Porta, R., Esposito, C., Gentile, V., Mariniello, L., Peluso, G., & Metafora, S. (1990) *Int. J. Pept. Protein Res. 35*, 117–122.
Pucci, P., Malorni, A., Marino, G., Metafora, S., Esposito, C., & Porta, R. (1988) *Biochem. Biophys. Res. Commun. 154*, 735–740.
Wagner, C. L., & Kistler, W. S. (1987) *Biol. Reprod. 36*, 501–510.
Williams-Ashman, H. G. (1984) *Mol. Cell. Biochem. 58*, 51–61.
Williams-Ashman, H. G., & Canellakis, Z. N. (1980) *Physiol. Chem. Phys. 12*, 457–472.

Characterization of a Partially Denatured State of a Protein by Two-Dimensional NMR: Reduction of the Hydrophobic Interactions in Ubiquitin†

Margaret M. Harding,‡,§ Dudley H. Williams,*,‡ and Derek N. Woolfson‡,∥
University Chemical Laboratory, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, U.K., and Department of Biochemistry, University of Cambridge, Tennis Court Road, Cambridge CB2 1QW, U.K.
Received February 21, 1990; Revised Manuscript Received October 19, 1990

ABSTRACT: A stable, partially structured state of ubiquitin, the A-state, is formed at pH 2.0 in 60% methanol/40% water at 298 K. Detailed characterization of the structure of this state has been carried out by 2D NMR spectroscopy. Assignment of slowly exchanging amide resonances protected from the solvent in the native and A-state shows that gross structural reorganization of the protein has not occurred and that the A-state contains a subset of the interactions present in the native state (N-state). Vicinal coupling constants and NOESY data show the presence of the first two strands of the five-strand β-sheet that is present in the native protein and part of the third β-strand. The hydrophobic face of the β-sheet in the A-state is covered by a partially structured α-helix, tentatively assigned to residues 24–34, that is considerably more flexible than the α-helix in the N-state. There is evidence for some fixed side-chain–side-chain interactions between these two units of structure. The turn-rich area of the protein, which contains seven reverse turns and a short piece of 3_{10} helix, does not appear to be structured in the A-state and is approaching random coil.

Ubiquitin is a cytoplasmic protein of 76 amino acids (8565 Da) found in all eukaryotic cells (Goldstein et al., 1975); it participates in a wide variety of cellular activities (Hershko et al., 1982, 1984; Chin et al., 1982). The X-ray crystal structure of human erythrocytic ubiquitin has been determined (Vijay-Kumar et al., 1985, 1987) and shows the extremely compact nature of the protein with approximately 87% of the polypeptide chain involved in hydrogen-bonded secondary structure (Figure 1) involving a 5-strand mixed β-sheet, 3.5 turns of an α-helix, a short piece of 3_{10} helix, and 7 reverse turns. The β-sheet has a characteristic left-handed twist, and the α-helix fits into the concavity formed by the sheet. There are several unusual features of the structure, namely two β-bulges, one of which is parallel, a highly contorted turn-rich area of 21 residues that contains four reverse turns, and a short piece of 3_{10} helix that forms two interlocked type III reverse turns. ^{1}H NMR[1] assignments of human ubiquitin have been carried out by two groups using sequential (Weber et al., 1987) and main-chain-directed (Di Stefano & Wand, 1987) structure analysis. Both solution structures, which were carried out at pH 4.7 and 5.8, respectively, agree with the solid-state structure. The relatively small size of ubiquitin, the X-ray crystal structure, and the assigned NMR spectrum provide a protein ideal for studying aspects of protein structure and folding by NMR spectroscopy.

Significant advances in protein engineering (Fersht, 1987) and the application of 2D NMR methods to protein structures (Wüthrich, 1989a,b) during the last decade have provided techniques that allow fundamental questions regarding protein structure and folding to be addressed. However, characterization of states other than the native or random-coil state of proteins has met with limited success. A knowledge of partially structured states, intermediate between the two extreme states, should provide insight into the relative strengths and different types of interactions that can stabilize a given section of the polypeptide chain and may provide information regarding the interactions responsible for directing the folding pathways of proteins.

Structural information about states present on the folding pathways of proteins has not generally been directly attainable as folding intermediates are invariably short-lived and not highly populated (Kim & Baldwin, 1982), and it has been necessary to develop methods whereby such kinetic intermediates can be trapped (Creighton, 1978; Biringer & Fink, 1982; Roder et al., 1988; Udgaonkar & Baldwin, 1988). More recently, much attention has been directed toward the so-called

† This work was supported in part by the U.K. Science and Engineering Research Council Molecular Recognition Initiative Programme and Pfizer Central Research, Sandwich. M.M.H. acknowledges receipt of a Ramsay Memorial General British Fellowship.
‡ University Chemical Laboratory.
§ Present address: Department of Organic Chemistry, University of Sydney, Broadway, NSW, 2006 Australia.
∥ Department of Biochemistry.

[1] Abbreviations: NMR, nuclear magnetic resonance; 1D and 2D, one and two dimensional; COSY, two-dimensional spin correlated spectroscopy; DQF, double-quantum filtered; NOE, nuclear Overhauser effect; NOESY, two-dimensional nuclear Overhauser enhancement spectroscopy; HOHAHA, two-dimensional homonuclear Hartmann–Hahn spectroscopy; DANTE, solvent suppression pulse sequence; CD, circular dichroism; MLEV, two-dimensional isotropic mixing experiments employing MLEV-17 or modified MLEV-17 mixing pulses; N-state, native state of the protein; U-state, unfolded state of the protein; A-state, partially folded state of the protein.

0006-2960/91/0430-3120$02.50/0

(a)

				5					10
Met	Gln	Ile	Phe	Val	Lys	Thr	Leu	Thr	Gly
				15					20
Lys	Thr	Ile	Thr	Leu	Glu	Val	Glu	Pro	Ser
				25					30
Asp	Thr	Ile	Glu	Asn	Val	Lys	Ala	Lys	Ile
				35					40
Gln	Asp	Lys	Glu	Gly	Ile	Pro	Pro	Asp	Gln
				45					50
Gln	Arg	Leu	Ile	Phe	Ala	Gly	Lys	Gln	Leu
				55					60
Glu	Asp	Gly	Arg	Thr	Leu	Ser	Asp	Tyr	Asn
				65					70
Ile	Gln	Lys	Glu	Ser	Thr	Leu	His	Leu	Val
				75					
Leu	Arg	Leu	Arg	Gly	Gly				

KEY : α-helix 3_{10}-helix reverse turn
β-sheet strand 1 (11-17), strand 2 (1-7), strand 3 (64-70), strand 4 (41-45), strand 5 (48-50)

(b)

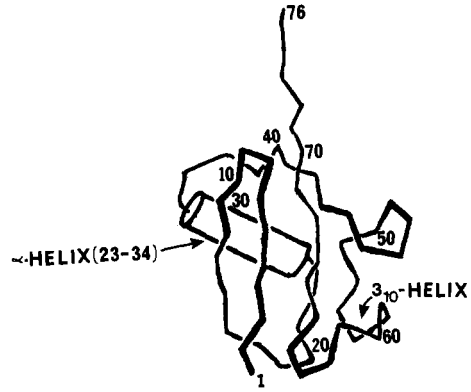

FIGURE 1: (a) Amino acid sequence of human erythrocytic ubiquitin showing the location of the secondary structure units. (b) Peptide backbone showing elements of the secondary structure.

"molten globule" state that has been detected for several small proteins (Dolgikh et al., 1981, 1985; Ohgushi & Wada, 1983; Baum et al., 1989; Goto & Fink, 1989). The term "molten globule" refers to a highly compact state of the protein that has a high content of secondary structure as estimated by CD but in which the side chains do not have fixed conformations. These species are equilibrium intermediates and are thus amenable to study in real-time experiments. Despite this, a full structural characterization of such partially structured states has not been possible to date, the most informative study being an NMR study on the molten globule state of α-lactalbumin (Baum et al., 1989). The amount of structural detail obtained in this study, however, was limited, primarily due to the poorly resolved amide proton resonances and broadened spectra.

During the course of structural studies of ubiquitin, we observed by NMR spectroscopy a partially structured stable state of the protein in 60% methanol at pH 2. This paper reports characterization of this state, termed the A-state, by 2D NMR spectroscopy. The presence of well-defined slowly exchanging amide resonances allowed us to obtain detailed information on the existing secondary and tertiary interactions. Studies of such structures provide valuable insight into the relative stability of regions within the protein and indicate the interactions that can be stabilized independently of the total global fold of the protein.

Materials and Methods

Ubiquitin, isolated from bovine red blood cells, was purchased from Sigma and used without further purification. NMR samples in 60% methanol were prepared by dissolving ubiquitin in 300 μL of CD_3OD and 200 μL of D_2O or H_2O to give solutions 1.5–2.5 mM in protein; the pH was then adjusted to 2. Those in 40% methanol used 200 μL of CD_3OD and 300 μL of D_2O and were 1–6 mM in protein; again the pH was adjusted after dissolution. All samples were prepared with isotopically labeled solvents; DCl and NaOD were purchased from Aldrich. The pH readings are uncorrected for deuterium isotope effects. Bovine α-lactalbumin was obtained from Sigma, and plastocyanin was prepared from spinach. Gel filtration was performed in a column of dimensions 100 × 2.5 cm packed with Sephadex G-50 (superfine) resin equilibrated with a 50 mM phosphate buffer at pH 7.

Proton NMR spectra were recorded with Bruker AM-400 and AM-500 spectrometers equipped with Aspect 3000 computers. Spectra were recorded over spectral widths of 4000 or 5000 Hz (at 400 and 500 MHz, respectively) with quadrature detection employed throughout. Chemical shifts were measured relative to internal dioxane at 3.73 ppm. Two-dimensional proton NMR spectra in D_2O were acquired in the phase-sensitive mode with quadrature detection in f_2 and with time proportional phase incrementation in f_1 (Marion & Wüthrich, 1983). Data sets resulting from 256 to 512 increments of t_1 were acquired and zero-filled to 1024 points, with each free induction decay composed of 2048 data points. Typically 64 or 96 transients were recorded for each increment of t_1 with a recycle delay of 1.2–2.0 s between transients. Solvent suppression in NOESY and COSY spectra was achieved by coherent presaturation during the recycle delay and mixing times. HOHAHA spectra were recorded by using a 7-kHz spin-locking field with an MLEV-17 sequence of 40–80 ms. Solvent suppression was achieved by a DANTE pulse sequence directly prior to a SCUBA preirradiation pulse to enhance the α resonances (Brown et al., 1988). Base-line correction was achieved with the sine modulation method. Data were subjected to shifted sine-bell or squared sine-bell weighting functions in f_1 and f_2.

Results

The Folded State: The N-State. The 400-MHz proton NMR spectrum of ubiquitin in D_2O is presented in Figure 2a. The chemical shift dispersion is typical of a globular protein of this size with chemical shifts spread over a much wider range of values than for the corresponding random-coil amino acids; the values reflect the local environment of the individual amino acids in the protein (Figure 1). The presence of slowly exchanging amide protons in the region 7–9 ppm reflects strong hydrogen bonding within the protein, which prevents exchange of the amide protons with solvent deuterions. Spectra recorded immediately after the protein was dissolved in D_2O show 26 amide resonances that exchange over a period of days to give a subset of 7 protons (Figure 2a) that are still present after weeks.

In terms of identifying and studying transient intermediates or partially structured states in the folding pathway of ubiquitin by NMR spectroscopy, several regions of the spectrum are particularly well suited for monitoring any small changes in chemical shift that may occur on disruption of the native structure. There are only four aromatic residues in the protein, Phe 4, Phe 45, Tyr 59, and His 68, and the spin systems of the histidine and tyrosine resonances are readily identified in the 1D spectrum (Figure 2a). The α proton of Phe 4 also resonates in a relatively downfield position at 5.6 ppm. The only other region of the spectrum not containing overlapping resonances is upfield of 0 ppm where the methyl doublet of Leu 50 (−0.17 ppm) and the γ proton of Ile 61 (−0.37 ppm) are located. These residues are located in different elements

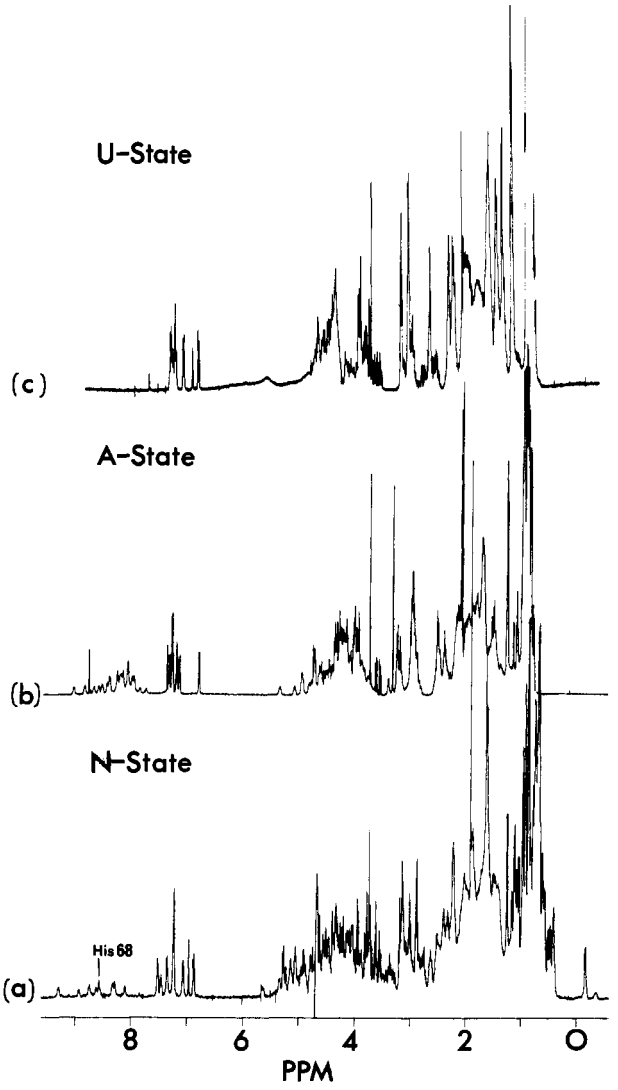

FIGURE 2: ^{1}H NMR spectra of ubiquitin: (a) 500 MHz, pH 5.8, D_2O, and 298 K after several months (N-state); (b) 500 MHz, pH 2, 60% CD_3OD/40% D_2O, and 298 K after 30 min (A-state); (c) 400 MHz, pH 2, urea (8 M), and 350 K (U-state).

of the secondary structure: Phe 4 (β-strand), Leu 50 (β-strand 5, methyl shielded by Tyr 59), Phe 45 (β-strand 4), Tyr 59 (reverse turn, part of the "turn-rich" area), Ile 61 (reverse turn, part of the "turn-rich" area, H^γ shielded by Phe 45), and His 68 (β-strand 3). Thus, disruption of the local tertiary and/or secondary structure in which these residues are located should be reflected in changes in chemical shifts of these residues and readily detected by NMR spectroscopy.

The Unfolded State: The U-State. Ubiquitin has been reported to be stable, with respect to unfolding, in the temperature range 23–80 °C and in the pH range 1.18–8.43 (Lenkinski et al., 1977; Cary et al., 1980). Titration experiments with urea and guanidinium chloride established that reversible unfolding/folding can be achieved by thermal means in 3 M guanidinium chloride or by variation of the concentration of guanidinium chloride (3–6 M) at room temperature. The protein retains its native structure at pH 5.8 in the temperature range 298–355 K but precipitates at higher temperatures. At pH lower than 2, the protein remains soluble in the range 298–368 K, and spectra were assigned to the U-state above 350 K (Figure 2c). Almost identical spectra of the U-state were obtained in urea (8 M) at 350 K (Figure 2c) and guanidinium chloride (6 M) at 298 K. The reduced spread of chemical shifts and the absence of peaks below 0

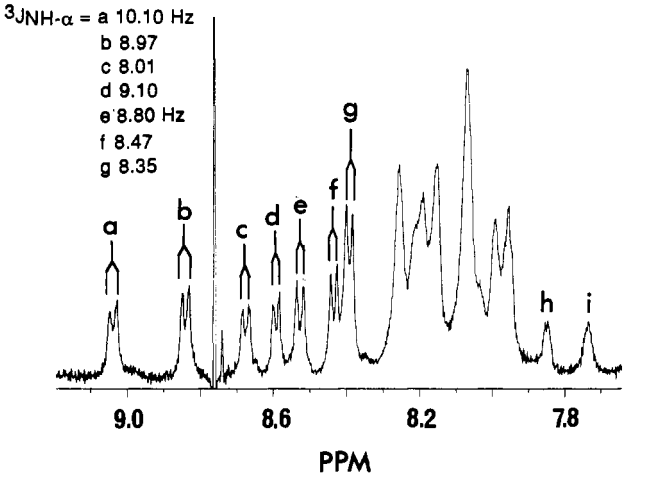

FIGURE 3: Resolution-enhanced section of the 500-MHz ^{1}H NMR spectrum of the A-state of ubiquitin.

ppm and between 4.5 and 6 ppm is immediately apparent, and the chemical shift dispersion approximates that predicted for an ideal random-coil polypeptide (Bundi & Wüthrich, 1979).

Methanol-Induced Conformational Changes: The A-State. Ubiquitin retains its native structure on changing the pH from 5.5 to 2.0, and with the exception of H4 of His 68, which shifts downfield due to protonation of the ring, the resonances show only small chemical shift differences. Titration of methanol into a sample at pH 2 resulted in the gradual appearance of a new species, termed the A-state, and only this species was detected in 60% methanol. During this experiment the pH of the sample was not adjusted for changes that may occur upon addition of methanol. However, subsequent samples were adjusted to pH 2 after dissolution of protein in the appropriate quantities of methanol and water (Materials and Methods).

The NMR spectrum of the A-state (Figure 2b), however, is considerably different from those of both the native and random-coil protein. The most notable features are the absence of signals upfield of 0 ppm and the presence of a large number of N–H resonances (Figure 3) that exchange over a period of ca. 24 h. This behavior is indicative of protons that are hydrogen bonded and hence implies the presence of some secondary structure in the protein under these conditions. Several peaks in the region 4.5–5.6 ppm, which are downfield of the expected values for residues in a random-coil polypeptide chain (Bundi & Wüthrich, 1979), also support the presence of a partially structured state.

Vicinal Spin–Spin Coupling Constants. Several of the N–H doublets in the spectrum of the A-state were resolved (Figure 3). Vicinal spin–spin couplings, $^3J_{NH\alpha}$ have been used in the assignment of the regular secondary structure in proteins (Wüthrich, 1986) with typical values as follows: α-helix, 3.9 Hz; 3_{10} helix, 4.2 Hz; antiparallel β-sheet, 8.9 Hz; parallel β-sheet, 9.7 Hz. The values obtained for the vicinal coupling constants in the resolved doublets (8–10 Hz) strongly suggest the presence of the β-sheet structure. Furthermore, the coupling constants of the two partially exchanged signals at 7.75 and 7.88 ppm (labeled h and i) are clearly considerably smaller and suggest some helical structure or at least that the amide protons giving rise to these signals are involved in a different type of interaction from those amide protons that give resolved doublets. The fact that these signals are at almost half-intensity compared to those arising from the putative β-sheet implies that N–H protons in this segment of the protein are more exposed to the solvent and suggests a greater degree of flexibility in this section of the structure. In addition, their

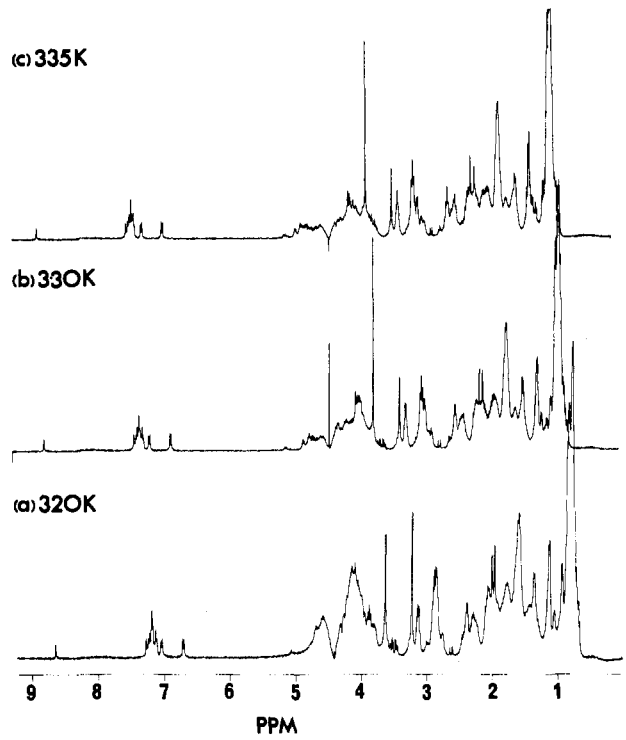

FIGURE 4: Temperature dependence of the 400-MHz ^1H NMR spectrum of ubiquitin in 60% CD_3OD/40% D_2O, pH 2 at (a) 320 K, (b) 330 K, and (C) 335 K.

lower chemical shifts (Figure 3) relative to the resolved doublets are consistent with their involvement in relatively weak hydrogen bonding, again in accord with their presence in a more flexible part of the structure.

Physical Properties. Preparation of soluble samples in 60% CD_3OD/40% D_2O in the pH range 5.5–2.0 was concentration dependent, and under these conditions the protein showed a strong tendency to aggregate at high concentrations. This behavior is consistent with some exposure of the hydrophobic residues to the solvent. In 60% methanol it is difficult to estimate what percentage of the unburied hydrophobic residues could be tolerated. However, at the concentrations at which the NMR experiments were carried out, line-width measurements were consistent with a monomeric unit of protein (see later).

Variable-Temperature NMR. The NMR spectrum assigned to the U-state of ubiquitin at pH 2 (Figure 2c) represents the averaged spectrum for the many conformational states that are rapidly equilibrating in the nonstructured polypeptide chain. A variable-temperature experiment was carried out on ubiquitin in 60% CD_3OD/40% D_2O, pH 2, in order to see if the secondary structure present could be further disrupted and the protein completely unfolded. This would allow characterization of the U-state in conditions similar to those in which the A-state exists.

Spectra were recorded in the range 320–335 K (Figure 4). Gradual sharpening of the entire spectrum occurred, particularly in the region from 4 to 5 ppm, and some compression of the aliphatic region was also observed. However, even at 335 K (Figure 4c), resonances were still detected downfield of 4.5 ppm. Increasing the methanol content to 71% induced no significant changes, and recooling to 300 K (not shown) gave essentially the same spectrum of the A-state. The spectra obtained support the idea that the protein is approaching the random coil with increasing temperature but, on the basis of chemical shifts, some ordered structure is still present even at 335 K.

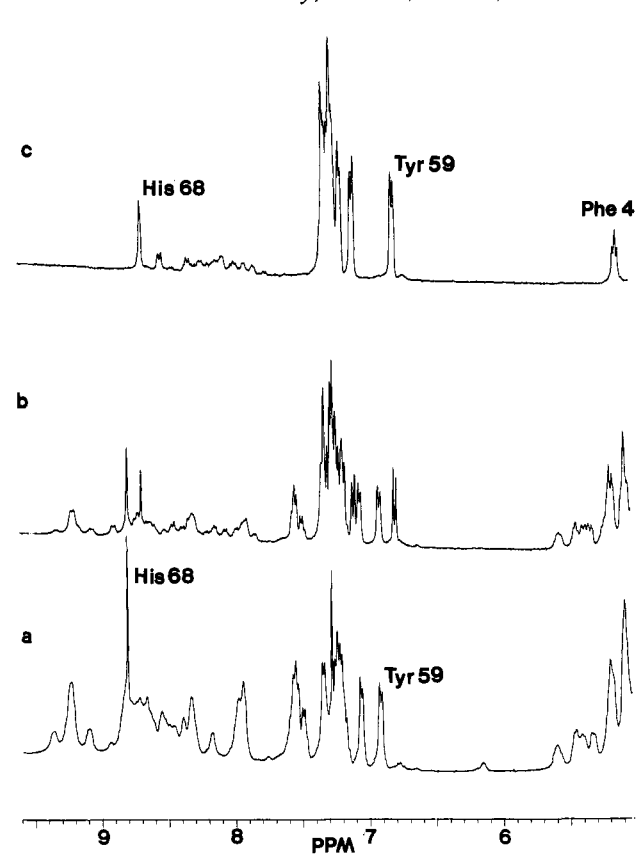

FIGURE 5: Downfield region of the 400-MHz ^1H NMR spectrum of ubiquitin in 40% CD_3OD/60% D_2O, pH 2, (a) at 300 K, in the N-state; (b) at 308 K, near the midpoint of the N–A unfolding transition; and (c) at 320 K, in the A-state.

Equilibrium Mixtures of the N- and the A-States. At lower concentrations of methanol (30%–60%) and at 298 K equilibrium mixtures of the N- and the A-states of ubiquitin are observed. The positions of these equilibria are both temperature (Figure 5) and concentration dependent. Analysis of this system by 1D NMR is proving useful in understanding the thermodynamics and kinetics of the unfolding transition. It has also allowed the question of whether or not the A-state is aggregated to any degree to be addressed.

The concentration dependence of the NMR spectrum of a ubiquitin sample at pH 2 and at 300 K in 40% CD_3OD has the following features: the relative amounts of N- and A-states as judged by resonance intensities change from approximately 30:1 to 1:1 in the concentration range 6–1 mM protein (spectra not shown). This experiment provides evidence for some aggregation of the N-state relative to the A-state under these conditions. Indeed, this conclusion is consistent with the broader resonances of the Tyr 59 protons (ortho to the hydroxyl group, 6.91 ppm) in the N-state in 40% methanol relative to the same proton resonance (6.8 ppm) in the A-state (Figure 5b). A demonstration that the A-state is monomeric by determining its effective molecular weight with gel filtration was not possible because the gel collapsed when run under the conditions required to populate only the A-state (60% methanol). However, all the available NMR evidence is consistent with the A-state being monomeric. We have made measurements of the line widths of all those proton resonances (His 68, 2-H; Tyr 59, 3,5-H; and Phe 4, Cα-H) that are resolved in spectra of the A- and N-states. The line widths of these resonances give us probes of the correlation time for tumbling (and hence information on the presence or absence of associated species) in that part of the A-state that we later conclude to be disordered (residues 37–62) and the portion we conclude

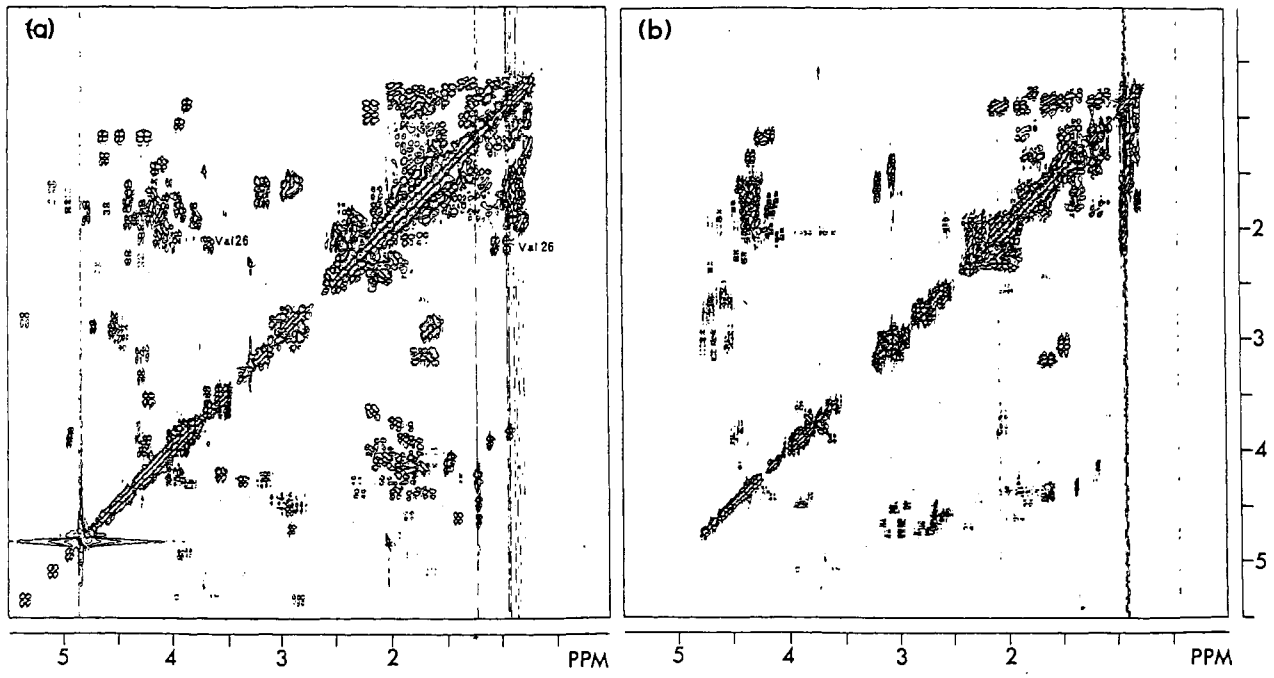

FIGURE 6: Upfield region of the 400-MHz DQF-COSY spectra of ubiquitin in (a) the A-state 60% CD_3OD/40% D_2O, pH 2, at 298 K; and (b) the U-state, in urea (8 M) at 350 K.

to retain the structure characteristic of some of the β-sheet of folded ubiquitin. For the N-state of ubiquitin in D_2O (4 mg of protein in 0.5 mL), the resonance widths at half-height are 3.5 Hz (his 68), 13.2 Hz (Tyr 59), and 26.4 Hz (Phe 4). For the A-state at the same concentration of sample (4 mg, 200 μL of D_2O, 300 μL of CD_3OD, pH 2.0) run immediately afterward on the same instrument, the resonance line widths at half-height are 3.2 Hz (His 68), 12.9 Hz (Tyr 59), and 26.2 Hz (Phe 4). In addition to this, we find that the N-state of ubiquitin in water runs as a monomer in gel filtration; the protein is resolved from plastocyanin (11 kDa) and α-lactalbumin (14.6 kDa) and elutes after them in these experiments. This result, together with the similar line widths of the widely distributed proton resonances, gives support to the conclusion (in the absence of a direct determination of molecular weight) that the A-state is monomeric.

Refolding Experiment. The limited chemical shift dispersion in certain regions of the spectrum of the A-state did not allow unambiguous assignment of many resonances. Similar problems of poor resolution due to widespread overlap and substantial line widths of resonances were reported in the assignment of the molten globule of guinea pig α-lactalbumin, and methods were developed to exploit the well-resolved spectrum of the native protein by using pH-jump experiments to label the amide protons protected from the solvent in the molten globule state (Baum et al., 1989). The problems of assignment of the A-state of ubiquitin are much less severe than in the case of α-lactalbumin, as the spectrum is relatively sharp and the amide protons are resolvable. However, an experiment similar to that used by Baum et al. proved to be extremely informative in assignment of the slowly exchanging amide protons in the A-state.

The protein, in the A-state, was prepared in 60% CD_3OD/40% D_2O at pH 2.0 and kept at room temperature for 30 min, i.e., a sufficient period of time to allow all amide protons not protected from the solvent to exchange with deuterons. The methanol was evaporated off by centrifugation under vacuum (to give the refolded protein at pH 2.0), and the pH was raised to 4.7, allowing direct comparison of the spectrum with the literature data. This experiment allowed immediate identification of the amide resonances protected from the solvent in the A-state *and* in the folded protein, i.e., Ile 3, Phe 4, Val 5, Lys 6, Thr 7, Ile 13, Leu 15, Val 17, Val 26, Ala 28, Glu 34, Leu 67, Leu 69, and Val 70. The most striking result from this experiment is that 11 of these residues are a subset of the slowly exchanging amides that form the extended β-sheet structure in the N-state (Figure 10a; Weber et al., 1987) and 3 residues are from the α-helix of the N-state. This result is strong evidence for conservation of a significant portion of the β-sheet in the A-state as well as a section of α-helix. The fact that resonances are detected in this experiment is clear evidence that the A-state contains a subset of the interactions present in the N-state. However, this experiment provides no information on any structure that is present in the A-state but not in the N-state.

Assignment of the A-State. The upfield regions of the DQF-COSY spectra of the A-state and the unfolded protein (8 M urea, 350 K) are presented in Figure 6. The large difference between these 2D spectra emphasizes that the A-state is not a random coil. A combination of DQF-COSY, HOHAHA, and NOESY spectra of the protein at pH 2.0 in both 60% CD_3OD/40% D_2O and 60% CD_3OD/40% H_2O was used to assign the spectrum of the A-state.

The assignment of all spin systems in the A-state was not possible due to the severe peak overlap in some regions of the spectrum. However, several spin systems, notably Phe 4 Hα, could be immediately identified by comparison of the spectra of the A-state and the N-state and served as starting points for determining the location of adjacent spin systems through a combination of COSY, HOHAHA, and NOESY experiments.

HOHAHA spectra were essential for unambiguous assignment of the slowly exchanging amide resonances (Figure 7). In freshly prepared samples of the A-state, 15–17 amide resonances were detected by integration of the 1D spectrum; ca. 10 of these peaks (bracketed in Figure 7) exchanged much more quickly with the solvent than the 8 well-resolved peaks. For comparison, in freshly prepared samples of the N-state of ubiquitin, 26 slowly exchanging amides are detected that exchange with different rates over a period of weeks (un-

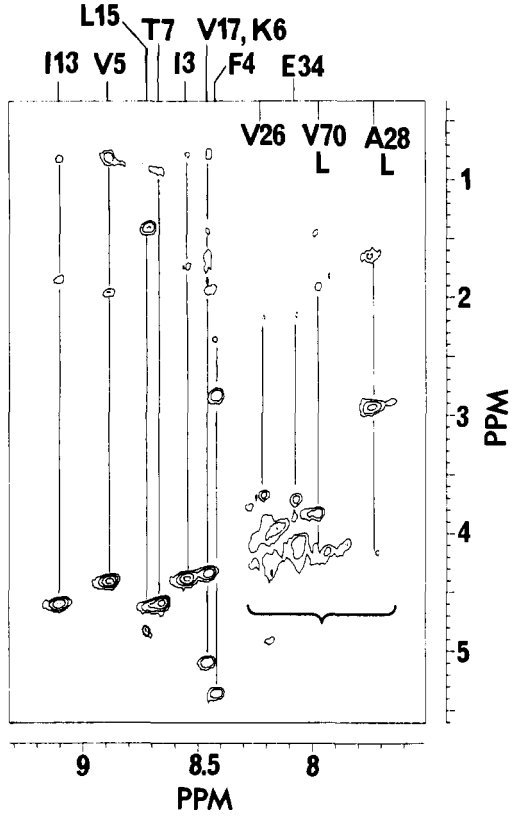

FIGURE 7: Section of the 400-MHz HOHAHA (MLEV 60 ms) spectrum on the A-state showing assignment of the slowly exchanging amide protons. Other peaks assignable at a lower contour level for spin systems are in the bracketed region. Unnumbered leucine spin systems are those from this type of residue that are as yet unassigned in a sequence-specific manner in the A-state.

published data). The refolding experiments (see above) identified the amide resonances protected in the A- and the N-states. Thus, the spectrum shown in Figure 7 must contain *at least* these 14 resonances. With the aid of NOESY experiments, the eight downfield spin systems (8.2–9.2 ppm) were assigned as were Val 26, Val 70, and Glu 34. The spin systems of Ala 28, Leu 67, and Leu 69 could not be unambiguously assigned from this spectrum, and their assignments (Figure 7), based on chemical shifts, are tentative. There are clearly several more unassigned peaks that resonate between 7.6 and 8.3 ppm. These peaks were not detected in the refolding experiment, either because they are hydrogen bonded in a structure not present in the N-state or because they have exchanged during the time course of the experiment (refolding and acquisition of the 2D spectrum).

The subset of eight more slowly exchanging protons in the A-state (Figure 7) was assigned to two strands of the β-sheet structure (Figure 8) by a NOESY experiment. These protons gave rise to $d_{N\alpha}$ cross-peaks as indicated. In addition, the interstrand $d_{\alpha\alpha}$ NOEs between the α protons of Thr 14/Phe 4, Lys 6/Thr 12, and Gln 2/Glu 16 provide conclusive evidence that the strands are the *same* two β-strands that are present in the N-state and confirm the results of the refolding experiment. All hydrogen-bonded amides defined in this section of the β-sheet exchange slowly with the solvent. The presence of a third β-strand parallel to residues 1–7 is supported by a $d_{N\alpha}$ interstrand cross-peak (Phe 4/Thr 66) and the slowly exchanging amide resonances of Phe 4, Lys 6, Leu 67, and Leu 69.

In the refolding experiment, three residues that are located in the α-helix in the N-state were detected (Val 26, Ala 28, and Glu 34). However, in the NOESY experiment acquired

Table I: Chemical Shifts of Assigned Residues of the A-State (60% Methanol, 298 K) of Ubiquitin

amino acid	NH	Hα	Hβ	other
Gln 2		4.93	1.64, 1.80	2.18
Ile 3	8.54	4.37	1.73	0.76
Phe 4	8.42	5.38	2.85	
Val 5	8.89	4.41	1.98	0.83
Lys 6	8.46	5.14	1.42, 1.66	1.73
Thr 7	8.65	4.61	0.92	
Thr 12		4.81	3.82	0.83
Ile 13	9.09	4.61	1.84	0.83
Thr 14		4.83	3.83	1.04
Leu 15	8.72	4.61	1.40	
Glu 16		3.95	1.60, 1.82	
Val 17	8.42	4.32	1.73	0.76
Val 26	8.21	3.69	2.18	1.05, 0.92
Glu 34	8.04	3.72	1.84, 1.65	2.12
Thr 66		4.38		0.71
His 68		4.73	2.93	
Val 70	7.92	4.15	1.92	0.83

in D_2O/CD_3OD (Figure 8) no interresidue d_{NN} NOEs, in particular NOEs characteristic of α-helical structure (Wüthrich, 1986), were detected in support of the possible existence of an α-helix in the A-state. In this experiment, cross-peaks were detected only for the amide signals that were at full or almost full intensity at the start of the experiment and did not decrease in intensity substantially during the time frame of the experiment (Figure 3). Hence, a NOESY spectrum was also recorded in 60% CD_3OD/40% H_2O, pH 2 (Figure 9). In contrast to the NOESY experiment carried out in fully deuterated solvents, the spectrum showed d_{NN} cross-peaks arising from the amide protons present at low intensity or not observed in fully deuterated solvents. This spectrum provides conclusive evidence for the presence of an α-helix in the A-state. The result clearly supports the conclusions reached from the measurement of vicinal coupling constants and from the refolding experiments, i.e., the existence of a loosely structured α-helical segment in the A-state.

The assignment of the full spin system of Val 26 (Figure 7) was used as a starting point to tentatively assign the helical segment of the A-state by assuming $d_{NN}(i, i+1)$ connectivities (Wüthrich, 1986). In principle, these assignments can be confirmed by a HOHAHA experiment in 60% CD_3OD/40% H_2O. However, severe peak overlap in both the NOESY and HOHAHA experiments did not permit determination of connectivity patterns and unequivocal assignment of a spin system to a given residue in all cases. The number of cross-peaks and the connectivity patterns between them (Figure 9), as well as the results of the refolding experiment, support a proposal for the presence of an α-helical segment involving residues 24–34. The chemical shifts of the spin system of Val 26 agree well with the chemical shifts observed in the N-state. In particular, the two methyl peaks of this residue resonate downfield of all other aliphatic methyl groups (see Figure 6a), and these shifts provide evidence for its presence in an α-helix in both the A- and N-states. The α-helix in the A-state must be considerably more flexible than the α-helix in the N-state, the reduced rigidity of the helix allowing N–H exchange with the solvent to occur in D_2O/CD_3OD. The chemical shifts of the resonances assigned, thus far, in the A-state are summarized in Table I.

Discussion

The A-state of ubiquitin exists under conditions of low pH in aqueous methanol. Ubiquitin contains 11 acidic residues (5 Asp, 6 Glu) and 10 basic residues (7 Lys, 2 Arg, 1 His) whose charge–charge interactions will be altered at low pH.

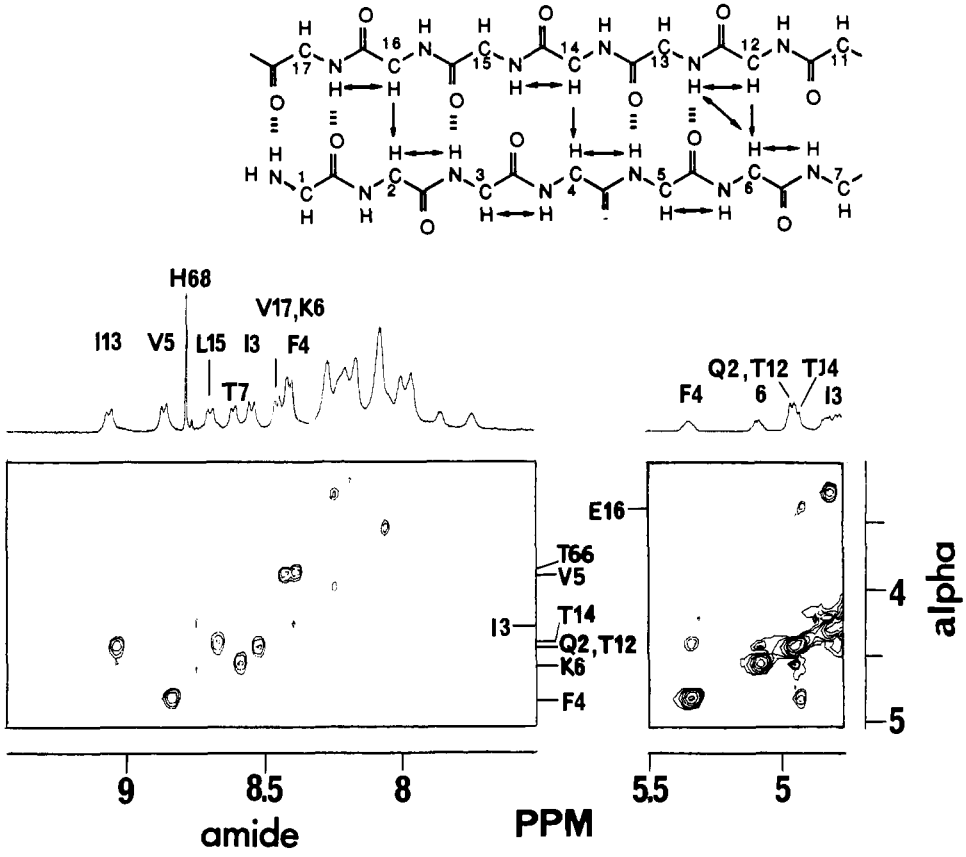

FIGURE 8: Section of the 500-MHz phase-sensitive NOESY (t_m 200 ms) showing assignment of the β-sheet structure in the A-state.

In particular, the salt bridges between Lys 11/Glu 34 and Lys 27/Glu 48, which serve to "anchor" the α-helix to the β-sheet, will be broken. Aqueous methanol should serve to weaken hydrophobic interactions, as the driving force for burying the nonpolar residues is greatly diminished, compared to water (Biringer & Fink, 1982). Thus, the stabilization of units of the secondary structure and the disruption of tertiary interactions are favored under these conditions.

Alcohol-induced changes of ubiquitin in which the native structure is converted to a form consistent with 50% helical structure as estimated by circular dichroism have been reported (Wilkinson & Mayer, 1986). Indeed, addition of alcohols is a common technique for increasing the helical content of proteins and peptides (Tanford, 1968; Timasheff, 1970). The results of this work show that, relative to the N-state, the helical content of ubiquitin under the conditions employed has decreased. Apparent discrepancies between the amount of secondary structure detected by NMR and circular dichroism in the molten globule state of α-lactalbumin have previously been noted (Baldwin, 1989), and these results emphasize that in complex systems the interpretation of circular dichroism data may not always be straightforward.

The β-sheet region of the A-state and the supporting NOE data are presented in Figure 10b. The data show clearly that the first two strands of the five-strand β-sheet of the N-state are conserved in the A-state. The A-state thus contains a subset of the interactions present in the N-state, and gross structural reorganization of the protein to another unrelated structure has not occurred.

While complete assignment of the α-helical section of the A-state was not possible, an α-helical region that is considerably more flexible than the α-helix in the N-state is present and most probably serves to cover the hydrophobic face of the β-sheet. The increased flexibility in this region of the polypeptide chain probably occurs as the hydrophobic face of the β-sheet would be expected to tolerate a higher degree of exposure to the solvent in 60% methanol. Some evidence for interaction of the helix with the hydrophobic face of the sheet is provided by chemical shifts of Val 26. The two methyl groups of this residue are significantly downfield on the other aliphatic resonances. The fact that two distinct methyl peaks are observed indicates that the side chain adopts a specific conformation. No direct evidence for the side chain interacting with the β-sheet was obtained in NOESY experiments. However, it is resonable to assume that the α-helix will adopt a conformation similar to that in the N-state, given that the fully assigned section of the A-state is indeed a subset of the secondary structure in the N-state.

The absence of signals upfield of 0 ppm indicates disruption of the turn-rich segment of the protein; this is not unreasonable considering the large number of polar residues in this region of the protein. Furthermore, disruption of this "side" of the protein probably can occur without gross effect on the first strands of the β-sheet. The poor chemical shift dispersion in certain regions of the spectra of the A-state prevented complete assignment of all residues by the sequential assignment method (Wüthrich, 1986), which is generally the method of choice for the assignment of protein spectra. The observed chemical shifts are entirely consistent with the partially structured nature of the A-state. This is well illustrated in the spectra in Figure 9b. The high concentration of peaks around 8 ppm presumably arises from amide protons in the unfolded region of the protein, and these chemical shifts approximate those of a random-coil protein (Bundi & Wüthrich, 1979). Further assignment of resonances of the A-state is in progress with magnetization-transfer experiments on samples containing mixtures of the A-state and N-state (Figure 5); in addition, the energetics and kinetics of the folding transition are being studied with similar

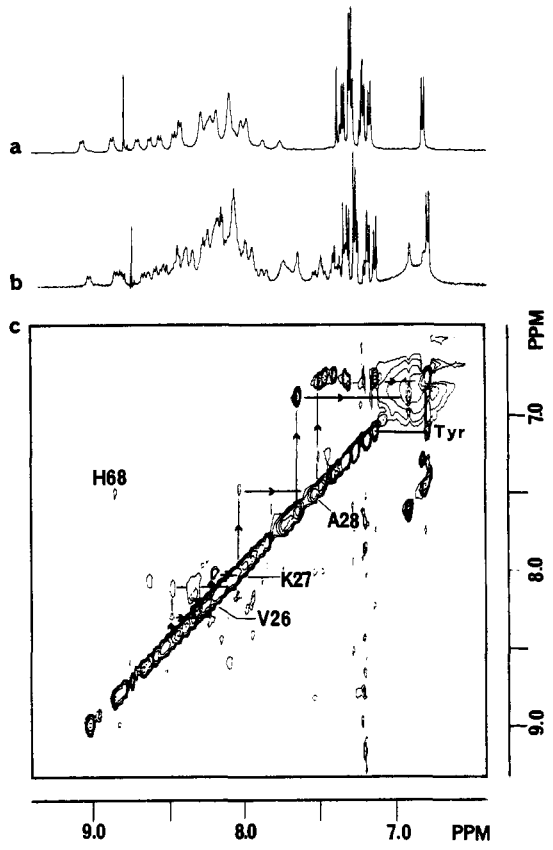

FIGURE 9: 500-MHz spectra of (a) the A-state in 60% CD_3OD/40% D_2O, pH 2; (b) the A-state in 60% CD_3OD/40% H_2O, pH 2; and (c) a section of the phase-sensitive NOESY (t_m 200 ms) of the A-state in 60% CD_3OD/40% H_2O, pH 2, showing d_{NN} NOEs.

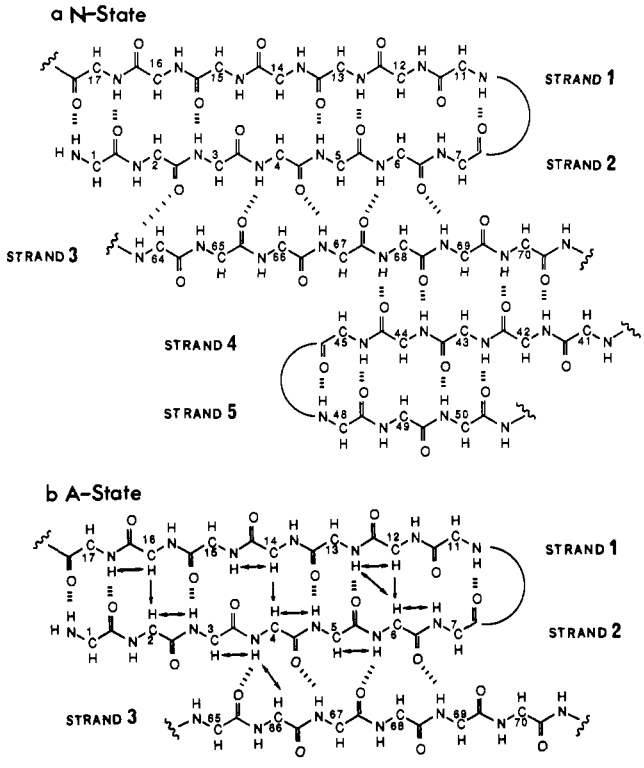

FIGURE 10: (a) β-Sheet structure in the N-state (Vijay-Kumar et al., 1987); (b) β-sheet structure in the A-state of ubiquitin.

experiments. Studies of this type should then establish the relevance, or otherwise, of this partially structured state to the folding pathway of ubiquitin. In particular, this system may allow the role of noncovalent interactions, exercised from the folded subdomain observed in the A-state, in directing the folding of the rest of the protein to be assessed.

CONCLUSIONS

A partially structured state of ubiquitin, the A-state, is formed in 60% CD_3OD/40% D_2O at pH 2. Characterization of this state by 2D NMR has allowed the following conclusions to be made: (1) Addition of methanol serves to disrupt the turn-rich area and the 3_{10} helix of the N-state, and residues 37–62 are essentially random coil. (2) The first three strands of the five-strand β-sheet of the N-state are largely conserved; this part of the A-state retains a high degree of stability, and the residues in this region of the A-state give rise to slowly exchanging amide resonances. (3) The α-helix of the N-state is also present in the A-state; reduction of its hydrophobic and ionic interactions with the neighboring β-sheet reduces its stability compared to the N-state. The amide protons in this helical segment exchange more rapidly with the solvent than the amide protons in the β-sheet section. (4) In the temperature range 300–335 K, the A-state approaches random coil, but some structure is still present at 335 K. (5) The formation of the A-state from the U-state could be relatively fast since, given the presence of favorable β-turns, the formation of both β-sheets and α-helices should be kinetically favorable. As such, the A-state could plausibly lie on the folding pathway, with the A to N transition a late step in folding. This transition would define aspects of a tertiary structure that are heavily reliant upon hydrophobic and van der Waals interactions and that on average are more remote from each other than the residues involved in β-sheet and α-helix formation.

ACKNOWLEDGMENTS

We thank Phil A. Evans for helpful discussions and the MRC Mill Hill NMR facility for access to the 500-MHz spectrometer. We are indebted to Tom Frenkiel and Chris Bauer for help with the HOHAHA and solvent suppression experiments described in the text.

REFERENCES

Baldwin, R. L. (1989) *Trends Biochem. Sci. 14*, 291–294.
Baum, J., Dobson, C. M., Evans, P. A., & Hanley, C. (1989) *Biochemistry 28*, 7.
Biringer, R. G., & Fink, A. L. (1982) *Biochemistry 21*, 4748.
Brown, S. C., Weber, P. L., & Mueller, L. (1988) *J. Magn. Reson. 17*, 166.
Bundi, A., & Wüthrich, K. (1979) *Biopolymers 18*, 285.
Cary, P. D., King, D. S., Crane-Robinson, C., Bradbury, E. M., Rabbani, A., Goodwin, G. H., & Johns, E. W. (1980) *Eur. J. Biochem. 112*, 577.
Chin, D. T., Kuehl, L., & Rechsteiner, M. (1982) *Proc. Natl. Acad. Sci. U.S.A. 77*, 1365.
Creighton, T. E. (1978) *Prog. Biophys. Mol. Biol. 33*, 231.
Di Stefano, D. L., & Wand, A. J. (1987) *Biochemistry 26*, 7272.
Dolgikh, D. A., Gilmanshin, R. I., Biazhnikov, E. V., Semisotnov, G. V., Venyaminov, S. Y., & Ptitsyn, O. B. (1981) *FEBS Lett. 136*, 311.
Dolgikh, D. A., Abaturov, L. V., Bolotina, I. A., Brazhnikov, E. V., Bychkova, V. E., Bushuev, V. N., Gilmanshin, R. I., Lebedev, Y. O., Seminsotnov, G. V., Tiktopulo, E. I., & Ptitsyn, O. B. (1985) *Eur. Biophys. J. 13*, 109.
Fersht, A. (1987) *Biochemistry 26*, 8031.
Goldstein, G., Scheid, M., Hammerling, U., Boyse, E. A., Schlesinger, D. H., & Niall, H. D. (1975) *Proc. Natl. Acad. Sci. U.S.A. 72*, 11.
Goto, Y., & Fink, A. L. (1989) *Biochemistry 28*, 945.

Hershko, A., Heller, H., Eytan, E., Kaklij, G., & Rose, I. A. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 7021.
Kim, P. S., & Baldwin, R. L. (1982) *Annu. Rev. Biochem. 51*, 549.
Lenkinski, R. E., Chen, D. M., Glickson, J. D., & Goldstein, G. (1977) *Biochim. Biophys. Acta 494*, 126-130.
Marion, D., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 113*, 967.
Ohgushi, M., & Wada, A. (1983) *FEBS Lett. 164*, 21.
Roder, H., Elove, G. A., & Englander, S. W. (1988) *Nature 335*, 700.
Tanford, C. (1968) *Adv. Protein Chem. 23*, 121.
Timasheff, S. N. (1970) *Acc. Chem. Res. 3*, 62.
Udgaonker, J. B., & Baldwin, R. L. (1988) *Nature 335*, 694.
Vijay-Kumar, S., Bugg, C. E., Wilkinson, K. D., & Cook, W. J. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 3582.
Vijay-Kumar, S., Bugg, C. E., Wilkinson, K. D., & Cook, W. J. (1987) *J. Mol. Biol. 194*, 531.
Weber, P. L., Brown, S. C., & Mueller, L. (1987) *Biochemistry 26*, 7282.
Wilkinson, K. W., & Mayer, A. N. (1986) *Arch. Biochem. Biophys. 250*, 390.
Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York.
Wüthrich, K. (1989a) *Acc. Chem. Res. 22*, 36.
Wüthrich, K. (1989b) *Science 243*, 45.

Chemical Synthesis, Purification, and Characterization of Two Inflammatory Proteins, Neutrophil Activating Peptide 1 (Interleukin-8) and Neutrophil Activating Peptide 2†

Ian Clark-Lewis,*,‡ Bernhard Moser,§ Alfred Walz,§ Marco Baggiolini,§ George J. Scott,∥ and Ruedi Aebersold‡

The Biomedical Research Centre and Department of Biochemistry, University of British Columbia, Vancouver, British Columbia, Canada V6T 1W5, The Theodor-Kocher Institute, University of Bern, P.O. Box 99, Bern, Switzerland, and The SCIEX Corporation, Thornhill, Ontario, Canada L3T 1P2

Received June 14, 1990; Revised Manuscript Received November 26, 1990

ABSTRACT: Two recently identified pro-inflammatory proteins, namely, neutrophil activating peptide 1 (NAP-1) [also termed interleukin-8 (IL-8)] and NAP-2, were chemically synthesized, purified, and characterized. The fully protected NAP-1/IL-8 (72 residues) and NAP-2 (70 residues) peptide chains were assembled by automated solid-phase methods with average stepwise yields of 99.5 and 99.3%, resulting in overall chain assembly yields of 70 and 62%, respectively. Deprotection resulted in crude products, which were allowed to fold by air oxidation, and were purified by two cycles of reverse-phase high-pressure liquid chromatography, yielding 27 mg of NAP-1/IL-8 and 22 mg of NAP-2. Purity was established by reverse-phase high-pressure liquid chromatography and isoelectric focusing, and the primary structures of the purified products were verified by using mass spectrometry and Edman sequencing methods. Synthetic and recombinant NAP-1/IL-8 were equally active on human neutrophil granulocytes as determined by measuring the induction of cytosolic free calcium, elastase release, and chemotaxis. Synthetic NAP-2 was equivalent to purified natural NAP-2 in the elastase release and calcium mobilization assays, but it was consistently less potent (3–5-fold) as a stimulus of chemotaxis, perhaps indicative of additional chemotactic components in the natural preparation. The results indicate that by chemical synthesis these cytokines can be obtained in purity and quantities suitable for further structural analysis, as well as functional studies both in vivo and in vitro. The ability to rapidly generate analogues with unambiguous primary structure suggests that this will be the method of choice for an in-depth study of structure–function relationships within this family of inflammatory cytokines.

The recruitment and activation of neutrophil leukocytes are central events in the acute inflammatory response. Recently, several members of a family of endogenous human proteins that are potent mediators of these events have been identified (Baggiolini et al., 1989; Matsushima & Oppenheim, 1989). The first protein with these activities to be identified was variously termed neutrophil activating factor (Walz et al., 1987), monocyte-derived neutrophil activating peptide (Schroder et al., 1987), and monocyte-derived neutrophil chemotactic factor (Yoshimura et al., 1987a,b). More recently, the terms neutrophil activating peptide 1 (NAP-1)[1] and interleukin-8 (IL-8) were proposed (Westwick et al., 1989; Leonard, 1990). Other proteins with sequence similarity that suggests structural similarity, including neutrophil activating peptide 2 (NAP-2) (Walz & Baggiolini, 1989) and melanoma

† This work was supported by the Protein Engineering Network Centre of Excellence (PENCE) and the Medical Research Council (MRC) of Canada. I.C.-L. and R.A. are recipients of MRC scholarships. Additional support was from the Swiss National Science Foundation (Grant 31-25200-88).

* To whom correspondence should be addressed.
‡ University of British Columbia.
§ University of Bern.
∥ The SCIEX Corp.

[1] Abbreviations: NAP, neutrophil activating peptide; IL-8, interleukin-8; PBP, platelet basic protein; HPLC, high-pressure liquid chromatography; *t*-Boc, *tert*-butyloxycarbonyl.

0006-2960/91/0430-3128$02.50/0