# A Partially Unfolded Structure of the Alkaline-Denatured State of Pepsin and Its Implication for Stability of the Zymogen-Derived Protein

**ARTICLE** *in* BIOCHEMISTRY · MAY 2000

Impact Factor: 3.02 · DOI: 10.1021/bi991923d · Source: PubMed



**6 AUTHORS**, INCLUDING:

Yuji O Kamatari
Gifu University
**47** PUBLICATIONS **867** CITATIONS
SEE PROFILE

Hironari Kamikubo
Nara Institute of Science and Technology
**86** PUBLICATIONS **1,581** CITATIONS
SEE PROFILE

Kuniaki Nagayama
The Graduate University for Advanced Studies
**217** PUBLICATIONS **8,820** CITATIONS
SEE PROFILE

# A Partially Unfolded Structure of the Alkaline-Denatured State of Pepsin and Its Implication for Stability of the Zymogen-Derived Protein

Takashi Konno,*,‡ Yuji O. Kamatari,§ Naoki Tanaka,|| Hironari Kamikubo,⊥ Christopher M. Dobson,§ and Kuniaki Nagayama#

*Center for Brain Experiment and Molecular Physiology, National Institute for Physiological Sciences, Okazaki 444-8585, Japan, Oxford Centre for Molecular Sciences, University of Oxford, Oxford, OX1 3QT, United Kingdom, Department of Polymer Science and Engineering, Kyoto Institute of Technology, Matsugasaki, Sakyo, Kyoto 606-0962, Japan, and Institute of Materials Structure Science, High Energy Accelerator Research Organization, Tsukuba, 305-0801, Japan*

*Received August 17, 1999; Revised Manuscript Received November 16, 1999*

ABSTRACT: Pepsin, a gastric aspartic proteinase, is a zymogen-derived protein that undergoes irreversible alkaline denaturation at pH 6−7. Detailed knowledge of the structure of the alkaline-denatured state is an important step in understanding the mechanism of the formation of the active enzyme. An extensive analysis of the denatured state at pH 8.0 was performed using a variety of techniques including $^1$H nuclear magnetic resonance spectroscopy and solution X-ray scattering. This analysis indicates that the denatured state under these conditions has a compact and globular conformation with a substantial amount of secondary and tertiary structures. The data suggest that this partially structured species has a highly folded region and a flexible region. The NMR measurements suggest that the folded region contains His53 and is located at least partly in the N-terminal lobe of the protein. The alkaline-denatured state experiences a further reversible denaturation step at higher pH or on heating; the midpoints of the unfolding transition are pH 11.5 (at 25 °C) and 53.1 °C (at pH 8.0), respectively. The present findings suggest that the proteolytic processing of pepsinogen has substantially modified the ability of the protein to fold, such that its folding process cannot progress beyond the partially folded intermediate of pepsin.

To understand fully the conformational behavior of a protein, it is necessary to define not only the structure of its native state but also that of various denatured states (*1, 2*). Knowledge of the latter is essential for understanding protein stability, as this depends on the difference between the Gibbs free energies of the native and the denatured states (*3, 4*). Denatured states are also important because they are the starting points of the process by which folding to the active form of the protein native state is achieved following biosynthesis (*5*−*7*). Recent studies have also revealed the biological significance of denatured states in processes such as aggregation (*8*−*10*), chaperone binding (*11, 12*), and transport across membranes (*13, 14*). The structures of denatured states have been investigated extensively using a range of techniques including spectroscopic, scattering, and other physical methods as well as protein engineering. A variety of different types of denatured state have been identified, differing in their overall dimensions and the extent of residual secondary and tertiary structures (*3, 4, 15*−*20*).

Most denatured states studied so far are in equilibrium with the native state. There are, however, several proteins that appear not to refold once the native structure is lost. Examples of proteins in this category of systems include several zymogen-derived enzymes produced by proteolytic processing after biosynthesis (*21*), which have been found to unfold irreversibly (*22*−*26*). Understanding the structural principles of this type of protein should give insights into the factors defining protein structure in general. In addition, zymogen-derived proteins are physiologically important since they play essential roles in blood coagulation and fibrinolysis, in the complement reaction of the immune response, in hormone activation, and, in the case of collagen and viruses, in supramolecular assembly (*21, 27*). Several lines of evidence suggest that the structures of zymogen-derived proteins in their denatured states have unusual features. First, as the native fold of the protein is destabilized, the denatured states can be found under mild conditions. Second, critical regions of the folded structure may be disrupted by the proteolytic processing, and the cooperativity of the folded state may be reduced. Third, since the primary structure of the protein is modified by proteolysis, the most stable fold of the protein may be different from that of the zymogen.

Recently, partially denatured states of zymogen-derived serine proteases including α-lytic protease (αLP; *28*)[1] and low molecular weight urokinase-type plasminogen activator (LMW uPA; *29*) were found to play important roles in their folding and stability. α-Lytic protease requires the respective pro-sequence for folding; without the pro-sequence, it is

* Correspondence should be addressed to this author at the Center for Brain Experiment, National Institute for Physiological Sciences, Myodaiji, Okazaki 444-8585, Japan. Telephone: +81-564-55-7830, FAX: +81-564-52-7913, e-mail: konno@nips.ac.jp.

‡ Center for Brain Experiment, National Institute for Physiological Sciences.

§ Oxford Centre for Molecular Sciences, University of Oxford.

|| Department of Polymer Science and Engineering, Kyoto Institute of Technology.

⊥ Institute of Materials Structure Science, High Energy Accelerator Research Organization.

# Molecular Physiology, National Institute for Physiological Sciences.

10.1021/bi991923d CCC: $19.00 
Published on Web 03/17/2000

trapped in a partially denatured state. It has often been hypothesized that the folded conformation cannot be reached because a very high kinetic barrier separates the intermediate and the more stable native state (*25, 26*), but this was recently proved not to be the case at least for αLP. The native state of αLP under physiological conditions is only metastable, the partially denatured state being more stable (*28*). A characteristic partially denatured state has also been found for LMW uPA, whose structure under mildly denaturing conditions possesses a nativelike structure only in its N-terminal lobe (*29*). There is, however, still only limited information about the structures of the denatured states of the zymogen-derived proteins. The major aim of the present study is to investigate such states for porcine pepsin.

Porcine pepsin is a gastric aspartic proteinase (molecular weight = 34 550) that plays an integral role in the digestive processes of vertebrates. The pH optimum of its catalytic activity is less than 2.0 (*30*). It is derived from its zymogen, pepsinogen, by removal of 44 amino acids from its amino terminus, to give a single-chain enzyme with a very low p*I* and three disulfide bridges (*22, 31*). X-ray diffraction analysis shows that the substrate binding cleft is located between two homologous portions of the structure: the N-terminal lobe (residues 1−172) and the C-terminal lobe (residues 173−326). The secondary structure of both regions consists almost entirely of *β*-sheets (*32*; PDB code 3pep). Pepsin has long been known to undergo a conformational transition from the native (at acidic pH) to the denatured state in a narrow pH range (between 6 and 7). This irreversible alkaline denaturation process of pepsin has been studied by various techniques (*33*−*35*) and found to be almost completely irreversible (*22, 36*). By contrast, the unfolding of the zymogen, pepsinogen, has been found to be completely reversible under carefully controlled conditions (*37*).

Lin et al. have, however, demonstrated that alkaline-denatured pepsin can be refolded in the presence of the recombinant N-terminal lobe of pepsinogen (*36, 38*). They have further suggested that the irreversible alkaline denaturation of pepsin could result from unfolding of the N-terminal lobe caused by the ionization of buried carboxyl groups, and that the C-terminal lobe of pepsin could retain its structure (*36*). These studies have provided important insights into the principles of the structural events resulting in the formation of pepsin. These conclusions were, however, based on indirect structural information concerning the denatured state of the protein. In the present study, the structure of the alkaline-denatured state of pepsin has been extensively studied by a combination of circular dichroism (CD), fluorescence, and $^1$H nuclear magnetic resonance (NMR) spectroscopy as well as by solution X-ray scattering (SAXS) and size-exclusion chromatography. The SAXS experiments were employed to characterize the size and shape of the denatured protein chain, while NMR was used to gain information about the nature and stability of tertiary structure in the denatured state.

## EXPERIMENTAL PROCEDURES

*Materials.* Porcine pepsin of the highest grade of purity was purchased from the Sigma Chemical Co. (St. Louis, MO) and further purified using a S-200 gel chromatography column (Pharmacia) equilibrated with the buffer required for the subsequent experiments. Other chemicals were of reagent grade and purchased from Nacalai Tesque (Kyoto, Japan). All solutions except those used for NMR measurements contained 20 mM MOPS, the pH being adjusted carefully with small amounts of NaOH or HCl. The concentration of pepsin was determined using the extinction coefficient $\epsilon_{280}$ = 38 600 $M^{-1}$ $cm^{-1}$ (*39*).

*CD and Fluorescence Measurements.* CD spectra were measured with a Jasco J-720 WI spectropolarimeter (JASCO Co., Japan) using a quartz cell with a path length of 1 or 10 mm. The protein concentration was maintained at 6 and 15 $\mu$M for the far- and near-UV CD measurements, respectively. The temperature of the solutions was maintained by a JASCO thermal controller. Fluorescence spectra were measured with a F4500 fluorometer (Hitachi, Japan) using a quartz cell with a light path of 10 mm. The temperature of the solutions was maintained at 25 °C by means of a thermostatically controlled water bath. Excitation wavelengths were 295 and 350 nm for the tryptophan and ANS fluorescence measurements, respectively. The bandwidths for excitation/emission light were 5 nm/5 nm. Protein concentrations were maintained at 6 $\mu$M. The sample solutions for the ANS binding experiment contained 40 $\mu$M ANS.

*NMR Measurements.* For $^1$H NMR measurements, the protein was dissolved in 20 mM sodium phosphate or glycine buffer prepared with 95% $H_2O$/5% $D_2O$ or with 100% $D_2O$, with addition of sodium chloride or urea as required. 2,2-Dimethyl-2-silapentane-5-sulfonic acid (DSS) was added as an internal chemical shift reference. The sample pH was adjusted using NaOD or DCl. The pH value in $D_2O$ (denoted pH*) was that indicated by direct reading of the pH meter. The concentration of pepsin for the NMR measurements was 10 mg/mL, except for the 1D nuclear Overhauser effect (NOE) difference measurements which were performed at 33 mg/mL. NMR measurements were performed at 25 °C on home-built NMR spectrometers belonging to the Oxford Centre for Molecular Sciences, with operating frequencies of 500.1, 599.7, 600.2, and 750.0 MHz. All experiments except the 1D NOE difference measurements were performed using the water-gate pulse sequence (*40*) for water suppression. 1D steady-state NOE difference spectra were obtained by applying 100 ms long selective presaturation pulses ($\nu_1$ = 7.1 Hz) alternately to specific resonances and to regions of the spectrum in which no resonances occur.

*SAXS Measurements.* Solution X-ray scattering experiments were carried out at the solution scattering station (SAXS camera) installed at BL-10C, the Photon Factory, Tsukuba, Japan. The sample-to-detector distance was about 90 cm, calibrated by meridional diffraction of dried chicken collagen. The X-ray wavelength was 1.488 Å. The sample cell was 50 $\mu$L in volume with 15 $\mu$m thick quartz windows, and had a 1 mm X-ray path length. The temperature of the sample was maintained by a thermostatically controlled water

[1] Abbreviations: CD, circular dichroism; NMR, nuclear magnetic resonance; SAXS, small-angle X-ray scattering; ANS, 8-anilino-1-naphthalenesulfonic acid; DSS, 2,2-dimethyl-2-silapentane-5-sulfonic acid; 1D, one-dimensional; $N_P$, the native state of pepsin at pH 5.6 and 25 °C; $I_P$, the alkaline-denatured state of pepsin at pH 8.0 and 25 °C; $D_{PU}$, the urea-denatured state of pepsin at pH 8.0, 4 M urea, and 25 °C; $D_{PH}$, the denatured state of pepsin at pH 8.0 and 70 °C; $D_{PA}$, the denatured state of pepsin at pH 12.0 and 25 °C; $R_G$, radius of gyration; NOE, nuclear Overhauser effect; αLP, α-lytic protease; LMW uPA, low molecular weight urokinase-type plasminogen activator.

bath. Protein concentrations were varied within the range of 2.5−10 mg/mL, and the measurement time for each scattering experiment was 5 min. Scattering data from solutions containing no pepsin were obtained for each solution condition to define the background. X-ray scattering intensities in the small-angle region are given as $I(Q) = I(0) \exp(-R_G^2Q^2/3)$, where $Q$ and $I(0)$ are the momentum transfer and the intensity at a scattering angle of 0°, respectively (*41*). $Q$ is defined by $Q = 4\pi \sin \vartheta/\lambda$, where $2\vartheta$ and $\lambda$ are the scattering angle and the wavelength of the X-ray, respectively. The radius of gyration ($R_G$) was obtained from the slope of the Guinier plot, $\log_e [I(Q)]$ against $Q^2$ (*41*). The regression lines were fitted to the data within $R_GQ < 1.3$ (*43, 44*). The distance distribution function, $P(r)$, was calculated using the indirect Fourier transform method of Moore (*45*).

*Size-Exclusion Chromatography.* Size-exclusion chromatography was performed using the SMART system with an S-200 column (Pharmacia, Sweden) monitored by absorption at 280 nm. Equilibrium and elution buffers contained 20 mM MOPS (pH 8.0) and 0.1 M NaCl. The temperature of the column was maintained at 25 °C and the flow rate of the buffer at 60 $\mu$L/min.

## RESULTS

*CD, ANS Fluorescence, and Chromatographic Analysis of the Alkaline-Denatured State.* At pH 5.6, where pepsin is in its native state, the far-UV CD spectrum has a shape typical of a $\beta$-sheet-rich protein (Figure 1A, thin solid line). At pH 8.0, where pepsin is denatured, the spectrum indicates a partial loss of secondary structure and the emergence of intensity characteristic of random coil regions of structure (Figure 1A, boldface solid line; *22*, *32*, *35*). The transition between the two spectra occurs in the pH range of 6−7, and is not reversible when the pH is decreased from 8.0 to below 6.0 (data not shown). Addition of 4 M urea at pH 8.0 leads to a further change in the spectral shape, typical of a highly unstructured conformation (Figure 1A, broken line). For convenience, we denote the native state as $N_P$, the alkaline-denatured state as $I_P$, and the denatured state in the presence of 4 M urea as $D_{PU}$, respectively (Table 1). CD spectra of $N_P$, $I_P$, and $D_{PU}$ were also measured in the near-UV region. The spectrum of the protein in the $I_P$ state was intermediate between those of the protein in the $N_P$ and $D_{PU}$ states, indicating that some tertiary interactions involving aromatic residues remain in the pepsin molecule in the $I_P$ state (data not shown).

Additional experiments were carried out to characterize the nature of $I_P$. ANS is known to bind to compact denatured states known as molten globules (*46*−*48*). However, the $I_P$ state of pepsin shows no binding affinity for ANS (Figure 1B), even in the presence of 0.2 M NaCl. The aggregation state of pepsin in the alkaline-denatured state was examined by size-exclusion gel chromatography. This is an important issue because the irreversibility of the unfolding of pepsin could originate from aggregation. In the $I_P$ state (pH 8.0 and 25 °C) in 1 mg/mL protein, however, the gel chromatographic pattern has a clear single peak at a retention time of 25.3 min (solid line in Figure 1C) and is unchanged even when the sample is incubated at 25 °C for 24 h (broken line in Figure 1C). This chromatographic pattern is not affected by addition of 0.2 M NaCl or by changes in the concentration of pepsin in the range of 0.1−10 mg/mL (data not shown). Overall, $I_P$ at pH 8.0 appears not to have molten-globule like character, and to have little tendency to form aggregated structures.

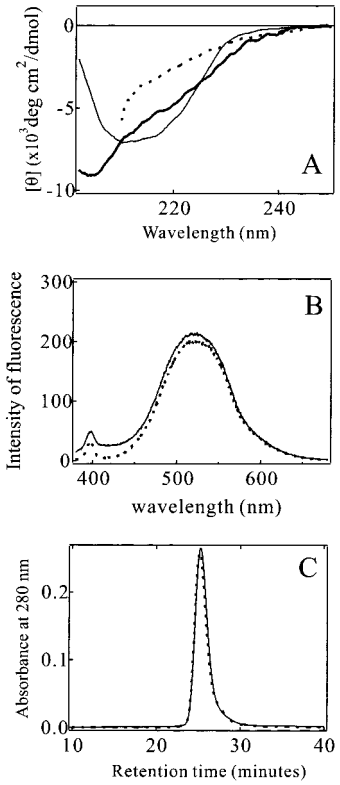


FIGURE 1: (A) Far-ultraviolet CD spectra of pepsin at 25 °C, pH 5.6 (thin solid line); pH 8.0 (boldface solid line); 4 M urea and pH 8.0 (broken line). (B) Fluorescence spectra of ANS (40 $\mu$M) at pH 8.0. The concentration of pepsin is 0 (broken line) and 6 $\mu$M (solid line). (C) Size-exclusion gel chromatography of the alkaline-denatured state of pepsin at a concentration of 1 mg/mL after incubation for 0 (solid line) and 24 h (broken line) at pH 8.0, 25 °C.

Table 1: Estimates of $R_G$ Values for the Different States of Pepsin

| | $R_{G,I}$ (Å)[a] | $R_{G,II}$ (Å)[b] | $R_G/R_{G(N)}$[c] |
|---|---|---|---|
| pH 5.6 ($N_P$) | 23.9 | 24.4 ± 0.5 | 1 |
| pH 8.0 ($I_P$) | 30.6 | 30.9 ± 1.4 | 1.28 |
| pH 8.0, 4 M urea ($D_{PU}$) | 40.2 | 39.9 ± 2.6 | 1.68 |

[a] $R_{G,I}$ was determined by a linear fit of the $R_G^2$ vs [pepsin] plot. [b] $R_{G,II}$ was determined from the scattering curve at infinite dilution. [c] $R_G/R_{G(N)}$ is the ratio of the $R_G$ value of each state relative to that of the native state.

*$^1$H NMR Analysis of the Alkaline-Denatured State.* Figure 2 shows the aliphatic proton regions of typical $^1$H NMR spectra of pepsin in the $N_P$, $I_P$, and $D_{PU}$ states. The vertical scale of the spectra was normalized such that the total signal area in the −1 to 3.5 ppm range is the same in all cases. The spectrum of $N_P$ shows considerable signal dispersion characteristic of a close-packed native conformation (Figure 2A). Many resolved and shifted peaks can be observed between −0.7 and 0.2 ppm, which are typical of ring-current-shifted signals from regions of rigid tertiary structure containing aliphatic and aromatic residues (*49*). The spectrum of $D_{PU}$ is typical of a highly denatured conformation of a protein, where almost all the proton signals of the amino acid residues are found close to their random-coil positions

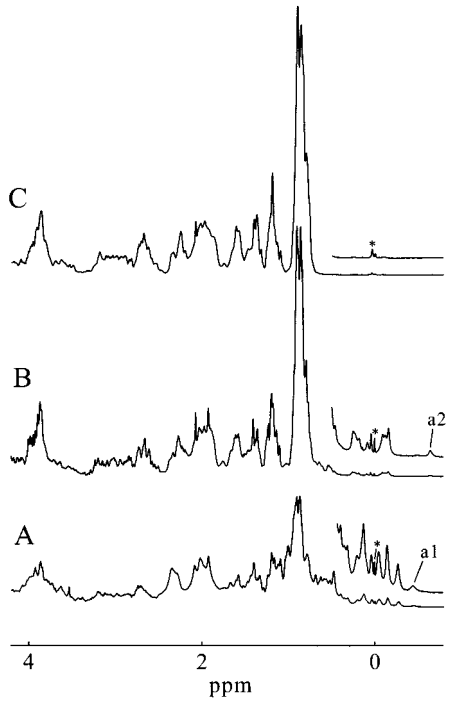


FIGURE 2: 750 MHz $^1$H NMR spectra of pepsin in $H_2O$ solution at pH 5.0 ($N_P$; A), pH 8.1($I_P$; B), and pH 8.1 in the presence of 4 M urea ($D_{PU}$; C). The vertical scale of each spectrum is normalized to the total signal intensity of the resonances in the range −1 to 3.3 ppm. The signals labeled "a1" in (A) and "a2" in (B) were used to compare their signal intensities (see Results). The signals labeled with an asterisk are those of DSS.

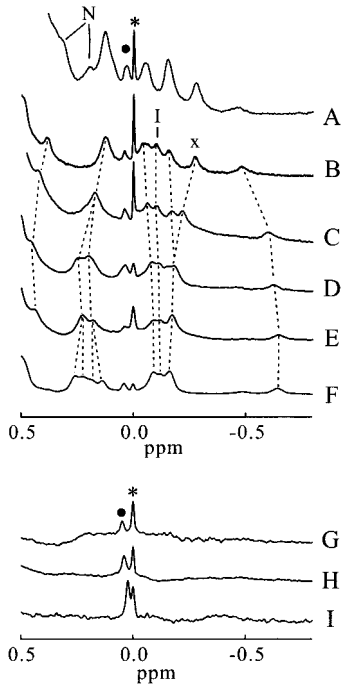


FIGURE 3: Upfield region of 600 MHz $^1$H NMR spectra of pepsin. Panels A−F are spectra of $N_P$ at pH 5.0 (A), $I_P$ at pH 5.2 (B), $I_P$ at pH 7.0 (C), $I_P$ at pH 8.1 (D), $I_P$ at pH 9.6 (E), and $I_P$ at pH 11.0 (F). The labels "N" (in A) and "I" (in B) indicate signals arising from the $N_P$ and $I_P$ states, respectively. The signal labeled "x" in (B) was saturated in the experiments shown in Figure 4. Broken lines connect the corresponding signals of the different pH values; their connectivities were established by more detailed pH titration experiments (data not shown). Panels G−I show spectra for the highly denatured state of pepsin, $D_{PU}$ at 4 M urea and pH 8.1 (G), $D_{PA}$ at pH 12.0 (H), and $D_{PH}$ at pH 8.0 and 70 °C (I). The signals labeled with an asterisk and (●) are those of DSS and an impurity, respectively. The vertical scale of each spectrum is normalized to the total signal intensity of the resonances in the range −1 to 3.3 ppm.

(Figures 2C and 3G). There is no evidence for residual tertiary structure in this urea-denatured state. The spectrum of $I_P$, however, has an appearance intermediate between that of $N_P$ and that of $D_{PU}$. The spectral pattern in most regions (for example, the methyl proton region between 0.5 and 1.5 ppm) appears more similar to that of $D_{PU}$ than to that of $N_P$ (Figure 2B), but there is clear evidence for well-resolved shifted signals in the −0.7 to 0.3 ppm range (inset of Figure 2B).

The intensities of the individual upfield-shifted signals in the spectrum of $I_P$ are approximately the same as those of peaks in the spectra of $N_P$ within experimental error. For example, the ratio of the area of the peak labeled "a1" of $N_P$ (Figure 2A) to that denoted "a2" of $I_P$ at pH 8.0 (Figure 2B) is 1.10 ± 0.3. This ratio, estimated using the spectrum of $I_P$ at pH 10.0 instead of that at pH 8.0, is also close to unity, 1.08 ± 0.2. The comparable intensity of the individual upfield-shifted signals for $N_P$ and $I_P$ indicates that all the pepsin molecules in the $I_P$ condition contain nativelike residual structure. On the other hand, the spectral properties of $I_P$ in the methyl region between 0.7 and 1.2 ppm characteristic of groups in unstructured regions of protein are much greater in $I_P$ than $N_P$. These results indicate that the alkaline-denatured state of pepsin is partially disordered but still contains persistent structure involving clusters of aromatic and aliphatic residues. These features of the NMR spectrum of $I_P$, including the amplitude of the upfield-shifted signals, are unchanged in the range of pH 8−11 (Figure 3D−F), supporting the conclusion that the conformational properties of pepsin are substantially unchanged over this pH range.

To compare the upfield-shifted signals of $I_P$ between −0.7 and 0.3 ppm with those of $N_P$, the pH of a solution of $I_P$ was gradually decreased to pH 5.2 (Figure 3B,C) or increased to pH 11.0 (Figure 3E,F). Acidification of $I_P$ to pH 5.2 did not result in recovery of the overall spectrum of $N_P$ (data not shown); this is consistent with the well-established irreversibility of the alkaline denaturation of pepsin. Gradual shifts of the upfield-shifted signals were, however, evident (Figure 3B−F). By comparing the spectrum of $I_P$ at pH 5.2 with that of $N_P$, the difference can be seen. For example, the signal labeled "I" (Figure 3B) in the spectrum of $I_P$ is not evident in the spectrum of $N_P$ while the signals labeled "N" in Figure 3A are present only in the spectrum of $N_P$. Other signals are closely similar in the spectra of both $N_P$ and $I_P$. The results indicate that some of the residual tertiary structure of $I_P$ is largely similar to the structure found in $N_P$, but marked difference exists in other regions. This supports the conclusion that the structure of pepsin in the $I_P$ state is not the same as that of $N_P$, and that $I_P$ does not refold to give a structure identical to $N_P$ when the pH is reduced.

Pepsin has only 1 histidine (His53) in its 326 residues (*32*). The expected chemical shift of the C2 proton of histidine in a random coil at pH 5 is about 8.6 ppm, and the sharp singlet is usually well separated from other proton resonances in $D_2O$ (*50*). Furthermore, histidine residues titrate typically close to pH 7 (the chemical shift changes from about 8.6 ppm to about 7.7 ppm), and are the only residue to titrate in this pH range under normal circumstances (*27, 50*). The

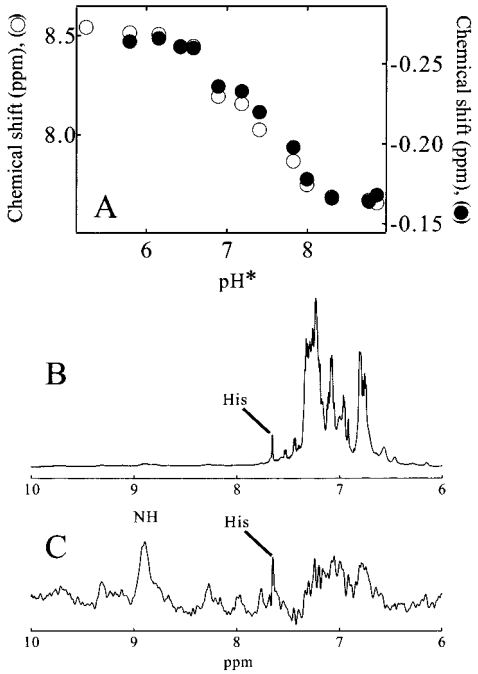


FIGURE 4: (A) pH-dependent changes in the chemical shift of the C2 proton of His53 (○) and the upfield-shifted signal "x" of Figure 3B (●) in $D_2O$ solution. (B) 1D NMR spectrum of the alkaline-denatured state $I_P$ of pepsin in $D_2O$ solution at pH* 8.9. (C) NOE difference spectrum resulting from saturation of the signal labeled "x" in Figure 3B. The label "His" indicates the C2 proton signal of His53 of pepsin.

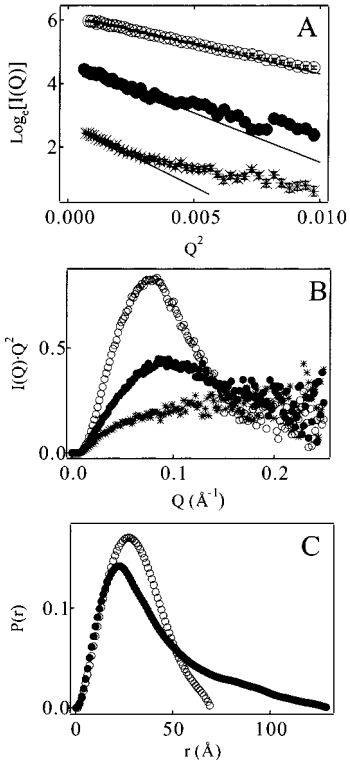


FIGURE 5: Guinier plots (A), Kratky plots (B), and $P(r)$ functions (C) of pepsin at pH 5.6 (○), pH 8.0(●), and pH 8.0 in the presence of 4 M urea (★) at 25 °C, 5 mg/mL pepsin. For clarity, each plot in (A) is shifted on the vertical axis.

signal labeled "His" in Figure 4B is the only one which satisfies all these conditions, and so we could readily assign the C2 proton signal of His53. The tertiary structure of $I_P$ was probed by examining the properties of this His C2 proton. First, we examined the pH dependence of the chemical shift of the His C2 signal and the upfield-shifted signals at −0.8 to 0.3 ppm of the $I_P$ state in $D_2O$ solution (Figure 4A). The midpoint of titration curves of the chemical shifts for His C2 and the signal labeled "x" in Figure 3B are almost identical, suggesting that the pH-dependent behavior of this upfield-shifted signal is directly coupled with the titration of His53. If this is actually the case, the methyl proton in the rigid folded cluster of $I_P$ should be in close contact with the histidine residue. The other upfield-shifted signals also show similar titration behavior although the magnitude of the shift change varies for individual resonances.

More direct structural information relating His53 to the upfield-shifted signals was obtained from 1D NOE difference spectra. NOE experiments were performed by presaturating the upfield-shifted signal of $I_P$ labeled "x" in Figure 3B. One of the signals enhanced by magnetization transfer from the methyl signal is the His53 C2 proton (Figure 4C), indicating that this single histidine has close contacts with the methyl protons involved in the rigid tertiary structure present in $I_P$. Other signals showing NOE enhancements are located between 8.0 and 9.7 ppm (Figure 4C), which we believe arise from exchangeable NH protons because of their chemical shift, line width, and time-dependent signal intensities in $D_2O$ (data not shown). Though these experiments were performed a few hours after sample preparation in $D_2O$ solutions, we observed residual signals from these exchangeable NH protons. The expected exchange rate for an amide hydrogen in a random coil state at pH* 8.9 and 25 °C is between $3 \times 10^2$ and $2 \times 10^5$ $min^{-1}$ (*51*). The exchange rates of the NH protons observed in the NOE difference experiments are therefore very much slower than those of NH protons in an unstructured protein. This is further evidence for the existence of stable residual structure in the alkaline-denatured state of pepsin. NOE experiments involving presaturation of upfield-shifted signals of $I_P$ other than "x" also show the enhancement of the His53 signal and various NH proton signals. The conclusion from the results of the pH titration and the 1D NOE experiments is that the tertiary structure of $I_P$ contains side chains which contribute to the upfield signals observed between −0.7 and 0.3 ppm, together with that of His53.

*SAXS Analysis of the Alkaline-Denatured State.* SAXS analysis has been extensively used to characterize the molecular dimensions of denatured proteins (*43*, *52*−*56*). Guinier plots of SAXS data give estimates of the radius of gyration ($R_G$) which is informative about the size and compactness of polypeptide chains (*42*). Figure 5A shows Guinier plots of $N_P$, $I_P$, and $D_{PU}$ for 5 mg/mL solutions of pepsin. The plots exhibit a linear region at low $Q$ (the Guinier region). The $R_G$ of pepsin in the different states was estimated at a variety of protein concentrations (see Experimental Procedures). Linear extrapolation of the plots of $R_G^2$ as a function of the concentration of pepsin gave values of $R_G$ extrapolated to zero protein concentration (the second column in Table 1). $R_G$ at infinite dilution was also estimated from the scattering curve by a separate extrapolation of each point in reciprocal space from the series of scattering data at different protein concentrations. The values of $R_G$ obtained using the two different methods were very similar to each other (Table 1). By averaging the results from the two

methods, the $R_G$ value of pepsin in the $N_P$ state was found to be 24.2 ± 0.5 Å. On alkaline denaturation at 8.0, it increased to 30.8 ± 1.2 Å (i.e., by about 30%) [$R_G/R_{G(N)}$ in Table 1]. Addition of 4 M urea ($D_{PU}$) increased $R_G$ by about 70% to a value of 40.0 ± 2.1 Å.

The Kratky plot, $I(Q)\cdot Q^2$ versus $Q$, is a useful means of describing the overall shape of a polymer molecule in solution (*41*), and has been widely used to study the structures of non-native states of proteins (*43, 53*−*56*). In essence, the Kratky plot shows a clear peak at $Q \sim 0.1$ for a globular conformation, while it has a plateau and then increases monotonically for an expanded unfolded conformer. The Kratky plot for the $N_P$ state has a well-defined peak at $Q = 0.08$, and a pattern typical of a native state of a protein, while that for the protein in the presence of 4 M urea ($D_{PU}$) at pH 8.0 is typical of a highly denatured chainlike conformation (Figure 5B). The plot for the protein at pH 8.0 in the absence of urea ($I_P$) shows intermediate character. It is disordered relative to $N_P$ but still possesses a globular component, represented by a broad peak at $Q = 0.09$.

A real-space representation of the information in the SAXS scattering curve can be obtained by a Fourier inversion of the scattered intensity to generate the distance distribution function, $P(r)$ (*45*). The information contained in this representation is the length distribution of atom-weighted interatomic vectors in the molecule. The $P(r)$ function of $N_P$ and $I_P$ shows a single peak, but that of $I_P$ has a very long tail extending to large dimensions (Figure 5C). Alkaline denaturation has therefore induced a large expansion of the pepsin structure. The maximum dimension ($d_{max}$) was obtained from the point where the $P(r)$ function approaches zero, and was ~69 and ~130 Å for $N_P$ and $I_P$, respectively. The value for the native state is similar to that estimated from the high-resolution X-ray crystallographic structure (67.4 Å; ref *32*).

*Denaturation of the Alkaline-Denatured State.* The stability of the residual structure in the alkaline-denatured state was examined as a function of pH and temperature. Transitions were monitored by far- and near-UV CD, fluorescence, and NMR measurements. An increase in pH from 8.0 to 12.0 results in changes in all the spectral properties (Figure 6A for far-UV CD). The transition curves of the far- and near-UV CD spectra, as well as the peak position of tryptophan fluorescence, are shown in Figure 6B, indicating that pepsin loses its residual structure concurrently and cooperatively in a pH range between 11.0 and 12.0; the transitions detected by the three probes occur in the same pH range within experimental error (Figure 6B). The loss of the rigid tertiary structures of $I_P$ by pH was also examined by NMR. The NMR spectra at pH values up to 11.0 were essentially the same as that at pH 8.0 (data not shown), but large changes occur between pH 11.0 and 12.0 (Figure 3H). All the characteristic signals of the rigid tertiary structure in the NMR spectrum are absent at pH 12.0. Increase in the solution temperature at pH 8.0 also induces cooperative and concurrent changes of the far- and near- UV CD spectra in the temperature range of 40−60 °C (Figure 6A,C). The NMR spectrum at 70 °C also shows the absence of the rigid tertiary structure in the heat-denatured state (Figure 3I). The pH- and heat-induced spectral changes were found to be essentially reversible (e.g., Figure 6A). For convenience, the highly denatured states at pH 8.0 and 70 °C or at pH 12.0

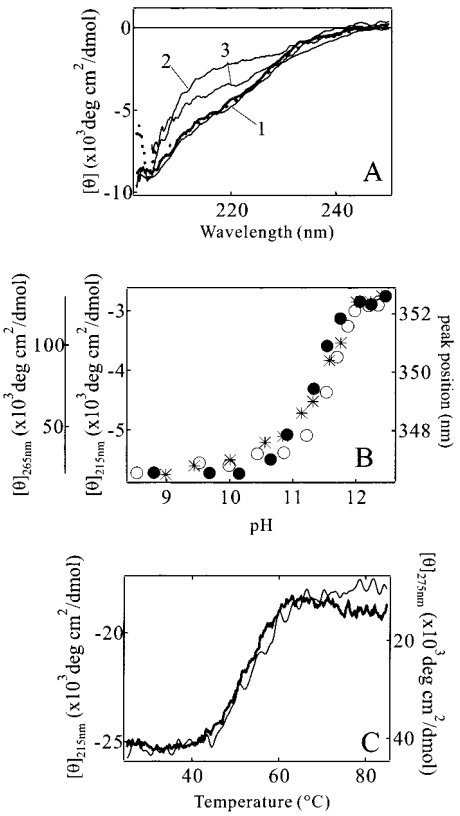


FIGURE 6: (A) Far-ultraviolet CD spectra of pepsin. Spectra at pH 8.0, 25 °C (thin solid line 1); pH 12.0, 25 °C (thin solid line 2); pH 8.0, 70 °C (thin solid line 3); pH 8.0 at 25 °C after exposure to pH 12.0 (boldface solid line); 25 °C at pH 8.0 after exposure to 70 °C (broken line). (B) Changes in ellipticity at 215 nm (open circles) and 275 nm (filled circles), and the wavelength of the maximum in the tryptophan fluorescence (star symbols) as a function of pH at 25 °C. (C) Changes in ellipticity at 215 nm (boldface line) and 265 nm (thin line) as a function of temperature at pH 8.0.

are abbreviated as $D_{PH}$ or $D_{PA}$, respectively. By the criteria of the far-UV CD spectra, the amount of secondary structures decreases in the order $N_P > I_P > D_{PH} > D_{PA} \sim D_{PU}$ (Figures 1 and 6A).

The multiple structural probes show almost identical transition curves for the change from $I_P$ to $D_{PA}$ to those from $I_P$ to $D_{PH}$ (Figure 6B,C). This suggests that no intermediates are populated in either of these denaturation processes, and that the transitions have effective two-state character. With this assumption, the transition curves in Figure 6B,C were analyzed using the equations (*57*):

$$\frac{D_{PA}}{I_P} = \frac{\alpha}{1-\alpha} = \frac{(10^{-pH_m n})}{(10^{-pHn})}$$

$$\frac{D_{PH}}{I_P} = \frac{\alpha}{1-\alpha} = \exp\left[-\frac{\Delta H}{R}(1/T - 1/T_m)\right]$$

where α, $n$, $pH_m$, $\Delta H$, and $T_m$ are the fraction of denatured protein, the number of protons associated with the alkaline denaturation step, the pH of the midpoint of the transition, the enthalpy of heat denaturation, and the denaturation temperature, respectively. A nonlinear least-squares algorithm gives the best-fit parameters listed in Table 2; the parameters determined from different experimental probes give almost identical values. The midpoints of the pH and thermal

Table 2: Parameters for the Denaturation of $I_P$[a]

| | far-UV | near-UV | fluorescence |
|---|---|---|---|
| alkaline denaturation | | | |
| $pH_m$ | 11.6 | 11.3 | 11.6 |
| $n$ | 2.64 | 1.67 | 1.84 |
| heat denaturation | | | |
| $\Delta H$ | 205 kJ | 195 kJ | nd[b] |
| $T_m$ | 52.6 °C | 53.6 °C | nd[b] |

[a] Fitted to a two-state model (see Results). [b] nd: not determined because of a large variation in the base line as the temperature was varied.

Table 3: Conformational States of Pepsin in Aqueous Solution[a]

| | secondary structure | tertiary structure | SAXS pattern |
|---|---|---|---|
| $N_P$ | ++ | ++ | native-state pattern |
| $I_P$ | + | + | compact-denatured pattern |
| $D_{PA}$ | − | − | nd[b] |
| $D_{PH}$ | ± | − | nd[b] |
| $D_{PU}$ | − | − | highly denatured pattern |

[a] The residual structure in large (++), significant (+), or marginal (±) amounts, or its absence (−). [b] nd: not determined because of technical difficulties.

unfolding transitions are 11.5 ± 0.1 (at 25 °C) and 53.1 ± 0.5 °C (at pH 8.0), respectively.

## DISCUSSION

The present analysis has shown that the alkaline-denatured state of pepsin ($I_P$) is partially structured rather than being fully unfolded. The CD spectrum of $I_P$ and its changes with pH, temperature, and addition of urea (Figures 1 and 6) confirm the presence of extensive residual secondary structure. Residual tertiary structure is also clearly evident from the presence of well-resolved and shifted NMR signals (Figure 2B). The SAXS analysis also demonstrates that the $I_P$ state contains a globular component (Figure 5B), and its overall dimensions are intermediate between those of $N_P$ and those of $D_{PU}$ ($R_G$ in Table 1). This partially denatured state $I_P$ is present at pH 8.0, where the precursor of pepsin, pepsinogen, still retains fully its native structure (*37*). The CD and NMR properties are highly stable in the pH range from 8 to 11, while at the extremes of this pH range abrupt transitions occur (Figures 3 and 7B). This demonstrates that the conformation of pepsin in the $I_P$ state is highly stable over a wide pH range and is distinct from the conformations of $N_P$ and $D_{PA}$. The fact that the intensities of the individual upfield-shifted signals of NMR of $I_P$ are closely similar to those of $N_P$ indicates that effectively all the pepsin molecules in the $I_P$ condition contain the same residual structure. Because of the limited accuracy of the area estimates of the NMR signals, however, we cannot completely exclude the possibility that a small fraction of molecules has a different conformation from that of the majority. But it is clear that the data are not consistent with the extensive unfolding of a significant proportion of the molecules to leave a heterogeneous distribution of conformations.

A well-established form of a compact partially folded denatured state is the molten globule state (*2, 47, 48*). However, the rigid tertiary structure observed by NMR is not characteristic of molten globule states. The partially folded region of $I_P$ is also highly stable, since very slowly exchanging NH protons were found in this structure (Figure 4C), again not characteristic of molten globules. In addition, the fluorescence dye ANS does not bind to the $I_P$ state (Figure 1B) although ANS binding to the incompletely collapsed hydrophobic core is an indicator of molten globule character (*46*−*48*). These results therefore suggest that the globular component of $I_P$ identified by SAXS (Figure 5B) originates from a nativelike region of the protein chain, not from a molten-globule-like conformation. The fact that some regions of the NMR spectrum of $I_P$ have similarities to that of $D_{PU}$ is, however, an indication that part of the protein chain is in a highly unfolded structure (Figure 2B). Thus, the alkaline-denatured state of pepsin appears to have a tightly folded region and a flexible unstructured region. This conclusion is also consistent with the shape of the $P(r)$ function of $I_P$, which has a peak characteristic of a relatively compact structure and a very extended tail to much higher dimensions (Figure 5C).

Pepsin is a two-domain protein composed of N- and C-terminal lobes (*32*). Energy calculations have predicted that the latter is more stable at high pH than the former (*36*), suggesting that the folded part of $I_P$ is located in the C-terminal lobe. However, the 1D NOE and pH titration data obtained in the present study (Figure 4) suggest that the rigid folded part of $I_P$ contains the histidine residue at position 53 located in the N-terminal domain of the molecule (*37*). The structure of pepsin in the region of His53 in the native state is composed of two short α-helices and several aromatic residues, providing an explanation for the chemical shift dispersion and the slowly exchanging amide protons in the $I_p$ state (*32*). This finding suggests that it could be the C-terminal domain that is unfolded in the $I_p$ state. It is possible, however, that the residual structures of $I_P$ reside in both the N- and C-terminal lobes. Furthermore, the folded structure appears to contain some elements of non-native tertiary structure ("I" in Figure 3B), which suggests that the partially folded structure of $I_P$ could be achieved not only by maintaining some regions of the native conformation but also by some conformational rearrangements. Recently, a partially denatured state has also been found for a serine protease, low molecular weight urokinase-type plasminogen activator (LMW uPA; *29*). LMW uPA inactivated with a covalently attached peptide denatures with an increase in temperature or by addition of denaturants in two discrete steps, and a partially structured state is populated in mildly denaturing conditions. This intermediate state of LMW uPA possesses a highly stable region of nativelike structure in its N-terminal lobe (*29*), a finding similar to that described here for the $I_P$ state of pepsin. The native states of pepsin and LMW uPA both have a two-domain structure composed of N- and C-terminal lobes (*58*) although the fold of each lobe of LMW uPA has a six-strand $\beta$-barrel fold and is distinct from the orthogonal arrangement of $\beta$-sheets in the structure of pepsin (*32*).

Table 3 summarizes the structural properties of the different conformational states of pepsin. In addition to the irreversible transition from $N_P$ to $I_P$ (*22, 36*), $I_P$ experiences further alkaline and heat denaturation. The residual structure of $I_P$ is highly stable, and the midpoints of the pH and heat denaturation transition are 11.5 ± 0.1 (at 25 °C) and 53.1 ± 0.5 °C (at pH 8.0), respectively (Table 2). The thermodynamic distinctions between $D_{PA}$, $D_{PH}$, and $D_{PU}$ are, however,

not clearly defined, and the three highly denatured states may simply differ in the detailed destabilization of regions of structure within the conformational ensemble. $D_{PA}$ is more highly denatured at the level of the secondary structure than $D_{PH}$, probably because of a very large density of negatively charged residues of pepsin at high pH (*31*, *32*). The *n* value of the transition, $I_P \leftrightarrow D_{PA} + nH^+$ (Table 2), implies that approximately two ionizable groups are involved in this second alkaline denaturation process. The denaturation of $I_P$ occurs reversibly, in contrast to the transition between $N_P$ and $I_P$. The present data suggest at least part of the residual structure in the $I_P$ state appears to be present in the N-terminal lobe, as indicated by the NMR measurements (Figure 4). In addition, the present analysis has failed to detect any substantial structural changes in the pH range from 8.0 to 11.0. Instead, all the measurements show that the structure of $I_P$ is preserved over the in the higher pH range of 11−12 (Figures 3 and 6).

The present findings, along with previous findings (*22*, *23*), suggest that the proteolytic processing of pepsinogen has substantially modified the ability of the protein to fold, such that it appears to be unable to progress beyond the partially folded intermediate $I_p$. At pH 8.0 and 25 °C, the most stable form of the precursor of pepsin, pepsinogen, is the native state, and this can be reached from the unfolded state without difficulty (*37*). The requirement of a pro-sequence for proper folding is well established for several serine proteases such as subtilisin (*23*) and α-lytic protease (αLP; *25*). Such a situation could occur either because the native state is not at the minimum of the free energy landscape or, alternatively because a very high free energy barrier separates the denatured and native states. It has often been suggested that the latter is the case (*25*, *26*), but recently the former situation has been demonstrated to be the reason for the inability of αLP to regain its native state in the absence of pro-region (*28*). The proteolytic removal of the pro-sequence from pro-αLP destabilizes the native state relative to the partially denatured state. The native state of pepsin, $N_P$, may also be metastable relative to $I_P$ even at pH values below 6 although this has not yet been established. Further comparative studies of denatured states and folding kinetics of various zymogen-derived proteins should, however, provide deeper insights into the roles of pro-regions in folding and into the principles that determine protein structures in general.

## ACKNOWLEDGMENT

The X-ray scattering measurements were performed with permission from the Program Advisory Committee of the Photon Factory (Proposal No. 97G127). We thank Dr. J. A. Jones for his help on the 1D NOE experiments and Dr. M. Kataoka and Dr. A. K. Chamberlain for their reading of the manuscript. We are much indebted to Maggie McCammon, Dr. Helena Hernandez, and Professor Carol V. Robinson for their support in preparing the manuscript. We also thank Prof. K. Imoto for his encouragement throughout this work.

## REFERENCES

1. Dill, K. A., and Shortle, D. (1991) *Annu. Rev. Biochem. 60*, 795−825.
2. Dobson, C. M. (1992) *Curr. Opin. Struct. Biol. 2*, 6−12.
3. Tanford, C. (1968) *Adv. Protein Chem. 23*, 121−217.
4. Privalov, P. L. (1989) *Annu. Rev. Biophys. Biophys. Chem. 18*, 47−69.
5. Fersht, A. R. (1993) *FEBS Lett. 325*, 5−16.
6. Ptitsyn, O. B. (1995a) *Curr. Opin. Struct. Biol. 5*, 74−78.
7. Dobson, C. M., Sali, A., and Karplus, M. (1998) *Angew. Chem. 37*, 868−893.
8. Sipe, J. D. (1992) *Annu. Rev. Biochem. 61*, 947−975.
9. Prusiner, S. B., Scott, M. R., DeArmond, S. J., and Cohen, F. E. (1998) *Cell 93*, 337−348.
10. Dobson, C. M. (1999) *Trends Biol. Sci. 24*, 329−313.
11. Sigler, P. B., Xu, Z., Rye, H. S., Burston, S. G., Fenton, W. A., and Horwich, A. L. (1998) *Annu. Rev. Biochem. 67*, 581−608.
12. Ellis, R. J. (1999) *Curr. Biol. 9*, R137−139.
13. Bychkova, V. E., Pain, R. H., and Ptitsyn, O. B. (1988) *FEBS Lett. 238*, 231−234.
14. Beermann, B. B., Hinz, H. J., Hofmann, A., and Huber, R. (1998) *FEBS Lett. 423*, 265−269.
15. Udgaonkar, J. B., and Baldwin, R. L. (1988) *Nature 335*, 694−699.
16. Matouschek, A., Kellis, J. T., Jr., Serrano, L., and Fersht, A. R. (1989) *Nature 340*, 122−126.
17. Shortle, D. (1992) *Q. Rev. Biophys. 25*, 205−250.
18. Dyson, H. J., and Wright, P. E. (1996) *Annu. Rev. Phys. Chem. 47*, 369−395.
19. Konno, T. (1998) *Protein Sci. 7*, 975−982.
20. Dobson, C. M., and Hore, P. J. (1998) *Nat. Struct. Biol. 5, Suppl.*, 504−507.
21. Neurath, H., and Walsh, K. A. (1976) *Proc. Natl. Acad. Sci. U.S.A. 73*, 3825−3832.
22. Bovey, F. A., and Yanari, S. S. (1960) in *The Enzyme III* (Boyer, P. D., Ed.) pp 63−91, Academic Press, New York.
23. Ikemura, H., Takagi, H., and Inoue, M. (1987) *J. Biol. Chem. 262*, 7859−7864.
24. Zhu, X. L., Ohta, Y., Jordan, F., and Inouye, M. (1989) *Nature 339*, 483−484.
25. Baker, D., Shiau, A. K., and Agard, D. A. (1993) *Curr. Opin. Cell Biol. 5*, 966−970.
26. Eder, J., and Fersht, A. R. (1995) *Mol. Microbiol. 16*, 609−614.
27. Creighton, T. E. (1993) *Proteins. Structure and molecular properties*, W. H. Freeman, New York.
28. Sohl, J. L., Jaswal, S. S., and Agard, D. A. (1998) *Nature 395*, 817−819.
29. Nowak, U. K., Cooper, A., Saunders, D., Smith, R. A., and Dobson, C. M. (1994) *Biochemistry 33*, 2951−2960.
30. Foltmann, B. (1981) *Essays Biochem. 17*, 52−84.
31. Davies, D. R. (1990) *Annu. Rev. Biophys. Biophys. Chem. 19*, 189−215.
32. Sielecki, A. R., Fedorov, A. A., Boodhoo, A., Andreeva, N. S., and James, N. G. (1990) *J. Mol. Biol. 214*, 143−170.
33. Buzzel, A., and Sturtevant, J. (1952) *J. Am. Chem. Soc. 74*, 1983−1987.
34. Edelhoch, H. (1957) *J. Am. Chem. Soc. 79*, 6100−6109.
35. Favilla, R., Parisoli, A., and Mazzini, A. (1997) *Biophys. Chem. 67*, 75−83
36. Lin, X. L., Loy, J. A., Sussman, F., and Tang, J. (1993) *Protein Sci. 2*, 1383−1390.
37. Ahmad, F., and McPhie, P. (1978) *Biochemistry 17*, 241−246.
38. Lin, X. L., Lin, Y. Z., Koelsch, G., Gustchina, A., Wlodawer, A., and Tang, J. (1992) *J. Biol. Chem. 267*, 17257−17263.
39. Aoki, K., Taguchi, H., Shindo, Y., Yoshida, M., Ogasahara, K., Yutani, K., and Tanaka, N. (1997) *J. Biol. Chem. 272*, 32158−32162.
40. Piotto, M., Saudek, V., and Sklenar, V. (1992) *J. Biomol. NMR 2*, 661−665.
41. Glatter, O., and Kratky, O. (1982) *Small-Angle X-ray Scattering*, Academic Press, New York.
42. Guinier, A., and Fournet, B. (1955) *Small-Angle Scattering of X-rays*, John Wiley, New York.
43. Konno, T., Kataoka, M., Kamatari, Y., Kanaori, K., Nosaka, A., and Akasaka, K. (1995) *J. Mol. Biol. 251*, 95−103.

44. Olah, G. A., Trakhanov, S., Trewella, J., and Quiocho, F. A. (1993) *J. Biol. Chem. 268*, 16241−16247.
45. Moore, P. B. (1980) *J. Appl. Crystallogr. 13*, 168−175.
46. Goto, Y., and Fink, A. L. (1989) *Biochemistry 28*, 945−952.
47. Kuwajima, K. (1989) *Proteins: Struct., Funct., Genet. 6*, 87−103.
48. Ptitsyn, O. B. (1995b) *Adv. Protein Chem. 47*, 83−229.
49. Wüthrich, K. (1976) *NMR in biological research, peptides and proteins*, North-Holland Publishing, Amsterdam.
50. Bundi, A., and Wuthrich, K. (1979) *Biopolymers 18*, 285−297.
51. Bai, Y., Milne, J. S., Mayne, L., and Englander, S. W. (1993) *Proteins: Struct., Funct., Genet. 17*, 75−86.
52. Sosnick, T. R., and Trewhella, J. (1992) *Biochemistry 31*, 8329−8335.
53. Damaschun, G., Damaschun, H., Gast, K., Misselwitz, R., Müller, J. J., Pfeil, W., and Zirwer, D. (1993) *Biochemistry 32*, 7739−7746.
54. Kataoka, M., Nishii, I., Fujisawa, T., Ueki, T., Tokunaga, F., and Goto, T. (1995) *J. Mol. Biol. 249*, 215−228.
55. Konno, T., Kamatari, Y. O., Kataoka, K., and Akasaka, K. (1997a) *Protein Sci. 6*, 2242−2249.
56. Kamatari, Y. O., Konno, T., Kataoka, M., and Akasaka, K. (1996) *J. Mol. Biol. 259*, 512−523.
57. Tanford, C. (1970) *Adv. Protein Chem. 24*, 1−95.
58. Spraggon, G., Phillips, C., Nowak, U. K., Ponting, C. P., Saunders, D., Dobson, C. M., Stuart, D. I., and Jones, E. Y. (1995) *Structure 3*, 681−691.

BI991923D