Structure and Stability of the N-Terminal Domain of the Ribosomal Protein L9: Evidence for Rapid Two-State Folding †

ARTICLE *in* BIOCHEMISTRY · FEBRUARY 1998

Impact Factor: 3.02 · DOI: 10.1021/bi972352x · Source: PubMed

4 AUTHORS, INCLUDING:

Judith Boice
Alexion
23 PUBLICATIONS **946** CITATIONS
SEE PROFILE

Robert Fairman
Haverford College
78 PUBLICATIONS **3,733** CITATIONS
SEE PROFILE

Structure and Stability of the N-Terminal Domain of the Ribosomal Protein L9: Evidence for Rapid Two-State Folding†

Brian Kuhlman,‡ Judith A. Boice,§,|| Robert Fairman,⊥ and Daniel P. Raleigh*,‡,#

Department of Chemistry and Graduate Programs in Biophysics and in Molecular and Cellular Biology, State University of New York at Stony Brook, Stony Brook, New York 11794-3400, Bristol-Myers Squibb Pharmaceutical Research Institute, P.O. Box 4000, Princeton, New Jersey 08543-4000, and Department of Molecular, Cell and Developmental Biology, Haverford College, Haverford, Pennsylvania 19041

Received September 22, 1997; Revised Manuscript Received November 18, 1997

ABSTRACT: The N-terminal domain, residues 1−56, of the ribosomal protein L9 has been chemically synthesized. The isolated domain is monomeric as judged by analytical ultracentrifugation and concentration-dependent CD. Complete ^{1}H chemical shift assignments were obtained using standard methods. 2D-NMR experiments show that the isolated domain adopts the same structure as seen in the full-length protein. It consists of a three-stranded antiparallel β-sheet sandwiched between two helixes. Thermal and urea unfolding transitions are cooperative, and the unfolding curves generated from different experimental techniques, 1D-NMR, far-UV CD, near-UV CD, and fluorescence, are superimposable. These results suggest that the protein folds by a two-state mechanism. The thermal midpoint of folding is 77 ± 2 °C at pD 8.0, and the domain has a $\Delta G°_{folding} = 2.8 \pm 0.8$ kcal/mol at 40 °C, pH 7.0. Near the thermal midpoint of the unfolding transition, the 1D-NMR peaks are significantly broadened, indicating that folding is occurring on the intermediate exchange time scale. The rate of folding was determined by fitting the NMR spectra to a two-state chemical exchange model. Similar folding rates were measured for Phe 5, located in the first β-strand, and for Tyr 25, located in the short helix between strands two and three. The domain folds extremely rapidly with a folding rate constant of 2000 s^{-1} near the midpoint of the equilibrium thermal unfolding transition.

The ribosomal protein L9 forms an interesting bilobal structure with a compact N-terminal domain connected by a long solvent-exposed α-helix to a compact C-terminal domain (*1*−*3*) (Figure 1). The N-terminal domain of L9 is an example of an important class of layered sheet−helix structures which are found in a number of other ribosomal proteins and in small RNA binding domains (*4*−*6*). The N-terminal domain is one of the simpler examples of this structure. It lacks disulfides, does not bind to metal ions or cofactors, and is relatively small. This type of fold can be generalized in the sense that the topology of the motif is the same if the short connecting helix, which lies atop the three-stranded sheet, is replaced by a β-strand or loop. Work by Efimov and Thornton has stressed the common topology of these folds (*7*−*9*). This so-called ABCD motif is found in a large number of proteins, but relatively little is known about the interactions which stabilize this structure or how it folds.

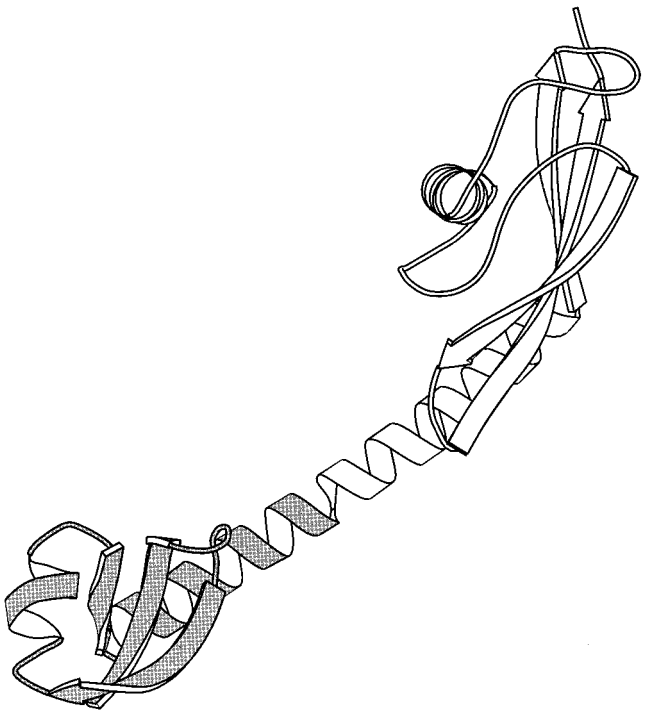

FIGURE 1: Ribbon diagram of L9 from *Bacillus stearothermophilus* [Drawing made using the progam MOLSCRIPT (*48*)]. The N-terminal domain studied here consists of the shaded region, residues 1−56.

† The NMR facility at SUNY Stony Brook is supported by grants from the NSF (CHE8911350, CHE9413510) and from the NIH (1S10RR554701). This work was supported by NSF Grant MCB 9600866 to D.P.R. D.P.R. is a Pew Scholar in the Biomedical Sciences. B.K. was supported by a GAANN Fellowship from the Department of Education during the initial stages of this project.

* Author to whom correspondence should be addressed. Telephone: 516-632-9547. Fax: 516-632-7960. E-mail: draleigh@ccmail.sunysb.edu.

‡ Department of Chemistry, State University of New York at Stony Brook.

§ Bristol-Myers Squibb Pharmaceutical Research Institute.

|| Current address: Department of Biochemistry, Merck Research Laboratories, P.O. Box 2000 RY50-105, Rahway, NJ 07065.

⊥ Haverford College.

Graduate Programs in Biophysics and in Molecular and Cellular Biology, State University of New York at Stony Brook.

S0006-2960(97)02352-0 CCC: $15.00
Published on Web 01/13/1998

The intact L9 protein also offers an interesting system for the study of the folding and assembly of multidomain proteins. The structure of the full-length L9 protein is intermediate between that of multidomain proteins in which structurally independent domains are connected by flexible linkers and multidomain proteins in which there are more extensive and intimate interdomain contacts (*10*−*12*). The two terminal domains of L9 are expected to fold in the absence of the other domain, but it is also likely that their folding and stability are not entirely independent. Our approach to understanding the folding of L9 is to study the isolated domains with the ultimate goal of comparing the folding and stability of the individual domains with the folding and stability of the intact protein. We have previously examined a peptide corresponding to the central helix of L9 and have shown that it has high intrinsic stability. This peptide is greater than 85% helical below room temperature (*13*), but at higher temperatures it begins to unfold and is less than 30% helical at 60 °C. Studies on full-length L9 show that the central helix is stable to 60 °C, indicating that it must be stabilized by interactions with the two terminal domains (*3*). In this paper we report the results of a study of the N-terminal domain of L9 in which we examine the structure, stability, and folding of the isolated domain. Our construct consists of residues 1−56 and is denoted as NTL9.

The small size of this domain, 56 residues, makes it an attractive target for biophysical studies. Until recently the only small monomeric proteins known to fold to a unique tertiary structure contained disulfide bonds or ligand binding sites. There are now a few examples of small proteins which adopt a stable tertiary structure. Their small size makes them more tractable for computational studies and also makes them attractive targets for de novo protein design.

There has been increased interest in studying small proteins which fold by a two-state mechanism (*14*, *15*). These proteins offer simple systems with which to probe the basic steps required in protein folding. There is also considerable interest in the early stages of protein folding and in experimental studies of proteins which can fold quickly. A number of recent developments have led to new methods for studying the early steps in folding. Potentially one of the most useful techniques is in fact one of the oldest, NMR[1] line-shape analysis. For proteins that are known to fold by a two-state mechanism and that fold on the appropriate time scale (10^{-4} to 10^{-3} s), it is possible to measure folding rates directly from NMR line-shape analysis (*16*, *17*). Using dynamic NMR to measure rates of protein folding is especially attractive since conventional stopped-flow experiments cannot directly detect events occurring on the submillisecond time scale. In the work described here, dynamic NMR is used to measure submillisecond folding rates for the N-terminal domain of the ribosomal protein L9.

[1] Abbreviations: CD, circular dichroism; CSI, chemical shift index; Da, dalton; DQF-COSY, double quantum filtered correlated spectroscopy; FMOC, 9-flourenylmethyloxycarbonyl; NMR, nuclear magnetic resonance; HPLC, high-pressure liquid chromatography; NOE, nuclear Overhauser effect; NOESY, nuclear Overhauser effect spectroscopy; PAL, 5-(4′-FMOC-aminomethyl-3′,5′-dimethoxyphenoxy)valeric acid; TBTU, 2-(1*H*-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium tetrafluoroborate; TFA, trifluoroacetic acid; TOCSY, total correlated spectroscopy; TSP, 3-(trimethylsilyl) propionate.

MATERIALS AND METHODS

Peptide Synthesis and Purification. NTL9 was synthesized on a 0.22 mmol scale by solid-phase methods using FMOC-protected amino acids and TBTU-mediated amide coupling on a Millipore 9050 Plus automated peptide synthesizer with standard reaction cycles. All β-branched amino acids and all residues which were coupled to β-branched residues were double coupled. Each coupling, except for the last, was followed by a capping step using acetic anhydride. Amino acid side chains were protected as follows: Asn and Cys, trityl group; Asp, *tert*-butyl ester; Lys, *tert*-butoxycarbonyl group; Ser and Thr, *tert*-butyl ether. Use of a resin with a PAL linker generated an amidated C-terminus following cleavage from the resin with 91% TFA/3% anisole/3% thioanisole/3% ethanedithiol. NTL9 was purified using reverse phase HPLC with a C4 column. An A−B gradient was used where the A eluent was 0.1% aqueous TFA and the B eluent was 0.1% TFA/90% 2-propanol/9.9% H_2O. Electrospray mass spectrometry (expected molecular weight 6218, observed molecular weight 6217) and nuclear magnetic resonance (NMR) confirmed the identity of the pure product. Protein concentrations were determined by tyrosine absorbance using an extinction coefficient of 1280 M^{-1} cm^{-1} at 280 nm for the protein in 6 M guanidine hydrochloride, pH 7.0 (*18*).

Sedimentation Equilibrium. A solution of NTL9 was dialyzed against 10 mM Na_2HPO_4, 100 mM NaCl at pH 7.4. Experiments were performed at 25 °C with a Beckman XL-A analytical ultracentrifuge, using rotor speeds of 30 000, 40 000, and 50 000 rpm. The concentration of NTL9 was 280 μM. Experiments were carried out using 12 mm pathlength, six-channel, charcoal-filled Epon cells with quartz windows. Data were collected using continuous radial scanning at 285 nm. Partial specific volumes were calculated from the weighted average of the partial specific volumes of the individual amino acids (*19*). The data were globally fit with both a single-species model with the molecular weight treated as a fitting parameter and to monomer−*n*mer equilibria with the molecular weight of the monomer held fixed. The HID program from the Analytical Ultracentrifugation Facility at the University of Connecticut was used for the fitting analysis.

Circular Dichroism (CD) Spectroscopy. CD spectroscopy was performed using an Aviv 62A DS spectrometer equipped with a Peltier temperature control unit.

Nuclear Magnetic Resonance (NMR) Spectroscopy. NMR experiments were performed on a Bruker Instruments AMX 600 spectrometer and Varian Instruments Inova 500 and 600 MHz spectrometers. All spectra were internally referenced to TSP at 0.0 ppm. 2D spectra were taken of 4 mM NTL9 at 25 °C in 90% H_2O/10% D_2O at pH 5.0 and in 100% D_2O at pD 5.0 (corrected meter reading). NOESY (mixing time = 50, 100, 150, and 210 ms), TOCSY, and DQF-COSY experiments were used to make assignments (*20*). $^3J_{HN\alpha}$ coupling constants were determined from absorptive and dispersive peak separations in a DQF-COSY spectrum (*21*). Additional 2D spectra, NOESY, DQF-COSY, and TOCSY, were obtained at 55 °C. Chemical shift indices (*22*) were calculated using the random coil chemical shifts from Wüthrich (*20*).

Thermal Denaturation. Thermal denaturation of NTL9 was followed by 1D NMR and by CD at 222 and 280 nm. The NMR and CD samples were in D_2O, 100 mM NaCl, 10 mM Tris at pD 8.0. Protein concentrations were 2.5 mM for NMR, 40 μM for far-UV CD, and 600 μM for near-UV CD. The temperature within the NMR sample was calibrated with ethylene glycol, and the temperature within the CD cuvette was calibrated with an external thermocouple. The experimental data, plots of ellipticity versus temperature or chemical shifts versus temperature, were fit with the program Kaleidagraph (Abelbeck Software) to eq 1.

$$f(T) = \frac{\alpha_N + \beta_N T + (\alpha_D + \beta_D T)e^{-[\Delta G^\circ_{D-N}(T)]/RT}}{1 + e^{-[\Delta G^\circ_{D-N}(T)]/RT}} \quad (1)$$

where

$$\Delta G^\circ_{D-N}(T) = \Delta H^\circ_{D-N}(T_m)\left(1 - \frac{T}{T_m}\right) - \Delta C^\circ_p\left[(T_m - T) + T\ln\left(\frac{T}{T_m}\right)\right] \quad (2)$$

f is chemical shift or ellipticity, and T is temperature. α_N, β_N, α_D, and β_D are parameters that define the ellipticity or chemical shift of the native (N) and denatured states (D). α and β describe a line with a slope equal to β and a y intercept equal to α. T_m is the transition midpoint, and $\Delta H^\circ_{D-N}(T_m)$ is the change in enthalpy at T_m. The value of ΔC°_p was set to 0.57 kcal mol^{-1} K^{-1}. This value was calculated from known correlations between a protein's size and its heat capacity (*23*). Changing ΔC°_p by $\pm$0.20 kcal mol^{-1} K^{-1} changes the calculated T_m values by less than 1 °C. Fraction folded at a given temperature was determined by using $f(T)$, α_N, β_N, α_D, and β_D. Thermal denaturation was 95−100% reversible in 100 mM NaCl, provided the protein was kept at high temperatures for less than 1 h.

Chemical Denaturation. Chemical denaturation of NTL9 was followed with CD at 222 nm and with tyrosine fluorescence on a Perkin-Elmer LS-5B luminescence spectrometer. The excitation wavelength for fluorescence was 279 nm, and the emission wavelength was 304 nm. The samples were prepared in 10 mM MOPS, 100 mM NaCl, at pH 7.0 with varying amounts of urea at 25 °C. The protein concentration was 70 μM. The same sample was used for fluorescence and CD. The concentration of urea was determined with refractometry (*24*). Kaleidagraph was used to fit the data to eq 1 with ΔG°_{D-N} now equal to eq 3

$$\Delta G^\circ_{D-N}([\text{urea}]) = m[\text{urea}] + \Delta G^\circ_{D-N}(0\text{ M}) \quad (3)$$

$\Delta G^\circ_{D-N}(0$ M) is the stability of the protein in the absence of denaturant. Fraction folded at a given urea concentration was determined by using $f([\text{urea}])$, α_N, β_N, α_D, and β_D.

Folding Rates from NMR Line-Shape Analysis. The ^{1}H NMR line shape of the 3,5 resonance of Tyr 25 (6.5 ppm, Figure 5) and the 2,6 resonance of Phe 5 (7.5 ppm, Figure 5) were fit to a two-site exchange model (*17, 25*). Intensity as a function of frequency [$I(\nu)$] depends on the populations of the native state and the denatured state (p_N, p_D), the apparent transverse relaxation times of the resonance in the native and denatured state (T_{2N}, T_{2D}), the folding and unfolding rate constants (k_f, k_u) and the frequencies of the native and denatured state resonances (ν_N, ν_D). The following equations describe how these parameters are related (*17, 25*).

$$I(\nu) = C_0\frac{\left\{P\left[1 + \tau\left(\frac{p_D}{T_{2N}} + \frac{p_N}{T_{2D}}\right)\right] + QR\right\}}{P^2 + R^2} \quad (4)$$

where P, Q, and R are given by

$$P = \tau[(1/T_{2N}T_{2D}) - 4\pi^2(\Delta\nu)^2 + \pi^2(\delta\nu)^2] + (p_N/T_{2N}) + (p_D/T_{2D}) \quad (5)$$

$$Q = \tau[2\pi\Delta\nu - \pi\delta\nu(p_N - p_D)] \quad (6)$$

$$R = 2\pi\Delta\nu\{1 + \tau[(1/T_{2N}) + (1/T_{2D})]\} + \pi\delta\nu\tau[(1/T_{2N}) - (1/T_{2D})]\pi\delta\nu(p_N - p_D) \quad (7)$$

$\delta\nu$, $\Delta\nu$, are τ given by

$$\delta\nu = \nu_N - \nu_D \quad (8)$$

$$\Delta\nu = [(\nu_N + \nu_D)/2] - \nu \quad (9)$$

$$\tau = 1/(k_u + k_f) \quad (10)$$

p_N and p_D can be expressed in terms of τ, k_f, and k_u

$$p_N = k_f\tau \quad (11)$$

$$p_D = k_u\tau \quad (12)$$

C_0 is a normalization constant that depends on the concentration of protein used in the experiment. The line shapes were fit using a nonlinear least-squares fitting routine (Sigmaplot, SPSS Inc.) in order to determine values for k_f, k_u, and C_0. The other parameters were determined from the spectra of folded and unfolded protein. For the fits to the phenylalanine resonance it was not possible to determine the frequency of the denatured state resonance from the spectra because of resonance overlap. For this reason, ν_D was treated as an independent variable and p_N was set equal to the value derived from the fit of the tyrosine resonance.

RESULTS

Sedimentation experiments indicate that NTL9 is monomeric in aqueous solution (Figure 2) at a concentration of 280 μM. The apparent molecular mass derived from a global fit to the centrifugation data is 5890 $\pm$ 150 Da which is within 5% of the known actual mass, 6219 Da. The difference between the expected and observed molecular mass is possibly due to nonideal effects. Fitting the data to a dimer gives a much poorer fit with large nonrandom residuals (data not shown). The thermal unfolding curves as monitored by CD or NMR are also independent of concentration. Temperature melts of NTL9 at 40 μM, 600 μM, and 2.5 mM are superimposable, and the 1D NMR spectra of 2 and 4 mM NTL9 are identical; this indicates that NTL9 remains a monomer up to a concentration of at least 4 mM.

NMR experiments show that the structure of the isolated domain is similar to the conformation of the domain in the

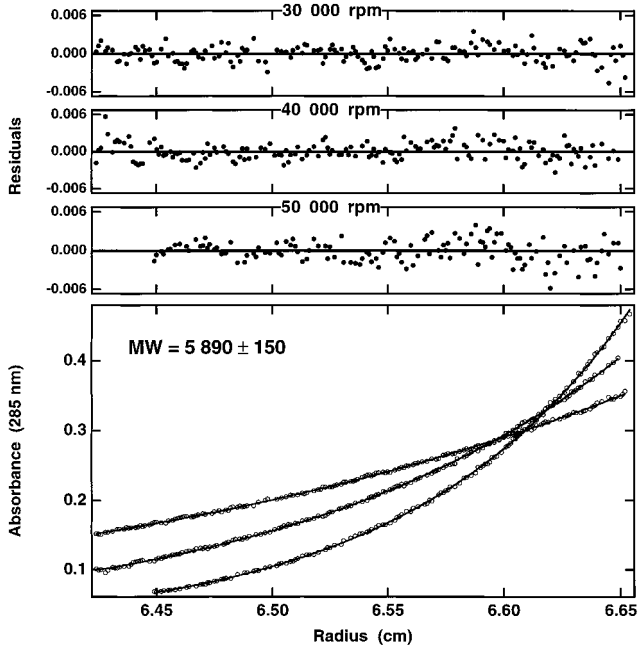

FIGURE 2: Analytical ultracentrifugation of NTL9 shows that it is monomeric. A single species fit to a molecular weight of 5890 is shown. This is within 5% of the true molecular weight, 6219. The residuals for three separate rotor speeds are shown.

full-length protein. Complete ^{1}H chemical shift assignments are provided in the Supporting Information. The observed chemical shifts, chemical shift indices (CSI), $^3J_{HN\alpha}$ coupling constants, and nuclear Overhauser effects (NOEs) are all consistent with the secondary structure and tertiary structure of the domain in full-length L9 (Figure 3a). The α-proton chemical shifts in NTL9 are on average within 0.02 ppm of those in L9, with only 4 of the 56 differing by more than 0.05 ppm (*2*). The C-terminal helix, residues 41−56, is intact except for fraying toward the C-terminal end of the helix (Figure 3a, Table 1). The chemical shift indices for the last two residues of the helix, Gln 55 and Arg 56, are zero, and there are no *i*, *i* + 3 NOEs to these residues. Interestingly, *i*, *i* + 4 NOEs rather than *i*, *i* + 3 NOEs are observed in the region corresponding to the short helix (residues 26−32). These NOEs are consistent with the full-length L9 structure in which residues 26−31 form a $\pi_{\alpha R}$ (*26*) turn with an *i*, *i* + 5 hydrogen bond between Ala 26 and Phe 31.

All the expected NOEs between β-strands are observed. These include NOEs between α-protons on adjacent strands: Ile 18−Ile 4, Asn 20−Lys 2, Ala 36−Phe 5, and Glu 38−Val 3 (Figure 3b). These cross-strand NOEs show that the register of the β-strands in the isolated domain is the same as in the intact protein. There are also several tertiary NOEs which are consistent with the full-length structure. The side chains of Phe 31 and Phe 5 form part of the hydrophobic core between the short helix and the β-sheet and show the expected NOEs with Met 1, Val 3, and Gly 13. The β-sheet and the C-terminal helix pack against each other, with Leu 44 and Leu 47 from the helix making contact with Ile 4, Ile 18, and Ile 37 from the sheet. In this case the expected NOEs are between methyl groups on these residues, and the poor chemical shift dispersion prevents unambiguous identification of these NOEs. The chemical shifts reported for the methyl groups of Ile 4, Ile 18, and Ile 37 in the full-length protein are all within 0.04 ppm of the values observed for NTL9, and in the case of Ile 4 and Ile 37 the shifts deviate

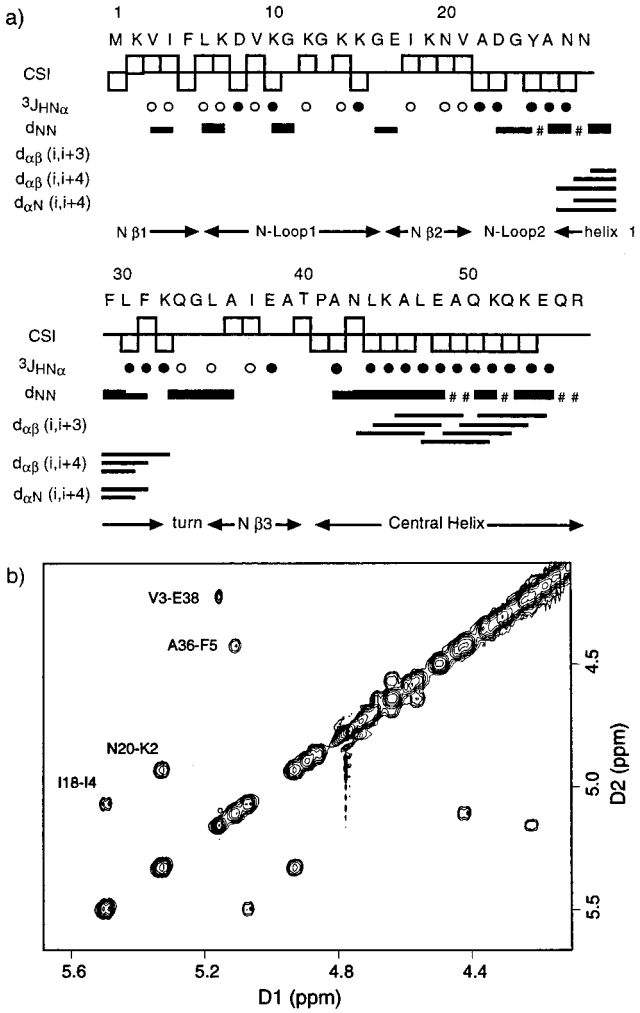

FIGURE 3: (a) Sequential NOEs for NTL9 show that the structure of the isolated domain is similar to its structure in full length L9. The intensity of the NOE is represented by the thickness of the shaded line. # denotes an NOE which could not be detected because the expected cross-peak is too close to the diagonal. CSI refers to chemical shift index; below the horizontal indicates a CSI of −1, and above the horizontal indicates +1. $^3J_{HN\alpha}$ couplings are shown as a filled circle if they are below 6 Hz, and an empty circle if they are above 8 Hz. Elements of secondary structure are labeled as defined in the combined NMR, X-ray structure (*2*). Spectra were taken at 25 °C, pH 5.0. (b) A portion of the 2D NOESY spectrum of NTL9 at 25 °C, pD 5.0. NOEs are shown between α-protons on adjacent β-strands.

from random-coil chemical shifts by 0.11−0.20 ppm (*2*). These results suggest that the C-terminal helix packs against the β-sheet in similar fashion in NTL9 and L9.

The one region in which the structure of the isolated domain may deviate from that of the intact protein is the loop (residues 10−16) that connects the first and second strands of NTL9. The loop contains only lysines and glycines and has the sequence ^{10}Lys-Gly-Lys-Gly-Lys-Lys-Gly^{16}. In the L9 structure Lys 14 has a ϕ angle of −51° but in NTL9 the measured $^3J_{HN\alpha}$ coupling constant is 9.7 Hz which corresponds to a ϕ angle between −90 and −150°. In NTL9 there is a medium-strength NOE between a β proton on Glu 17 and the amide proton on Lys 14. In the structure of L9 these groups are over 7 Å apart.

Thermal denaturations of NTL9 were followed by near- and far-UV CD as well as by NMR (Figure 4). NTL9 has one tyrosine, Tyr 25, which has a near-UV CD signal with

Table 1: NMR Data Provide Evidence for the Fraying of the C-Terminal Helix of NTL9[a]

residue	$^3J_{HN\alpha}$ (25 °C) (Hz)	$^3J_{HN\alpha}$ (55 °C) (Hz)	NH−NH (25 °C)[b]	NH−NH (55 °C)[b]	Cα(55−25 °C) (ppm)
Ala 42	3.6	2.8	strong	medium	−0.02
Asn 43	6.1		strong	strong	
Leu 44	5.0	6.1	strong	strong	−0.02
Lys 45	4.3	4.6	strong	strong	−0.01
Ala 46	4.8	4.4	strong	strong	−0.03
Leu 47	5.1	5.0	strong	strong	−0.03
Glu 48	4.4	5.0	*	*	0.01
Ala 49	4.4	5.1	*	*	0.00
Gln 50	5.5	6.0	strong	none	0.04
Lys 51	5.6	5.8	*	*	0.04
Gln 52	5.3	6.0	strong	none	0.05
Lys 53	5.7	6.6	strong	none	0.02
Glu 54	5.9	6.6	*	*	0.02
Gln 55	6.4	6.6	*	*	0.04
Arg 56	6.8	7.3			0.01

[a] Values of the $^3J_{HN\alpha}$ coupling constants and the intensity of the amide-to-amide sequential NOEs measured at 25 and 55 °C are tabulated for residues 42−56. The difference between the Cα-proton chemical shifts measured at 55 and 25 °C is also listed. [b] Strength of NOEs between backbone amides on adjacent residues. "Strong" corresponds to 1.8−3.0 Å, and "medium" corresponds to 1.8−4.0 Å. "None" indicates the absence of an NOE. An asterisk indicates that a possible NOE is too close to the diagonal to be observed.

a maximum at 278 nm and a mean residue ellipticity of −40 mdeg cm^2 dmol^{-1}. The signal decays to zero as the protein unfolds. For the temperature melts followed by CD the protein was at high temperatures for less than 1 h, and unfolding was greater than 95% reversible. The T_m measured by near-UV CD (77 ± 2 °C) and far-UV CD (78 ± 2 °C) are the same.

The equilibrium unfolding transition can also be monitored with NMR. In almost all studies reported to date, the conversion between the folded and unfolded states is slow on the chemical shift time scale. As a result, separate peaks are observed for the folded and unfolded states, and the fraction folded can be determined from the relative intensities. In contrast, NTL9-folds on the fast to intermediate exchange time scale (Figure 5). In this case the observed chemical shifts are a population weighted average of the folded and unfolded chemical shifts and can be used to determine the fraction folded.

A minor complication is caused by the single proline at position 41 which has a trans peptide bond in the folded protein, but can have a cis or trans peptide bond in the unfolded protein. The proline is located immediately before the start of the long central helix, and examination of the structure indicates that a cis proline cannot be accommodated in the native state. Only a single set of resonances is observed for the folded protein, and a strong NOE is observed from the α-proton of residue 40 to the δ-protons of proline 41 indicating that the Thr 40−Pro 41 peptide bond exists exclusively in the trans conformation in the folded state. This implies that the fraction of the molecules which contain a cis proline peptide bond in the unfolded state must first isomerize to the trans form before they are competent to fold. The cis−trans isomerization rate for a prolyl peptide bond is known to be slow on the chemical shift time scale (*27*, *28*). As a result the small population of unfolded protein with a cis proline will be in slow exchange with the major unfolded form and hence with the folded state. Unfolded protein in slow exchange is probably the reason for the peak

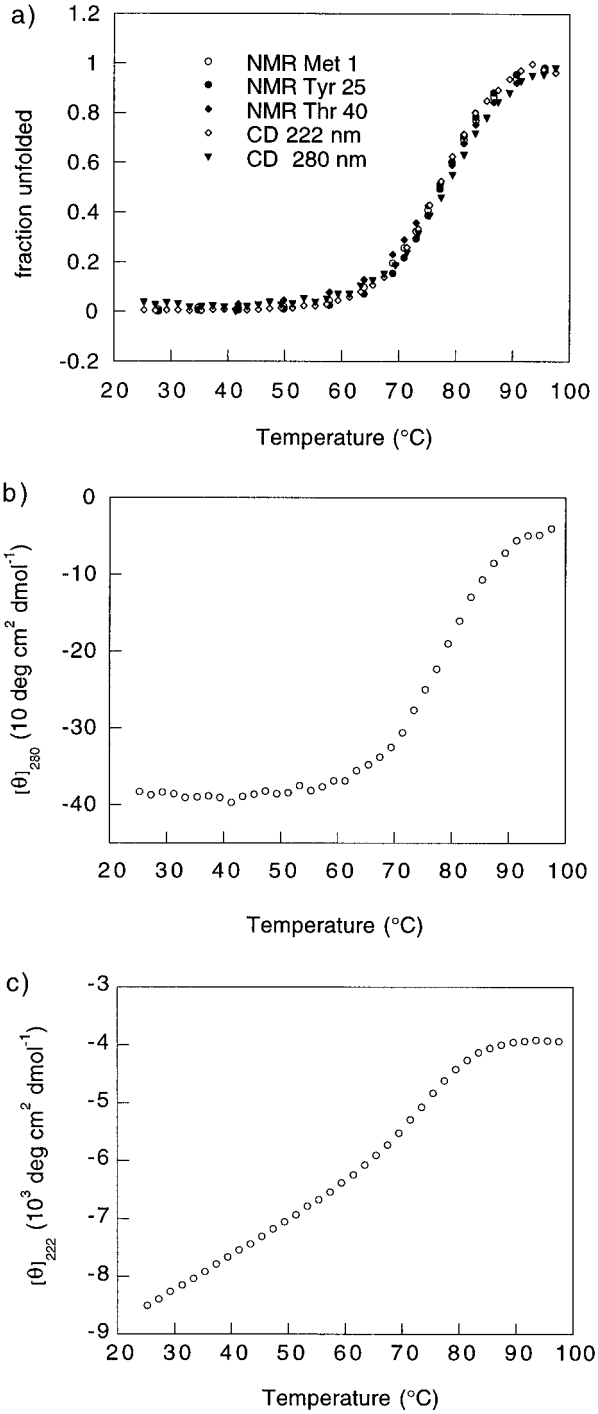

FIGURE 4: Temperature denaturation of NTL9 in 100 mM NaCl, 10 mM Tris, 100% D_2O, pD 8.0. (a) Fraction unfolded of NTL9 as a function of temperature as determined by several different techniques. Equation 1 was used to determine fraction unfolded. (b) Near-UV CD signal at 280 nm as a function of temperature. (c) Far-UV CD signal at 222 nm as a function of temperature.

which grows in at the unfolded chemical shift of Tyr 25 (6.9 ppm in Figure 5). There is apparently a second minor peak of even lower intensity that appears as a small shoulder on the cis peak. This small peak likely arises from protein that is irreversibly unfolded at high temperatures. For this experiment NTL9 was at high temperatures for over 2 h, at which point thermal folding is not completely reversible. NTL9 contains four asparagines, and deamidation of asparagines is known to occur with time constants on the order of hours at high temperatures (*29*). Protein that is irreversibly

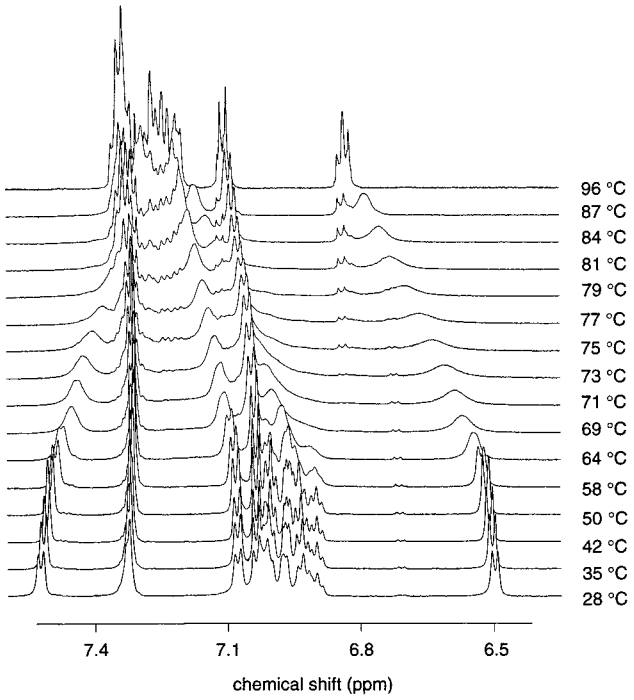

FIGURE 5: 1D NMR of NTL9 at different temperatures in 100 mM NaCl, 10 mM Tris, 100% D_2O, pD 8.0. The small peak growing in with temperature probably arises from the small fraction of unfolded molecules which contain a cis prolyl peptide bond.

unfolded will not influence the measured chemical shifts and will not affect the fraction folded that is measured.

The 3,5 protons of Tyr 25, the methyl protons of Met 1, and the methyl protons of Thr 40 all give rise to well-resolved peaks that could be followed as a function of temperature using 1D NMR. Equation 1 was used to fit the temperature dependence of the chemical shifts of these three peaks as well as the temperature-dependent ellipticities determined from the CD studies. For the NMR experiments, fraction folded was adjusted to take into account the small amount of unfolded protein in slow exchange due to isomerism about the prolyl peptide bond. This correction changed the measured T_m by less than 1 °C. Fraction unfolded was plotted as a function of temperature (Figure 4a). The five curves overlay well. The experimentally determined midpoints for the transition are as follows: 77 ± 2 °C Tyr 25, 77 ± 2 °C Thr 40, 77 ± 2 °C Met 1, 77 ± 2 °C far-UV CD; and 78 ± 2 °C near-UV CD. These results strongly suggest that there are no significantly populated intermediates throughout the equilibrium unfolding transition.

The raw data for the thermal denaturation of NTL9 as followed by near- and far-UV CD are shown in panels b and c of Figure 4, respectively. One noticeable feature is the pronounced pretransition in the far-UV CD data. The CD signal at 222 nm is in large part due to the helixes in NTL9, and the decreased signal at 222 nm with temperature might be due to helix fraying or unwinding. Leu 47 is the last residue in the C-terminal helix to form part of the hydrophobic core of the N-terminal domain. Gln 50 and Lys 51 also contact the Lys−Gly loop between the first two β-strands of the N-terminal domain, while residues 52−56 do not make any contacts with the rest of the N-terminal domain. 2D NMR experiments were performed at 55 °C to determine if the C-terminal helix remains intact at higher temperatures. These studies show that the helix is fraying

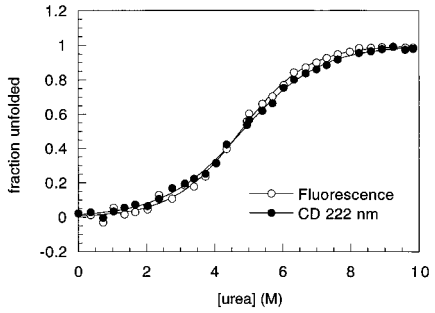

FIGURE 6: Urea denaturation of NTL9 as followed by tyrosine fluorescence and far-UV CD (222 nm). The peptide was in 10 mM MOPS, 100 mM NaCl, pH 7.0 at 40 °C. The curves were fit using eqs 2 and 3. The curve fit to the far-UV CD data gives a $\Delta G°$ (0 M urea, 40 °C) of 2.6 ± 0.8 kcal/mol, and the curve fit to the fluorescence data gives a $\Delta G°$ (0 M urea, 40 °C) of 3.0 ± 0.8 kcal/mol.

as the temperature is raised (Table 1). Residues 48−56 all have larger $^3J_{HN\alpha}$ coupling constants at 55 °C, and the chemical shifts of the α-protons of residues 48−56 have all moved toward the random coil values. Also the *i*, *i* + 1 NH−NH NOEs are no longer observed for Gln 50−Lys 51, Gln 52−Lys 53, and Lys 53−Glu 54.

Chemical denaturation of NTL9 was followed with far-UV CD and tyrosine fluorescence. The only tyrosine in the protein, Tyr 25, has an emission maximum at 304 nm and an excitation maximum at 279 nm. The emission maximum is still 304 nm in unfolded NTL9, but the fluorescence intensity decreases 2−3-fold when the protein unfolds. Urea denaturation curves determined by the two techniques are superimposable (Figure 6), suggesting that the chemically induced unfolding transition is also two-state. The curve fit to the far-UV CD data gives a $\Delta G°$ (0 M urea, 40 °C) of 2.6 ± 0.8 kcal/mol and an *m* value of 0.54 ± 0.10 kcal mol^{-1} M^{-1} with a midpoint at 4.8 M urea. The curve fit to the fluorescence data gives a $\Delta G°$ (0 M urea, 40 °C) of 3.0 ± 0.8 kcal/mol and an *m* value of 0.63 ± 0.10 kcal mol^{-1} M^{-1} with a midpoint at 4.8 M urea. The errors are large because the transition is broad as is typical with small proteins, and therefore it is difficult to determine the four parameters which define the pre- and post-transitions. This uncertainty accounts for the slight divergence in the two fitted curves (Figure 6). The experiment was conducted at 40 °C because a urea denaturation of NTL9 at 25 °C is difficult to fit since there is no post-transition.

m values represent the change in free energy per amount of denaturant added and are directly related to the steepness of a chemical denaturation curve. Myers et al. have determined an empirical correlation between changes in accessible surface area (ΔASA) on folding and *m* values (*23*). Using the structure of L9 and the access module in the program Whatif, we calculate that the ΔASA of unfolding for NTL9 is 4200 Å^2 (*30*). The correlation proposed by Myers et al. for proteins with no disulfide cross-links predicts an *m* value of 0.59 kcal mol^{-1} M^{-1} at 25 °C. This calculated value is similar to the values observed with far-UV CD and fluorescence at 40 °C, 0.54 ± 0.1 and 0.63 ± 0.1 kcal mol^{-1} M^{-1}, respectively.

All of the available data indicate that the folding of NTL9 is two-state. The temperature-dependent NMR spectra show

Table 2: NMR Line-Shape Analysis[a]

temp (°C)	k_f (Y25) (s^{-1})	k_u (Y25) (s^{-1})	k_f (F5) (s^{-1})	k_u (F5) (s^{-1})	$k_f/(k_u + k_f)$ (Y25)	fraction folded (CD)
69.0	2000	319	1780	284	0.86	0.81
71.0	1930	452	2150	504	0.81	0.77
73.1	1840	623	1750	593	0.75	0.70
75.2	1790	918			0.66	0.63
77.3	1600	1170			0.58	0.55
79.4	1480	1660			0.47	0.46
81.4	1360	2270			0.37	0.38
83.4	1290	3080			0.30	0.30
86.7	1040	4580			0.19	0.19

[a] The rate constants for folding and unfolding are listed for fits to two resonances, the 3,5 protons of Tyr 25 and the 2,6 protons of Phe 5. The fraction of molecules that are folded as determined from kinetic results [$k_f/(k_u + k_f)$] is compared with the fraction folded determined from equilibrium unfolding experiments monitored with near-UV CD.

that the rate of folding is such that intermediate exchange line shapes are observed. Consequently it is possible to use NMR line-shape analysis to determine rates of protein folding and unfolding (*17*). Two resonances in the 1D NMR spectra are well-enough resolved to use this technique to measure folding rates, the 3,5 protons on Tyr 25 (6.5 ppm, Figure 5) and the 2,6 protons on Phe 5 (7.5 ppm, Figure 5). The results are listed in Table 2. The fits to the two resonances give rate constants within 12% of each other. This is within the range of the estimated uncertainty. To estimate the reliability of the fit, the effect on the calculated spectra of varying both the folding rate (k_f) and the unfolding rate (k_u) was calculated for the tyrosine resonance at 71 °C (Figure 7). The results are shown in Figure 7, which displays the residuals between the experimental data and the calculated line shape. The fits are very sensitive to changes in k_f and k_u when they are changed independently of each other. When the ratio of k_f to k_u is kept constant, the residuals only become significantly larger when k_f is changed by more than 12−14%.

The protein folds very rapidly with a folding rate constant of 2000 s^{-1} at 71 °C. Over the temperature range examined the rate of folding decreases as the temperature is raised, indicating non-Arrenhius behavior. A non-Arrhenius temperature dependence of the folding rate has been observed for other proteins near the midpoint of their thermal unfolding transitions (*31*−*34*). One explanation for the non-Arrhenius temperature dependence is that there is a change in heat capacity between the unfolded state and the transition state ensemble and that as the temperature is raised the unfolded state becomes destabilized relative to the transition state ensemble (*35*, *36*).

If the N-terminal domain of L9-folds via a two-state mechanism, the equilibrium constant determined from the kinetic data should agree with the equilibrium constant determined from the equilibrium thermal unfolding experiments. Using the rates of folding and unfolding derived from line-shape analysis, it is possible to calculate the fraction of molecules folded at a given temperature. In Table 2 these values are compared with results from equilibrium thermal unfolding experiments followed by near-UV CD. The results agree to within 7% of each other and are consistent with a two-state folding mechanism.

DISCUSSION

The data presented here show that the N-terminal domain of L9 is able to fold in isolation. The domain is monomeric,

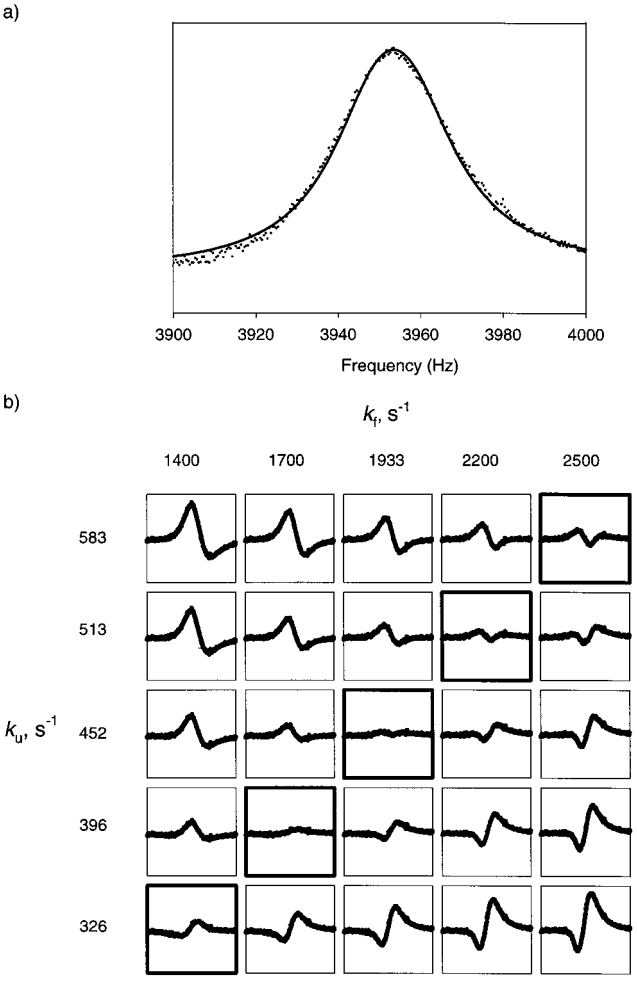

FIGURE 7: Line-shape analysis of the NMR resonance from the 3,5 protons of Tyr 25 at 71 °C. The spectrum was fit using a two-state exchange model (see Materials and Methods). (a) The observed resonance (points) and the simulated spectrum (line) that gave the best fit to the data. The measured rate of folding was 1930 s^{-1}, and the measured rate of unfolding was 450 s^{-1}. (b) The residuals between simulated and observed data for different combinations of k_f and k_u. The diagonal squares outlined in bold all have the same equilibrium constant, k_f/k_u.

and the structure of the isolated domain is similar to the structure of the N-terminal domain in full-length L9. NTL9 has high thermostability with a T_m of 77 °C in 100 mM NaCl, 100% D_2O at pD 8.0. The T_m for this domain in the intact L9 protein is not known, but a temperature melt of L9 followed by far-UV CD has been reported (*3*). L9 unfolds between 70 and 100 °C. Since CD provides a global view of the secondary structure it is not possible to independently follow the unfolding transition of the domains. The high intrinsic thermal stability of both the N-terminal domain and the intact protein make it difficult to judge if the domain is stabilized by interactions with the rest of the protein. Studies on the isolated central helix of L9 indicate that it is exceptionally stable and is greater than 85% helical at 1 °C, but is less than 30% helical at 60 °C (*13*). These results show that the central helix must be stabilized by interactions with the N- and C-terminal domains.

All of the available evidence indicates that the N-terminal domain folds rapidly by a two-state mechanism. We have shown that NMR line-shape analysis can be used to measure the folding and unfolding rates of NTL9. Use of this technique to measure folding rates is appealing because

conventional stopped-flow equipment cannot directly measure events occurring on the submillisecond time scale. Currently there is considerable interest in developing techniques to measure folding events occurring on the submillisecond time scale (*37*, *38*). These methods include ultrarapid mixing (*39*), laser-induced temperature jumps (*40*, *41*), photochemical triggers (*42*), as well as NMR (*16*, *17*). The advantages of the NMR technique are that it is experimentally simple and the temperature and solvent conditions can be easily varied. The limitation of the technique is that it is conducted under equilibrium conditions and requires a detectable population of folded and unfolded protein.

During the last two years several proteins have been shown to fold on the submillisecond time scale. These include an all-α-helical structure derived from phage λ repressor (*16*, *17*), the all-α-helical yeast acyl-coenzyme A binding protein (*43*), an all-β-sheet cold shock protein (*44*), the mixed α + β structure of NTL9 described here, and another α + β fold protein, procarboxypeptidase A2 (*45*). All of these proteins are small, of less than 100 residues, but clearly a specific fold does not appear to be required for fast folding, nor does small size alone appear to be sufficient to ensure such rapid folding (*46*, *47*).

The results reported here show that NTL9 will be a useful protein for further protein folding studies. Its small size means that it can be chemically synthesized, which allows for the addition of nonnatural amino acids and for isotopic labeling of specific residues. The combination of site-specific isotopic labeling with dynamic NMR line-shape analysis will allow folding rates to be measured on a residue-specific basis.

ACKNOWLEDGMENT

We thank Professor David Hoffman for helpful discussions, for providing the coordinates of L9, and for providing data before publication. Mass spectrometry was performed at the University of Illinois Mass Spectrometry Center.

SUPPORTING INFORMATION AVAILABLE

^{1}H assignments for NTL9 at 25 °C, pH 5.0 (2 pages). Ordering information is given on any current masthead page.

REFERENCES

1. Hoffman, D. W., Davies, C., Gerchman, S. E., Kycia, J. H., Porter, S. J., White, S. W., and Ramakrishnan, V. (1994) *EMBO J. 13*, 205−212.
2. Hoffman, D. W., Cameron, C. S., Davies, C., White, S. W., and Ramakrishnan, V. (1996) *J. Mol. Biol. 264*, 1058−1071.
3. Lillemoen, J., Cameron, C. S., and Hoffman, D. W. (1997) *J. Mol. Biol. 267*, 482−493.
4. Wolin, S. L., and Walter, P. (1991) *Curr. Opin. Struct. Biol. 1*, 251−257.
5. Nagai, K. (1996) *Curr. Opin. Struct. Biol. 6*, 53−61.
6. Burd, C. G., and Dreyfuss, G. (1994) *Science 265*, 615−621.
7. Efimov, A. V. (1994) *FEBS Lett. 355*, 213−219.
8. Orengo, C. A., and Thornton, J. M. (1993) *Structure 1*, 105−120.
9. Efimov, A. V. (1995) *J. Mol. Biol. 245*, 402−415.
10. Price, N. C. (1994) in *Mechanisms of Protein Folding* (Pain, R. H., Ed.) Oxford University Press, Oxford, U.K.
11. Jaenicke, R. (1991) *Biochemistry 30*, 3147−3161.
12. Abkevich, V. I., Gutin, A. M., and Shakhonovich, E. I. (1995) *Protein Sci. 4*, 1167−1177.
13. Kuhlman, B., Yang, H. Y., Boice, J. A., Fairman, R., and Raleigh, D. P. (1997) *J. Mol. Biol. 270*, 640−647.
14. Jackson, S. E., and Fersht, A. R. (1991) *Biochemistry 30*, 10436−10443.
15. Fersht, A. R. (1997) *Curr. Opin. Struct. Biol. 7*, 3−9.
16. Burton, R. E., Huang, G. S., Daugherty, M. A., Fullbright, P. W., and Oas, T. G. (1996) *J. Mol. Biol. 263*, 311−322.
17. Huang, G. S., and Oas, T. G. (1995) *Proc. Natl. Acad. Sci. U.S.A. 92*, 6878−6882.
18. Brandts, J. F., and Kaplan, L. J. (1973) *Biochemistry 12*, 2011−2024.
19. Cohn, E. J., and Edsall, J. T. (1943) *Proteins, Amino Acids and Peptides as Ions and Dipolar Ions*, Reinhold Publishing Corp., New York.
20. Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, John Wiley & Sons, New York.
21. Kim, Y., and Prestegard, J. H. (1990) *J. Magn. Reson. 84*, 9−13.
22. Wishart, D. S., Sykes, B. D., and Richards, F. M. (1992) *Biochemistry 31*, 1647−1651.
23. Myers, J. K., Pace, C. N., and Scholtz, J. M. (1995) *Protein Sci. 4*, 2138−2148.
24. Bower, V. E., and Robinson, R. A. (1963) *J. Phys. Chem. 67*, 1524−1527.
25. Sandstrom, J. (1982) *Dynamic NMR Spectroscop.*, Academic Press, London.
26. Rajashankar, K. R., and Ramakumar, S. (1996) *Protein Sci. 5*, 932−946.
27. Raleigh, D. P., Evans, P. A., Pitkeathly, M., and Dobson, C. M. (1992) *J. Mol. Biol. 228*, 338−342.
28. Grathwohl, C., and Wüthrich, K. (1981) *Biopolymers 20*, 2623−2633.
29. Geiger, T., and Clarke, S. (1987) *J. Biol. Chem. 262*, 785−794.
30. Vriend, G. (1990) *J. Mol. Graphics 8*, 52−56.
31. Alexander, P., Orban, J., and Bryan, P. (1992) *Biochemistry 31*, 7243−7248.
32. Tan, Y. J., Oliveberg, M., and Fersht, A. (1996) *J. Mol. Biol. 264*, 377−389.
33. Schindler, T., and Schmid, F. (1996) *Biochemistry 35*, 16833−16842.
34. Chen, B., Baase, W. A., and Schellman, J. A. (1989) *Biochemistry 28*, 691−699.
35. Baldwin, R. L. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 8069−8072.
36. Dill, K. A., and Chan, H. S. (1997) *Nat. Struct. Biol. 4*, 10−19.
37. Eaton, W. A., Thompson, P. A., Chan, C.-K., Hagen, S. J., and Hofrichter, J. (1996) *Stucture 4*, 1133−1139.
38. Eaton, W. A., Munoz, V., Thompson, P. A., Chan, C.-K., and Hofrichter, J. (1997) *Curr. Opin. Struct. Biol. 7*, 10−14.
39. Chan, C.-K., Hu, Y., Takahashi, S., Rousseau, D. L., Eaton, W. A., and Hofrichter, J. (1997) *Proc. Natl. Acad. Sci. U.S.A. 94*, 1779−1784.
40. Williams, S., Causgrove, T. P., Gilmanshin, R., Fang, K. S., Callender, R. H., Woodruff, W. H., and Dyer, R. B. (1996) *Biochemistry 35*, 691−697.
41. Ballew, R. M., Sabelko, J., and Gruebele, M. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 5759−5764.
42. Jones, C. M., Henry, E. R., Hu, Y., Chan, C.-K., Luck, S. D., Bhuyan, A., Roder, H., Hofrichter, J., and Eaton, W. A. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 11860−11864.
43. Krageland, B. B., and Poulsen, F. M. (1996) *J. Mol. Biol. 256*, 187−200.
44. Schindler, T., Herrler, M., Marahiel, M. A., and Schmid, F. X. (1995) *Nat. Struct. Biol. 2*, 663−673.
45. Viguera, A. R., Villegas, V., Aviles, F. X., and Serrano, L. (1997) *Folding Des. 2*, 23−33.
46. Briggs, M. S., and Roder, H. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 2017−2021.
47. Khorasanizadeh, S., Peters, I. D., Butt, T. R., and Roder, H. (1993) *Biochemistry 32*, 7054−7063.
48. Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946−950.

BI972352X