Bovine Serum Fetuin Is Unfolded through a Molten Globule State†

Changqing Wang,‡ Ioan Lascu,§ and Anna Giartosio*,||

Institute for Structural Biology and Drug Discovery, Virginia Biotechnology Research Park, Richmond, Virginia 23298, University of Bordeaux-2, IBGC-CNRS, 33077 Bordeaux, France, and Department of Biochemical Sciences "A.Rossi Fanelli", University of Rome La Sapienza, and CNR Center of Molecular Biology, Rome, Italy

Received September 15, 1997; Revised Manuscript Received March 23, 1998

ABSTRACT: The reversible heat- and GuHCl-induced unfolding of bovine serum fetuin (BSF) has been studied by differential scanning calorimetry, circular dicroism, tryptophan fluorescence, and size-exclusion chromatography. We show here that thermal unfolding of BSF occurs in two distinct steps corresponding to transitions from the native (N) to an intermediate (I) and from the intermediate to the unfolded state (U). The N↔I transition is highly cooperative and can well be accounted for by a two-state mechanism. The I↔U transition is also cooperative but to a lesser extent than the N↔I transition. CD spectra show that the protein in the I state retains nearly all of the native secondary structure and has a largely disrupted tertiary structure. However, the hydrophobic environment of the single tryptophan residue is not changed, and some compactness is retained in the I state. The structural properties of this intermediate state are apparently characteristic of a molten globule. The GuHCl-induced unfolding is also a two-step process with an I state around 2 M GuHCl. Although the structural features of the denaturant-induced I state are somewhat different from those of the heat-induced I state, the unfolding free energies $\Delta G^\circ_{N\leftrightarrow I}$ and $\Delta G^\circ_{I\leftrightarrow U}$ as well as $\Delta G^\circ_{N\leftrightarrow U}$ obtained from these two methods are comparable. We argue that the observed two-state N↔I transition is due to the melting of the tertiary packing, while leaving quasi-intact the secondary structure and some long-range interactions in the I state. These long-range interactions, together with the secondary structural elements, would be responsible for the observed cooperativity of the I↔U transition.

Proteins may unfold through intermediate(s) which retain(s) a substantially high content of secondary structure but little or no tertiary structure known as "molten globules" (*1*, *2*). The molten globule has attracted much attention in recent years not only because it provides clues for understanding the classical two-state unfolding mechanism observed for many proteins but also because it is believed to be identical to the partially folded intermediate transiently accumulated in the early stage of folding (*1*−*4*). However, two basic questions regarding the molten globule state have not yet been answered. First, is the molten globule a state in which the protein is partly unfolded, i.e., some domains are unfolded while others remain native, or is it a state in which the global native tertiary structure is lacking, i.e., in a nativelike "format" with fluctuating side chains? There has not been a clear delineation between these two different states. The recent argument that classical examples of molten globule proteins (α-lactalbumin, equine lysozyme, staphylococcal nuclease, and apomyoglobin) are in fact in a partly folded form even made questionable the real existence of the molten globule state (*5*). Second, is the molten globule a state thermodynamically distinct from the unfolded one? For some proteins, it undergoes a cooperative two-state transition, indicating that it is thermodynamically separated from the unfolded state (*6*−*9*). For others, its unfolding was shown to be a gradual process without detectable heat absorption (*10*−*11*). Theoretical calculations also led to different conclusions (*12*−*14*). The ambiguity comes from the fact that the molten globule state was usually observed under extreme conditions (e.g., low or high pH, high temperature, or moderate concentration of denaturant) and no detailed thermodynamic analyses coupled with the observed structural properties could be performed.

BSF[1] is a single-chain glycoprotein abundant in calf serum. Each protein molecule contains six carbohydrate chains (*15*) which comprise 24% of its total molecular weight (48 000). We have previously shown that this protein is thermally denatured through a two-state transition with very low specific enthalpy (*16*). We suggested that this was due to the incomplete unfolding of this protein under the conditions studied. In the present work, we extended our study by performing experiments over a wider range of temperature, pH, and GuHCl concentration. We found in the unfolding process an intermediate state which has the characteristics of the molten globule. With calorimetric, optical, and chromatographic methods combined with a proper thermo-

† Supported by Marie Curie Research Fellowship 930174 from the European Union to C.W. and by MURST (Ministero dell'Università e della Ricerca Scientifica e Tecnologica, Italy).

* Correspondence should be addressed to this author at the Dipartimento di Scienze Biochimiche "A.Rossi Fanelli", Università La Sapienza, Piazzale A.Moro 5, 00185 Roma, Italy. FAX: 39-6-4440062. Telephone: 39-6-49910576. E-mail: giartosio@axcasp.caspur.it.

‡ Institute for Structural Biology and Drug Discovery.

§ University of Bordeaux-2.

|| University of Rome La Sapienza, and CNR Center of Molecular Biology.

[1] Abbreviations: BSF, bovine serum fetuin; GuHCl, guanidine hydrochloride; CD, circular dichroism; DSC, differential scanning calorimetry; SEC, size-exclusion chromatography.

S0006-2960(97)02301-5 CCC: $15.00
Published on Web 05/21/1998

dynamic analysis, we have structurally and energetically characterized such a state.

MATERIALS AND METHODS

Protein and Instrumentation. BSF (grade VII) was purchased from Sigma and further purified as described (*15*) or by gel filtration on Sephacryl S200 HR. A stock solution of 10 mg/mL protein was prepared in 50 mM sodium phosphate, pH 6.0. All experiments were performed in this buffer if not otherwise stated. An extinction coefficient of $E_{278} = 5.3$ for a 1% solution (*17*) was used for determining protein concentration.

DSC experiments were performed with a MicroCal MC-2 differential scanning microcalorimeter (MicroCal Inc., Nothhampton, MA) as described previously (*16*). Protein concentration was 6 mg/mL, and scan speed was 60 °C/h. Reversibility of thermal unfolding was checked by repeated scans of the sample. CD spectra were recorded using a JASCO J-710 spectropolarimeter (JASCO) with a quartz cuvette of 1 mm and 1 cm path, respectively, in the far- and near-UV region. Protein concentration was 0.15 and 2 mg/mL, respectively. Each sample was scanned 5−10 times with a scan speed of 20 nm/min and bandwidth of 1 nm. The scans for each sample were averaged and corrected by subtracting a buffer base line. Fluorescence measurements were performed with a Spex Max spectrofluorometer (Spex Industries, Edison, NJ) on solutions of 0.05 mg/mL protein. Fluorescence spectra were recorded between 300 and 430 nm with an excitation wavelength of 280 nm. Before making measurements of denaturant-induced unfolding, protein samples were incubated at defined denaturant concentrations at 4 °C overnight to achieve equilibrium. The reversibility of GuHCl-induced unfolding was rigorously checked by performing renaturation studies monitored with CD and fluorescence methods. The protein was first unfolded in a buffer containing 50 mM sodium phosphate and 8 M GuHCl, pH 6.0, at 4 °C overnight. Refolding was initiated by diluting the unfolded protein into the same buffer without GuHCl. The refolded samples were prepared so that the final protein concentration was 0.15 and 0.05 mg/mL, respectively, for far-UV CD and fluorescence, while the GuHCl concentration varies between 0.32 and 5.76 M. SEC experiments were performed with the BioRad BioLogic system on a Superose 12 column. The column was calibrated with proteins of known molecular weight and preequilibrated with various concentrations of GuHCl; 100 μL of 5 mg/mL native or GuHCl-denatured protein solution was applied to the column and eluted at 0.5 mL/min at 25 °C to study unfolding or refolding, respectively. Unfolding of the protein was found completely reversible by overlapping of the denaturation curve with the renaturation curve.

Data Analyses. DSC thermograms were analyzed using the "Origin" software provided by MicroCal. van't Hoff enthalpy, ΔH_{vH}, was determined according to the equation:

$$\Delta H_{vH} = \frac{4RT_m^2 C_{pm}}{\Delta H} \quad (1)$$

where R is the gas constant, C_{pm} the excess heat capacity at the half transition temperature T_m, and ΔH the calorimetric enthalpy. Unfolding free energy extrapolated to 25 °C, $\Delta G°$, was calculated using the equation:

$$\Delta G(T) = \Delta H\left(1 - \frac{T}{T_m}\right) + \Delta C_p\left[T - T_m - T\ln\left(\frac{T}{T_m}\right)\right] \quad (2)$$

where ΔC_p is the difference in heat capacity before and after the thermal transition.

Raw data for heat-induced unfolding monitored by 285 and 220 nm CD and by the fluorescence emission maximum wavelength were converted to the apparent fraction of native protein, F_N, versus temperature (Figure 4) according to the equation:

$$F_N = \frac{(\theta_U + m_u T) - \theta}{(\theta_U + m_u T) - (\theta_N + m_N T)} \quad (3)$$

where θ is the spectroscopic property being measured at temperature T, and θ_N and θ_U are the intercepts and m_N and m_U the slopes of the pre- and posttransitional base lines of the raw data, respectively. The F_N-T plot was then fitted to a two-state model using the equation:

$$F_N = \frac{1}{1 + e^{[\Delta H(T/T_m - 1) + \Delta C_p(T_m - T - T\ln(T/T_m))]/RT}} \quad (4)$$

Initial estimates of T_m, ΔH, and ΔC_p were made according to values of DSC experiments. $\Delta G°$(25 °C) was calculated from eq 2.

Raw data curves of GuHCl-induced equilibrium unfolding monitored at 220 nm CD (Figure 5), fluorescence intensity, and maximum emission wavelength (Figure 6) were fit to a two-state model using the equation described by Santoro and Bolen (*18*):

$$\theta = \frac{(\theta_N + m_N[D]) + (\theta_U + m_U[D])e^{-[(\Delta G°(H_2O) + m[D])/RT]}}{1 + e^{[(\Delta G°(H_2O) + m[D])/RT]}} \quad (5)$$

where θ, θ_N, θ_U, m_N, and m_U were defined as above. $\Delta G°$-(H_2O) is the free energy in the absence of denaturant, m the slope of the unfolding curve, and [D] the GuHCl concentration. All data sets were also tried to fit to a three-state model as described in ref *19*. Inclusion of an intermediate in the fitting procedure did not improve the fitting results. We assume therefore that each curve represents a two-state transition.

Data conversion and curve fitting were performed with the program JMP 3.1.5 (SAS Institute, Inc., Cary, NC) in the computer center of Virginia Commonwealth University, Richmond, VA.

RESULTS

DSC-Monitored Thermal Unfolding. In a previous work, we have shown that BSF is thermally denatured through a two-state transition with a specific enthalpy considerably lower than those reported for other small globular proteins (*16*). This could be explained by the incomplete unfolding of this protein under our conditions. In the present work, we extended the temperature range in which the thermal unfolding was performed. Figure 1a shows the DSC curve of BSF in 50 mM phosphate buffer, pH 6.0. After the main transition at 59.5 °C, the excess heat capacity does not remain

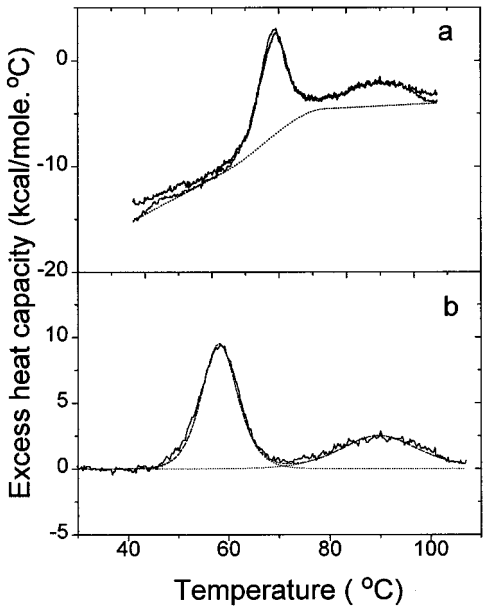

FIGURE 1: Differential scanning calorimetry of BSF in 50 mM sodium phosphate, pH 6.0. Protein concentration is about 5 mg/mL; scan speed is 60 °C/h. Panel a: Experimental curve (solid line) and base line (dotted line) drawn as described in the text. The dashed line represents a second scan of the sample cooled to 4 °C after heating to 105 °C in the first scan. Panel b: Experimental curve after subtraction of the base line (solid line) and deconvolution of the excess heat capacity function into two two-state transitions (dotted lines).

constant as expected for a completely unfolded state, but increases with increasing temperature, resulting in a second broad heat absorption peak in the temperature range 80−110 °C.

A broad diffuse heat absorption peak has been observed for unfolding of the molten globule (*7*, *9*, *20*) or molten globular-like state (*21*, *22*) of many proteins. We were interested in investigating if this intermediate state is a heat-induced molten globule and what structural and energetic properties it might possess.

A reliable thermodynamic analysis of DSC data requires that the unfolding process be reversible and not kinetically controlled. This was checked by repeated scans and by changing the protein sample concentration. We found that both transitions are highly reversible in that they recovered nearly all the peak area from the first scan. This is true even for a third cycle of scan. Changing the sample concentration from 1.0 to 8.0 mg/mL does not affect either the transition enthalpy (ΔH) or the transition temperature (T_m), indicating that the unfolding process is not kinetically controlled.

Deconvolution of the heat capacity function needs a proper base line to be drawn. A key parameter in drawing the base line is the C_p value of the protein in the I state. As this value will influence ΔC_ps of both N↔I and I↔U transitions, we have determined it more accurately by performing a series of DSC studies in the presence of different concentrations of GuHCl and at different pHs. Both $\Delta H_{N\leftrightarrow I}$ and $T_{mN\leftrightarrow I}$ change with GuHCl concentration. Plotting ΔH_{NvI} versus $T_{mN\leftrightarrow I}$ gives a linear function of $\Delta H_{N\leftrightarrow I} = -68.9 + 2.7T_{mN\leftrightarrow I}$ (Figure 2). The slope of this plot, 2.7 kcal/(mol·°C), represents the difference between the heat capacities of the protein in the N and I states ($\Delta C_{pN\leftrightarrow I}$). Using this value and extrapolating the heat capacities of the native (10−45 °C) and unfolded (95−110 °C) states into the transition region, a progress base line is drawn (Figure 1a).

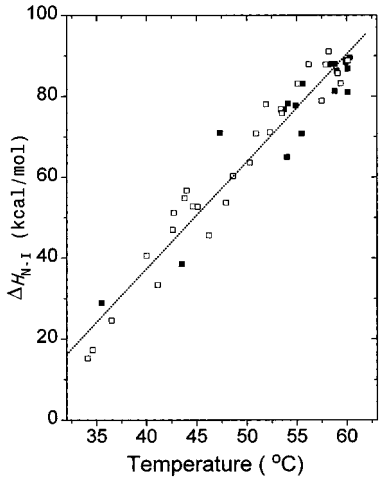

FIGURE 2: Dependence of N↔I transition enthalpy, $\Delta H_{N\leftrightarrow I}$, on transition temperature, $T_{mN\leftrightarrow I}$. Data pairs were obtained from DSC experiments in different concentrations of GuHCl (□) and at different pHs (■). The dotted line is the linear fit of the two data sets.

Deconvolution was done by fitting the curve to a minimal number of two-state transitions after subtracting the progress base line. The fitting results are shown in Figure 1b and the thermodynamic parameters listed in Table 1. From Table 1 we see that the first peak fits to a two-state transition quite well. The difference between the calorimetric and van't Hoff enthalpies does not exceed 3%. This result is consistent with our previous study which also showed a two-state transition. In contrast, the second transition can only be approximated to a two-state model. The fitting error is greater in respect to the N↔I transition, exceeding 15% (Table 1). This may reflect the fact that greater fluctuations in $T_{mI\leftrightarrow U}$ and $\Delta C_{pI\leftrightarrow U}$ affect the calculations of $\Delta H_{I\leftrightarrow U}$, ΔH_{vH}, and hence the $\Delta H_{I\leftrightarrow U}/\Delta H_{vH}$ ratio.

Protein stability was defined by Becktel and Schellman (*23*) as a function of the free energy of unfolding, ΔG, versus temperature. At a given temperature T, ΔG was calculated according to eq 2. Free energy values extrapolated to 25 °C, $\Delta G°_{N\leftrightarrow I}$ and $\Delta G°_{I\leftrightarrow U}$, are given in Table 1.

CD and Fluorescence-Monitored Thermal Unfolding. Thermal unfolding of BSF was monitored by circular dicroism spectroscopy in both the far- and near-UV ranges, and by the intrinsic fluorescence of the single tryptophan of the protein. Figure 3a shows the CD spectra in the wavelength range of 190−250 nm at 25 and 70 °C. The spectrum of native protein has a minimum at 208 nm and a small shoulder at 220 nm. The essential absence of the 222 nm minimum suggests that the protein has a low α-helix content. At 70 °C, which corresponds to the posttransitional temperature of the N↔I transition on DSC scans, neither the form nor the magnitude of the spectrum is substantially changed upon heating. This indicates that the heat-induced N↔I transition observed on DSC does not lead to destruction of the secondary structure of BSF. Included in Figure 3a is also a spectrum of the protein in the presence of 7 M GuHCl, showing complete unfolding.

Figure 3b shows the CD spectra in the near-UV range. The spectrum of native protein (25 °C) is characterized by a negative band with a minimum at 269 nm and two maxima

Table 1: Thermodynamic Parameters Determined from DSC and Near-UV CD Experiments

	$T_{mN\leftrightarrow I}$ (°C)	$\Delta H_{N\leftrightarrow I}$ (kcal/mol)	$\Delta C_{pN\leftrightarrow I}$ [kcal(mol·°C)]	$\Delta G°_{N\leftrightarrow I}$ (kcal/mol)	$\Delta H_{N\leftrightarrow I}/\Delta H_{vH}$	$T_{mI\leftrightarrow U}$ (°C)	$\Delta H_{I\leftrightarrow U}$ (kcal/mol)	$\Delta C_{pI\leftrightarrow U}$ [kcal/(mol·°C)]	$\Delta G°_{I\leftrightarrow U}$ (kcal/mol)	$\Delta H_{I\leftrightarrow U}/\Delta H_{vH}$
DSC[a]	59.5 ± 0.5	87.7 ± 4.3	2.66 ± 0.18	4.07 ± 0.24	1.02 ± 0.14	91.5 ± 2.1	45.7 ± 3.7	0.78 ± 0.25	3.29 ± 0.31	1.15 ± 0.32
CD[b]	58.9 ± 1.2	85.0 ± 8.2	2.70 ± 0.22	3.84 ± 0.31	1.0					

[a] DSC parameters, $T_{mN\leftrightarrow I}$, $\Delta H_{N\leftrightarrow I}$, $T_{mI\leftrightarrow U}$, $\Delta H_{I\leftrightarrow U}$, and $\Delta C_{pI\leftrightarrow U}$, were obtained from deconvoluting eight independent experimental curves; $\Delta C_{pN\leftrightarrow I}$ was determined from the $\Delta H_{N\leftrightarrow I}-T_m$ plot (Figure 2). $\Delta G°_{N\leftrightarrow I}$ and $\Delta G°_{I\leftrightarrow U}$ were calculated according to eq 2. [b] CD parameters, $T_{mN\leftrightarrow I}$, $\Delta H_{N\leftrightarrow I}$, and $\Delta C_{pN\leftrightarrow I}$, were determined from four experiments by fitting the $\theta_{285}-T$ curves to the two-state transition; $\Delta G°_{N\leftrightarrow I}$ was calculated according to eq 2.

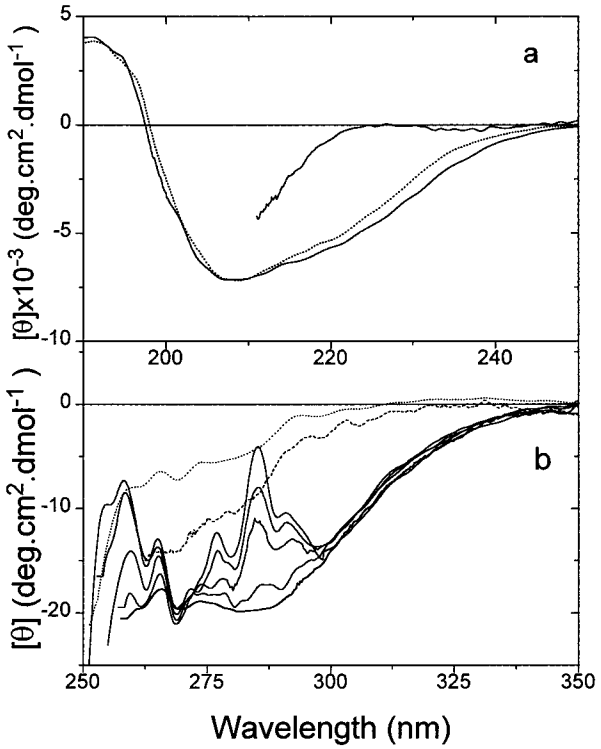

FIGURE 3: Heat effect on CD spectra of BSF in the far- and near-UV regions. Panel a: Far-UV spectrum of native protein at 25 °C (dotted line), the intermediate state at 70 °C (solid line), and the unfolded state in 7 M GuHCl (short solid line). Panel b: Near-UV spectra at different temperatures. From top to bottom: 25, 55, 59, 65, and 73 °C. Dotted line: sample preheated to 100 °C for 2 min and measured at 80 °C; dashed line, protein in 7 M GuHCl (see text).

at 258 and 285 nm. The negative band results from the asymmetric interactions of the protein tryptophan, tyrosines, and disulfide bonds. Heating the sample decreases both the 258 and the 285 nm maxima. At 25 °C, the molar ellipticity at 285 nm is − 4.1 deg·cm²·dmol⁻¹. At 55 °C, it is lower than this value, suggesting the onset of disruption of the tertiary structure. With increasing temperature, the 285 nm signal decreases sharply and reaches a constant value (−21.3 deg·cm²·dmol⁻¹) at about 73 °C. The parallel decrease of the 258 and 285 nm signals results in loss of the main CD structural bands and makes the spectrum monotonic in this region. Included in Figure 3b are also a spectrum of the sample taken at 80 °C, after being kept at 100 °C for 2 min, and a spectrum taken at 25 °C in 7 M GuHCl.

Figure 4 shows the apparent fraction of protein in the native state, F_N, monitored at 285 and 220 nm CD, and by the maximum wavelength of fluorescence emission, as a function of temperature. The 285 nm data set can well be fitted to a two-state model according to eq 4. The derived parameters $T_{mN\leftrightarrow I}$ and $\Delta C_{pN\leftrightarrow I}$ as well as the calculated $\Delta G°$-(25 °C) are given in Table 1 for comparison with those

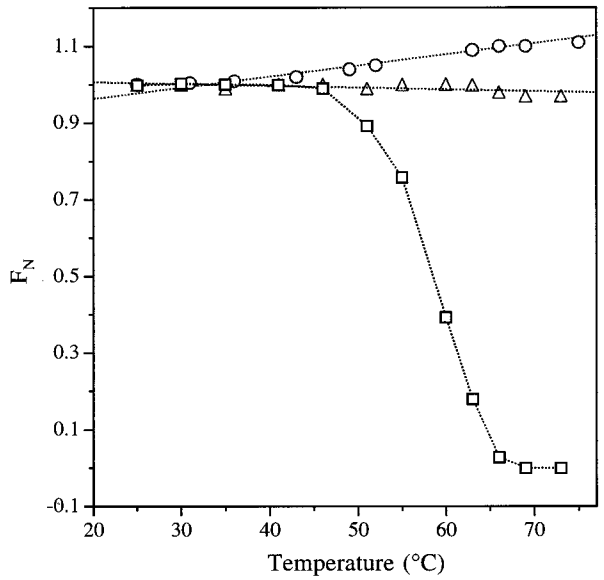

FIGURE 4: Apparent fraction of the protein in the native form monitored by ellipticity at 285 nm (□) and 220 nm (○) and by fluorescence emission maximum wavelength (△), as a function of temperature. The 285 nm data set was fitted to a two-state transition (dotted line) according to eq 4, as described under Materials and Methods. The straight dotted lines for the 220 nm CD and fluorescence data sets are shown only for visual effect.

obtained from DSC studies. Inspection of the data reveals close similarity between the parameters obtained from DSC and near-UV CD experiments. In contrast, the molar ellipticity at 220 nm does not change significantly in this temperature range. This indicates that the first transition observed on DSC scans is a transition from the native to an intermediate state, which involves only the melting of the tertiary structure but not of the secondary structure. Interestingly, the N↔I transition does not lead to exposure of the single tryptophan residue to a less hydrophobic environment, as judged by the constancy of the fluorescence emission maximum. The interactions responsible for such a wavelength constancy will be discussed later.

GuHCl-Induced Unfolding Monitored by CD at 25 °C. GuHCl-induced unfolding monitored in the near-UV region appears to have a complicated pattern. The negative band decreases in the GuHCl concentration range of 0−1.0 M, increases in the range of 1.0−2.0 M, and levels off at 2.5 M (Figure 5, top panel). The decrease in the negative band at low GuHCl concentration is probably caused by binding of the GuHCl ions to the protein molecules which results in a more intense interaction between the aromatic residues (*8*). The absence of a sigmoidal pattern of the curve prevents a detailed analysis. However, C_m, at which the transition is half-completed, can be estimated to be equal or less than 1.5 M. Beyond 2 M GuHCl, the tertiary structure of the protein seems to be completely destroyed.

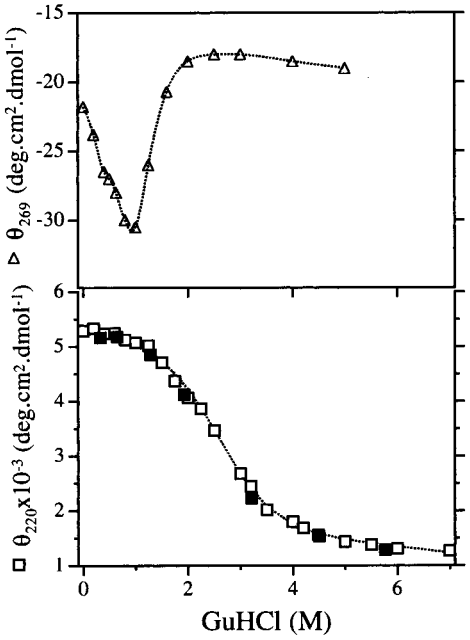

FIGURE 5: GuHCl-induced unfolding of BSF in 50 mM phosphate, pH 6.0, monitored by CD. Top panel: Near-UV data at 269 nm (△). The dotted line is only for visual effect and does not conform to any model. Bottom panel: Far-UV data for unfolding (□) and refolding (■) at 220 nm. The dotted line shows the two-state curve fitted as described under Materials and Methods.

Table 2: Thermodynamic Parameters for GuHCl-Induced Unfolding of BSF

	C_m[b](M)	m[a] [kcal/(mol·M)]	$\Delta G°_{N\leftrightarrow I}$[a] (kcal/mol)	$\Delta G°_{I\leftrightarrow U}$[a] (kcal/mol)
fluorescence	1.5 ± 0.2	−2.4	3.7 ± 0.3	
emission max wavelength	2.5 ± 0.2	−1.2		3.0 ± 0.3
near-UV CD	≤1.5[c]			
far-UV CD	2.5 ± 0.1	−1.2		3.0 ± 0.2
SEC	≥2.1[c]			

[a] Values obtained from fitting the raw datacurves in Figure 5 and Figure 6 to a two-state model according to eq 5. [b] Calculated from the equation: $C_m = -\Delta G°(H_2O)/m$. [c] Estimated value. Parameters were obtained from 5 independent experiments for fluorescence and 3 experiments for CD.

Unfolding of the secondary structure was monitored in the 190−250 nm region in the GuHCl concentration of 0−7 M. The molar ellipticity at 220 nm is plotted as a function of GuHCl concentration in Figure 5, bottom panel. The raw data curve was fitted to a two-state model, and the values of C_m, m, and $\Delta G°(H_2O)$ are given in Table 2. Comparison of the C_m values of the near- and far-UV CD curves reveals that the unfolding of tertiary structure precedes that of secondary structure by 1.0 M GuHCl. The noncoincidence of the two transitions clearly indicates that an intermediate exists during unfolding.

GuHCl-Induced Unfolding Monitored by Fluorescence. Fluorescence intensity and maximum emission wavelength of the tryptophan residues of proteins are sensitive probes for protein denaturation. BSF contains a single tryptophan at position 51 (*24*). The fluorescence emission spectrum of BSF at an excitation wavelength of 280 nm is dominated by tryptophan emission and has a maximum emission wavelength of 340 nm. Figure 6 shows the fluorescence intensity change and emission maximum wavelength shift with

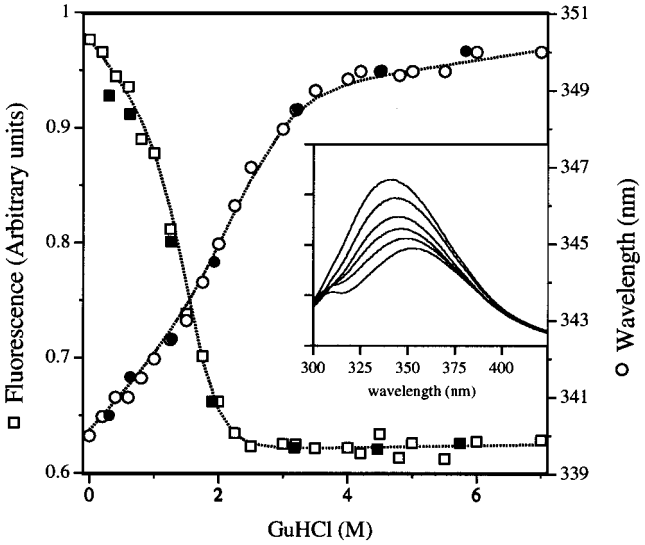

FIGURE 6: Intrinsic tryptophan fluorescence (□) and maximum emission wavelength (○) as a function of GuHCl concentration. The raw data sets were fitted to two-state models according to eq 5. The filled symbols represent data obtained by refolding as described under Materials and Methods. Inset: Fluorescence spectra of BSF at different concentrations of GuHCl. From top to bottom: 0, 0.4, 1.0, 1.25, 2.5, and 7.0 M.

increasing GuHCl concentration. Addition of GuHCl results in a decrease in emission intensity and a concomitant red shift of the maximum wavelength. The intensity reaches a minimum at about 2.5 M GuHCl while the emission maximum continues to shift to longer wavelengths till it levels off at and beyond 4 M GuHCl.. Curve fitting according to a two-state model described in eq 5 gave C_m values of 1.5 and 2.5 M, respectively, for the fluorescence and peak shift curves. The two different C_m values suggest they monitor different unfolding events. This should not be true if only tryptophan contributes to the fluorescence spectrum, but BSF contains 18 other aromatic residues that could influence the observed emission spectrum. The noncoincidence of the two C_m values again suggests the existence of an intermediate.

GuHCl-Induced Unfolding Monitored by SEC. A protein in the unfolded state has a much larger molecular dimension than in the native state. The molten globule state has an expanded dimension but retains substantial compactness. Changes in molecular size can easily be detected with a FPLC system. This technique has been used to characterize the intermediate state and to differentiate the molten globule from the unfolded state (*25*). Figure 7 shows the GuHCl dependence of elution volume of BSF at 25 °C. Increase in GuHCl concentration leads the protein to elute with smaller elution volume, i.e., with larger molecular size. Only one protein peak was detected, its elution volume decreasing with increasing GuHCl concentration. This is indicative of a reversible transition, fast on the time scale of the SEC experiment. The initial increase in the elution volume is not due to alteration of the column properties (GuHCl concentrations were varied in a random fashion to avoid this type of artifact) but is due to shrinking caused by binding of GuHCl ions to the protein molecules, the same phenomenon observed in the near-UV CD experiments (see above). Unfolding occurs between 1 and 4 M GuHCl. The curve does not show an evident biphasic type as expected for a

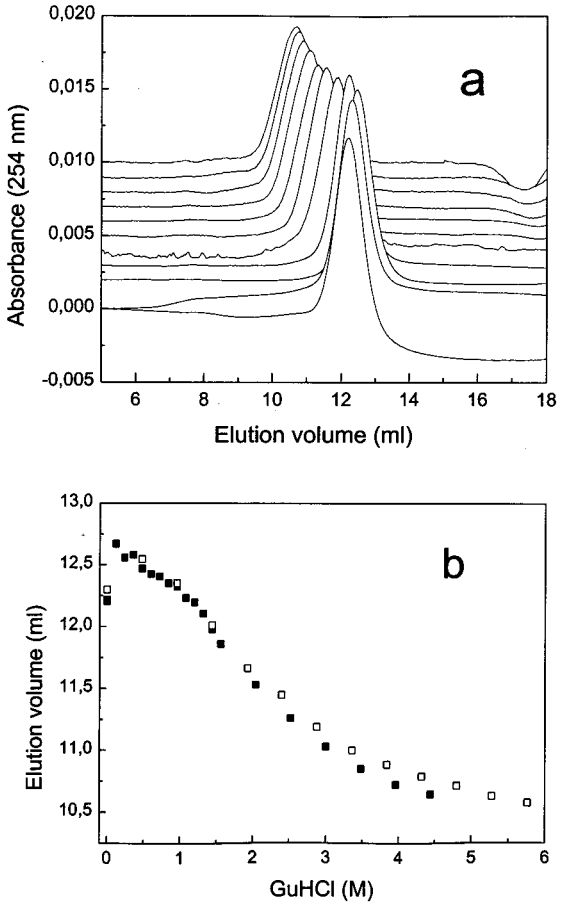

FIGURE 7: BSF denaturation by and renaturation from GuHCl studied by SEC. Panel a: Elution profiles from a Superose 12 column during denaturation. 100 μL of BSF was applied to the column equilibrated with various concentrations of GuHCl in 50 mM phosphate, pH 6.0 at 25 °C. GuHCl concentration was between 0 M (bottom curve) and 4.4 M (top curve) in steps of 0.48 M. The absorbance offset is for clarity. Panel b: Elution volumes obtained during denaturation (filled symbols) and renaturation (open symbols) as a function of GuHCl concentration.

three-state unfolding. This is not surprising because a molten globule may have a nativelike compactness and the difference in molecular dimensions between the native and the molten globule states might be too small to be detected by SEC. These problems occurring at low denaturant concentration prevent a global theoretical analysis of the denaturation process. Nevertheless, it can be clearly seen that the major part of molecular expansion occurs beyond 2 M GuHCl, at a concentration where the tertiary structure is completely lost. A compact intermediate lacking tertiary packing is consistent with a molten globule.

DISCUSSION

Heat-Induced Molten Globule State and Its Structural Features. Thermal unfolding of BSF apparently occurs in two steps as demonstrated by the presence of two distinct heat absorption peaks on the DSC thermograms. The first transition occurs at 59.5 ± 0.5 °C with an enthalpy of 87.7 ± 4.3 kcal/mol. These values are practically identical with those determined in the previous study at pH 6 (*16*). The second transition occurs at a much higher temperature (91.5 ± 2.1 °C) than the first one and was not detected previously because the scans were stopped at lower temperature.

The two well-separated thermal transitions suggest that an intermediate state is highly populated at temperatures between the two transitions. Two structural aspects of this I state should be considered. First, CD spectra in the far-UV region indicate that the I state retains most, if not all, of the native secondary structure. At 70 °C where the first transition is completed, the spectral intensity does not decrease as expected for the unfolding of a protein but slightly increases instead (Figures 3a and 4). A more pronounced CD signal in the peptide region upon transition to the molten globule state has been observed for other proteins (*2*, *26*). It was attributed to clustered aromatic residues which contribute not only to spectra in the near-UV region but also to spectra in the far-UV region (*27*, *28*). Second, the I state retains some higher order interactions although its global native tertiary structure is lacking. This is indicated by the CD spectrum in the near-UV region at 70 °C. At this temperature, the spectrum loses its fine structure with abolishment of the two spectral maxima at 258 and 285 nm but still retains substantial intensity in this region (Figure 3b). This spectrum is surely not that of a protein in the completely unfolded form because it can further be decreased and abolished by heating to higher temperature or by addition of a high concentration of GuHCl (Figure 3b, spectra at 80 °C and 7 M GuHCl). Consistent with this result is the fluorescence study. The native protein has a maximum emission wavelength of 340 nm; the N↔I transition apparently does not lead to a greater exposition to the solvent (Figure 4). The substantial near-UV CD signal as well as the constant emission maximum wavelength at 70 °C strongly suggests the existence of some tertiary packing, at least around the tryptophan and some of the aromatic residues. The observation of such a molten globule state with some tertiary packing is not unusual. Although the classical molten globule state is defined as lacking native tertiary structure, recent studies have clearly shown that some native tertiary interactions exist in the molten globule state of proteins such as myoglobin (*29*), cytochrome *c* (*30*), and equine lysozyme (*31*, *32*).

GuHCl-Induced Intermediate. GuHCl-induced unfolding is also a two-step process manifested by the noncoincidence of C_m values obtained by different optical methods. The first step (N↔I transition) occurs at 1.5 M GuHCl and is characterized by a decrease in fluorescence intensity, a red shift of the emission maximum (Figure 6), and the abolishment of native tertiary structure (Figure 5a). The second step (I↔U transition) involves a further shift of the emission maximum (Figure 6), a progressive decline of far-UV ellipticity (Figure 5b), and a large expansion of the molecular dimension (Figure 7). It is important to note that the abolishment of the tertiary structure in the I state does not lead to a large expansion of the molecular dimension of the protein. This is consistent with other studies which have shown that the dimension of the molten globule is usually no larger than 15% of the native protein (*2*).

Energetics of I↔U Transition. An interesting observation of the present study is the cooperativity of the transition from the molten globule to the unfolded state. Although some theoretical calculations predict that this transition should be a gradual process without detectable heat absorption, many experimental evidences have shown the contrary (*7*, *8*, *13*, *20*). In our case, the molten globule state of BSF is clearly

separated from the unfolded state by an evident, though broad, heat absorption peak (Figure 1). Data analysis gives a $\Delta H_{I\leftrightarrow U}$ of 45.7 kcal/mol and a $\Delta C_{pI\leftrightarrow U}$ of 0.78 kcal/(mol·°C). The nonzero values of $\Delta H_{I\leftrightarrow U}$ and $\Delta C_{pI\leftrightarrow U}$ undoubtedly indicate that this molten globule state is thermodynamically different from the unfolded state. It should be noted from Table 1 that the $\Delta H_{I\leftrightarrow U}/\Delta H_{vH}$ ratio is 1.2 ± 0.3. The larger than unity value suggests that the I↔U transition may not strictly be a two-state process. This is in contrast to the results from GuHCl-induced unfolding studies. In all cases, the transition curves can be well approximated to a two-state model. One of the reasons for this discrepancy may be the uncertainty in determining the $\Delta H_{I\leftrightarrow U}/\Delta H_{vH}$ ratio from DSC scans. As this transition occurs at a very high temperature, fluctuations in $T_{mI\leftrightarrow U}$ and C_p of the unfolded protein would greatly influence $\Delta H_{I\leftrightarrow U}$, $\Delta C_{pI\leftrightarrow U}$, and hence the $\Delta H_{I\leftrightarrow U}/\Delta H_{vH}$ ratio. Another explanation might be that the I↔U transition observed on DSC is not exactly the same as that induced by GuHCl. Whereas thermal unfolding results in exposure and solvation of nonpolar groups, GuHCl-induced unfolding also involves interactions of GuHCl ions with protein residues and water molecules (*33*, *34*). This may reflect the difference between free energies obtained from the two methods. While ΔG°s are comparable for heat and GuHCl-induced unfolding, the former is 10% larger than the latter.

At present, the forces that stabilize the molten globule state are not entirely clear. Nonspecific hydrophobic interactions were supposed to play a central role in maintaining a hydrophobic core, around which are held together the secondary structural elements and the fluctuating side chains of the protein. This is the basis of a theoretical prediction for a noncooperative MG↔U phase transition (*12*, *14*). However, if some long-range interactions participate in stabilizing the molten globule state, then the situation may be very different. In the case that the free energy of disruption of long-range interactions is comparable to that of the secondary structural elements, one should expect a cooperative transition rather than a gradual monotonic process. The differences in the content of long-range interactions should reasonably explain the differences in the cooperativity observed for molten globule states of many proteins. In this sense, the molten globule state of BSF may be stabilized by a large number of long-range interactions, in addition to the core-maintaining hydrophobic forces.

It is of importance to note the similarity between the molten globule state of BSF and folding intermediates of many other proteins. According to Semisotnov et al. (*35*), the early folding intermediate of carbonic anhydrase occurs with simultaneous screening of a hydrophobic core from solvent and formation of some tight packing. Jackson et al. (*36*) studied the effect of mutations on the folding core of chymotrypsin inhibitor 2 and suggested the core is formed and stabilized by tightly packed side chains rather than only by loose hydrophobic interactions. Similar evidences were obtained by Fersht (*37*) and by computer simulation studies (*38*, *39*). The formation of some tertiary packing in the early stage of folding may be very important in resolving the conformational search problem. An initial tertiary packing may stabilize the folding intermediate in a nativelike "format", allowing the fluctuating side chains to explore a large but restricted region of conformational space and to finally reach the native position without getting trapped in misfolded conformations.

Another potential stabilizing factor for the stability of the molten globule state of BSF is the presence of disulfide bonds. BSF contains 12 cysteines and makes 6 disulfide bonds (*17*). Study of the role of these disulfide bonds is complicated by aggregation upon unfolding of the reduced protein (not shown). At present, it is not known if these bonds stabilize the molten globule state as they stabilize the native structure. The high stability, in terms of $T_{mI\leftrightarrow U}$ and $\Delta H_{I\leftrightarrow U}$, of the molten globule is rather unusual, and participation of specific tertiary interactions, including disulfide bonds, in stabilizing such a state seems likely. A molten globule stabilized by higher order interactions is relevant to studies in protein folding intermediates with nativelike tertiary packing. In this context, BSF may serve as a model system.

REFERENCES

1. Kuwajima, K. (1989) *Proteins: Struct., Funct., Genet. 6*, 87−103.
2. Ptitsyn, O. B. (1992) in *Protein Folding* (Creighton, T. E., Ed.) pp 243−300, W. H. Freeman & Co., New York.
3. Baldwin, R. L. (1993) *Curr. Opin. Struct. Biol. 3*, 84−91.
4. Jennings, P. A., and Wright, P. E. (1993) *Science 262*, 892−896.
5. Privalov, P. L. (1996) *J. Mol. Biol. 258*, 707−725.
6. Kuroda, Y., Kidokoro, S., and Wada, A. (1992) *J. Mol. Biol. 223*, 1139−1153.
7. Gittis, A. G., Stites, W. E., and Lattman, E. E. (1993) *J. Mol. Biol. 232*, 718−724.
8. Hagihara, Y., Tan, Y., and Goto, Y. (1994) *J. Mol. Biol. 237*, 336−348.
9. Nishii, I., Kataoka, M., and Goto, Y. (1995) *J. Mol. Biol. 250*, 223−238.
10. Yutani, K., Ogasahara, K., and Kuwajima, K. (1992) *J. Mol. Biol. 228*, 347−350.
11. Griko, Y. V., Freire, E., and Privalov, P. L. (1994) *Biochemistry 33*, 1889−1899.
12. Shakhnovich, E. I., and Finkelstein, A. V. (1989) *Biopolymers 28*, 1667−1680.
13. Xie, D., Bhakuni, V., and Freire, E. (1991) *Biochemistry 30*, 10673−10678.
14. Bryngelson, J. D., and Wolynes, P. G. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 7524−7528.
15. Green, E. D., Adel, G., and Baenziger, J. U. (1988) *J. Biol. Chem. 263*, 18253−18268.
16. Wang, C. Q., Eufemi, M., Turano, C., and Giartosio, A. (1996) *Biochemistry 35*, 7299−7307.
17. Spiro, R. G. (1963) *J. Biol. Chem. 238*, 644−649.
18. Santoro, M. M., and Bolen, D. W. (1988) *Biochemistry 27*, 8063−8068.
19. Garvey, E. P., and Mathews, C. R. (1989) *Biochemistry 28*, 2083−2093.
20. Hamada, D., Kidokoro, S., Fukada, H., Takahashi, K., and Goto, Y. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 10325−10329.
21. Carra, J. H., Anderson, E. A., and Privalov, P. L. (1994) *Protein Sci. 3*, 952−959.
22. Griko, Y., Freire, E., Privalov, G., Van Dael, H., and Privalov, P. L. (1995) *J. Mol. Biol. 252*, 447−459.
23. Becktel, W. J., and Schellman, J. A. (1987) *Biopolymers 26*, 1859−1877.
24. Brown, W. M., Dziegielewska, K. M., Saunders, N. R., Christie, D. L., Nawratil, P., and Müller-Esterl, W. (1992) *Eur. J. Biochem. 205*, 321−331.
25. Uversky, V. N., and Ptitsyn, O. B. (1996) *J. Mol. Biol. 255*, 215−228.

26. Jagannadham, M. V., and Balasubramanian, D. (1985) *FEBS Lett. 188*, 326−330.
27. Manning, M. C., and Woody, R. W. (1989) *Biochemistry 28*, 8609−8613.
28. Chakrabarry, A., Kortemme, T., Padmanabhan, S., and Baldwin, R. L. (1993) *Biochemistry 32*, 5560−5565.
29. Kay, M. S., and Baldwin, R. L. (1996) *Nat. Struct. Biol. 3*, 439−445.
30. Marmorino, J. L., and Pielak, C. J. (1995) *Biochemistry 34*, 3140−3143.
31. Van Dael, H., Haezebrouck, P., Morozova, L., Arico-Muendel, C., and Dobson, C. M. (1993) *Biochemistry 32*, 11886−11894.
32. Morozova, L. A., Haynie, D. T., Arico-Muendel, C., Van Dael, H., and Dobson, C. M. (1995) *Nature Struct. Biol. 2*, 871−875.
33. Tanford, C. (1968) *Adv. Protein Chem. 23*, 121−275.
34. Makhatadze, G. I., and Privalov, P. L. (1992) *J. Mol. Biol. 226*, 491−505.
35. Semisotnov, G. V., Rodionova, N. A., Kutyshenko, V. P., Ebert, B., Blank, J., and Ptitsyn, O. B. (1987) *FEBS Lett. 224*, 9−13.
36. Jackson, S. E., el Masry, N., and Fersht, A. R. (1993) *Biochemistry 32*, 11270−11278.
37. Fersht, A. R. (1995) *Proc. Natl. Acad. Sci. U.S.A. 92*, 10869−10873.
38. Abkevich, V. I., Gutin, A. M., and Shakhnovich, E. I. (1994) *Biochemistry 33*, 10026−10036.
39. Shakhnovich, E. I., Abkevichi, V. I., and Ptitsyn, O. B. (1996) *Nature 379*, 96−98.

BI9723010