Identification of Initiation Sites for T4 Lysozyme Folding Using CD and NMR Spectroscopy of Peptide Fragments †

ARTICLE *in* BIOCHEMISTRY · JUNE 2000

Impact Factor: 3.02 · DOI: 10.1021/bi000070i · Source: PubMed

4 AUTHORS, INCLUDING:

John Wade
The Florey Institute of Neuroscience and Men…
292 PUBLICATIONS **6,702** CITATIONS
SEE PROFILE

Michael J Mcleish
Indiana University-Purdue University Indiana…
105 PUBLICATIONS **1,803** CITATIONS
SEE PROFILE

Identification of Initiation Sites for T4 Lysozyme Folding Using CD and NMR Spectroscopy of Peptide Fragments†

Lidia V. Najbar,‡ David J. Craik,*,§ John D. Wade,|| and Michael J. McLeish*,‡,⊥

Victorian College of Pharmacy, Monash University, 381 Royal Parade, Parkville, Victoria 3052, Australia, Centre for Drug Design and Development, University of Queensland, Brisbane 4072, Australia, and Howard Florey Institute, University of Melbourne, Parkville 3052, Australia

Received January 12, 2000; Revised Manuscript Received March 20, 2000

ABSTRACT: Using CD and 2D ^{1}H NMR spectroscopy, we have identified potential initiation sites for the folding of T4 lysozyme by examining the conformational preferences of peptide fragments corresponding to regions of secondary structure. CD spectropolarimetry showed most peptides were unstructured in water, but adopted partial helical conformations in TFE and SDS solution. This was also consistent with the ^{1}H NMR data which showed that the peptides were predominantly disordered in water, although in some cases, nascent or small populations of partially folded conformations could be detected. NOE patterns, coupling constants, and deviations from random coil Hα chemical shift values complemented the CD data and confirmed that many of the peptides were helical in TFE and SDS micelles. In particular, the peptide corresponding to helix E in the native enzyme formed a well-defined helix in both TFE and SDS, indicating that helix E potentially forms an initiation site for T4 lysozyme folding. The data for the other peptides indicated that helices D, F, G, and H are dependent on tertiary interactions for their folding and/or stability. Overall, the results from this study, and those of our earlier studies, are in agreement with modeling and HD−deuterium exchange experiments, and support an hierarchical model of folding for T4 lysozyme.

One of the major unsolved problems in molecular biology is the mechanism by which proteins fold into their native structure. It is now well established, by stopped-flow CD[1] experiments and pulsed hydrogen-exchange methods combined with two-dimensional NMR spectroscopy, that many proteins fold via at least partially structured intermediates, referred to as "molten globules" (*1*−*3*). These partially folded states are characterized by substantial contents of native secondary structure, but no ordered tertiary structure. However, such secondary structure is formed rapidly, and the above techniques are not capable of providing information on the processes occurring prior to their formation, i.e., the processes whereby protein folding is initiated.

One approach used to examine these earlier events has been to study peptide fragments of proteins. Short model peptides (*4*), and peptide fragments corresponding to regions of secondary structure of many proteins, have, under certain conditions, been shown to adopt preferences for folded conformations in solution (*5*). Detection of nativelike structure in short peptides occurring on a submillisecond time scale, and in the absence of tertiary interactions, suggests that these regions of secondary structure form early in the folding process (*5*, *6*). Indeed, CD and two-dimensional NMR studies on helical and turn regions derived from myohemerythrin (*7*), plastocyanin (*8*), myoglobin (*9*), cytochrome *c* (*10*), thermolysin (*11*), and phage 434 Cro protein (*12*) have been used to detect both stable and nascent secondary structure in short peptide fragments in which tertiary interactions were lacking. Further evidence comes from the more recent studies on the helical mutants from barnase (*13*, *14*), helical peptides of cytochrome *c* and bovine calcium binding protein (*15*), and helical and turn segments of hemoglobin of *C. thummi* (*16*). In addition to identifying the stretches of proteins that may be responsible for the initiation of folding, the latter studies have provided information on the role of local shorter range interactions that may be involved in the folding process.

In the majority of the investigations, the helical and turn conformations adopted by the peptides in aqueous solution were transient in nature and, presumably, only marginally stable. These transient structures often show preference toward some folded conformers which are in equilibrium

† The peptide synthesis studies at the Howard Florey Institute were supported by an Institute Block Grant from the National Health and Medical Research Council of Australia. We thank Monash University Research Training and Support Branch for the financial assistance of a Postgraduate Publications Award for L.V.N. during the preparation of this paper. D.J.C. is an Australian Research Council Senior Research Fellow.

* Authors to whom correspondence should be addressed. Author contact information (M.J.M.): Tel (734) 615 1787, Fax (734) 615 3079, E-mail mcleish@umich.edu.

‡ Victorian College of Pharmacy.

§ Centre for Drug Design and Development.

|| Howard Florey Institute.

⊥ Current address: College of Pharmacy, University of Michigan, 428 Church St., Ann Arbor, MI 48109.

[1] Abbreviations: T4L, bacteriophage T4 lysozyme; CD, circular dichroism; NMR, nuclear magnetic resonance; TFE, 2,2,2-trifluoroethanol; SDS, sodium dodecyl sulfate; TSP, 3-(trimethylsilyl)propionic acid-*2,2,3,3-d*$_4$ sodium salt; NOE, nuclear Overhauser effect; NOESY, two-dimensional nuclear Overhauser spectroscopy; TOCSY, total correlation spectroscopy; COSY, two-dimensional correlated spectroscopy; DQF-COSY, double-quantum-filtered COSY; DIPSI, decoupling in the presence of scalar interactions; FID, free induction decay; 2D, two-dimensional; HD, hydrogen−deuterium.

10.1021/bi000070i CCC: $19.00
Published on Web 04/21/2000

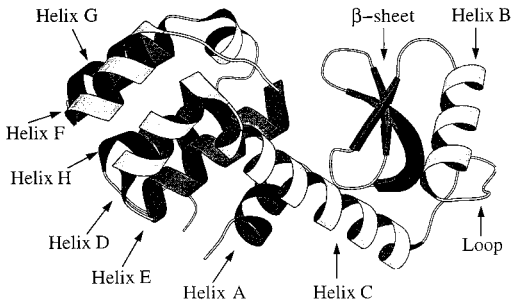


FIGURE 1: The X-ray crystal structure of T4 lysozyme (*64, 65*) highlighting all the regions of secondary structure. The diagram was created using Molscript (*66*).

with the unfolded state, and have been termed 'nascent' helices and turns (*17*). Their solution structure in water represents ensembles flickering in and out of any well-defined conformation. This is mainly due to the fact that these peptides fold on a short time scale and require longer range tertiary interactions for their stability. However, if these nascent structures have an inherent propensity for native structure formation, in the presence of cosolvents such as 2,2,2-trifluoroethanol (TFE) they can be stabilized into well-ordered conformations (*18*). In many cases (including several of the aforementioned studies), helical, turn, and sheet regions have been shown to be stabilized by TFE into their native conformation (*7, 19*−*21*). However, in some cases, particularly in loop and/or β-sheet regions of native protein structure, non-native helical conformations have been stabilized by TFE (*22*−*25*). In some cases, this was interpreted as being evidence for nonhierarchical folding (*24, 25*), although close examination indicates that the non-native structure brought about by TFE is highly flexible, without stable helical H-bonds (*23, 26, 27*). By contrast, the H-bonds in TFE-stabilized β-hairpins and the nativelike β-structure are stable and nativelike (*26*−*28*). Micellar SDS has also been used with success to stabilize peptides with helical propensity (*29*−*31*).

T4 lysozyme is a small protein (164 residues; Figure 1) which folds in a highly cooperative manner. The enzyme consists of two domains, and while there is little evidence for the independent unfolding and refolding of each domain (*32*), there is good evidence that the C-domain refolds first and then stabilizes the folding of the N-terminal domain (*32*−*34*). We have been interested in identifying those regions of secondary structure in T4 lysozyme (Figure 1) which may act as initiation sites in the folding of this protein. Toward that end, we designed a series of peptides encompassing all elements of T4L secondary structure (Table 1). Using CD and NMR spectroscopy, we have studied the solution conformations of peptides corresponding to helix A (*35*), helix B (*36*), helix C (*37*), and the β-sheet region (*38*). These peptides were primarily unstructured in water but, to varying degrees, adopted relatively stable conformations in the presence of TFE and SDS micelles. On the basis of these studies, we have proposed that helices A and C and, possibly, the β-sheet may provide initiation sites for the refolding of the rest of the protein (*35, 37, 38*).

In the present study, we report our findings on the five remaining helical fragments of T4 lysozyme. Peptides corresponding to T4 lysozyme helices D, E, F, G, and H, as well as the loop region, were synthesized. CD and NMR spectroscopy have been used to determine the structure of the peptides in both TFE and SDS micelles. This study completes our dissection of T4 lysozyme and allows us to comment on the possible folding pathway, as well as on the utility of TFE and SDS micelles as solvents for studies of this type.

MATERIALS AND METHODS

Peptide Synthesis. The peptides corresponding to α-helical fragments derived from T4 lysozyme (Table 1) were assembled by continuous flow Fmoc solid-phase synthesis using Pepsin KA resin as solid support, and *O*-pentafluorophenyl esters of N^{α}-Fmoc-amino acids (*39*). Side chain protection was afforded by *tert*-butyl ester and ether (Asp, Glu, Ser) and by BOC (Lys). N^{α}-Fmoc deprotection during the assembly was with 20% piperidine in DMF. Following completion of the synthesis, the dried peptide−resin was cleaved and deprotected by 2 h reaction with 95% aqueous TFA. The peptides were purified by conventional preparative RP-HPLC on Vydac C18 using 0.1% TFA-based buffers. The purified products eluted as single peaks on both analytical HPLC and capillary electrophoresis. Peptide concentrations were determined by amino acid analysis. LYS(1−13), which encompasses helix A, was available from a previous study (*35*).

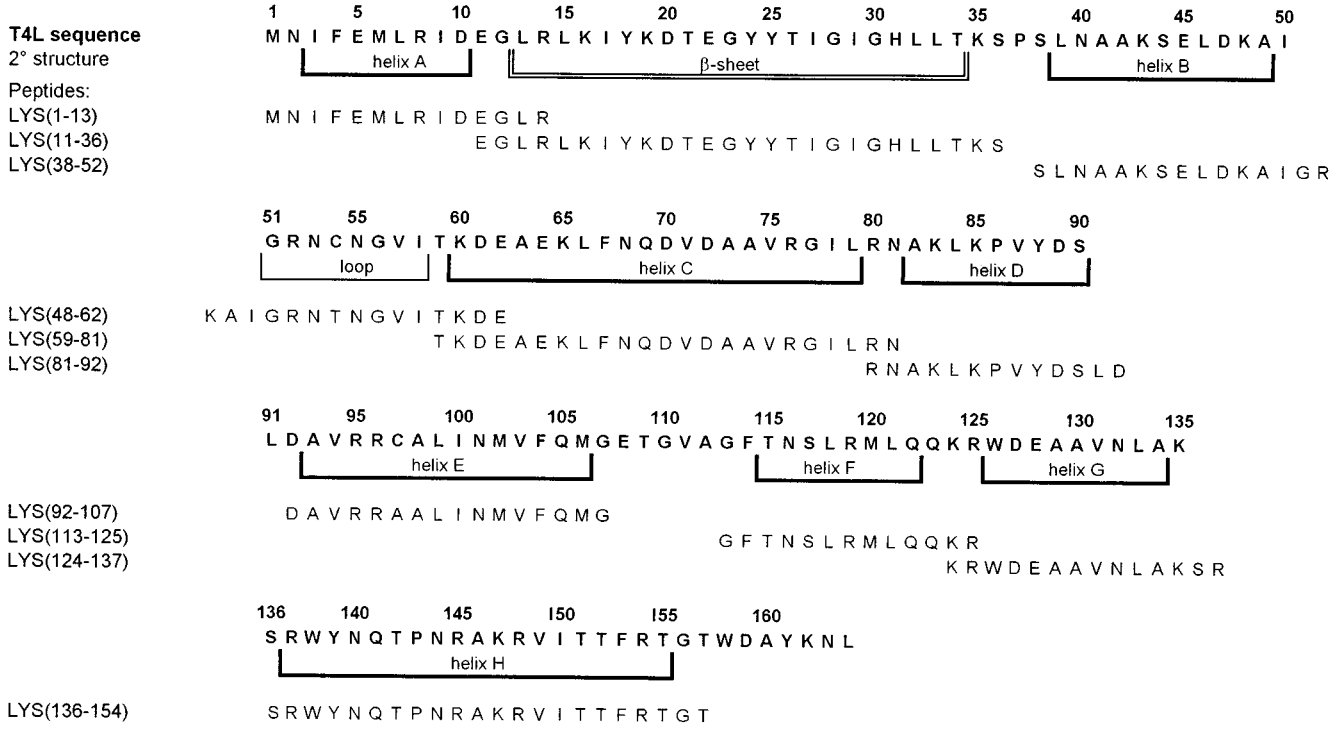
Note that, while the peptides are named to indicate the residue numbers in the T4L sequence, throughout the text the individual peptide residues are numbered according to the residue position in the amino acid sequence of the synthesized peptide, rather than the actual T4L residue numbers.

CD Spectropolarimetry. All peptides were examined in the concentration range of 25−500 μM. The samples were prepared in 10 mM potassium phosphate buffer at pH 4.8. The exception was LYS(1−13), which, although soluble in TFE and 200 mM SDS, was difficult to solubilize in phosphate buffer. Therefore, this peptide was dissolved in TFE and diluted with buffer to the appropriate TFE concentration. The CD spectra were recorded on an AVIV 60DS spectropolarimeter at 288 K which had been calibrated using *d*-10-camphorsulfonic acid. Cells having a path length of 0.1 or 0.02 cm were employed, and were maintained at the required temperature using a Lauda circulating water bath. Spectra were an average of five scans recorded with a bandwidth of 1 nm, a 0.25 nm step size, and a 0.2 s time constant. Following baseline correction, the observed ellipticity was converted to a mean residue ellipticity, $[\theta]$ (deg·cm^2·mol^{-1}), using the relationship $[\theta] = \theta \times l/(lcN)$, where θ is the observed ellipticity, l is the path length in millimeters, c is the molar concentration, and N is the number of residues in the peptide. The spectra were then smoothed using a third-order polynomial function.

^{1}H NMR Spectroscopy. Peptide solutions were prepared to a concentration of 1−2 mM in a volume of 0.65 mL (pH 3.5−5.0). The solvents employed included 90% H_2O/10% D_2O, 15% TFE/80% H_2O/5% D_2O, 50% TFE/45% H_2O/5% D_2O, 50% TFE/50% D_2O, and 200 mM deuterated SDS micelles in 90% H_2O/10% D_2O. For the slow exchange experiments, samples were prepared in 100% D_2O, 50% TFE /50% D_2O, or 200 mM SDS-d_{25}/100% D_2O.

^{1}H NMR spectra were recorded on Bruker AMX 300, 500, and 600 MHz spectrometers. All two-dimensional (2D)

Table 1: Amino Acid Sequence of T4 Lysozyme Showing Secondary Structure and Sequences of the Peptides



spectra were acquired at temperatures ranging from 278 to 303 K, with time proportional phase incrementation (TPPI) for quadrature detection in the F1 dimension (*40*). Typical 2D experiments included DQF-COSY (*41*), TOCSY (*42*, *43*) using a DIPSI spin-locking sequence (*44*), and NOESY (*45*). The mixing times ranged from 65 to 120 ms in TOCSY and from 100 to 250 ms in NOESY experiments.

For DQF-COSY experiments, the H_2O resonance was suppressed by gated irradiation during the relaxation delay (1.5−2 s), with the transmitter offset placed to coincide with the H_2O resonance. Solvent suppression for NOESY and TOCSY experiments was achieved using a 1-1 binomial pulse sequence in place of the final 90° pulse (*46*), in combination with mild presaturation. Spectra were generally acquired with 4K complex data points in the F2 dimension and 256−700 increments in the F1 dimension, and were of 32 scans each (between 64 and 72 scans for NOESY).

All data were processed using UXNMR (Bruker) and FELIX (Hare Research Inc.) software. For 2D experiments, the t_1-dimension was zero-filled to 2048 real data points, and $\pi/2$ phase-shifted, sine-squared window functions were applied. In the cases of severe spectral overlap, $\pi/4$ or $\pi/8$ phase-shifted sine-bell-squared window functions were applied. Polynomial baseline correction was applied in selected regions of the spectra. Chemical shifts were referenced to internal 2,2-dimethyl-2-silapentane-5-sulfonate (DSS) or 3-(trimethylsilyl)propionic acid-*2,2,3,3-d*$_4$ sodium salt (TSP) at 0.0 ppm.

RESULTS

Peptide Synthesis. The peptides LYS(81−92), LYS(92−107), LYS(113−125), LYS(124−137), and LYS(136−154) were designed to encompass helices D, E, F, G, and H, while LYS(48−62) included the loop region (Table 1). Residues flanking each helix were included to minimize end-fraying effects and to enhance water-solubility. The peptides were prepared with both termini unprotected, and, in general, the synthesis and purification was straightforward. The exception was the helix E peptide, LYS(92−107), which was relatively insoluble and difficult to purify. The initial peptide contained a free cysteine residue, but when this was replaced by an alanine, the peptide was more amenable to purification. To avoid similar problems with the loop peptide, LYS(48−62), the cysteine at position 7 was replaced by a threonine. These substitutions were considered to be reasonable as similar changes were made in the "pseudo-wild type" (C54T C97A) T4 lysozyme which has been adopted as the standard for use in T4L folding studies (*47*, *48*).

CD Spectropolarimetry. The CD spectra of the peptides, in water, 50% TFE, and 200 mM SDS, are shown in Figure 2. With the exception of LYS(92−107), all of the peptides exhibited minima below 200 nm, indicative of a random conformation (*49*). LYS(92−107), which represents helix E, showed a broad minimum around 215 nm, indicative of β-structure, possibly due to the formation of small soluble aggregates, along with some contribution from helical conformers (*49*). Upon addition of 50% TFE, the majority of the peptides showed a shift in the minimum toward 208 nm, with some ellipticity also developing at 222 nm. This is consistent with a shift toward an α-helical conformation (*49*) and is in accord with previous studies in which T4 lysozyme-based peptides were readily able to adopt helical conformations in the presence of TFE (*35*−*38*). LYS(92−107) had the highest helical content of all the peptides, whereas the loop peptide, LYS(48−62), and the helix D peptide, LYS-(81−92), were exceptions in that both maintained their minima at around 200 nm and showed little evidence of helical character.

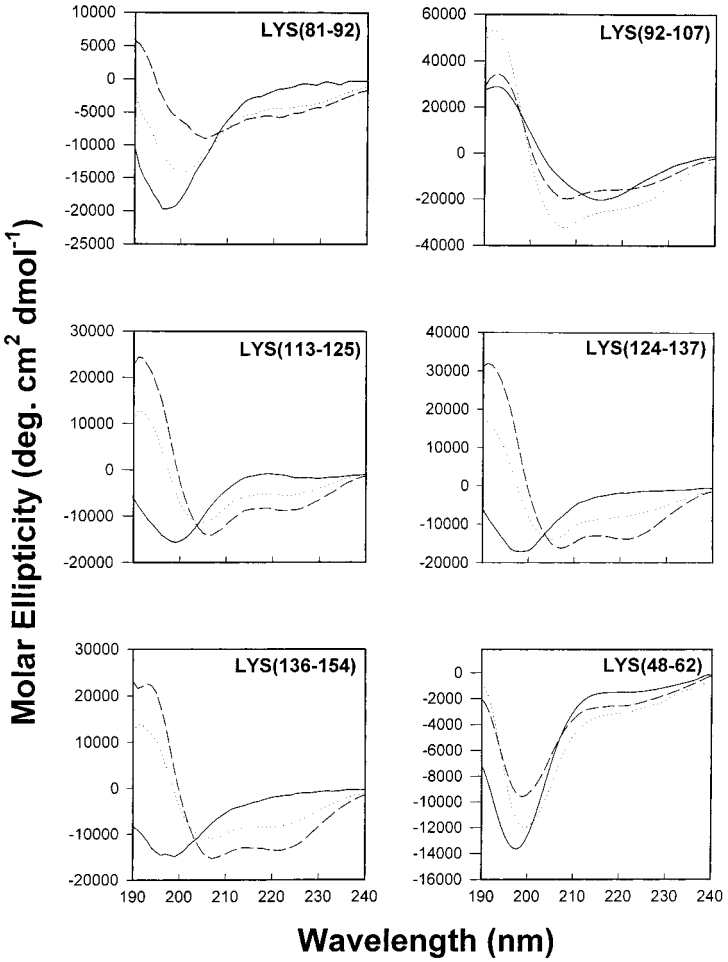


FIGURE 2: UV−CD spectra of each of the peptides in 10 mM phosphate buffer, pH 4.8 (solid lines), in 50% TFE (dotted lines), and in 200 mM SDS (dashed lines). Peptide concentrations were between 25 and 150 μM. Spectra were obtained at 288 K.

In SDS micelles, which have been used to provide a hydrophobic environment and as membrane mimetics (*31*), a major shift toward a helical conformation was also observed for all peptides except LYS(48−62) and LYS(81−92). Interestingly, for the helix F, G, and H peptides, the helicity in SDS micelles was greater than that observed in 50% TFE. This is in contrast to earlier studies of the helix A and C peptides, in which the greatest helicity was observed in 50% TFE (*35, 37, 38*), but similar to results for helix B (*36*).

Assuming that the absorption at 222 nm is almost exclusively due to α-helix (*49*), it is possible to calculate the α-helicity using the equation of Chen et al. (*50*):

$$[\theta]_\lambda = (f_H - ik/N)[\theta]_{H\lambda\infty}$$

where $[\theta]_\lambda$ is the observed mean residue ellipticity at wavelength λ, f_H is the fraction of helix, i is the number of helical segments (1 in this case), k is a wavelength-dependent constant, N is the number of residues, and $[\theta]_{H\lambda\infty}$ is the maximum mean residue ellipticity for a helix of infinite length. Table 2 shows the helicity of each peptide, expressed as a percentage, in water, 50% TFE, and 200 mM SDS.

Broadly speaking, the peptides can be divided into three groups. The first, which includes the helix D and loop peptides, shows little or no helical content under any conditions. This result for the loop peptide is comforting as both TFE and SDS are generally reported to stabilize rather than induce secondary structure (*7, 19, 31*), and the loop peptide would not be expected to have any propensity for helical conformation. The second group, which includes the helix C, helix E, and sheet peptides, shows the greatest helicity in 50% TFE although the helical conformation is also stabilized in SDS. The third group contains the helix B, F, G, and H peptides and exhibits considerably greater helical content in SDS micelles. The helix A peptide is a little difficult to characterize, showing about equal helicity in both TFE and SDS, but, given that its helicity in TFE is much greater than other peptides in the last group, it probably belongs in the second group.

Table 2: CD Determination of α-Helical Content[a,b]

peptide	region	KPO_4, pH 4.8	50% TFE	200 mM SDS
LYS(1−13)	helix A	13[c]	37	42
LYS(11−36)	β-sheet	0[d]	33[d]	18[d]
LYS(38−52)	helix B	7[e]	30[e]	50[e]
LYS(48−62)	loop	4	9	7
LYS(59−81)	helix C	11[f]	55[f]	32[f]
LYS(81−92)	helix D	5	13	18
LYS(92−107)	helix E	45	70	46
LYS(113−125)	helix F	3	17	27
LYS(124−137)	helix G	5	24	43
LYS(136−154)	helix H	5	24	40

[a] CD data were collected at 15 °C, with peptide concentrations between 25 and 250 μM. [b] The percent helicity was calculated using the method of Chen et al. (*50*). [c] Determined using 5% TFE to enhance peptide solubility. [d] From Najbar et al. (*38*). [e] From Najbar et al. (*36*). [f] From McLeish et al. (*37*).

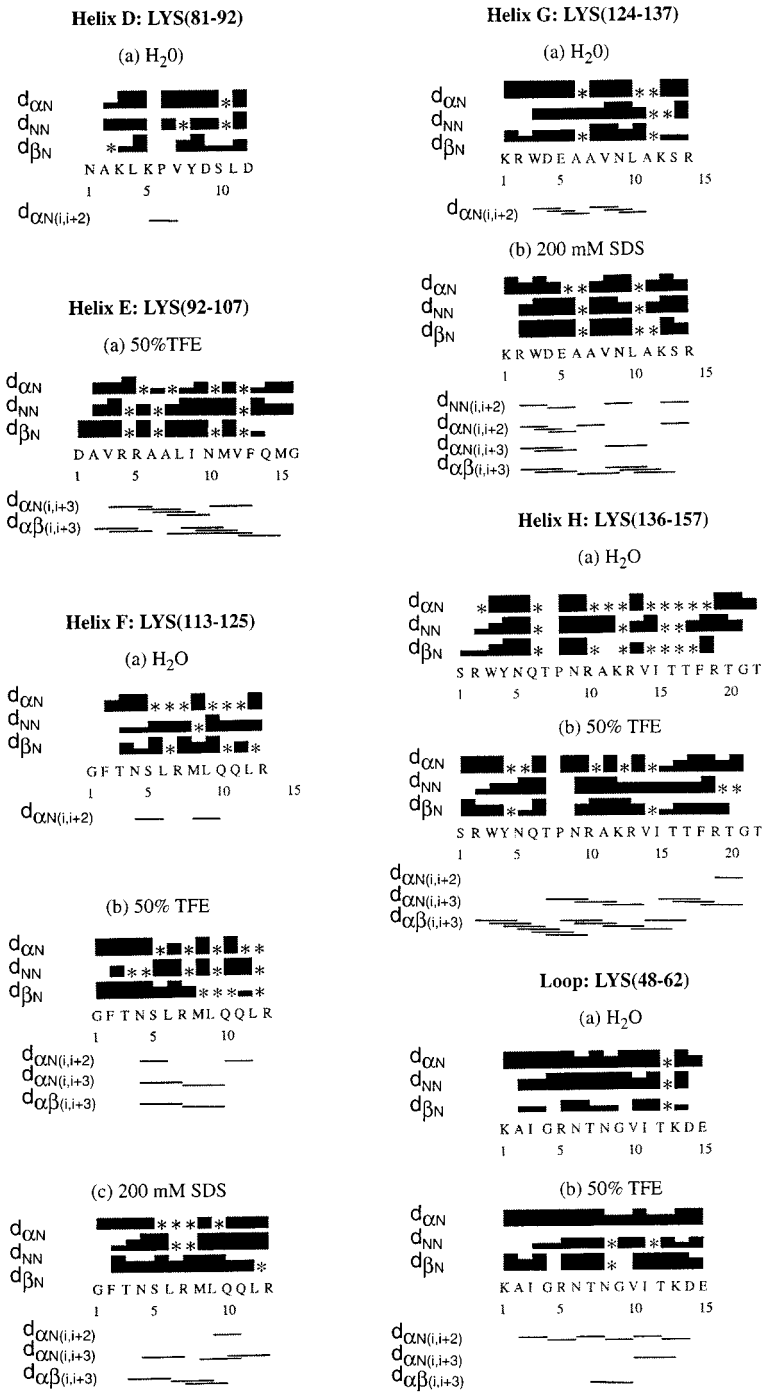


FIGURE 3: Summary of sequential and medium-range NOEs observed for all the peptides in different solvents. Approximate NOE intensities are indicated by the height of the bars. Overlapping and ambiguous NOEs are indicated by an asterisk. Note that the individual peptide residues are numbered according to the residue position in the amino acid sequence of the synthesized peptide, rather than the actual T4L residue numbers.

NMR Spectroscopy. The peptides were examined in three solvents: H_2O, 50% TFE, and SDS micelles. Peptide concentrations of 1−2 mM were used in the NMR studies, and although most peptides were soluble in all three solvents, some micelle and TFE samples required a brief period of sonication to ensure total dissolution. NMR assignments were made using a combination of TOCSY, DQF-COSY, and NOESY methods. Broader line widths were often observed in the SDS micelle spectra (and sometimes in TFE spectra) due to the slower correlation time of the peptides in this solvent. A summary of the observed NOEs for the peptides in different solvent systems is provided in Figure 3.

Chemical shifts provide a valuable indication of the secondary structure adopted by peptides. Upfield chemical shift deviations from their "random coil" values are observed for αH protons in α-helical conformations, and downfield αH shifts are observed for residues existing in a *β*-sheet or turn conformation (*51*, *52*). Figure 4 summarizes the secondary shifts (i.e., deviations from random coil values) for all of the peptides.

Helix D: LYS(81−92). In the native protein, this is the shortest helix; it follows and is adjacent to helix C (residues 60−80) and is the first of a series of helices in the "all α" domain. NOESY spectra for LYS(81−92) were initially

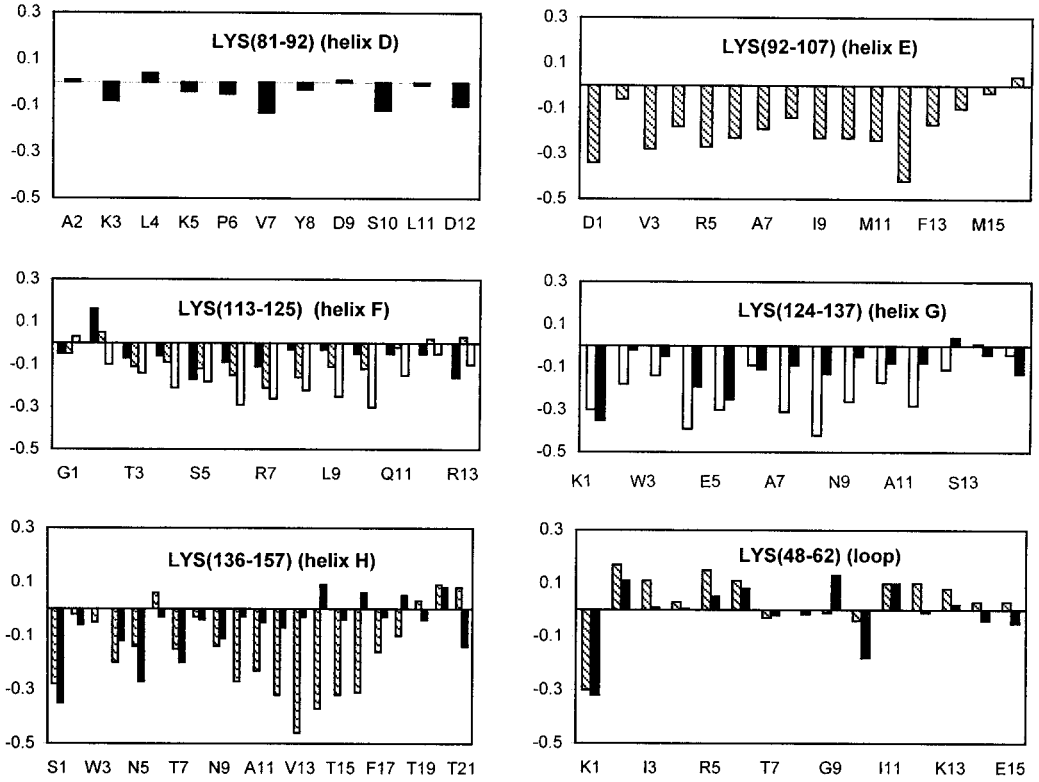


FIGURE 4: Secondary shifts (ppm) for all of the T4 lysozyme peptides in various solvents. Secondary shifts were calculated by subtracting random coil shifts (*67*) from the observed αH proton shifts. The negative secondary shifts of magnitude greater than 0.1 ppm are generally indicative of helical structures. Filled bars correspond to spectra recorded in water, hatched bars to 50% TFE, and open bars to 200 mM SDS. Individual peptide residues are numbered according to the residue position in the amino acid sequence of the synthesized peptide.

obtained in H_2O at 283 K; however, the spectra were of poor quality. The addition of 15% TFE greatly improved the chemical shift dispersion. Complete NMR assignments were made in 15% TFE, and the NOEs observed are presented in Figure 3. The presence of strong $d_{\alpha N}(i,i+1)$, weaker $d_{\alpha N}(i,i)$, and medium intensity $d_{NN}(i,i+1)$ NOEs, as well as the absence of any medium-range NOEs characteristic of α-helical populations, suggests that this peptide is predominantly unstructured in aqueous solution (in agreement with the CD data). This is supported by $^3J_{NH-H\alpha}$ coupling constants which were in the range of 6−8 Hz. Deviations from random coil chemical shifts also support this (Figure 4). Because of the relative lack of structure, no NMR assignments were made for this peptide in 50% TFE and SDS micelles.

Helix E: LYS(92−107). In T4L, helix E shows the highest protection from deuterium exchange (*53*). It is a relatively hydrophobic helix with most side chains buried within the protein interior. In the crystal, this helix makes contacts with a number of residues in helix H (which spans residues 137−155) and helix A (residues 3−8), and with some residues within the β-sheet (residues 13−34).

In H_2O and SDS micelles, the broad line widths and poor signal-to-noise ratio observed in the 1D spectra were suggestive of aggregation. However, assignments were made in 50% TFE, where the NOESY spectrum showed good chemical shift dispersion. The observation of strong $d_{NN}(i,i+1)$ NOEs in comparison to relatively weaker $d_{\alpha N}(i,i+1)$ NOEs, together with the presence of many medium-range $d_{\alpha N}(i,i+3)$ and $d_{\alpha\beta}(i,i+3)$ NOEs and upfield αH shifts from random coil values (as seen by the consistent series of negative secondary shifts in Figure 4), suggests a helical conformation for this peptide. This is supported by $^3J_{NH-H\alpha}$ coupling constants for a number of residues which were in the range of 4−5 Hz. The observation of helical structure for LYS(92−107) is in agreement with the CD data, and the contrast with the results obtained above for the helix D peptide is consistent with the notion that TFE stabilizes helices only in peptides with intrinsic propensity for helix formation (*20*).

Helix F: LYS(113−125). In the intact protein, the majority of the side chains of this helix are exposed to the solvent. For the peptide fragment in H_2O, strong $d_{\alpha N}(i,i+1)$ together with weaker $d_{\alpha N}(i,i)$ and $d_{NN}(i,i+1)$ NOEs were observed. The presence of one strong $d_{NN}(i,i+1)$ NOE and two weak $d_{\alpha N}(i,i+2)$ NOEs indicates that some populations of loosely folded turn conformers are present in the structural ensemble but, overall, the peptide is predominantly in a random coil in aqueous solution. This proposal is supported by the small secondary shifts in Figure 4 and by the CD results. By contrast with the data in H_2O, the presence of a series of stronger $d_{NN}(i,i+1)$ NOEs together with medium-range $d_{\alpha N}(i,i+3)$ and $d_{\alpha\beta}(i,i+3)$ NOEs in the NOESY spectrum of this peptide in 50% TFE suggests the presence of helical populations within the conformational ensemble. Helical populations in this solvent are also supported by $^3J_{NH-H\alpha}$ coupling constants in the range 5−6 Hz, a number of slowly exchanging NH protons, and chemical shifts generally upfield from random coil values, as well as the CD data.

Figure 5 shows regions of the NOESY spectrum of helix F in SDS micelles at 298 K, and highlights the presence of a large number of $d_{NN}(i,i+1)$ NOEs. The types of NOEs observed are similar to those in TFE, and again are strongly suggestive of helical conformations. The coupling constants of residues Phe2, Thr3, and Lys12 are in the range 5−6 Hz,

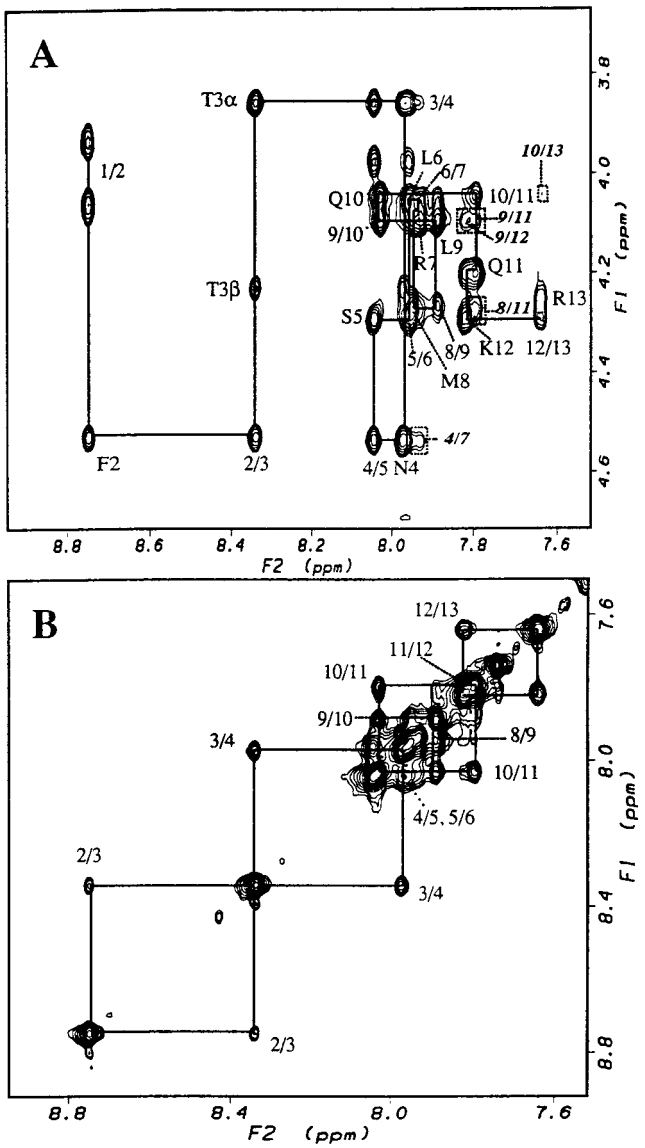


FIGURE 5: (A) CαH−NH and (B) NH−NH regions of a NOESY spectrum (mixing time = 250 ms) for LYS(113−125) in 200 mM SDS micelles at 298 K, 600 MHz. Medium-range NOEs are indicated in boxes.

consistent with helical conformations, the remainder being between 7 and 8 Hz. The latter coupling constants suggest some degree of conformational averaging within the population of helical forms. Chemical shift deviations from random coil values indicate that LYS(113−125) is more helical in SDS micelles than in TFE, an observation supported by the CD data in Table 2.

Helix G: LYS(124−137). In the folded protein, this amphipathic helix is one of the longer helices and is located in the "all α" domain. In H_2O, the presence of medium intensity $d_{NN}(i,i+1)$ NOEs and a number of weaker $d_{\alpha N}(i,i+2)$ NOEs, together with small upfield chemical shift deviations from random coil values, indicates that, at best, only nascent helical conformers may be present. Relatively intense $d_{\alpha N}(i,i+1)$ NOEs imply that random coil conformations predominate in this solvent.

Since CD spectropolarimetry showed a low helical content (about 15%) for this peptide in TFE, further NMR experiments were only performed in SDS micelles, where approximately 30% helix was detected by CD. Although many medium-range NOEs could not be observed in SDS micelles due to spectral overlap, the fact that most of the $d_{\alpha,N}(i,i)$ intensities were stronger than $d_{\alpha N}(i,i+1)$ NOEs, and that strong $d_{NN}(i,i+1)$ and $d_{\alpha\beta}(i,i+3)$ NOEs (refer to Figure 3) were observed throughout the whole peptide, supports the presence of a significant content of helical conformations for LYS(113−125) in SDS. The secondary shifts in Figure 4 also show a significant enhancement of helical conformations in SDS relative to water.

Helix H: LYS(136−157). This region is partially buried and partially solvent-exposed, and makes numerous contacts with helix E and also contacts some residues in the β-sheet segment. Spectra were acquired in both H_2O and 15% TFE, but as the chemical shift dispersion was better in the 15% TFE spectra, chemical shift assignments were determined in that solvent. Although the NOESY spectrum was quite overlapped, strong $d_{\alpha N}(i,i+1)$ NOEs suggested that the peptide is predominantly random coil. No medium-range NOEs characteristic of α-helical populations were observed. The generally small secondary shifts also supported the conclusion that the peptide is predominantly random coil.

The NOESY spectrum in 50% TFE showed better chemical shift dispersion than the H_2O or 15% TFE NOESY spectra. The presence of helical populations is well supported by strong $d_{NN}(i,i+1)$ NOEs and many $d_{\alpha N}(i,i+3)$ and $d_{\alpha\beta}$-$(i,i+3)$ NOEs throughout the peptide sequence. This conclusion is further supported by the marked enhancement of negative secondary shifts (Figure 4) upon addition of TFE.

In SDS micelles, broad line widths were observed in the 1D spectrum at pH 7.5 and 4.3. Although 2D spectra were recorded, they were of poor quality, making the assignment very difficult. However, a number of strong $d_{NN}(i,i+1)$ NOEs were observed (data not shown), and although these NOEs could not be specifically assigned, it could be concluded that folded and possibly helical populations are present for this peptide in SDS micelles.

Loop: LYS(48−62). The loop region of T4 lysozyme is mostly solvent-exposed, and it is located between helices B and C. This loop-derived peptide provides a control, as it should remain unstructured in the presence of structure-stabilizing solvents such as TFE or SDS micelles, a prediction borne out by the CD studies. The NOE data in Figure 3 and the chemical shift data in Figure 4 further support this conclusion.

DISCUSSION

The CD and NMR studies on the peptide fragments described here and elsewhere (*35*−*38*) have identified essentially three classes of peptide. The first comprises LYS-(48−62) and LYS(81−92), corresponding to the loop region and helix D of T4 lysozyme, respectively, and shows no evidence of secondary structure formation in any solvent. The lack of structure for the loop peptide is to be expected given the suggestion that only peptides with a propensity for helix/sheet formation will become structured in the presence of TFE or SDS micelles (*19*, *20*, *31*). The inability of LYS(81−92) to maintain any secondary structure in either solution implies that helix D does not provide an initiation site for the folding of T4 lysozyme, but rather it requires tertiary contacts for its formation and/or stability.

The second group comprises LYS(113−125), LYS(124−137), and LYS(136−157), corresponding to helices F, G,

and H, respectively. This group shows little or no structure in aqueous solution, but displays some evidence for helical conformations in TFE or SDS micelles. However, these helical conformations exist in equilibrium with a predominant ensemble of unfolded conformers. These peptides behave similarly to LYS(38−51), the helix B peptide (*36*), in that they show more helical character in SDS micelles than they do in 50% TFE (Table 2). Clearly, this group shows some propensity for helix formation, but it is unlikely to be sufficient for any of these peptides to form an initiation site for T4 lysozyme folding.

From this study, the sole representative of the third group is LYS(92−107), the peptide encompassing helix E, a region which the HD-exchange studies of Lu and Dahlquist (*53*) had identified as being perhaps the first to fold. As can be seen from the CD spectrum (Figure 2), the peptide shows no evidence of helix formation in aqueous solution but does have a tendency to aggregate. A similar spectrum is observed for LYS(1−13), the helix A peptide, which requires addition of a small amount of TFE to permit a CD spectrum to be obtained (McLeish, unpublished data). Both NMR and CD data show that these peptides readily form helices in 50% TFE, as do the data for LYS(59−81), the peptide encompassing helix C (*37*). Further, these peptides all exhibit a greater degree of helix formation in TFE than in SDS micelles. According to the framework model of protein folding (*54*−*56*), short peptides encompassing regions that are involved in the early stages of folding of the parent protein should form stable nativelike secondary structure without the requirement of tertiary interactions (*5*). Given the ability of this group of peptides to attain relatively stable helical structure, it would appear that these regions are most likely to provide initiation sites for the folding of T4 lysozyme. To some extent, the sheet peptide, LYS(11−36), also fits into this group as it is readily able to adopt a helical conformation (albeit non-native) in TFE and, to a lesser extent, in SDS. However, this peptide can also adopt a β-sheet conformation in nonmicellar SDS, as well as in low concentrations of TFE, and it appears that a small population of 'nativelike' turn persists, even in aqueous solution (*38*). The importance of solution environment on the transition between helical and sheet forms of this peptide has been described in considerable detail (*38*).

As a corollary to the framework model, it might be expected that those regions folding first may provide the bulk of the tertiary interactions required to stabilize the late-forming elements of secondary structure. Figure 6 shows the number of contacts between the side chains of those regions of secondary structure having intramolecular separations of $\leq$5 Å. Consistent with predictions that it provides a nucleation site, in the crystal helix E is buried within the protein core and makes 22 intramolecular contacts with helix H alone, as well as a number of contacts with helices A, F, and G and the β-sheet. As predicted, segments corresponding to helices A, E, and C and the β-sheet make a considerably greater number of close contacts than do the other segments. The exception is helix H, which makes as many as 50 contacts with other elements in the folded protein, including those with helix E. In one of the earliest studies on T4L folding, Desmadril and Yon (*32*) suggested that Trp138 (located in helix H) forms part of a nucleation center upon hydrophobic interaction with Met102 and Met106 in helix

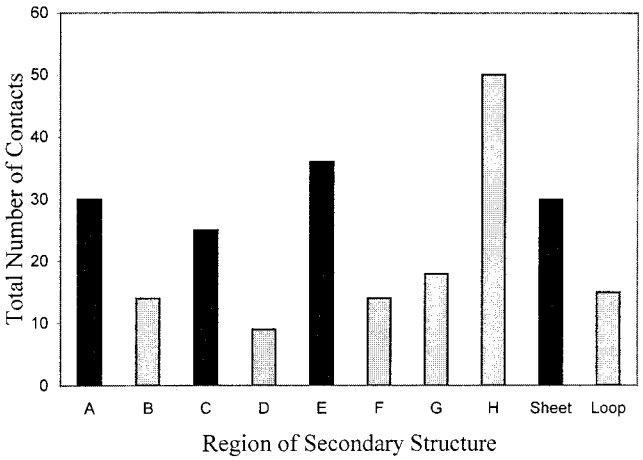


FIGURE 6: A graphical representation of the total number of contacts (<5 Å) in each region of secondary structure in T4 lysozyme (*65*). Regions thought to act as initiation sites are shaded in black.

E. Given so many contacts, it seems that the mutual interaction of helix E and helix H must be important in the early stages of folding. However, in the HD-exchange studies of Lu and Dahlquist (*53*), helix H showed only mild protection from exchange. This lack of protection from exchange is somewhat surprising as Bahar et al. (*57*) have suggested that slow-exchanging regions are usually located in regions with large numbers of tertiary contacts. Still, they further suggested that chain connectivity and the topology of tertiary contacts in the folded state will also affect the observed behavior (*57*). The lack of protection from exchange observed for helix H may well be brought about by these factors. However, given the apparent lability of the amide protons, and the relatively low level of structure attained by the peptide in TFE, it does not appear that helix H forms a nucleation site. Conversely, helix H does make a large number of tertiary contacts, and, given this apparent contradiction, it remains unclear what role the helix E/H interactions play in the folding process. What is clear is that the more solvent-exposed helices B, F, and G, as well as helix D and the loop, have relatively fewer tertiary interactions, consistent with the observations of Bahar et al. (*57*), and consistent with being formed in the final stages of folding.

Most previous studies employing peptide fragments to study protein folding have focused on the structure of peptides in aqueous solution and in TFE. To provide an alternative means of stabilizing secondary structure, in this study we have extended the range of solvents to include SDS micelles. TFE is thought to stabilize helices either by direct association (*58*, *59*), leading to the stabilization of intramolecular hydrogen bonds, or by an indirect mechanism (*26*, *60*) in which the peptide bonds are stabilized by weakening the interaction between water and peptide CO and NH groups. Ultimately, both mean that the formation of helical structure will be driven by intramolecular forces; i.e., the extent of helix formation in TFE will be sequence-driven. SDS, on the other hand, not only provides a hydrophobic environment that can stabilize ordered conformations (*61*), thereby providing stabilization of some sequence-driven folding, but also provides the possibility for the intermolecular interaction of peptide side chains with either polar headgroups or hydrophobic tails (*62*). In this way, it was

thought that SDS micelles in effect may provide a surface, akin to a partially folded protein, for peptides to fold on.

Given this line of thought, it is gratifying to see that the peptides which did show conformational preferences could be split into two groups. The first group, those peptides with the greatest propensity for helix formation, showed most structure in TFE. The second group showed a greater ability to maintain a helical conformation in SDS micelles, and it is not unreasonable to assume that this resulted in large part from the interactions of the peptide side chains with the micelle. This group of peptides represents those regions which require few tertiary interactions to maintain structure, but which do not have a great intrinsic drive for helix formation. For peptides of this type, it would seem that SDS micelles, rather than TFE, may be the solvent of choice for determining helical propensity.

The NMR studies of Lu and Dahlquist (*53*) provided the first indication that helix A, helix E, helix C, and the β-sheet region are the initial structural elements formed during the folding of T4 lysozyme. Subsequently, Xie and Freire (*63*) employed a thermodynamic search algorithm to identify structural determinants for the molten globule intermediates of several globular proteins. For T4 lysozyme, their calculations predicted that, in equilibrium folding intermediates, the β-sheet region of the N-terminal domain, helices A and E of the C-terminal domain, and helix C, linking the domains, would be folded while the rest of the protein would be unfolded. The results of our current investigation, and those of our previous studies on T4 lysozyme-derived peptides (*35*−*37*), support these predictions, although they do provide an additional suggestion that β-sheet formation may occur via an α-helical intermediate (*38*). Recently, Llinas and Marqusee (*33*) prepared fragments corresponding to the individual subdomains of T4 lysozyme. They found that the isolated C-terminal subdomain could form a stable structure but that the N-terminal subdomain was unstable and concluded that the C-terminal domain was crucial for conferring stability to the N-terminal subdomain. More recently, studies using methionine and alanine substitutions have shown that the formation of nativelike structure in the C-terminus is the rate-limiting step in T4lL folding (*34*). In toto, it would appear that the peptide fragment, H−D exchange NMR, structure calculation, and mutagenesis approaches are affording consistent results and provide considerable evidence that T4 lysozyme folds in an hierarchic manner.

SUPPORTING INFORMATION AVAILABLE

Chemical shift assignment tables for the T4 lysozyme peptides in different solvents (12 pages). This information is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Kuwajima, K. (1989) *Proteins: Struct., Funct., Genet. 6*, 87−103.
2. Baldwin, R. L. (1993) *Curr. Opin. Struct. Biol. 3*, 84−91.
3. Woodward, C. K. (1994) *Curr. Opin. Struct. Biol. 4*, 112−116.
4. Oas, T. G., and Kim, P. S. (1988) *Nature 336*, 42−48.
5. Dyson, H. J., and Wright, P. E. (1991) *Annu. Rev. Biophys. Biophys. Chem. 20*, 519−538.
6. Wright, P. E., Dyson, H. J., and Lerner, R. A. (1988) *Biochemistry 27*, 7167−7175.
7. Dyson, H. J., Merutka, G., Waltho, J. P., Lerner, R. A., and Wright, P. E. (1992) *J. Mol. Biol. 226*, 795−817.
8. Dyson, H. J., Sayre, J. R., Merutka, G., Shin, H.-C., Lerner, R. A., and Wright, P. E. (1992) *J. Mol. Biol. 226*, 819−835.
9. Reymond, M. T., Merutka, G., Dyson, H. J., and Wright, P. E. (1997) *Protein Sci. 6*, 706−716.
10. Wu, L. C., Laub, P. B., Elove, G. A., Carey, J., and Roder, H. (1993) *Biochemistry 32*, 10271−10276.
11. Jiménez, M. A., Bruix, M., Gonzalez, C., Blanco, F. J., Nieto, J. L., Herranz, J., and Rico, M. (1993) *Eur. J. Biochem. 211*, 569−581.
12. Padmanabhan, S., Jimenez, M. A., and Rico, M. (1999) *Protein Sci. 8*, 1675−1688.
13. Kippen, A. D., Arcus, V. L., and Fersht, A. R. (1994) *Biochemistry 33*, 10013−10021.
14. Kippen, A. D., Sancho, J., and Fersht, A. R. (1994) *Biochemistry 33*, 3778−3786.
15. Pintar, A., Chollet, A., Bradshaw, C., Chaffotte, A., Cadieux, C., Rooman, M. J., Hallenga, K., Knowles, J., Goldeberg, M., and Wodak, S. (1994) *Biochemistry 33*, 11158−11173.
16. Czisch, M., Liebers, V., Bernstein, R., Chen, Z., Baur, X., and Holak, T. A. (1994) *Biochemistry 33*, 9420−9427.
17. Dyson, H. J., Rance, M., Houghten, R. A., Wright, P. E., and Lerner, R. A. (1988) *J. Mol. Biol. 201*, 201−217.
18. Buck, M. (1998) *Q. Rev. Biophys. 31*, 297−355.
19. Segawa, S. I., Fukuno, T., Fujiwara, K., and Noda, Y. (1991) *Biopolymers 31*, 497−509.
20. Sönnichsen, F. D., Van Eyk, J. E., Hodges, R. S., and Sykes, B. D. (1992) *Biochemistry 31*, 8790−8798.
21. Cox, J. P. L., Evans, P. A., Packman, L. C., Williams, D. H., and Woolfson, D. N. (1993) *J. Mol. Biol. 234*, 483−492.
22. Fan, P., Bracken, C., and Baum, J. (1993) *Biochemistry 32*, 1573−1582.
23. Buck, M., Schwalbe, H., and Dobson, C. M. (1995) *Biochemistry 34*, 13219−13232.
24. Hamada, D., Kuroda, Y., Tanaka, T., and Goto, Y. (1995) *J. Mol. Biol. 254*, 737−746.
25. Shiraki, K., Nishikawa, K., and Goto, Y. (1995) *J. Mol. Biol. 245*, 180−194.
26. Luo, P., and Baldwin, R. L. (1997) *Biochemistry 36*, 8413−8421.
27. Schönbrunner, N., Wey, J., Engels, J., Georg, H., and Kiefhaber, T. (1996) *J. Mol. Biol. 260*, 432−445.
28. Blanco, J. F., Jiménez, A., Pineda, A., Rico, M., Santoro, J., and Nieto, J. L. (1994) *Biochemistry 33*, 6004−6014.
29. Wu, C.-S. C., and Yang, J. T. (1978) *Biochem. Biophys. Res. Commun. 82*, 85−91.
30. Mammi, S., and Peggion, E. (1990) *Biochemistry 29*, 5265−5269.
31. Rizo, J., Blanco, F. J., Kobe, B., Bruch, M. D., and Gierasch, L. M. (1993) *Biochemistry 32*, 4881−4894.
32. Desmadril, M., and Yon, J. M. (1984) *Biochemistry 23*, 11−19.
33. Llinás, M., and Marqusee, S. (1998) *Protein Sci. 7*, 96−104.
34. Gassner, N. C., Baase, W. A., Lindstrom, J. D., Lu, J., Dahlquist, F. W., and Matthews, B. W. (1999) *Biochemistry 38*, 14451−14460.
35. McLeish, M. J., Nielsen, K. J., Wade, J. D., and Craik, D. J. (1993) *FEBS Lett. 315*, 323−328.
36. Najbar, L. V., Craik, D. J., Wade, J. D., Lin, F., and McLeish, M. J. (1995) *Biochim. Biophys. Acta 1250*, 163−170.
37. McLeish, M. J., Nielsen, K. J., Najbar, L. V., Wade, J. D., Lin, F., Doughty, M. B., and Craik, D. J. (1994) *Biochemistry 33*, 11174−11183.
38. Najbar, L. V., Craik, D. J., Wade, J. D., Salvatore, D., and McLeish, M. J. (1997) *Biochemistry 36*, 11525−11533.
39. Wade, J. D., Layden, S. S., Lambert, P. F., Kakouris, H., and Tregear, G. W. (1994) *J. Protein Chem. 13*, 315−321.
40. Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, John Wiley & Sons, New York.
41. Rance, M., Sörensen, O. W., Bodenhausen, G., Wagner, G., Ernst, R. R., and Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 117*, 479−485.

42. Bax, A., and Davis, D. G. (1985) *J. Magn. Reson. 65*, 355−360.
43. Braünschweiler, L., and Ernst, R. R. (1983) *J. Magn. Reson. 53*, 521−528.
44. Cavanagh, J., and Rance, M. (1992) *J. Magn. Reson. 96*, 670−678.
45. Kumar, A., Ernst, R. R., and Wüthrich, K. (1980) *Biochem. Biophys. Res. Commun. 95*, 1−6.
46. Plateau, P., and Gueron, M. (1982) *J. Am. Chem. Soc. 104*, 7310−7311.
47. Matsumura, M., Becktel, W. J., and Matthews, B. W. (1988) *Nature 334*, 406−410.
48. Wetzel, R., Perry, L. J., Baase, W. A., and Becktel, W. J. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 401−405.
49. Woody, R. W. (1985) *Circular dichroism of peptides*, Vol. 7, Academic Press, New York.
50. Chen, Y.-H., Yang, J. T., and Chau, K. H. (1974) *Biochemistry 13*, 3350−3359.
51. Wishart, D. S., Sykes, B. D., and Richards, F. M. (1991) *J. Mol. Biol. 222*, 311−333.
52. Wishart, D. S., Sykes, B. D., and Richards, F. M. (1992) *Biochemistry 31*, 1647−1651.
53. Lu, J., and Dahlquist, F. W. (1992) *Biochemistry 31*, 4749−4756.
54. Ptitsyn, O. B. (1973) *Dokl. Akad. Nauk SSSR 210*, 1213−1215.
55. Kim, P. S., and Baldwin, R. L. (1982) *Annu. Rev. Biochem. 51*, 459−489.
56. Kim, P. S., and Baldwin, R. L. (1990) *Annu. Rev. Biochem. 59*, 631−660.
57. Bahar, I., Wallqvist, A., Covell, D. G., and Jernigan, R. L. (1998) *Biochemistry 37*, 1067−1075.
58. Jasanoff, A., and Fersht, A. R. (1994) *Biochemistry 33*, 2129−2135.
59. Hirota, N., Mizuno, K., and Goto, Y. (1998) *J. Mol. Biol. 275*, 365−378.
60. Cammers-Goodwin, A., Allen, T. J., Oslick, S. L., McClure, K. F., Lee, J. H., and Kemp, D. S. (1996) *J. Am. Chem. Soc. 118*, 3082−3090.
61. Wu, C.-S. C., Hachimori, A., and Yang, J. T. (1982) *Biochemistry 21*, 4556−4562.
62. Zhong, L., and Johnson, W. C., Jr. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 4462−4465.
63. Xie, D., and Freire, E. (1994) *J. Mol. Biol. 242*, 62−80.
64. Remington, S. J., Anderson, W. F., Owen, J., Ten Eyck, L. F., Grainger, C. T., and Matthews, B. W. (1978) *J. Mol. Biol. 118*, 81−98.
65. Weaver, L. H., and Matthews, B. W. (1987) *J. Mol. Biol. 193*, 189−199.
66. Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946−950.
67. Wishart, D. S., Bigam, C. G., Holm, A., Hodges, R. S., and Sykes, B. D. (1995) *J. Biomol. NMR 5*, 67−81.

BI000070I