Scott, R. W., Bergman, B. L., Bajpai, A., Hersh, R. T., Rodriguez, H., Jones, B. N., Barreda, C., Watts, S., & Baker, J. B. (1985) *J. Biol. Chem. 260*, 7029-7034.
Shelburne, F. A., & Quarkfordt, S. H. (1977) *J. Clin. Invest. 60*, 944-950.
Shing, Y., Folkman, J., Sullivan, R., Butterfield, C., Murray, J., & Klagsbrun, M. (1984) *Science 223*, 1296-1299.
Sprengers, E. D., & Kluft, C. (1987) *Blood 69*, 381-387.
Stathakis, N. E., & Mosesson, M. W. (1977) *J. Clin. Invest. 60*, 855-873.
Stein, P. L., van Zonneveld, A. J., Pannekoek, H., & Strickland, S. (1989) *J. Biol. Chem. 264*, 15441-15445.
Suzuki, K., Nishioka, J., Kusumoto, H., & Hashimoto, S. (1984) *J. Biochem. 95*, 187-195.
Thorsen, S., Philips, M., Selmer, J., Lecander, I., & Astedt, B. (1988) *Eur. J. Biochem. 175*, 33-39.
Tollefsen, D. M., & Blank, M. K. (1981) *J. Clin. Invest. 68*, 589-596.
Tollefsen, D. M., Majerus, D. W., & Blank, M. K. (1982) *J. Biol. Chem. 257*, 2162-2169.
Travis, J., & Salvesen, G. (1983) *Annu. Rev. Biochem. 52*, 655-709.
van Mourik, J. A., Leeksma, O. C., Reinders, J. H., de Groot, P. G., & Zandbergen-Spaargearen, J. (1985) *J. Biol. Chem. 260*, 11300-11306.
Villanueva, G. B., & Allen, N. (1983) *J. Biol. Chem. 258*, 14048-14053.
Wagner, O. F., de Vries, C., Hohmann, C., Veerman, H., & Pannekoek, H. (1989) *J. Clin. Invest. 84*, 647-655.
Willems, C., Astaldi, G. C. B., de Groot, P. G., Janssen, M. C., Gonsalves, M. D., Zeijlemaker, W. P., van Mourik, J. A., & van Aken, W. G. (1982) *Exp. Cell Res. 139*, 191-197.
Wiman, B., Almquist, A., Sigurdardottir, O., & Lindahl, T. (1988) *FEBS Lett. 242*, 125-128.

Detection and Characterization of Intermediates in the Folding of Large Proteins by the Use of Genetically Inserted Tryptophan Probes†

Corinne J. Smith,*,‡ Anthony R. Clarke,‡ William N. Chia,‡ Laurie I. Irons,§ Tony Atkinson,§ and J. John Holbrook‡

Molecular Recognition Centre and Department of Biochemistry, University of Bristol School of Medical Sciences, University Walk, Bristol BS8 1TD, U.K., and Division of Biologics, Public Health Laboratory Service Centre for Applied Microbiology and Research, Salisbury, Wiltshire SP4 0JG, U.K.

Received June 18, 1990; Revised Manuscript Received October 8, 1990

ABSTRACT: L-Lactate dehydrogenase from *Bacillus stearothermophilus* was rebuilt by using site-directed mutagenesis to produce an enzymically active, tryptophan-less enzyme by replacing all the wild-type tryptophans (80, 150, and 203) by tyrosines. Nine single tryptophan-containing active enzymes were constructed from this enzyme by genetically replacing one of the tyrosines 36, 85, 147, 190, 203, 237, 248, 279, or 285 by tryptophan. The equilibrium and the time-resolved tryptophan fluorescence intensity and anisotropy were used to report unfolding events in guanidine hydrochloride (GHCl) monitored from these nine defined positions. Three structural transitions, half complete at 0.55, 1.7, and 2.8 M GHCl, were identified and defined four folding intermediates, I (native), II (expanded monomer 1), III (expanded monomer 2), and IV (random coil), stable at 0, 1, 2.2, and 4 M GHCl, respectively. Intermediate II is a globular monomer. All the probed α-helices and most of the β-structure was intact. There was an increase in the rate but not the extent of the mobilities of six of the probed tryptophan side chains, indicating loss of tertiary structure. Circular dichroism (CD) showed all the secondary structure to be intact. Intermediate III is monomeric and still globular, but the tryptophan anisotropy indicated an increase mobility at positions 36, 85, 190, 203, 279, and 285. Helix α-B is further disrupted but helices α-1F, α-2G, and α-3G were still rigid. CD showed half the secondary structure to be still intact. Intermediate IV is a random coil in which all tryptophans have complete rotational freedom and the helix CD signal is lost. Intermediates II and III both have characteristics ascribed to a molten globule state. The time-resolved fluorescence anisotropy of a tryptophan on a helix in folding intermediates has a correlation time of 2-3 ns and a residual anisotropy of 0.08. Probes close together in the primary sequence (279 and 285) reported identical unfolding events. The nine probes suggest partial structures for the two monomeric intermediates in folding of *B. stearothermophilus* lactate dehydrogenase, which can be compared to two kinetic intermediates proposed by Jaenicke [(1987) *Prog. Biophys. Mol. Biol. 49*, 117-237].

There is increasing evidence [Ptitsyn, 1981; cf. Ptitsyn (1987) and Kuwajima (1989)] that proteins fold to their active conformation via a series of distinct intermediate structures in a defined pathway. A study of the process of folding requires characterization of the structures of intermediates at different stages along the pathway. A well-established method of trapping intermediates (Creighton, 1978) uses the formation of disulfide bonds between cysteine residues as proteins containing such bonds refold. This method has elucidated features of the folding process in bovine pancreatic trypsin inhibitor and the structure of one intermediate species has been characterized by NMR spectroscopy. Another approach makes use of the exchange reaction between the backbone amide

† This work was supported by the Science and Engineering Research Council, U.K., Porton Industries Ltd., and NATO (travel).
* Author to whom correspondence should be addressed.
‡ Molecular Recognition Centre, Bristol.
§ Public Health Laboratory Service, Salisbury.

0006-2960/91/0430-1028$02.50/0

protons of a folding protein and the solvent (Udgaonkar & Baldwin, 1988; Baum et al., 1989). In the case of ribonuclease A, this approach resulted in the isolation of an independently folded protein fragment, which was shown, by NMR, to have helical character in solution.

Despite these successes, the methods have limitations: only a subset of proteins contain disulfide linkages and proteins of more than 100 residues cannot be easily used in NMR studies. Thus there remains a need for a method of defining folding intermediates of larger proteins where the presence of a number of well-positioned disulfide linkages is not essential. We propose here a technique, applicable to any cloned and expressed protein, that combines site-directed mutagenesis and measurement of intrinsic tryptophan fluorescence. One minor limitation is that the protein should not contain tryptophan residues important for folding or function. To evaluate the advantages and limitations of the approach, the folding of a large, multisubunit protein, lactate dehydrogenase, was investigated. Three stable intermediates were identified and their structures partially established from the equilibrium and time-resolved tryptophan fluorescence of the protein. Structural information was obtained from fluorescence signals from nine different lactate dehydrogenase constructions, each of which contained only a single tryptophan at a known position in the sequence of the enzyme. The nine probe sites were widely distributed throughout the protein tertiary structure and thus the signal changes during unfolding could be unambiguously assigned to defined positions in the original three-dimensional structure.

The kinetics of refolding and reassociation of unfolded lactate dehydrogenases have been elucidated by Jaenicke [for a review, see Jaenicke (1987)] and are described by a five-state kinetic model:

$$4M \xrightarrow[\text{fast}]{} 4M' \xrightarrow[k_1]{} 4M'' \xrightarrow[\text{fast}]{} 2D \xrightarrow[k_2]{} T$$

(where M is a random chain, M′ and M″ are two monomeric folding intermediates, D = dimer, and T = tetramer). k_1 is 1.45×10^3 s^{-1} and k_2 is 0.5×10^{-4} M^{-1} s^{-1} for pig muscle LDH renatured from 6 M GHCl.[1] Our results for equilibrium folding are also consistent with two monomeric folding intermediates existing prior to the dimerization step between 0 and 1 M GHCl. The *Bacillus stearothermophilus* lactate dehydrogenase is stable and enzymically active as a dimer. For the bacterial enzyme the dimer to tetramer step is already fully characterized in enzyme directly isolated from *B. stearothermophilus* cells [see Clarke et al. (1989) for a brief review].

Materials and Methods

Genetic Manipulation Procedures. In general these used methods described by Maniatis et al. (1982). Restriction endonucleases, DNA ligases, and other DNA-modifying enzymes were obtained from Boehringer Corporation Ltd. and used in the buffers and under the conditions recommended by the supplier. Oligonucleotide site-directed mutagenesis was performed essentially as described by Carter et al. (1985). Oligodeoxynucleotides for site-directed mutagenesis and DNA sequencing were synthesized by using an Applied Biosystems 380B or a Du Pont Coder 300 DNA synthesizer. The LDH mutants used here were generated from a tryptophanless gene in which the three wild-type tryptophan codons (80, 150, 203) were changed to tyrosine. The template is bacteriophage M13mp8 (Messing & Vieira, 1982) with the *B. stearothermophilus* lactate dehydrogenase gene inserted between the *Eco*RI and *Pst*I sites, which was purified from *Escherichia coli* TG1 (*lac-pro*, *sup*E, *thi*, *hsd*D5, [F′*tra*D36, *pro*AB, *lac*I^Q, *lac*Z M15]). *E. coli* BMH71-18*mut*L (*lac-pro*, *sup*E, *thi*, *hsd*D5, [F′*pro*AB, *lac*I^Q, *lac*Z M15], *mut*L::Tn10) was used as a host for site-directed mutagenesis. *E. coli* TG1 was used as a host for transformation and expression of LDH genes in plasmid pKK223-3 (Brosius & Holy, 1984; Pharmacia P-L, Uppsala, Sweden) as described for pLDH41 (Barstow et al., 1986). At nine points throughout the protein (Figure 2) a tyrosine was changed to tryptophan to give nine mutant LDHs, each containing only a single tryptophan residue (Waldman et al., 1987a, 1988; Clarke et al., 1987; Atkinson et al., 1987). Thus in this paper a mutant described as 85W means a protein with a wild-type amino acid sequence except with tyrosines at positions 80, 150, and 203 and with residue 85 (originally a tyrosine) replaced by a tryptophan. To verify whether the mutant sequences were correct, nucleotide sequence determinations were made by the chain-termination procedure of Sanger et al. (1981), in part using a Du Pont Genesis 2000 automated sequencer (data not shown).

Purification of the LDH Mutants from E. coli. Cells were grown in 2xYT broth containing ampicillin, were harvested by centrifugation, and were resuspended in 50 mM TEA/NaOH buffer pH 6 and lysed by sonication. The LDH mutants were precipitated by adding 0.4 g of $(NH_4)_2SO_4$/mL of supernatant and recovered by centrifugation. The pellet was dissolved in a minimum volume of TEA buffer containing 0.15 M NaCl and dialyzed against 50 volumes of that buffer overnight at 4 °C. The dehydrogenases were purified by affinity chromatography on an 8 × 6 cm column of oxamate–agarose in 0.2 mM NADH (Clarke et al., 1985). The enzymes were eluted with 50 mM TEA adjusted to pH 9.3 containing 0.5 M NaCl (without NADH). The enzymes were finally purified by ion-exchange chromatography on Q-Sepharose fast flow column and eluted with a 0–1 M NaCl gradient and were stored as the precipitates formed on adding 0.4 g of $(NH_4)_2SO_4$/mL of eluate. The samples gave a single band of apparent M_r 33 000 on sodium dodecyl sulfate–polyacrylamide gel electrophoresis calibrated with pig LDH. The methods yielded 150 mg of LDH per liter of *E. coli* culture broth.

Steady-State Kinetics. Steady-state measurements were made by following the change in absorbance at 340 nm in the NADH/NAD$^+$ conversion. All assays were at 25 °C and in 100 mM TEA buffer, pH 6. Determinations of k_{cat} and K_M for pyruvate were made at saturating NADH concentrations (0.2 mM). We define k_{cat} as the rate of turnover of the enzyme at saturating coenzyme and substrate concentrations. The enzyme concentration used was 6 nM subunits. Temperature stability measurements were carried out by incubating 5 μM LDH for 20 min at each temperature from 50 to 90 °C. The samples were then rapidly cooled and assayed under the conditions described above.

Unfolding Conditions. The protein was unfolded by incubation with 0–4.5 M GHCl (Bethesda Research Laboratories, enzyme grade) in the presence of 5 mM β-mercaptoethanol in 50 mM triethanolamine hydrochloride buffer at pH 7 for at least 4 h at room temperature. The long time required for equilibrium meant that new samples had to be prepared for each measurement and it was difficult to obtain fluorescence intensities to better than ±4% or ±0.05 M GHCl.

Tyrosine Absorbance. The tyrosine absorbance of each mutant and wild-type LDH was measured at 287 nm for a range of GHCl concentrations and under the unfolding con-

[1] Abbreviations: LDH, L-lactate dehydrogenase; GHCl, guanidine hydrochloride; TEA, triethanolamine hydrochloride; CD, circular dichroism.

Table I: Time-Resolved Anisotropy (τ_c and r_∞) and Catalytic Constants for Each Single Tryptophan Mutant[a]

	[GHCl] (M)	τ_c (ns)	r_∞	χ^2	local structure	k_{cat} (s^{-1})	K_M (mM pyruvate)
36W	0	7.3	0.106	1.14	helix α-B	181	6
	1	2.5	0.082	1.08			
	2.2	2.2	0.060	1.05	GFVGASWVFALMN		
	4	3.1	0.022	1.00			
85W	0	4.7	0.099	1.06	turn	177	6
	1	1.9	0.094	1.16			
	2.2	2.1	0.070	0.98	DIYHGDWDDCRDA		
	4	2.2	0.017	1.06			
147W	0	2.1	0.093	1.04	helix α-1F	201	6
	1	3.3	0.088	1.10			
	2.2	3.1	0.092	1.02	PVDILTWATYKFS		
	4	2.3	0.023	1.06			
190W	0	8.1	0.068	1.09	β-sheet	244	6
	1	2.2	0.069	1.23			
	2.2	2.8	0.049	1.46	PQNVHAWIIGEHG		
	4	3.2	0.002	1.39			
203W	0	9.5	0.086	1.02	turn	210	10
	1	2.0	0.102	1.34			
	2.2	2.8	0.065	1.23	DTELPVWSQAYIG		
	4	2.5	0.015	1.25			
237W	0	3.7	0.065	1.18	helix α-2G	134	4
	1	2.7	0.074	1.20			
	2.2	2.7	0.079	1.24	NVRDAAWQIIEKK		
	4	2.1	0.021	1.25			
248W	0	2.1	0.075	1.10	helix α-3G	175	6
	1	2.1	0.071	1.02			
	2.2	2.4	0.074	1.16	KKGATYWGIAMGL		
	4	2.3	0.012	0.97			
279W	0	8.7	0.084	1.10	β-sheet	182	4
	1	3.1	0.080	1.03			
	2.2	4.2	0.064	1.25	AYLDGLWGERDVY		
	4	2.7	0.010	1.08			
285W	0	5.1	0.086	1.14	turn	105	3
	1	2.4	0.082	1.21			
	2.2	2.2	0.064	1.06	YGERDVWIGVPAV		
	4	2.9	0.011	1.06			

[a] For wild type k_{cat} is 243 s^{-1} and K_M is 6 mM pyruvate.

ditions described above. Ten micromolar subunits was required to provide an adequate absorbance signal.

Fluorescence Measurements. Equilibrium fluorescence measurements were carried out on a SLM 8000 single photon counting spectrofluorometer by using 2 μM subunits. The same instrument was used to measure the Stern–Volmer quenching constants (Weller, 1961) by sodium iodide on chosen unfolded states for each mutant. In these experiments aliquots of 5 M NaI containing 1 mM $Na_2S_2O_3$ (to prevent formation of I_3^-) were added to the unfolded protein and the tryptophan fluorescence was measured (excitation 297 nm, emission 345 nm). Time-resolved fluorescence anisotropy measurements were carried out on an Edinburgh Instruments Model 199S fluorescence decay time spectrometer by using the time-correlated single photon counting method with simultaneous acquisition of fluorescence and of the excitation lamp pulse (SAFE option) according to Birch and Imhof (1990). The flash lamp gas used was hydrogen and the full width at half-maximum of the pulse was 1.4 ns (excitation at 295 nm, emission selected by a Melles Griot 320-nm cuton filter WG320). These last two studies used 3 μM subunits. For time-resolved anisotropy the analyzing polarizer was rotated between horizontal and vertical every 30 s. The data analysis package provided by IBH Consultants, Glasgow, Scotland, was used to fit the anisotropy decays to a single exponential decay to a limiting value. No attempt was made to measure the long-time rate of anisotropy decay (Brownian motion >10 ns) as the lifetimes of most of these fluorophores (≤5 ns) were too short for a sensible estimate. The lifetimes are in Table I of the supplementary material but were predominantly a mixture of 3- and 8-ns exponentials.

Gel Filtration. The apparent molecular weight of unfolded species of lactate dehydrogenase was estimated by using a Bio-Rad Bio-Sil TSK-250 HPLC gel filtration column equilibrated with GHCl as described above with an initial protein concentration of 10 μM. The steel jacketed column was operated with mechanical injection within a fully automated Pharmacia FPLC apparatus, which allowed the elution volumes to be repeated to ±0.05 mL. The eluted protein peak was approximately 2.5 μM. The column was calibrated under the same conditions as the 0 M GHCl point by using Bio-Rad gel filtration molecular weight standard number 151–1901 containing thyroglobulin (v_e = 5.75 mL), IgG (7.70 mL), ovalbumin (8.95 mL), myoglobin (10.55 mL), and cyanocobalamin (12.20 mL).

Circular Dichroism. The circular dichroism spectra (205–300 nm) of unfolded protein at a series of concentrations of GHCl were measured with a JASCO-J-600 spectropolarimeter at a protein concentration of 0.9 $mg \cdot mL^{-1}$ in 0.2-mm path length cells and analyzed for helical content from the mean residue elipticity at 222 nm (Chen et al., 1972). The strong absorbance of GHCl below 205 nm prevented quantitative estimates of β-sheet content.

Results

Effect of Substitutions on Protein Stability and Activity. The decision to insert the tryptophan probes in positions normally occupied by tyrosine residues was made because both aromatics have similar hydrophobicity values (Wolfenden et al., 1979) and similar average buried volumes (Chothia, 1975). Positions were selected to have minimum destabilizing effect on the native protein. The stability of each mutant with respect

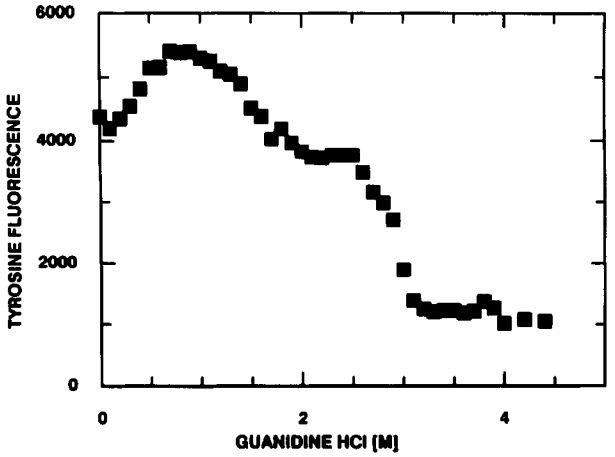


FIGURE 1: Equilibrium unfolding of tryptophan-less lactate dehydrogenase monitored by tyrosine fluorescence. Tyrosine fluorescence emission was measured at 303 nm with excitation at 278 nm. The protein concentration used was 2 μM subunits. The graph shows three transitions with midpoints at 0.5, 1.7, and 2.9 M GHCl.

to temperature was determined and the temperature at which half-activity was lost for all mutants was within 7 °C of the wild-type value. For each of the nine mutants prepared, the kinetic parameters k_{cat} and K_M were measured and compared to those of the wild-type enzyme. As shown in Table I the values were only slightly altered, indicating only small perturbations of the active site arose due to the substitutions. Some constructions were made to study conformation changes during the catalytic cycle of the enzyme (Waldman et al., 1987b), and the slightly altered enzymic properties will be treated in full elsewhere.

Each single tryptophan mutant contains 14 tyrosine residues widely dispersed throughout the protein. The tyrosine absorbance at 287 nm decreases sharply as the protein unfolds in GHCl and thus reports the unfolding seen from the average of many sites, rather than the specific site reported by each tryptophan. Supplementary material Figure 1 shows the change in tyrosine absorbance at 287 nm with respect to GHCl concentration for representative mutant proteins. Within experimental error all the mutants and wild-type protein showed the same change in absorbance as a function of GHCl concentration, suggesting that the substitution of tryptophan for tyrosine does not significantly perturb the unfolding pathway. This control is needed since Kellis et al. (1988) have shown that even conservative mutations such as isoleucine to valine at sensitive hydrophobic sheet/helix contacts can alter the midpoints of urea-unfolding transitions by 0.5 M urea in small globular proteins.

Gel Filtration of Partially Unfolded Lactate Dehydrogenases. The graph of elution volume versus GHCl concentration for wild-type enzyme (supplementary material Figure 2) shows that the elution volume increases from that of the dimer to that of the monomer between 0 and 1 M GHCl. Between 1 and 2.2 M GHCl there is no change in hydrodynamic volume. This implies that any changes observed by other techniques over this range must be due to rearrangements within the subunit. Over the 2.3–4 M concentration range there is a steady increase in molecular volume commensurate with the transition from a globular, compact structure to an extended random coil.

Variation of Tyrosine Fluorescence of the Tryptophanless Protein with GHCl Concentration. The original "template" protein contains 15 tyrosine residues widely distributed through the protein volume. The absence of tryptophan allows the tyrosine fluorescence to be measured and used as a sensitive

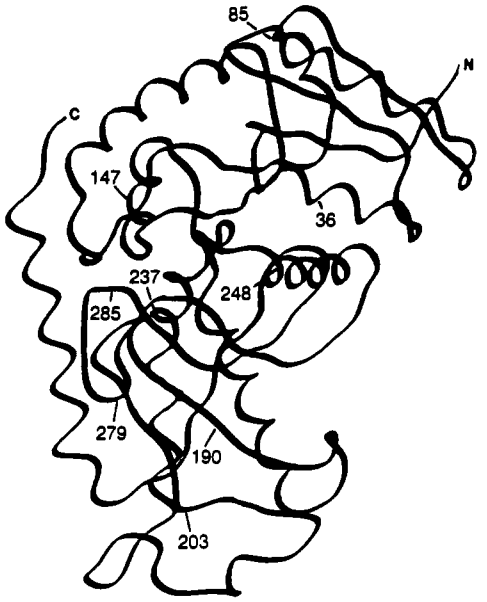


FIGURE 2: Positions of tryptophan probes in the lactate dehydrogenase subunit. The figure shows a diagram of a subunit of *B. stearothermophilus* LDH made by using the INSIGHT program (Biosym Corp.) installed on a Silicon Graphics Personal Iris workstation. The backbone of the protein is shown as a ribbon, and the positions of each single change of tyrosine to tryptophan are indicated by the residue number. The coordinates are from a partly refined structure of this protein crystallized from PEG-6000 with an R factor = 21% for 62 506 independent reflections in the 30–2.5-Å resolution range (D. B. Wigley, H. Muirhead, and J. J. Holbrook, personal communication).

means of establishing the general transitions undergone by LDH as it unfolds. Figure 1 shows the variation of the tyrosine fluorescence (excitation 278 nm, emission 303 nm) of this template enzyme with GHCl concentration: three main transitions are observed with midpoints at 0.5, 1.7, and 2.9 M GHCl. Three plateau regions are observed demonstrating that stable folding intermediates exist at 1.0, 2.2, and 4 M GHCl. To define these changes further, the series of single tryptophan mutants was used to identify which points in the LDH structure report a structural perturbation at each step.

Description of the Positions of the Tryptophan Probes in the Native Structure. Figure 2 shows the position of each single tryptophan in a lactate dehydrogenase subunit. 36W, 147W, 237W, and 248W are on helices (α-B, α-1F, α-2G, and α-3G, respectively) while 190W is on a β-sheet. 279W is on a β-sheet and 285W an extended chain each side of a surface loop. 85W is on a random chain close in space to 36W. 203W is on the surface in the dimer but in the tetramer is buried on the dimer–dimer P-axis interface. The *B. stearothermophilus* LDH polypeptide is numbered from 15 to correspond with conventional residue numbers of Eventoff et al. (1977) and the sequence immediately surrounding each site is given in Table I. The secondary structure is named according to Adams et al. (1972).

Variation of Tryptophan Fluorescence of Lactate Dehydrogenase Mutants with GHCl Concentration. Equilibrium unfolding curves of the nine single tryptophan containing proteins are shown in Figure 3. Unlike a majority of equilibrium denaturation curves that have been published for small monomeric proteins (Nojima et al., 1977; Ginsburg & Carrol, 1965; Anfinsen, 1972), the curves reveal several distinct transitions over the concentration range of the denaturing agent. The fluorescence intensities show changes in the local environment of a tryptophan residue at the positions 36, 85, 147, 190, 203, 237, 248, 279, and 285 in the protein sequence.

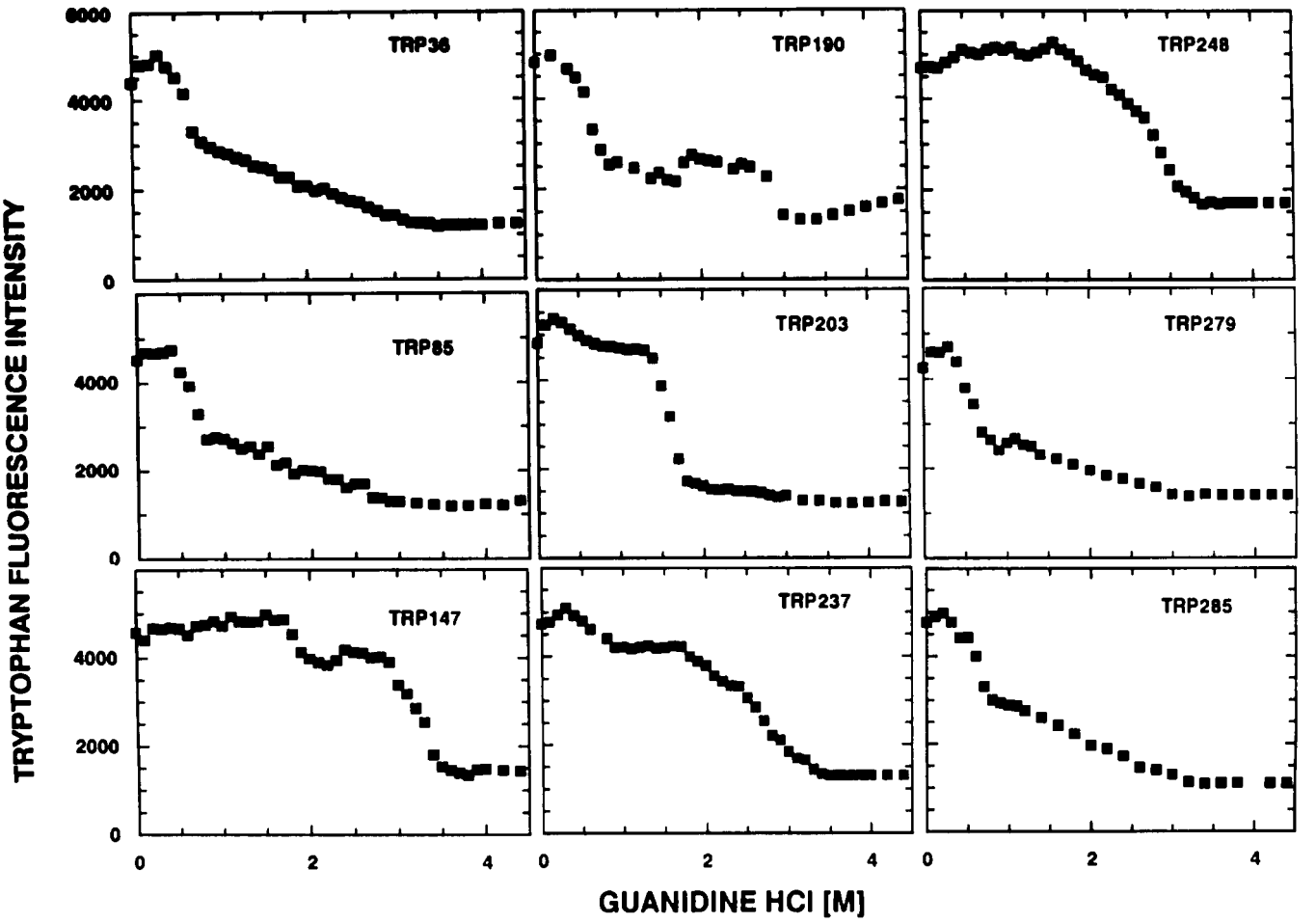


FIGURE 3: Equilibrium unfolding of single tryptophan lactate dehydrogenase mutants monitored by tryptophan fluorescence intensity. Tryptophan fluorescence was excited at 295 nm and measured at 345 nm. Fluorescence intensity was measured at a series of GHCl concentrations at a protein concentration of 2 μM subunits after equilibration for >4 h. The results show that several distinct changes occur during unfolding in different regions of the protein.

Some changes are seen by most of the probes measured; for example, the transition at 0.5 M GHCl is monitored by probes at positions 36, 85, 190, 203, 279, and 285. The transition at 1.7 M GHCl, on the other hand, is seen only by probes at positions 147, 190, and 203 (and by the tyrosine fluorescence of the tryptophan-less mutant), where the magnitude of the change at position 203 is by far the greatest of the three. Probes at 147, 237, and 248 monitor a transition at 2.8 M GHCl, each probe losing a similar proportion of fluorescence during this change. Like the tyrosine fluorescence the nine probes identify three distinct transitions during the unfolding of the protein. The possibility that a transition may be the sum of several events is not, however, ruled out (for example, the 2.8 M transition of Figure 1 might be a shallower transition at 2.6 M followed by a sharper one at 2.95 M GHCl). The first occurs between 0 and 1 M GHCl, the second between 1 and 2.2 M GHCl, and the third between 2.2 and 4 M GHCl. The concentrations marking the divisions between transitions have been taken as being those at which a stable folding intermediate exists at equilibrium. The curves also show a reproducible transition at 0.1 M GHCl. This is possibly due to the disruption of the small portion of tetramer (Clarke et al., 1985a) present at the 2–10 μM (subunit) protein concentration used in these experiments. The enzyme activity is undetectable in assays at above 0.2 M GHCl, but since we know the dimer is fully active in the absence of GHCl the activity loss induced by 0.2 M GHCl and the tetramer to dimer dissociation must be different events. The experiments in the following sections attempt characterization of the intermediates at 0, 1, 2.2, and 4 M GHCl.

Characterization of Stable Folding Intermediates. The intermediates were studied using a variety of techniques. Time-resolved fluorescence anisotropy provides information on both the speed of motion of the fluorophore and the amplitude of angular excursions (and thus the flexibility of the chain on which it rests). Stern–Volmer quenching studies show the solvent accessibility of each probe, while circular dichroism gives an estimation of the proportion of α-helix present in each intermediate.

Time-Resolved Fluorescence Anisotropy Measurements. Figure 4 shows the results of time-resolved fluorescence anisotropy experiments carried out on each mutant. A single exponential was fitted to the original data (an example of which is given in the supplementary material, Figure 3). This fit allows an estimation of the r_∞ to be made, which in turn is a measure of the local flexibility of the area surrounding the tryptophan probe. We define r_∞ as the anisotropy remaining after the initial fast phase of decay (at about 10–15 ns). The second phase of slow loss of anisotropy due to the tumbling of the molecule (up to 80 ns) was not fitted but was approximated by a straight line because the tryptophans had lost most of their fluorescence intensity by 20 ns. Figure 5 provides an explanation of how r_∞ relates to the flexibility of a side chain and the protein backbone. The correlation time of the anisotropy decay provides a measure of the speed of motion of the fluorophore. The degree of change of anisotropy (r_0–r_∞) reflects the conformational space available to the probe. Rather unexpectedly [from the model of Kinosita et al. (1977)]

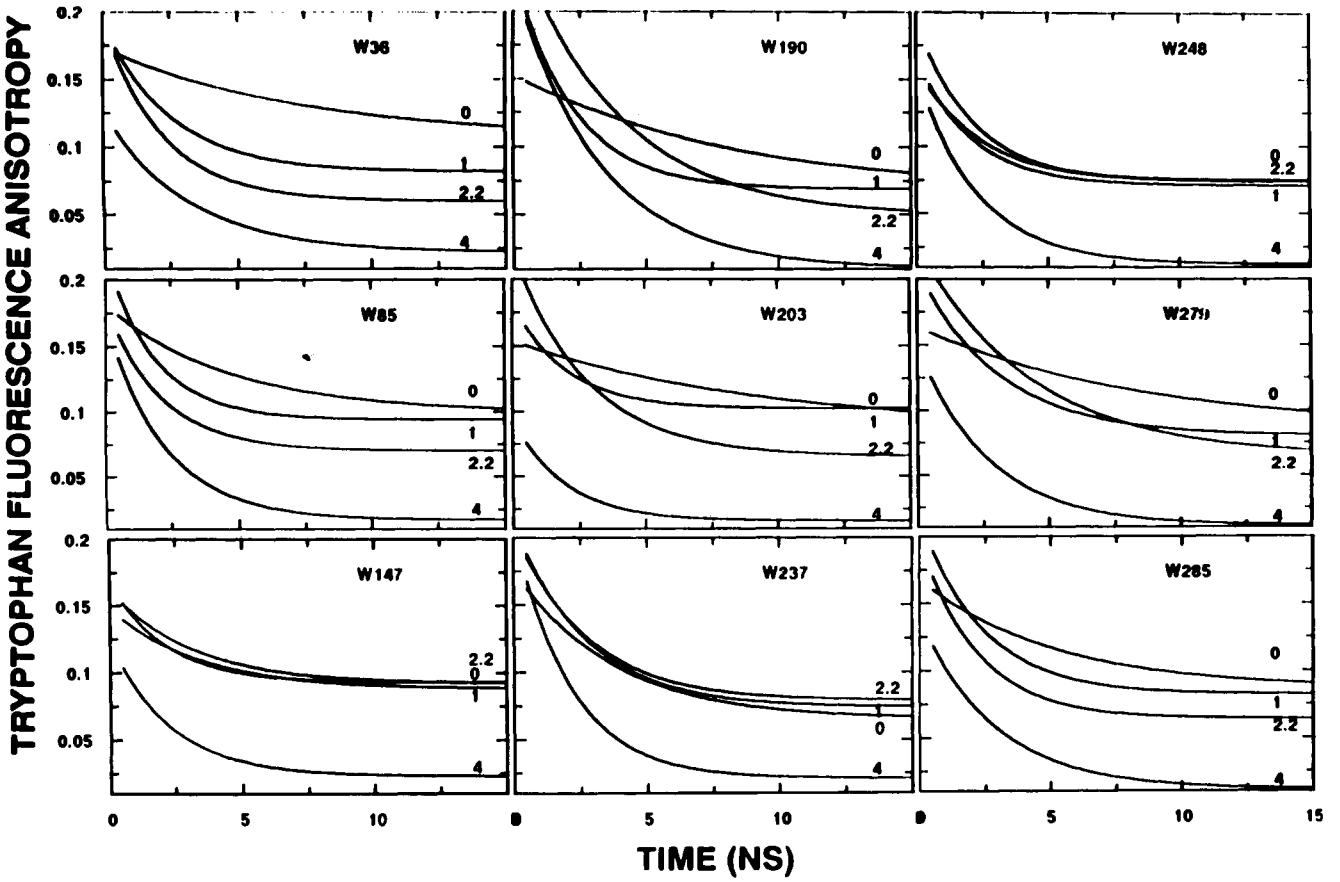
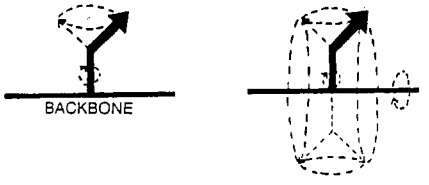


FIGURE 4: Time-resolved fluorescence anisotropy decays for nine single tryptophan mutants of lactate dehydrogenase. Each curve represents the decay in anisotropy with respect to time at 0, 1, 2.2, and 4 M GHCl. The data were fitted with a single correlation time in each case. A second, longer correlation time was apparent in some cases due to the tumbling of the molecule. The fluorescence lifetime of tryptophan is too short, however, to enable data to be collected over a time period sufficiently long for this second lifetime to be fitted. Tryptophan fluorescence was excited at 295 nm and measured at >320 nm. The protein concentration was 3 μM subunits.



A: r_C = 2-3 ns r_∞ = 0.06-0.09 B: r_C = 2-3 ns r_∞ = 0.00-0.02

FIGURE 5: Interpretation of extent of loss of anisotropy. A represents a side chain on its protein backbone in a partially denatured state. The side chain is free to rotate under these conditions, while the protein backbone remains rigid. In this situation anisotropy can only be partially lost before Brownian motion of the whole molecule is apparent. B represents a side chain and backbone for a protein in a totally denatured state. Under these conditions both side chain and backbone are free to rotate. In this mobile, random-coil state all anisotropy would be lost rapidly to give a value for r_∞ of zero.

the correlation times of all the tryptophans in the partially and fully unfolded proteins were in the range 2–4 ns and did not vary in any simple pattern with increasing unfolding (Table I). Only in the native state are these side chains appreciably immobilized with correlation times in the 5–10-ns region. Most strikingly, mutants 147W, 237W, and 248W, in which helices are labeled, show no increase in flexibility or speed of motion of the fluorophore until 4 M GHCl, where the anisotropy decays to zero very rapidly (with a correlation time of about 2.5 ns). This is in contrast to the mutant 36W labeled at the remaining helix α-B, which shows a steady, persistent gain in flexibility over the concentration range of GHCl. Hence three out of the four helices probed retain rigid nativelike characteristics in 2.2 M GHCl. This suggests the folding intermediates at 1 and 2.2 M GHCl have considerable helical content. The intermediate at 1 M GHCl is mainly distinguished from the native structure by an increase in rate of loss of anisotropy rather than a decrease in r_∞. The change is most marked at positions 36, 85, 190, 203, 279, and 285, which show restriction in the fast motion of the probe in the native state. Hence the speed of movement of the tryptophans is increased in all cases except for the very stable helices (147W, 237W, and 248W). We interpret this faster relaxation as a loss of steric restraint of the side-chain motion caused by an overall loosening of the tertiary structure of a molecule in which the secondary structure is essentially intact.

At 2.2 M GHCl the final anisotropy of all the mutants is greater than zero at 15 ns, well before decay due to Brownian motion begins to be seen. In most cases (that is, except for 147W, 237W, and 248W) the final anisotropy is lower than for 1 M GHCl. Hence the intermediate at this concentration is more flexible overall than that at 1 M GHCl but is considerably more rigid than the unfolded form at 4 M GHCl, which has r_∞ close to zero in all cases.

Stern–Volmer Quenching Experiments. A plot of Stern–Volmer quenching constant (Weller, 1961) versus GHCl concentration for each mutant protein is shown in the supplementary material, Figure 4. The error in the constants is ± 0.001 mM^{-1}. A correction was applied for the possible screening effect of GHCl by first measuring the quenching constant for *N*-acetyltryptophanamide in each of the four guanidine concentrations and then scaling up the results obtained for the protein according to the proportion by which the GHCl lowered the quenching constant for *N*-acetyltryptophanamide (the lowered values were 1.18, 1.41, and 1.42 for 1, 2.2, and 4 M GHCl). The accessibility of the probes

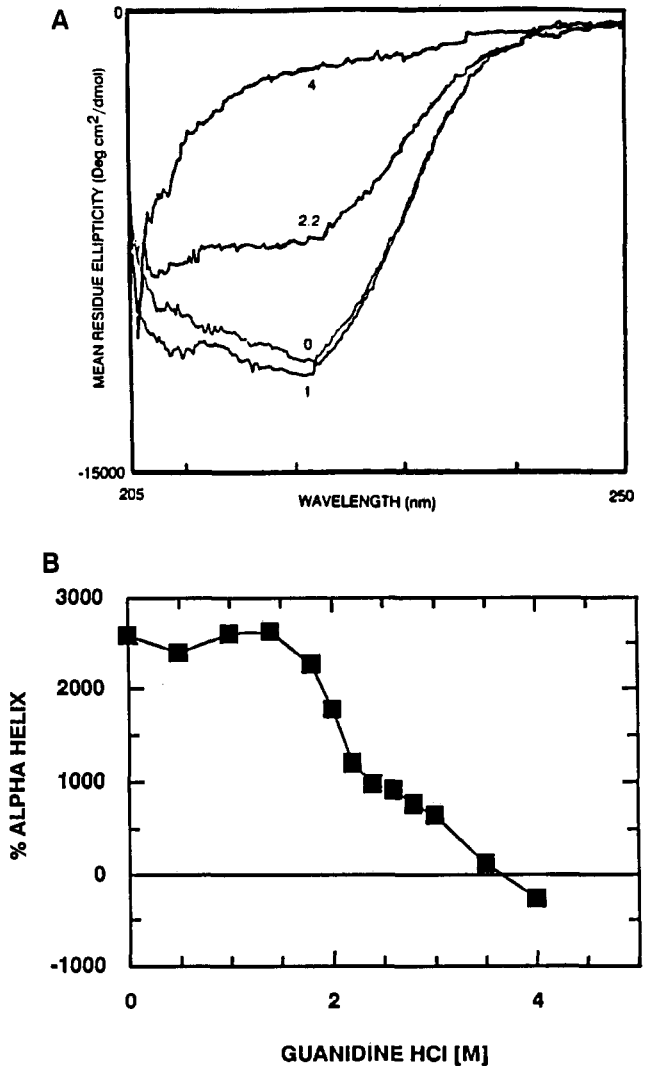


FIGURE 6: Circular dichroism spectra of wild-type LDH during unfolding in guanidine hydrochloride. (A) Protein (30 μM subunit) was incubated at 0, 1, 2.2, and 4 M GHCl and the circular dichroism spectrum scanned between 205 and 300 nm (only results below 250 nm shown). (B) The percent α-helix calculated from the mean residue ellipticity at 222 nm according to Chen et al. (1972) at a series of GHCl concentrations.

to solvent provides another means of deducing the nature of the intermediates. At 0 M GHCl most probes are inaccessible to solvent with the two exceptions of 237W and 248W, which are on surface helices. These appear to become buried at 1 M GHCl in contrast with all the other probes, which experience an increase in solvent accessibility at this concentration. At 2.2 M GHCl the overall trend is a further increase in solvent accessibility, although 147W shows no change and 203W appears to become more buried. At 4 M GHCl there is little further increase in solvent accessibility. The Stern–Volmer quenching constant might be expected to be the same for each tryptophan in 4 M GHCl; however, this is not the case. We have not been able to explain these differences by considering the local primary sequence around each probe (Table I). We thus consider the Stern–Volmer quenching constant as reporting relative, rather than absolute, accessibility to solvent.

These results provide further evidence for the existence of at least two folding intermediates with different structures between the unfolded chain and the native dimer. It is important to note that the intermediate at 2.2 M GHCl is almost as accessible to solvent as the fully unfolded protein. This accords with the property of the molten globule state referred to under Discussion. Gel filtration shows that the molecular volume changes little between 1 and 2.2 M GHCl, while the time-resolved fluorescence anisotropy data show that there is an increase in side-chain flexibility over this range. Thus the increase in solvent accessibility between 1 and 2.2 M GHCl must be due to a breakdown of structure within the protein globule and not due to the disruption of the globule itself. Hence the results define two folding intermediates, one at 1 M GHCl containing a large proportion of ordered structure within the protein globule and another at 2.2 M GHCl, which is more disordered and has lost some, but not all, ordered secondary structure.

Circular Dichroism Spectra. The circular dichroism spectra of the four unfolded states were recorded and are presented in Figure 6. Analysis of these data at 222 nm (Chen et al., 1972) shows the native state has 26% α-helix, in 1 M GHCl the protein has 25% α-helix, in 2.2 M GHCl the protein has 12% α-helix (45% of the native), and in 4 M GHCl the protein has −3% α-helix; that is, it is a random coil. Although it was not possible to quantitatively analyze for β-structure due to the absence of data at <205 nm, the curve at 210 nm was very similar for both wild type and intermediate I (1 M GHCl), suggesting that little loss of β-structure took place during this first unfolding transition. Figure 6B shows two regions where loss of helix content takes place, at 2 and 3.1 M GHCl.

Discussion

Before these experiments it was unclear whether single tryptophan probes would report events at some distance from the site of the probe or be sensitive only to their immediate environment. The latter might be expected from the known properties of the indole fluorophore. A mechanism for long-range effects on indole fluorescence is radiationless energy transfer, but simple proteins have no chromophores with a large overlap of the tryptophan emission spectrum in the 310–370-nm range and transfer from tryptophan to tyrosinate is very short range with a critical transfer distance of about 1.5 nm (Edelhoch et al., 1967). The experiments with lactate dehydrogenase indicate that these residues predominantly respond to local environment. For example, the probes W279 and W285 show similar changes in fluorescence intensity (Figure 3) and anisotropy (Figure 4) but are insensitive to the events that give large changes in other probes (for example, the change in fluorescence intensity of W203 at 1.7 M GHCl). The local environment may of course be provided by amino acid residues that are remote in the primary sequence but close in the tertiary structure. This is seen by the decrease in correlation time on going from native to the first unfolded structure by six of the nine tryptophan side chains (Table I). Thus a series of genetically inserted fluorescent probes can yield information on the local folding events at specific probe sites in a large oligomeric protein.

In addition to the random unfolded coil, two monomeric intermediates were studied, one stable at 1 M GHCl and the other at 2.2 M GHCl. The first is a monomer and shows little perturbation of backbone structure over the native protein although the speed of motion of the fluorophore side chains is increased, implying a relaxation of the tertiary structure contacts, which, in the native state, provide steric hindrance to the rotation of some tryptophan residues. Most regions probed are more open to solvent (as the Stern–Volmer quenching constants showed) than in the native state and the protein appears to have lost very little secondary structure (as is seen from time-resolved fluorescence anisotropy and circular dichroism). Hence it can be concluded that initial unfolding

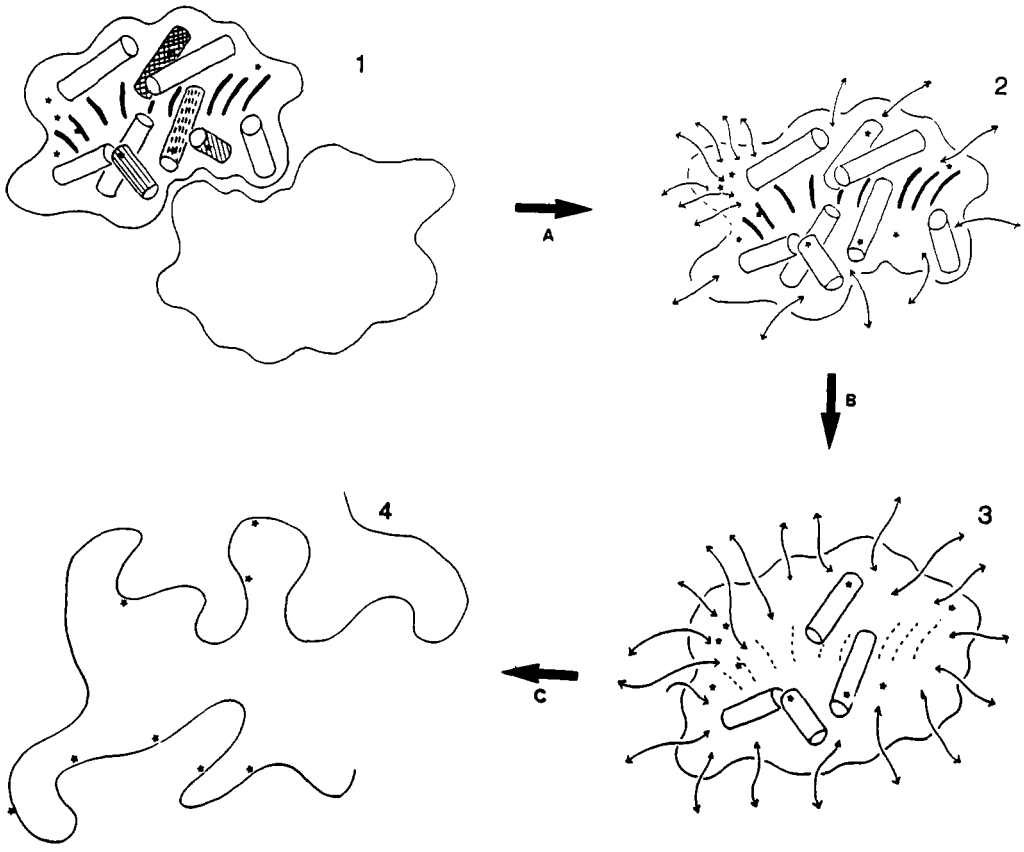


FIGURE 7: Diagram illustrating the structures formed during the unfolding of *B. stearothermophilus* lactate dehydrogenase by guanidine hydrochloride. The views are with the *P*-axis horizontal, the *Q*-axis extending from the paper, and the *R*-axis vertical. The five-sided stars represent the positions of the nine single tryptophan probes. Cylinders are the helices that were investigated, α-B [(\\\), 36W], α-1F [(XXX) 147W], α-2G [(=), 237W], and α-3G [(---) 248W]. The thick black bars indicate the extended β-sheet, which stabilizes the core of the subunit. 1 is the globular dimer stable up to about 0.3 M GHCl. Transition A is at 0.55 M GHCl and leads to a monomeric globule 2, stable at 1 M GHCl, with nearly all the secondary structure intact (except for α-B at the subunit–subunit contact surface). The double-headed arrows in 2 symbolize the increased exposure of the tryptophan probes to solvent. The helices are now further separated to show the lack of restraint to tryptophan side-chain rotation of the expanded tertiary structure. Transition B at 1.7 M GHCl leads to structure 3, stable at 2.2 M GHCl, which has even greater penetration by solvent, a greater mobility of β-sheet-carrying tryptophan probes, and an overall loss of 40% of helical structure. Helices α-1F, α-2G, and α-3G, which surround the active center, are still intact and indistinguishable from their native form. Transition C at 2.8 M GHCl is the final collapse to a random coil, stable at 4 M GHCl.

events cause an increase in the solvation of the molecule while leaving the secondary structure essentially intact. This structural transition at 0.7 M GHCl involves not only the increase in solvation of the molecule but also the dissociation of the dimer to the monomer, demonstrated by gel filtration. Our present measurements cannot distinguish whether one of these events precedes the other or the change is concerted.

The second intermediate at 2.2 M GHCl has approximately the same hydrodynamic volume as the first but contains only about 45% of the native α-helical structure. It is much more flexible than the first intermediate and has greater solvent accessibility (as witnessed by the overall increase in the Stern–Volmer quenching constants). At this concentration of GHCl, the probed secondary structures of the protein have only been partially lost. For example, the limiting anisotropy of all sites except the stable helices is decreased from its value in 1 M GHCl but still exhibits significant angular restriction compared to the random coil. This second intermediate has increased flexibility and solvent accessibility, while still retaining the properties of a compact protein globule with a similar hydrodynamic volume to the first intermediate. Unlike helices α-1F, α-2G, and α-3G, which were as stable as in the native state, helix α-B was destabilized. The maintenance of restricted angular motion in the probes 147W, 237W and 248W on the three stable helices is seen to correlate with the otherwise uninterpreted fluorescence intensities of the tryptophan probes: six of the tryptophan probes have the low fluorescence intensity of the unfolded state at 2 M GHCl. Only 147W, 237W, and 248W show a sharp loss of fluorescence intensity in the 2.9 M GHCl transition as these helices finally unwind. Interestingly, the three most stable helices probed (1F, 2G, and 3G) provide most of the stable secondary structure that in the native state surrounds the lactate dehydrogenase catalytic vacuole.

The final unfolded state is essentially a random chain. The α-helices, as measured by circular dichroism, are completely lost. This state has the greatest solvent accessibility of any intermediate. The chain is now so flexible that all the tryptophan probes have sufficient angular freedom to loose all their anisotropy within a few nanoseconds. The hydrodynamic volume, estimated from gel filtration, increases greatly over the range 2.5–4 M GHCl, as would be expected in the transition from globule to random coil.

Both the monomeric intermediates (II and III) that retain their structure have measured properties that are ascribed to the molten globule state as listed in the recent review of Kuwajima (1989). In both, the (tryptophan) amino acid side chains have lost restraint to angular motion that was present in the native tertiary structure; both have increased solvent accessibility. In the first all the secondary structure is intact. In the second about half is lost. With many approaches undertaken previously, it has been impossible to decide whether the secondary structure observed in a folding intermediate is the same as that present in the native conformation. By

knowing from where in the protein a particular signal arises we can answer this question in proteins that have no disulfides and that are far too large for NMR.

We have attempted to summarize our findings in a diagram (Figure 7), and this reveals that even with nine tryptophan probes we are some way from having surveyed even all the α-helices. It should, however, be possible to insert probes on the remaining helices and eventually have a complete description of the tertiary structures of folding intermediates.

ACKNOWLEDGMENTS

We thank D. R. Hartwell, K. Fantom, and D. Emery for assistance with the oligonucleotide synthesis and DNA sequencing and A. D. B. Waldman for preliminary experiments.

SUPPLEMENTARY MATERIAL AVAILABLE

Table of the multiple exponential fits to the time-resolved decays of the fluorescence intensities of nine single tryptophan mutants and four figures of the unfolding profile of lactate dehydrogenases reported by tyrosine absorbance, gel filtration of folded and partially folded lactate dehydrogenase, the fluorescence anisotropy decay for the single tryptophan mutant 237W, and Stern–Volmer quenching constants for the tryptophan fluorescence of each mutant (6 pages). Ordering information is given on any current masthead page.

Registry No. LDH, 9001-60-9; tryptophan, 73-22-3.

REFERENCES

Adams, M. J., Buehner, M., Chandrasekhar, K., Ford, G. C., Hackert, M. L., Liljas, A., Lentz, P. J. Jr., Rao, S. T., Rossmann, M. G., Smiley, I. E., & White, J. L. (1972) in *Protein–Protein Interactions* (Jaenicke, R., & Helmreich, E., Eds.) Springer Verlag, Berlin & New York.

Anfinsen, C. B. (1972) *Biochem. J. 128*, 737–749.

Atkinson, T., Barstow, D. A., Chia, W. N., Clarke, A. R., Hart, K. W., Waldman, A. D. B., Wigley, D. B., Wilks, H. M., & Holbrook, J. J. (1987) *Biochem. Soc. Trans. 15*, 991–993.

Barstow, D., Clarke, A. R., Chia, W. N., Wigley, D., Sharman, A. F., Atkinson, T., Minton, N. P., & Holbrook, J. J. (1986) *Gene 46*, 47–55.

Baum, J., Dobson, C. M., Evans, P. A., & Hanley, C. (1989) *Biochemistry 28*, 7–13.

Birch, D. J. S., & Imhof, R. E. (1990) in *Fluorescence Spectroscopy Vol. 1: Principles and Techniques* (Lakowitz, J. R., Ed.) Plenum Press, New York (in press).

Brosius, J., & Holy, A. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 6929–6933.

Carter, P., Bedouelle, H., & Winter, G. (1985) *Nucleic Acids Res. 13*, 4431–4443.

Chen, Y. H., Yang, J. T., & Martinez, H. M. (1972) *Biochemistry 11*, 4120–4131.

Chothia, C. (1975) *Nature 254*, 304–308.

Clarke, A. R., Atkinson, T., Campbell, J. W., & Holbrook, J. J. (1985a) *Biochim. Biophys. Acta 829*, 387–396.

Clarke, A. R., Waldman, A. D. B., Munro, I., & Holbrook, J. J. (1985b) *Biochim. Biophys. Acta 828*, 375–379.

Clarke, A. R., Wigley, D. B., Chia, W. N., Barstow, D., Atkinson, T., & Holbrook, J. J. (1986) *Nature 324*, 699–702.

Clarke, A. R., Barstow, D. A., Waldman, A. D. B., Wilks, H. M., Hart, K. W., Atkinson, T., & Holbrook, J. J. (1987) *Biochem. Soc. Trans. 15*, 1147–1148.

Clarke, A. R., Atkinson, T., & Holbrook, J. J. (1989) *Trends Biochem. Sci. 14*, 145–148.

Creighton, T. E. (1978) *Prog. Biophys. Mol. Biol. 33*, 231–297.

Edelhoch, H., Brand, L., & Wilchek, M. (1967) *Biochemistry 6*, 547–558.

Eventoff, W., Rossman, M. G., Taylor, S. S., Torff, H. J., Meyer, H., Keil, W., & Kiltz, H. H. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 2677–2681.

Ginsburg, A., & Carrol, W. R. (1965) *Biochemistry 4*, 2159–2174.

Jaenicke, R. (1987) *Prog. Biophys. Mol. Biol. 49*, 117–237.

Kellis, J. T., Jr., Wyberg, K., Sali, D., & Fersht, A. R. (1988) *Nature 333*, 784–786.

Kinosita, K., Jr., Kawato, S., & Ikegami, A. (1977) *Biophys. J. 20*, 289–305.

Kuwajima, K. (1989) *Proteins: Struct., Funct., Genet. 6*, 87–103.

Maniatis, T., Fritsch, E. F., & Sambrook, J. (1982) *Molecular Cloning—a Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Messing, J., & Vieira, J. (1982) *Gene 19*, 269–276.

Nojima, H., Ikai, A., Oshima, T., & Noda, H. (1977) *J. Mol. Biol. 116*, 429–442.

Ptitsyn, O. B. (1981) *FEBS Lett. 131*, 197–202.

Ptitsyn, O. B. (1987) *J. Protein Chem. 6*, 273–293.

Sanger, F., Coulson, A. R., Barrell, B. G., Smith, A. J. H., & Roe, B. A. (1980) *J. Mol. Biol. 143*, 161–178.

Udgaonkar, J. B., & Baldwin, R. L. (1988) *Nature 335*, 694–699.

Waldman, A. D. B., Clarke, A. R., & Holbrook, J. J. (1987a) *Protein Eng. 1*, 248–248.

Waldman, A. D. B., Barstow, D. A., Clarke, A. R., Gratton, E., Atkinson, T., & Holbrook, J. J. (1987b) *Biochem. Soc. Trans. 15*, 1158–59.

Waldman, A. D. B., Hart, K. W., Clarke, A. R., Wigley, D. B., Barstow, D. A., Atkinson, T., Chia, W. N., & Holbrook, J. J. (1988) *Biochem. Biophys. Res. Commun. 150*, 752–759.

Weller, A. (1961) *Prog. React. Kinet. 1*, 187–214.

Winter, G., Fersht, A. R., Wilkinson, A. J., Zoller, M., & Smith, M. (1982) *Nature 299*, 756–758.

Wolfenden, R. V., Cullis, P. M., & Southgate, C. C. F. (1979) *Science 206*, 576–577.