameters were dramatically reduced, and the appropriateness of applying the three-state model to D3D4 unfolding was confirmed. Similarly, it was obvious from the lack of coincidence of the absorbance and fluorescence data for sCD4 that unfolding of sCD4 was also not two-state, yet the fluorescence data alone resemble two-state unfolding (see Figure 3B) since, as with D3D4, there was a very low fluorescence change for the transition from the folded protein to the intermediate. Although fluorescence data were customarily obtained at a lower concentration than the absorbance data to avoid problems due to inner filter effects, one set of fluorescence measurements at the higher concentration showed the same results as measurements at the lower concentration, i.e., a better fit to the three-state equation. Furthermore, calorimetric measurements of sCD4 at a constant scan rate showed essentially no change in T_m over the concentration range 13−490 μM (unpublished results). This is a sensitive corroboration that aggregation is not occurring and is not responsible for the differences seen in the absorbance and fluorescence measurements.

Structural Interpretation of Unfolding Intermediates. The consistency of the sCD4 and D3D4 unfolding data with the three-state model indicates the presence in each case of an unfolding intermediate. As observed for many multidomain proteins, we expected the formation of an unfolding intermediate for sCD4 to involve unfolding of one or more domains. In order to aid the identification of the intermediates in sCD4 and D3D4 unfolding, the solvent accessibility

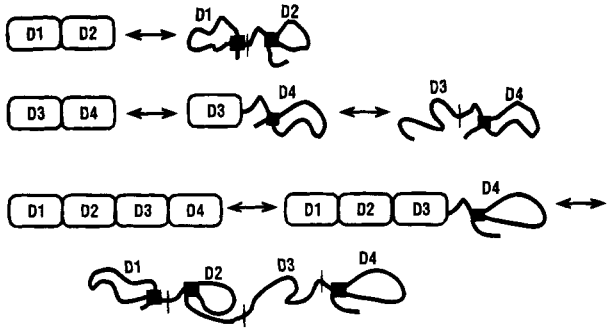

FIGURE 7: Cartoon model of the structural species present in the equilibrium unfolding of D1D2, D3D4, and sCD4. Rounded rectangles represent folded domains, and squiggly curves represent unfolded domains. Disulfide bonds in D1, D2, and D4 are indicated with squares. Short vertical lines mark the boundaries between domains.

of tryptophans and their location relative to other residues which could quench tryptophan fluorescence were examined in the human D1D2 and rat D3D4 crystal structures. For D3D4 unfolding, the majority of the absorbance change occurs during formation of the intermediate ($Z_A \approx 0.75$) whereas none of the fluorescence change occurs in this step ($Z_F \approx 0$). Unfolding of D4 to form the intermediate is consistent with the observed absorbance change, since, by looking at the overall solvent accessibility of tryptophans in the two domains, one would predict 60% of the absorbance change upon unfolding to arise from D4 and 40% to arise from D3. The lack of fluorescence change can also be explained by D4 unfolding, since conserved Trp343, the only buried tryptophan in D4 (8% solvent-accessible surface area) and therefore the expected source of the majority of the unfolding spectroscopic changes, is near a disulfide bond (~10 Å distance in crystal structure, <6 Å after potential segmental motion), and disulfide bonds are known to quench tryptophan fluorescence (Schiller, 1985). Since Trp343 is only two residues away from the disulfide bond (Figure 1), it is possible that the fluorescence is quenched in both the native and unfolded D4, resulting in the observed lack of fluorescence change. Note that D3 has no disulfide bonds nor are there any histidines, also known to quench tryptophan fluorescence, located close in primary sequence and spatial proximity to the buried tryptophan in D3. The unfolding model is schematically shown in Figure 7.

The first step of sCD4 unfolding also lacks a change in fluorescence, suggesting that D4 unfolding could also be the first step of sCD4 unfolding. The other possible quenched tryptophan is the partially buried tryptophan in D2 (33% solvent-accessible surface area), Trp157, which is also within two residues of a disulfide bond. However, since the absorbance change upon unfolding is assumed to be proportional to the sum of the buried tryptophans' surface area, the fractional absorbance change of the first step in sCD4 unfolding is most consistent with the absorbance change expected from D4 unfolding. Therefore, the unfolding data are consistent with the proposal that the intermediate in sCD4 unfolding consists of unfolded D4 and folded D1, D2, and D3. This unfolding model is depicted in the cartoon in Figure 7. This model is speculative, as is the model for D3D4 unfolding, because of the difficulty in correlating the individual environments of a sizable number of tryptophans with overall absorbance or emission changes. Nevertheless, examination of the available information reveals trends which allowed us to develop models which can be experimentally tested.

Because of the location of D4 adjacent to the transmembrane segment which is missing in sCD4, we cannot rule out that a missing interaction of D4 with the membrane surface is the basis for the greater sensitivity of D4 to urea denaturation which results in it being in an unfolded state in the intermediate. D4 itself is not truncated, however, and the crystal structure determination shows that D4 retains a domain structure similar to the other three domains, with greatest similarity to D2 (Brady et al., 1993; Lange et al., 1994).

Published Structural Information on sCD4 Domain Interfaces. Ultracentrifugation measurements give evidence that sCD4 is an extended rodlike molecule that is not bent in the middle since the shape component of the frictional ratio corresponds to an axial ratio of 5.8 to 7.2 in agreement with rough dimensions from sCD4 crystals of 25 × 25 × 125 Å (Kwong et al., 1990) and as expected from the linear end-to-end connection of 25 × (25−35) × (60−65) Å D1D2 (Ryu et al., 1990, 1994; Wang et al., 1990) and 25 × 35 × 60 Å D3D4 (Brady et al., 1993). There are, therefore, apparently no direct contacts between D1 and D3, D1 and D4, or D2 and D4.

sCD4 domains are closely related in structure to immunoglobulin domains, with D1 and D2 being most similar to IgG variable (V) and constant (C) domains, respectively. Although IgG κ light chain composed of a V and a C domain exhibited non-two-state unfolding, with almost independent unfolding of the two domains (Tsunenaga et al., 1987), we did not expect each domain in sCD4 to act independently, because of the difference in the size of the domain interfaces of sCD4 compared to immunoglobulins. For example, the hydrophobic interface between sCD4 D1 and D2 (Ryu et al., 1990) contains 616 Å^2 of buried surface area (determined with the QCPE Connolly routine) while the interdomain interface of V_H and C_H in Fab New is only 220 Å^2 (Saul et al., 1978). The large D1D2 interface appears to be mainly due to the continuous β strand that connects the two domains, a feature not present in an Ig light chain. Since a similar continuous strand connects rat D3 and D4, it is probable that in human sCD4, D3 and D4 possess a similarly large interface region. The interfacial area of rat D3D4 (Brady et al., 1993) is 694 Å^2 (determined with the QCPE Connolly routine). Finally, one model for the end-to-end connection of human D1D2 and rat D3D4 (Brady et al., 1993) buries 600 Å^2 of surface area at the D2/D3 interface. Our spectroscopic and calorimetric results indicate that in spite of these large, similarly-sized domain interfaces and the extended rodlike shape, sCD4 and D3D4 do not unfold in one step, but rather each apparently proceeds through a single thermodynamically stable intermediate.

Thermodynamic Evidence of Domain Interactions. There are several lines of evidence from the unfolding data that point to interactions between adjacent domains.

(A) D1/D2 Interactions. First is the apparent two-state unfolding of D1D2, which suggests that the interaction between D1 and D2 is sufficiently strong to couple their unfolding. It is possible that an intermediate is present in D1D2 unfolding but is not detectable. This would require coincidence or near-coincidence of all the thermodynamic parameters, namely, $\Delta G_{1H2O} \approx \Delta G_{2H2O}$, $m_{1D} \approx m_{2D}$, and $Z_a \approx Z_f$. It is very unlikely, however, that D1 and D2 are

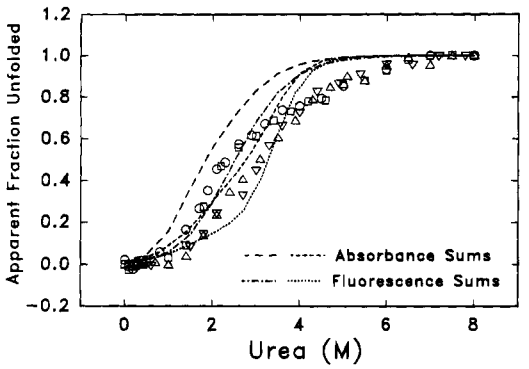

FIGURE 8: Comparison of experimental sCD4 unfolding data [Δabs (○, □, and ◇) and Δfluor (△ and ▽)] with the theoretical sums of D1D2 and D3D4 unfolding data. D1D2 and D3D4 results were summed in two ways. One type of sum shown with long dashes and dash−dot−dashes is simply a weighted sum of f_{app} for D1D2 plus f_{app} for D3D4 (essentially independent unfolding of these two parts). The other sums shown (medium dashed and dotted curves) were four-state fits obtained while the parameters were restrained to ranges determined from D1D2 and D3D4 values as described under Materials and Methods. The fit parameters obtained for the four-state fits shown in this figure are ΔG_{1H2O} = 2.5 kcal/mol, ΔG_{2H2O} = 2.5 kcal/mol, ΔG_{3H2O} = 2.6 kcal/mol, m_{1D} = 1.3 kcal mol^{-1} M^{-1}, m_{2D} = 1.1 kcal mol^{-1} M^{-1}, m_{3D} = 0.5 kcal mol^{-1} M^{-1}, Z_{1A} = 0.48, Z_{2A} = 0.39, Z_{1F} = 0.01, and Z_{2F} = 0.72.

noninteracting domains based on the β-strand connection and the large hydrophobic domain interface identified in the crystal structure.

(B) D2/D3 Interactions. The second line of evidence comes from looking at D1D2 and D3D4 unfolding results for signs of stabilization from an interdomain interaction between D2 and D3, since the fragments lack a domain interface which is present in the intact sCD4 molecule. This ability to compare unfolding with and without the D2/D3 interface allows us to most definitively ascertain the existence of interactions at this interface, of which the least is known from the crystal structure database. Nonadditivity can be seen simply from the fact that sCD4 unfolding is adequately fit to a three-state transition while unfolding with at least four states is predicted from the combination of the two-state unfolding of D1D2 and the three-state unfolding of D3D4. This inequality in the number of states suggests that there is an interaction between D2 and D3 and that D1D2 and D3D4 are not equivalent to the unfolding subunits of sCD4. Without measuring individual unfolding of the actual unfolding subunits of sCD4, the determination of any interaction energy is precluded.

The sum of the unfolding curves of the fragments will equal the unfolding of the whole if there is no interaction (Tsunenaga et al., 1987), but will be different from unfolding in the intact sCD4 molecule if an interaction energy (between D2 and D3) is present or if there are structural rearrangements due to an interaction. Figure 8 shows, along with experimental sCD4 unfolding data, two sums of D1D2 and D3D4 unfolding. The details of how these sums were generated are given under Materials and Methods. One type of sum is the mathematical addition of the f_{app}s for D1D2 and D3D4 unfolding after multiplying by fractional contribution terms. This sum simulates independent unfolding of the two fragments. The second type of sum is a fit to the four-state unfolding equation (eq 10) where one of the three equilibrium constants for this equation is drawn from D1D2 unfolding and two come from D3D4 unfolding. As one can see from Figure 8, unfolding of the isolated parts, D1D2 and D3D4, cannot be added by either method to obtain the unfolding profile of sCD4. These results further demonstrate that interdomain interactions likely exist between D2 and D3 in sCD4 which affect domain stability and/or cause structural rearrangements relative to the isolated components.

The interaction between D2 and D3 is expected to make the D1D2D3 unit in sCD4 unequivalent to isolated D1D2 plus the D3 portion of isolated D3D4. Such an unequivalency is confirmed in Table 1 by the measured thermodynamic parameters. We have proposed that D1D2D3 in sCD4 and D3 in D3D4 are the units which unfold in the I ⇌ U unfolding steps of the respective proteins, so their unfolding is described by m_{2D} and ΔG_{2H2O}. The thermodynamic parameters for D1D2D3 unfolding (sCD4's m_{2D} and ΔG_{2H2O}) were smaller than the corresponding parameters for D1D2 unfolding (D1D2's m_D and ΔG_{H2O}) or for D3 unfolding (D3D4's m_{2D} and ΔG_{2H2O}).

sCD4 proteolysis results have been cited as suggestive of a flexible juncture between D2 and D3 (Kwong et al., 1990; Ryu et al., 1990, 1994). In order to draw this conclusion, however, it would seem necessary to first have multiple reports of proteolysis in the D2/D3 interface by proteases of varying sequence specificity, and such evidence has not been reported. In all the reports of enzymatic digestion of CD4 [with endoproteinase Glu-C (V8), papain, trypsin, and chymotrypsin], multiple proteolysis products are generated, many at sites other than the D2/D3 interface (Richardson et al., 1988; Ibegbu et al., 1989). Only for chymotrypsin has a protelysis site near the D2/D3 interface been reported (Healey et al., 1990), but we have also identified other chymotrypsin proteolysis sites that are cleaved early in the time course of proteolysis in each of the sCD4 domains (D1-Trp62, D1-Leu74, D2-Leu116, D3-Phe229, and D4-Trp334; unpublished observations). A correlation has been seen between sites susceptible to proteolysis attack and segments of the polypeptide chain with higher mobility (Fontana et al., 1986), but there is no thermodynamic reason to equate higher mobility with lower stability, as has been assumed for the D2/D3 interface. Also, it has been shown from studies of the effect of single site mutations on the folding thermodynamics and hydrogen exchange kinetics that hydrogen exchange from the folded state is not correlated with global stability (Kim & Woodward, 1993; Kim et al., 1993). Our results indicate that the D2/D3 interface is more comprehensively described as an interface with substantial interdomain interactions, while the flexible exposed chymotrypsin-sensitive segment at the D2/D3 interface is evidently short.

(C) D3/D4 Interactions. A rationale was given earlier for suggesting that the same domain is unfolded in the intermediates of D3D4 and sCD4 and that this domain is D4. In the linear extrapolation model, the dependence of the free energy of denaturation on denaturant concentration is a constant, m_D. As listed in Table 1, m_{1D} for sCD4 (2.26 kcal mol^{-1} M^{-1}) is very similar to m_{1D} for D3D4 (1.93 kcal mol^{-1} M^{-1}). This fits with our hypothesis that the unfolded domain in sCD4 and D3D4 unfolding intermediates is the same. ΔG_{1H2O} for sCD4 is approximately 1.2 kcal/mol higher than ΔG_{1H2O} of D3D4. This increase in unfolding free energy is indicative of stabilization of D4 by its attachment to D1D2D3. This is not a measure of the D3/D4 interaction energy since the comparison is between two proteins which

contain the D3/D4 interface (sCD4 and D3D4). A D3/D4 interaction must be present, however, in order for D4 to sense the attachment of D1D2 to D3.

CONCLUSIONS

A single equilibrium intermediate is detected during both thermal and denaturant-induced unfolding of sCD4. An analysis of the published structures of D1D2 and D3D4 suggests the intermediate may contain unfolded domain 4 and folded domains 1, 2, and 3. This result is interesting since there are fewer reports for ligand binding sites on D4 or of mutations in D4 and antibody interactions with D4 which affect ligand binding.

We also conclude that there are significant interactions between each of the four domains of sCD4. The interaction energies cannot be quantitated, but we know they are strong enough to result in three-state unfolding of a multidomain protein which would pass through many more states during unfolding if each domain were to unfold independently. In both D1D2 and sCD4, D1 and D2 appear to unfold together. Since unfolding of D1D2 and D3D4 cannot be summed to obtain unfolding of intact sCD4, interdomain interactions likely exist between D2 and D3. Also, the enhanced stability of D4 in sCD4 relative to D4 in D3D4 suggests stabilizing interactions exist between D4 and the rest of the sCD4 molecule which are felt through the D3/D4 interface. The domain interaction energies do not have to be particularly strong in order to couple domains if the domains are similar in stability. But even weak interactions would have biological significance since they could be easily modulated and particularly sensitive to conformational changes. The structure of CD4 with its composition of four extracellular folded domains and consequently its multiplicity of domain connections seems to optimally combine rigidity and limited flexibility in a way that could be very useful for detecting and modulating multiple ligand binding.

In addition to these specifics about sCD4 unfolding, there are a couple of general conclusions we can draw from our measurements. First, we have shown the utility of thermodynamic measurements in providing information on the structural organization of a multidomain protein. This information is particularly useful since, as has been seen from other crystallographic studies, functionally important contacts between domains cannot be quantitated or even consistently ascertained from the crystal structure.

Second, we have also documented the importance of monitoring unfolding by both absorption and emission difference spectroscopy and the importance of applying a global fit which simultaneously determines thermodynamic parameters from both sets of measurements. Such precautions improve the chances of determining the correct thermodynamic model for unfolding.

ACKNOWLEDGMENT

We thank Drs. Lee Shorter and Robin Roman (SmithKline Beecham) for moral and scientific support during the early stages of this work. We also thank Mr. Barry Heard (Southern Research Institute) for his assistance with D3D4 purification and analysis of chymotrypsin proteolysis of sCD4. We also appreciate the assistance of Ms. Zi Yang (University of Alabama at Birmingham) with Figure 1.

REFERENCES

Arthos, J., Deen, K. C., Chaikin, M. A., Fornwald, J. A., Sathe, G., Sattentau, Q. J., Clapham, P. R., Weiss, R. A., McDougal, J. S., Pietropaolo, C., Axel, R., Truneh, A., Maddon, P. J., & Sweet, R. W. (1989) *Cell 57*, 469−481.

Bhakuni, V., Xie, D., & Friere, E. (1991) *Biochemistry 30*, 5055−5060.

Brady, R. L., Dodson, E. J., Dodson, G. G., Lange, G., Davis, S. J., Williams, A. F., & Barclay, A. N. (1993) *Science 260*, 979−983.

Brandts, J. F., Hu, C. Q., Lin, L.-N., & Mas, M. T. (1989) *Biochemistry 28*, 8588−8596.

Burkly, L. C., Olson, D., Shapiro, R., Winkler, G., Rosa, J. J., Thomas, D. W., Williams, C., & Chisholm, P. (1992) *J. Immunol. 149*, 1779−1787.

Carson, M. (1987) *J. Mol. Graphics 5*, 103−106.

Chuck, R. S., Cantor, C. R., & Tse, D. B. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 5021−5025.

Clark, S. J., Jefferies, W. A., Barclay, A. N., Gagnon, J., & Williams, A. F. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 1649−1653.

Clayton, L. K., Hussey, R. E., Steinbrich, R., Ramachandran, H., Husain, Y., & Reinherz, E. L. (1988) *Nature 335*, 363−366.

Clayton, L. K., Sieh, M., Pious, D. A., & Reinherz, E. L. (1989) *Nature 339*, 548−551.

Consler, T. G., & Lee, J. C. (1988) *J. Biol. Chem. 263*, 2787−2793.

Deen, K. C., McDougal, S., Inacker, R., Folena-Wasserman, G., Arthos, J., Rosenberg, J., Maddon, P. J., Axel, R., & Sweet, R. W. (1988) *Nature 331*, 482−484.

Demchenko, A. P. (1986) *Ultraviolet Spectroscopy of Proteins*, pp 65−67 and 247−251, Springer-Verlag, Berlin.

Dianzani, U., Shaw, A., Al-Ramadi, B. K., Kubo, R. T., & Janeway, C. A., Jr. (1992) *J. Immunol. 148*, 678−688.

Eftink, M. R. (1994) *Biophys. J. 66*, 482−501.

Filimonov, V. V., Pfeil, W., Tsalkova, T. N., & Privalov, P. L. (1978) *Biophys. Chem. 8*, 117−122.

Fleury, S., Lamarre, D., Meloche, S., Ryu, S.-E., Cantin, C., Hendrickson, W. A., & Sekaly, R.-P. (1991) *Cell 66*, 1037−1049.

Fontana, A., Fassina, G., Vita, C., Dalzoppo, D., Zamai, M., & Zambonin, M. (1986) *Biochemistry 25*, 1847−1851.

Fukada, H., Takahashi, K., & Sturtevant, J. M. (1987) *Biochemistry 26*, 4063−4068.

Garrett, T. P. J., Wang, J., Yan, Y., Liu, J., & Harrison, S. C. (1993) *J. Mol. Biol. 234*, 763−778.

Healey, D., Dianda, L., Moore, J. P., McDougal, J. S., Moore, M. J., Estes, P., Buck, D., Kwong, P. D., Beverley, P. C. L., & Sattentau, Q. J. (1990) *J. Exp. Med. 172*, 1233−1242.

Hu, C. Q., & Sturtevant, J. M. (1987) *Biochemistry 26*, 178−182.

Ibegbu, C. C., Kennedy, M. S., Maddon, P. J., Deen, K. C., Hicks, D., Sweet, R. W., & McDougal, J. S. (1989) *J. Immunol. 142*, 2250−2256.

Kim, K.-S., & Woodward, C. K. (1993) *Biochemistry 32*, 9609−9613.

Kim, K.-S., Fuchs, J. A., & Woodward, C. K. (1993) *Biochemistry 32*, 9600−9608.

Kwong, P. D., Ryu, S.-E., Hendrickson, W. A., Axel, R., Sweet, R. M., Folena-Wasserman, G., Hensley, P., & Sweet, R. W. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 6423−6427.

Lamarre, D., Ashkenazi, A., Fleury, S., Smith, D. H., Sekaly, R.-P., & Capon, D. J. (1989) *Science 245*, 743−746.

Lange, G., Lewis, S. J., Murshudov, G. N., Dodson, G. G., Moody, P. CE., Turkenburg, J. P., Barclay, A. N., & Brady, R. L. (1994) *Structure 2*, 469−481.

Lentz, B. R., Wu, J. R., Sorrentino, A. M., & Carleton, J. N. (1991) *Biophys. J. 60*, 942−951.

Maddon, P. J., Littman, D. R., Godfrey, M., Maddon, D. E., Chess, L., & Axel, R. (1985) *Cell 42*, 93−104.

Maddon, P. J., Moleneaux, S. M., Maddon, D. E., Zimmerman, K. A., Godfrey, M., Alt, F. W., Chess, L., & Axel, R. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 9155−9159.

Matthews, C. R., & Crisanti, M. M. (1981) *Biochemistry 20*, 784−792.

Mizukami, T., Fuerst, T. R., Berger, E. A., & Moss, B. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 9273−9277.
Moebius, U., Clayton, L. K., Abraham, S., Diener, A., Yunis, J. J., Harrison, S. C., & Reinherz, E. L. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 12008−12012.
Moebius, U., Pallai, P., Harrison, S. C., & Reinherz, E. L. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 8259−8263.
Montgomery, D., Jordan, R., McMacken, R., & Friere, E. (1993) *J. Mol. Biol. 232*, 680−692.
Moore, J. P., & Sweet, R. W. (1993) *Perspect. Drug Discovery Design 1*, 235−250.
Morjana, N. A., McKeone, B. J., & Gilbert, H. F. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 2107−2111.
Pace, C. N. (1986) *Methods Enzymol. 131*, 266−280.
Pace, C. N., Shirley, B. A., & Thomson, J. A. (1989) in *Protein Structure: a practical approach* (Creighton, T. E., Ed.) pp 311−330, Oxford University Press, Oxford, England.
Privalov, P. L. (1982) *Adv. Protein Chem. 35*, 1−104.
Ramsay, G., & Friere, E. (1990) *Biochemistry 29*, 8677−8683.
Richardson, N. E., Brown, N. R., Hussey, R. E., Vaid, A., Matthews, T. J., Bolognesi, D. P., & Reinherz, E. L. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 6102−6106.
Rowe, E. S., & Tanford, C. (1973) *Biochemistry 12*, 4822−4827.
Rudd, C. E., Trevillyan, J. M., Dasgupta, J. D., Wong, L. K., & Schlossman, S. F. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 5190−5194.
Ryu, S.-E., Kwong, P. D., Truneh, A., Porter, T. G., Arthos, J., Rosenberg, M., Dai, X., Xuong, N., Axel, R., Sweet, R. W., & Hendrickson, W. A. (1990) *Nature 348*, 419−426.
Ryu, S.-E., Truneh, A., Sweet, R. W., & Hendrickson, W. A. (1994) *Structure 2*, 59−74.
Santoro, M. M., & Bolen, D. W. (1988) *Biochemistry 27*, 8063−8068.
Saul, F. A., Amzel, L. M., & Poljak, R. J. (1978) *J. Biol. Chem. 253*, 585−597.
Schiller, P. W. (1985) in *The Peptides* (Gross, E., & Meienhofer, J., Eds.) Vol. 7, pp 115−164, Academic Press, Inc., New York.
Schmid, F. X. (1989) in *Protein Structure: a practical approach* (Creighton, T. E., Ed.) pp 251−285, Oxford University Press, Oxford, England.
Szpikowska, B. K., Beecham, J. M., Sherman, M. A., & Mas, M. T. (1994) *Biochemistry 33*, 2217−2225.
Tsunenaga, M., Goto, Y., Kawata, Y., & Hamaguchi, K. (1987) *Biochemistry 26*, 6044−6051.
Veillette, A., Bookman, M. A., Horak, E. M., & Bolen, J. M. (1988) *Cell 55*, 301−308.
Wang, J., Yan, Y., Garrett, T. P. J., Liu, J., Rodgers, D. W., Garlick, R. L., Tarr, G. E., Husain, Y., Reinherz, E. L., & Harrison, S. C. (1990) *Nature 348*, 411−418.
Warren, J. R., & Gordon, J. A. (1966) *J. Phys. Chem. 70*, 297−300.
Williams, A. F. (1987) *Immunol. Today 8*, 298−303.

BI942931R