Quaternary Structure of Aldolase Leads to Differences in Its Folding and Unfolding Intermediates†

Hai Pan‡ and David L. Smith*

Department of Chemistry, University of Nebraska, Lincoln, Nebraska 68588-0304

Received December 20, 2002

ABSTRACT: Pulsed hydrogen exchange mass spectrometry has been used to investigate folding of rabbit muscle aldolase, an α/β-barrel protein exhibiting the classic TIM structure. Aldolase unfolded in GdHCl refolded as the denaturant concentration was reduced by dialysis. Samples withdrawn during dialysis were pulse-labeled with deuterium to identify unfolded regions in structural forms highly populated during the folding process. Intact, labeled aldolase was digested into fragments, which were analyzed by HPLC electrospray ionization mass spectrometry to detect the H/D exchange along the aldolase backbone. For some concentrations of GdHCl, bimodal distributions of deuterium were found for most peptic fragments, indicating that regions represented by these fragments were either unfolded or folded in the intact polypeptide prior to labeling. The extent of folding was determined from these mass spectra, as well as by CD (220 nm) and enzymatic activity. These results show that folding to the active form involves three domains and two intermediates. Approximately 110 residues fold to highly compact forms in each step. These results also show that each folding domain includes widely separated regions of the backbone. When compared with the results of a previous study of aldolase unfolding, these results show that the folding and unfolding domains include most of the same residues. However, three short segments change domains depending on whether the process is folding or unfolding. These changes are attributed to the very stable quaternary structure of rabbit muscle aldolase.

Many of the current tenets of protein folding hold that the free energy surfaces joining unfolded and folded forms of proteins have the appearance of funnels that facilitate biased searches for the native state. These surfaces may have wells and ridges that lead to detectable populations of intermediates and define folding pathways. The ability to identify a few major features (i.e., wells and ridges) of folding and unfolding paths has raised the question of the similarity of folding and unfolding paths, and whether folding paths might be determined from unfolding paths (*1*−*3*). A related question asks whether regions in which hydrogen exchange is particularly slow might be the folding core of proteins (*4*, *5*).

Rabbit muscle aldolase is a homotetramer (M_r = 157 kDa) in which each subunit has 363 residues that form a parallel α/β-barrel (*6*, *7*). Recent advances in methodology made possible by joining mass spectrometry and amide hydrogen exchange (HXMS)[1] facilitate structural studies of such large proteins (*8*−*11*). With a moderately high molecular mass and a structure similar to the classical TIM barrel, aldolase is a good model for extending folding and unfolding studies to proteins that are too large to be analyzed easily by NMR. Previous reports from this laboratory showed that rabbit muscle aldolase unfolds in denaturants in three discrete steps (*12*, *13*). Furthermore, these results showed that aldolase remained a tetramer at denaturant concentrations that left nearly all of the backbone unfolded, as sensed by amide hydrogen exchange. Finding quaternary structure more resistant to denaturants than secondary and tertiary structure was surprising. This study was directed toward studying the role of this very stable quaternary structure on the populations and structures of intermediates formed during the refolding of aldolase.

The populations and structures of intermediates formed when aldolase was refolded from high concentrations of denaturant were determined by hydrogen exchange mass spectrometry. Because all attempts to refold aldolase by rapid dilution of the sample led to formation of large, soluble assemblies of the aldolase monomer, refolding was studied by slowly reducing the denaturant concentration through dialysis. Under these conditions, refolding occurred over a range of denaturant concentrations. Because the time required to establish equilibrium populations of unfolded and folded forms of the protein was shorter than the time required for dialysis, the populations measured for the various structural forms are assumed to be similar to equilibrium populations. These "near equilibrium" populations measured at different dialysis times indicate how the populations changed with denaturant concentration. In these experiments, the populations of molecules with specific regions of the aldolase backbone folded were determined from mass spectra of

† This work was supported by a grant from the National Institutes of Health (GM RO1 40384) and the Nebraska Center for Mass Spectrometry.

* To whom correspondence should be addressed. Telephone: (402) 472-2794. Fax: (402) 472-9402. E-mail: dsmith7@unl.edu.

‡ Current address: M/S 30W-3-A, Amgen Inc., One Amgen Center Drive, Thousand Oaks, CA 91320.

[1] Abbreviations: CD, circular dichroism; CID MS/MS, collision-induced dissociation mass spectrometry/mass spectrometry; ESIMS, electrospray ionization mass spectrometry; GdHCl, guanidine hydrochloride; H/D, hydrogen/deuterium; HXMS, hydrogen exchange mass spectrometry.

10.1021/bi027388q CCC: $25.00
Published on Web 04/19/2003

peptic fragments derived from the protein labeled intact. Analysis of these data led to identification of regions of similar stability. Collectively, these results show that refolding of aldolase leads to population of two intermediates with structures similar to those found previously for unfolding of the aldolase tetramer. The similarity of these intermediates indicates that aldolase follows similar unfolding and refolding pathways. However, there are small regions that appear to change domains when the starting point is changed, suggesting that the quaternary structure of aldolase can alter the unfolding−refolding pathways. Possible folding models for aldolase are discussed.

MATERIALS AND METHODS

Folding and Labeling of Aldolase. Rabbit muscle aldolase (Sigma Chemical Co.) was unfolded in 6 M GdHCl [H_2O, 50 mM phosphate buffer (pH 7.0) and 10 mM DTT, and 1 mM EDTA] for 24 h. Folding was initiated by decreasing the concentration of GdHCl through dialysis into a solution of the same composition, but with no GdHCl (1:100 solution volumes). Dialysis was performed in a cassette (Pierce model 66410ZZ) with a molecular mass cutoff of 10−12 kDa. The concentration of aldolase in the dialysis cassette was 10 μM.

Aliquots (50 μL) of the dialysate were withdrawn at various times from 5 min to 24 h and diluted 10-fold with D_2O buffer [50 mM phosphate (pD 7.0) and GdDCl]. Isotope exchange was quenched 10 s later by decreasing the pD to 2.5 and the temperature to 0 °C. The quench buffer was 100 mM phosphate (pH 2.3) in a 1:1 D_2O/H_2O mixture. The concentration of GdDCl in the deuterated buffer used to label unfolded regions of the protein matched the concentration of GdDCl in the dialysate at the time each aliquot was withdrawn. This solution was prepared by dialyzing 6 M GdDCl in D_2O for times that matched those of the aldolase samples. Aldolase labeled with deuterium was stored at −80 °C until it was analyzed by HPLC−ESIMS.

Two reference samples, designated "0% ref" and "100% ref" were used to adjust results for deuterium gain and loss during analysis. The 0% ref was prepared by adding unlabeled aldolase to a quench solution without pulse labeling. The 100% ref was prepared by adding aldolase incubated in D_2O at 37 °C for 24 h to a quench solution. Two additional reference samples, designated folded and unfolded references, were used to determine deuterium levels expected for pulse labeling native and unfolded aldolase. These samples were prepared by exposing native or completely unfolded aldolase to D_2O for 10 s prior to quenching H/D exchange. Unfolded aldolase was prepared by equilibrating aldolase in a 6 M GdHCl/H_2O mixture for 24 h. The deuterium levels in quench solutions and the dilution volumes were chosen so that the H:D ratios in the quenched samples were the same (approximately 0.5). This precaution facilitates adjustments for deuterium gains and losses occurring during digestion and analysis. Reported pH and pD values were taken directly from the pH meter without correction.

HPLC Mass Spectrometry. The labeled protein (250 pmol of monomer) was digested with pepsin for 5 min (S:E mass ratio of 1:1 by weight, 0 °C, pH 2.5) before analysis by HPLC−MS. The HPLC injector and column (Vydac C18, 1 mm × 50 mm) were maintained at 0 °C. An aliquot of the digest (250 μL) was injected into the column at a flow rate of 100 μL/min. After the mixture had been desalted with 2% acetonitrile for 2.5 min, the flow rate was reduced to 40 μL/min. Peptic fragments were eluted using a gradient from 2 to 40% acetonitrile in 8 min and detected on-line by a Micromass AutoSpec magnetic sector mass spectrometer equipped for electrospray ionization. Ions in the mass-to-charge range of m/z 300−1300 were recorded using a focal plane detector. All peptides were identified prior to the hydrogen exchange experiments by collision-induced dissociation (CID) tandem mass spectrometry (MS/MS) using a Finnigan LCQ ion trap mass spectrometer. Mass spectra of peptides exhibiting bimodal distributions of isotopes were deconvoluted into their constituent low- and high-mass envelopes using the computer program PeakFit (version 4.0, Jandel Scientific Co.). The areas of these envelopes were used to determine the populations of aldolase that were unfolded in specific regions. The uncertainty in the deuterium levels in these envelopes was typically 0.2 Da. Labeled peptides were analyzed by CID MS/MS using an ion trap mass spectrometer operated with a parent ion mass window width of 7 Da and a collision energy of 25%.

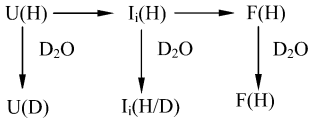
Aldolase Activity. The biological activity of aldolase, which was followed throughout dialysis in separate experiments that did not use D_2O, was determined by the extent of production of 3-phosphoglyceraldehyde from fructose 1,6-bisphosphate (*14*). The level of production of 3-phosphoglyceraldehyde was determined by the absorbance (240 nm) of hydrazone produced by its reaction with hydrazine. Because each assay required only 1 min, the gain in activity could be followed from the earliest dialysis times.

Circular Dichroism Measurements. The folding of aldolase at various times during dialysis was monitored by far-UV circular dichroism (222 nm) using a Jasco J-600 spectrophotometer. The refolding conditions were similar to those for the HX experiments. Because each analysis required only 20 s, the formation of secondary structure could be followed from the beginning.

RESULTS

Incubation of aldolase in 6 M GdHCl for 24 h unfolds and dissociates the homotetramer into its subunits (*15*, *16*). Folding was initiated by dialyzing away the GdHCl. Aliquots taken following various dialysis times were pulse-labeled with D_2O to mark unfolded regions of aldolase. The labeled protein was digested with pepsin into peptides whose deuterium levels and intermolecular distributions were determined by HPLC−MS (*17*, *18*). Both digestion and HPLC were performed under conditions that minimized isotope exchange at peptide amide linkages (*19*). However, deuterium located in rapidly exchanging positions in the side chains and at the N- and C-termini was replaced with protium during HPLC (*20*). As a result, the increased molecular masses of the peptides were direct measures of the number of deuteriums located at peptide amide linkages. This experimental procedure is illustrated in Scheme 1 to emphasize how folded states were detected by pulse labeling. This scheme shows that pulse labeling leads to complete exchange in unfolded regions and minimal exchange in folded regions. The 10 s labeling time used in these experiments was sufficient for complete isotope exchange

Scheme 1



at all peptide amide linkages in unfolded regions of the protein (*19*).

Results from four reference samples of aldolase were used to interpret the mass spectra of peptides derived from labeled aldolase. Two reference samples, called 0% ref and 100% ref, had no deuterium or were completely exchanged in D_2O prior to quenching, respectively. Deuterium levels found in these samples were used to determine the quantity of deuterium that was gained or lost during digestion and HPLC. This information was used to adjust deuterium levels found in other samples for similar artifactual gains and losses of deuterium that occurred during digestion and HPLC (*17*). Deuterium levels found in two additional reference samples, called "folded" and "unfolded", were used to determine the deuterium levels expected for native and unfolded aldolase after pulse labeling.

For the digestion conditions used in this study, 30 peptides covering 80% of the backbone gave strong mass spectra. Bimodal isotope patterns were found in the mass spectra of 27 peptic fragments. These bimodal isotope patterns are illustrated in Figure 1 for three peptic fragments, including residues 144−151, 171−187, and 327−336 taken from intact aldolase following various dialysis times. Spectra of these fragments from unfolded and folded references (top and bottom of Figure 1) give the deuterium levels detected in segments of aldolase when aldolase was either totally unfolded or in the native state. The molecular masses of these three peptic fragments derived from the 100% ref sample (data not shown) indicate that they had an average of 6.6, 10.8, and 5.3 excess deuteriums, respectively. Comparison of these values with deuterium levels expected for complete exchange at each peptide amide linkage shows that the deuterium recoveries for these peptic fragments were 73, 72, and 76%, respectively. Some of this deuterium was lost during digestion (5 min), and some was lost during HPLC (4−8 min). Similar levels of deuterium recovery were found for the other peptic fragments.

Mass spectra taken following dialysis times in the range of 30−100 min exhibit bimodal intermolecular distributions of deuterium. That is, these segments were deuterated in some molecules, but protiated in others. Mass spectra of fragments for shorter or longer dialysis times had single envelopes of isotope peaks, the signature for uniform structure. This bimodal distribution of deuterium showed that the backbone regions represented by these particular peptic

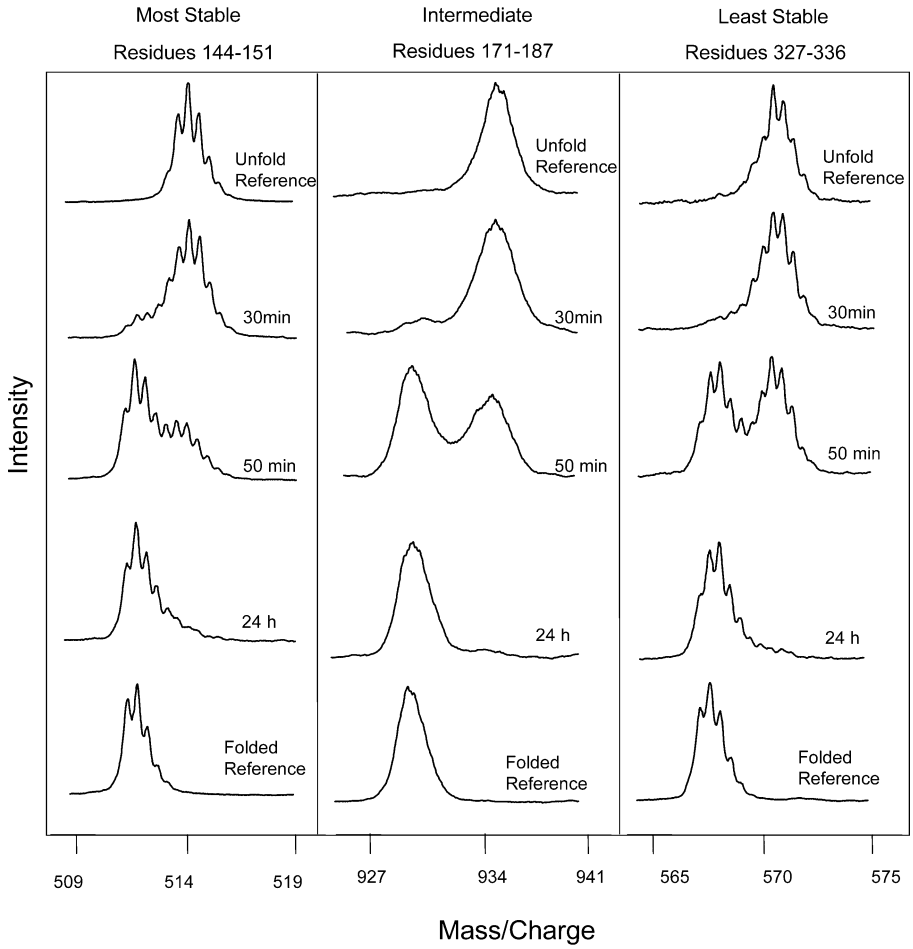


FIGURE 1: Mass spectra of three peptic fragments of aldolase following dialysis of the intact protein for various times in GdHCl and H_2O. Each sample was labeled for 10 s in GdDCl and D_2O prior to quenching and digestion with pepsin. Spectra of unfolded and folded references (top and bottom, respectively) indicate the deuterium levels found for unfolded and folded aldolase. Least Stable, Intermediate, and Most Stable refer to the relative stability of the domains, which was determined from the dialysis times required for half of the molecules to become folded in these regions.

Table 1: Folding of the Aldolase Backbone Represented by 27 Peptic Fragments[a]

Least Stable		Intermediate		Most Stable	
residues	$t_{1/2}$ (min)	residues	$t_{1/2}$ (min)	residues	$t_{1/2}$ (min)
58−63	54	31−49	48	1−17	38
68−78	53	31−53	47	144−151	37
84−105	55	171−187	46	145−151	37
106−130	55	178−187	46	162−170	35
299−307	60	180−187	48	204−211	36
299−310	60	229−238	47	251−256	36
311−326	56	247−250	48	255−269	38
327−336	53	284−298	44	256−269	36
357−363	58	286−298	45	257−269	36
mean	56		47		36
SD	2.5		1.6		1.0

[a] The dialysis times required for these segments in half of the molecules to become protected from hydrogen exchange were determined from bimodal isotope patterns (see Figure 1). Similarity in apparent folding times ($t_{1/2}$) was used to group the segments according to their stability.

fragments were folded in some molecules and unfolded in others. Over a dialysis period of 24 h, the concentration of GdHCl decreased from 6 to 0.06 M. Only the low-mass envelope (little deuterium, high level of protection) remained after 24 h, indicating that all of the molecules (>90%) had become folded. The average molecular masses of the populations represented by the high-mass envelope, the unfolded reference, and the 100% ref were similar, indicating that the amide hydrogens in this population had little or no protection from the solvent. Likewise, the average molecular masses of the low-mass envelope and the folded reference were similar, consistent with a high level of protection due to folding in this population. It is noted that the peak widths in mass spectra of peptides from the folded reference are slightly narrower than those for the refolded aldolase. Mass spectra of most of the peptic fragments of refolded aldolase also exhibited slightly broadened peaks.

The relative intensities of the low- and high-mass envelopes indicate the relative populations of molecules folded in the backbone regions represented by these peptic fragments. The results in Figure 1 show that the population of folded molecules gradually increased with dialysis time until all molecules were folded. The relative intensities of the low- and high-mass envelopes of these fragments were used to determine the folding status of specific regions of the aldolase backbone at various times during dialysis. The dialysis times required for half of the molecules to fold in regions represented by the peptic fragments are presented in Table 1. Close examination of these times showed that they could be divided into three groups, designated as Least Stable, Intermediate, and Most Stable. Because near equilibrium populations of folded domains prevailed throughout the dialysis, the increases in the populations of these domains with dialysis time indicated their relative stabilities. These groups will be termed folding domains to distinguish them from unfolding domains found in our previous studies of aldolase unfolding. Results for all 27 peptic fragments are presented in Figure 2, which is a plot of the percent of molecules that were folded in specific regions versus the dialysis time. Data points and error bars are the means and standard deviations for all of the peptic fragments comprising each unfolding−refolding domain. This figure also shows the GdHCl concentration as a function of the dialysis time.

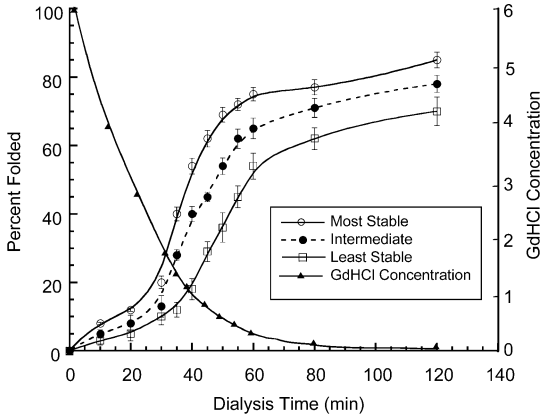


FIGURE 2: Percentage of aldolase molecules folded in regions represented by peptic fragments as a function of the dialysis time. Data points and error bars are the means and standard deviations for all of the peptic fragments comprising each folding domain (see Table 1). The GdHCl concentration during the dialysis is also included.

These results show that most of the protection from hydrogen exchange occurred when the GdHCl concentration was 1.0−1.5 M.

Mass spectra of two additional fragments (residues 152−161 and 152−162) had single envelopes of isotopes. For all dialysis times, the average mass of ions comprising these envelopes was equal to the average masses of ions in the 100% ref sample, indicating that all of the amide hydrogens in these regions had been replaced with deuterium during the 10 s exposure to D_2O. It is important to note that these segments in the folded reference were also completely deuterated. Examination of the three-dimensional structure of aldolase reveals that these fragments are derived from loop regions that exchange rapidly in folded aldolase (*8*). Because these hydrogens have very little protection in the folded protein, the deuterium levels are the same for folded and unfolded aldolase.

In addition to the 29 segments described above, the segment including residues 188−203 displayed a bimodal isotope pattern. However, the low-mass envelope shifted to a lower mass during dialysis, suggesting that this 15-residue segment transcended two folding domains. This interpretation is supported by the stabilities of the adjacent fragments of residues 178−187 and 204−211, which belong to the Intermediate and Most Stable folding domains, respectively (Table 1). To investigate this hypothesis further, CID MS/MS analysis was used as described previously (*21*). Deuterium levels in specific regions of this peptic fragment were determined from the masses of several b-series ions as a function of the dialysis time. These results (Figure 3) show that residues 200−203 folded considerably faster than residues 188−199. When compared with folding times of the three folding domains, these results show that residues 200−203 (half-life of 33 min) belong to the Most Stable domain, while residues 188−199 (half-life of 49 min) fold with the Intermediate domain. It follows that this particular transition between the Most Stable and Intermediate folding domains is located very close to residue 200.

Results described above illustrate aldolase folding, as sensed by the dialysis times when amide hydrogens along the backbone become protected from isotope exchange. Aldolase folding was also followed by the evolution of

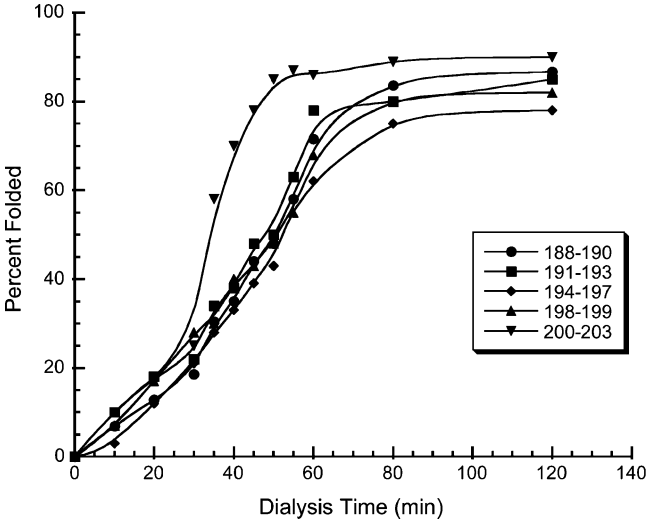


FIGURE 3: Percentage of aldolase molecules folded in regions represented by b-series ions produced by collision-induced dissociation of a peptic fragment (residues 188−203) cleaved from aldolase dialyzed for various times.

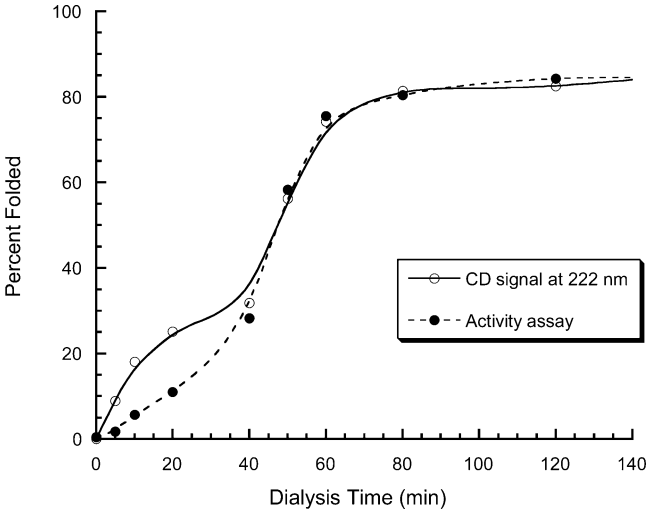


FIGURE 4: Aldolase folding progress with dialysis time determined by CD (222 nm) and enzymatic activity.

α-helices and enzymatic activity assessed in parallel experiments. The percent of aldolase that was folded at various dialysis times was estimated from CD (222 nm) signals obtained for the native aldolase and aldolase that was completely unfolded in 6 M GdHCl. The enzymatic activity of aldolase was also determined following various dialysis times. Results of these experiments (Figure 4) show that approximately 85% of both α-helix character and enzymatic activity of aldolase was regained after folding under the dialysis conditions described herein.

DISCUSSION

Two Intermediates Are Populated during Folding of the Aldolase Monomer. In this study, folding of aldolase monomers was studied by unfolding and dissociating the native tetrameric protein in 6 M GdHCl followed by slowly reducing the denaturant concentration by dialysis over a period of 24 h. This approach was necessary because all attempts to refold aldolase by rapidly diluting the GdHCl gave clear solutions, but inactive protein (unpublished results). All three methods used to follow folding in this study, amide hydrogen exchange, CD (222 nm), and enzymatic activity, showed that most of the refolding occurred after dialysis for 40−60 min. In addition, activity and CD measurements showed that aldolase regained 85% of its original form at the end of dialysis (Figure 4), indicating that most of the protein refolded to its native state.

After dialysis times of 30−60 min, mass spectra of most peptides had two envelopes of isotope peaks, indicating a bimodal distribution of deuterium. Correlation of deuterium levels in these envelopes with results for reference samples indicated that specific regions of the aldolase backbone were folded in some molecules and unfolded in others. These results also showed that H/D exchange in aldolase equilibrated in 6 M GdHCl was complete within 10 s. That is, the same deuterium levels were found in peptic fragments taken from the intact protein in both the unfolded and 100% ref samples. For the conditions used in these experiments (pH 7.0 and 22 °C), exchange into most unfolded segments should be 99% complete within 2 s (*19*). Finding complete exchange in the peptic fragments taken from unfolded aldolase pulse labeled for 10 s shows that any residual structure in aldolase equilibrated in 6 M GdHCl has a protection factor of less than 2−10. These results do not preclude the possibility of residual structure that is transparent to hydrogen exchange.

The deuterium levels and intermolecular distributions in the refolded aldolase, as indicated by the average mass and width of the low-mass envelope of isotope peaks, provide information about the structure of refolded aldolase. The deuterium levels found in the low-mass envelopes were generally equal to those found in native aldolase pulse labeled under the same conditions (i.e., the folded reference). These results show that both refolded and native aldolase had similar structures, as sensed by the rapidly exchanging amide hydrogens. However, it is noted that the width of the low-mass envelope was consistently greater for refolded aldolase than for native aldolase (see Figure 1). The broader intermolecular distribution of deuterium suggests that refolding under the conditions described herein leads to some structural heterogeneity in the refolded aldolase. This heterogeneity may reflect the approximately 15% of the protein that did not fold to an active form (Figure 4).

The times required for the folding domains in half of the molecules to fold were determined from the mass spectra of peptic fragments taken from the labeled protein following various dialysis times. Sorting the folding times of 28 fragments led to three groups of fragments with folding half-lives of 36, 47, and 56 min (Table 1). The average standard deviation of folding times within each group, 1−3 min, is much smaller than the differences (9−11 min) between the mean folding times of the groups, justifying the sorting process. Each domain has 9 or 10 peptides that cover approximately one-third of the entire backbone. These results show that refolding of aldolase from 6 M GdHCl involves at least three folding domains. Molecules with one (Most Stable) or two (Most Stable and Intermediate) domains folded are designated as the unfolding−folding intermediates. A three-dimensional representation of an aldolase monomer illustrating these folding domains is presented in Figure 5A, where the Most Stable, Intermediate, and Least Stable folding regions are colored red, green, and blue, respectively. The folding behavior of some regions remains undetermined because some peptic fragments failed to give quality mass

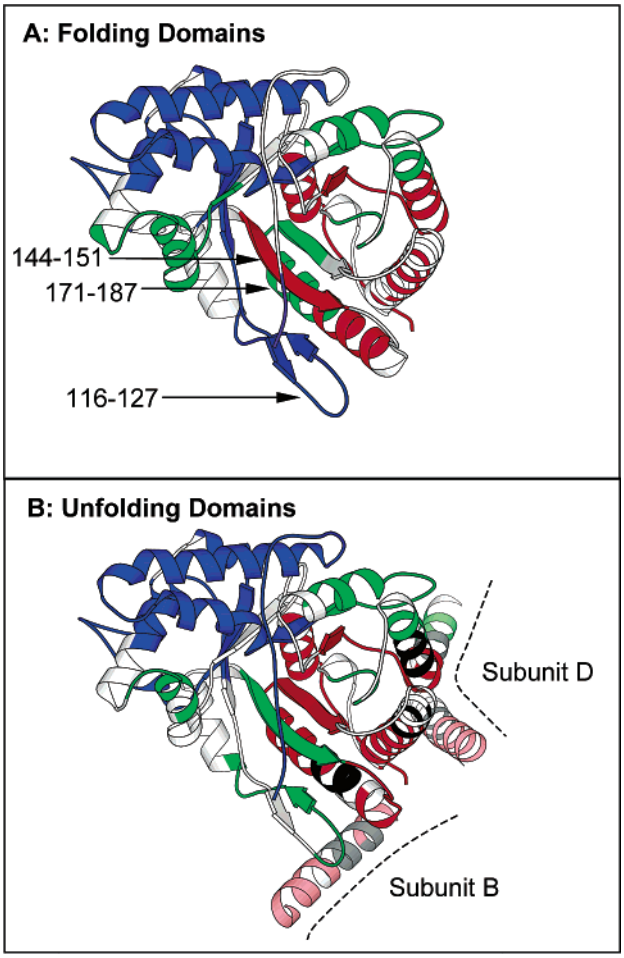


FIGURE 5: Three-dimensional views of aldolase monomers illustrating the folding (results presented herein) and unfolding (*12*) domains. (A) Most Stable, Intermediate, and Least Stable folding domains are colored red, green, and blue, respectively. Backbone regions for which the folding rates could not be determined, either because the corresponding peptic fragments were not found or because they did not give bimodal isotope patterns, are colored gray. Three segments found to change domains are labeled. (B) Most Stable, Intermediate, and Least Stable unfolding domains are colored red, green, and blue, respectively. Parts of two α-helices from adjacent subunits are indicated with pale colors. The three black regions are short segments of the aldolase backbone that are particularly resistant to unfolding.

spectra due to the low concentrations of aldolase used in this study and because the native form of some regions has little protection from H/D exchange.

The relationship between the primary structure of aldolase and its three folding domains is presented in Figure 6. These results show that the Most Stable folding domain includes at least four different regions of the aldolase backbone that are separated by one or more parts of other folding domains. The regions comprising the Most Stable folding domain include residues 1−17, 144−170, 204−211, and 255−269. These four widely dispersed regions must come together in a cooperative folding step. Results presented in Figure 6 also show that the Intermediate folding domain consists of at least four isolated regions (residues 31−53, 171−187, 229−238, and 284−298), while the Least Stable domain appears to consist of two regions (residues 58−130 and 299−363). These results show that the cooperative unfolding and refolding of each of the three folding domains of aldolase

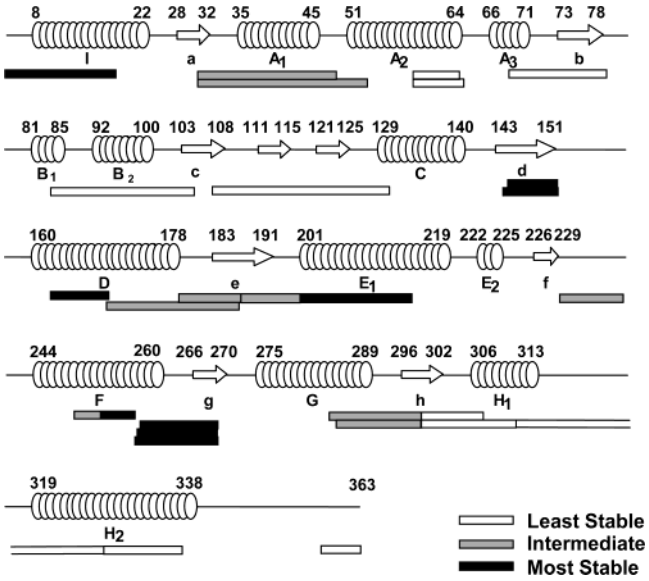


FIGURE 6: Primary structure of aldolase illustrating locations of secondary structural units, peptic fragments used in this study, and the three folding domains.

require docking multiple, noncontiguous regions of the backbone.

The locations of the α-helices and *β*-strands in aldolase are also presented in Figure 6. The four regions comprising the Most Stable folding domain include parts of α-helices D−F and I, as well as parts of *β*-strands d and g. Similarly, the four regions making up the Intermediate domain include parts of several helices (A_1, A_2, D, and F) as well as parts of three strands (a, e, and h). These results do not appear to support a simple model where all α-helices or all *β*-strands fold as a unit. Although some of the strands and helices are close in the primary structure, there is little evidence that the folding domains consist of specific structural motifs, such as a helix−strand−helix motif.

The unfolding and refolding behavior of rabbit muscle aldolase appears to be different from the unfolding and folding behaviors reported for other *β*-barrel proteins. Analysis of various constructs of the α-subunit of tryptophan synthase by different physical methods indicated that the N-terminus including $(\beta\alpha)_{1-6}$ folds initially (*22*−*24*). This folding pattern was also predicted by theoretical simulation using a lattice model for the unfolded protein and Monte Carlo dynamics (*25*). A similar folding pattern has been reported for phosphoribosylanthranilate isomerase (*26*, *27*). Unfolding studies of the α-subunit of luciferase also suggest that folding may start with the N-terminus (*28*). Another *β*-barrel protein, imidazole glycerol phosphate synthase (HisF), appears to fold through assembly of two units, $(\beta\alpha)_{1-4}$ and $(\beta\alpha)_{5-8}$ (*29*). These results, as well as those of the study on aldolase presented here, suggest that several different models may be required to describe the unfolding and refolding of α/*β*-barrel proteins, despite their high level of structural similarity.

The hierarchic model for protein folding starts with a general contraction, followed by formation of local helices and turns, followed by docking of these local, relatively unstable structures to form the stable tertiary structure (*30*, *31*). These results (see Figure 6) indicate that individual helices may unfold and refold as parts of two different

domains. That is, unfolding and refolding of individual helices may occur in two steps. For example, the N-terminal part of helix D folds with the Most Stable domain, while the C-terminal part of this helix folds with the Intermediate domain. Likewise, the N-terminal part of helix F folds with the Intermediate domain, while the C-terminal part of this helix folds with the Most Stable domain. These results show that formation of structure offering protection from hydrogen exchange occurred on a time scale that was much shorter than the labeling time, 10 s. Although it cannot be determined from these results whether the local structure was formed before, with or after formation of longer-range tertiary structure, finding that helices can fold in two highly resolved steps suggests that the tertiary forces play a significant role in establishing boundaries of folding domains. Similar results have been reported for apomyoglobin, where the B helix appears to fold in two different steps (*32*).

Quaternary Structure Affects the Structure of Equilibrium Intermediates. The urea-induced unfolding of aldolase has been studied extensively using H/D exchange and MS to identify and quantify intermediates under both kinetic and equilibrium conditions (*12*, *13*, *33*). These studies showed that aldolase has two unfolding intermediates (i.e., three unfolding domains) that may be highly populated under both kinetic and equilibrium conditions. Locations of the three unfolding domains within the three-dimensional structure of aldolase are indicated in Figure 5B, where the Most Stable, Intermediate, and Least Stable unfolding regions are colored red, green, and blue, respectively. Isotope exchange and CD showed that aldolase was largely unfolded in 3.5 M urea. However, isotope exchange in three short segments (residues 166−170, 204−206, and 251−255, black in Figure 5B) was not detected under these conditions. These three segments are parts of α-helices located in the subunit binding surface (*7*) where they interact with α-helices from other subunits (pale colors in Figure 5B). Contact between the subunits is primarily through hydrophobic side chains located in the flanking α-helices. The extremely slow hydrogen exchange in these segments was interpreted as evidence that aldolase maintains its quaternary structure long after it has lost most of its secondary and tertiary structure.

In these experiments, folding was initiated from 6 M GdHCl, which dissociates aldolase into monomeric subunits (*16*). Several lines of evidence indicate that folding occurred prior to assembly of monomers to form tetramers. For example, regions including the three short segments previously found to be diagnostic for the unfolded tetramer achieved protection at the same dialysis time as all other segments in the Most Stable folding domain (see the Most Stable column in Table 1). Had there been significant concentrations of unfolded tetramer, these segments would have achieved protection faster than other parts of the fast folding domain. Folding of aldolase prior to assembly of the tetramer is also supported by rapid mixing studies (unpublished results) and with earlier studies (*34*, *35*). Size-exclusion chromatography shows that approximately 24 h is required for formation of the tetramer (unpublished results).

It is interesting to note that these folding experiments gave primarily aggregates or inactive protein when the concentration of aldolase was greater than 10 μM. No aggregation was evident in the unfolding studies, which were performed with an aldolase concentration of 200 μM. These results suggest that the three short segments in the subunit binding surface may be responsible for aggregation in folding studies. Identification of specific regions responsible for aggregation by amide hydrogen exchange may be an important step in redesigning proteins to be resistant to aggregation.

Table 2: Segments of the Aldolase Backbone that Are in One Folding Domain, but in a Different Unfolding Domain

residues	found during folding	expected during unfolding	found during unfolding
116−127	Least Stable	Least Stable	Intermediate
144−151	Most Stable	Most Stable	Intermediate
171−187	Intermediate	Intermediate	Most Stable

Results presented in Figure 5 indicate that aldolase follows similar folding and unfolding pathways. That is, most of the residues comprising the Most Stable folding domain (red in Figure 5A) were the residues found in the Most Stable unfolding domain (red in Figure 5B). Likewise, most of the residues in the Least Stable folding domain (blue) were in the Least Stable unfolding domain (blue). There are, however, three regions (see Figure 5A) that do not fit this simple pattern. Results for these segments, which change domains when the folding process is reversed, are summarized in Table 2. The segment including residues 116−127 was part of the Least Stable folding domain (blue in Figure 5A), suggesting it would be part of the Least Stable unfolding domain. However, it was found as part of the Intermediate unfolding domain (green in Figure 5B). These residues, along with the adjacent segment of residues 106−115, form two poorly defined β-strands that are not part of the main β-barrel comprising the structural core of aldolase. The segment including residues 144−151, which forms one of the principal β-strands in the α/β-barrel, was part of the Most Stable folding domain (red in Figure 5A), suggesting that it would be part of the Most Stable unfolding domain. However, experiment showed that this segment was part of the Intermediate unfolding domain. The third segment found to change domains, residues 171−187, formed large parts of a β-strand and an adjoining α-helix. This segment was part of the Intermediate folding domain, but part of the Most Stable unfolding domain.

Finding two intermediates with similar structures in both unfolding and folding experiments suggests that these two processes use similar pathways. However, finding three relatively small regions that do not follow the expected folding−unfolding patterns indicates slightly different reaction paths. In these experiments, folding was initiated from 6 M GdHCl, a condition in which aldolase was completely unfolded and disassembled. Unfolding was performed in 1−4 M urea, a condition that largely unfolds aldolase, but does not disassemble the tetramer. Because of the unusual stability of the quaternary structure in aldolase, the unfolding experiments did not involve monomers. These results suggest that binding to α-helices in adjacent subunits alters the structure in the regions, including the three segments that changed domains, thereby changing the reaction paths slightly.

The contact between these three segments and other subunits has been estimated from differences in their accessible solvent areas when calculated for the monomer and tetramer. The average accessible solvent areas for residues 116−127 and 171−187 are reduced to approximately 50% in the tetramer, indicating that these segments

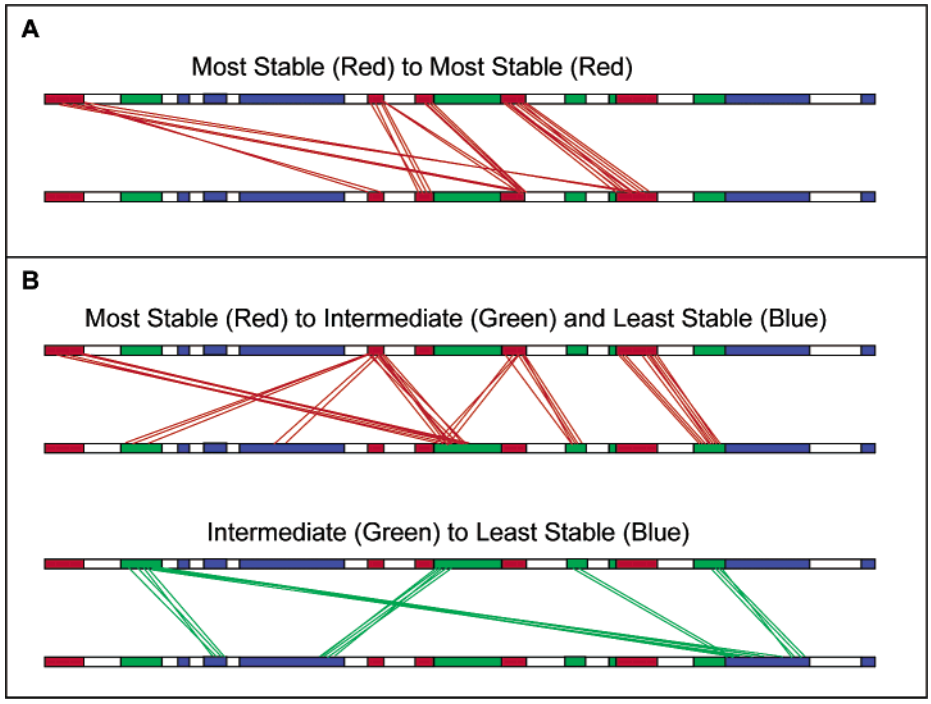


FIGURE 7: Contact map for backbone segments of aldolase based on its crystal structure (PDB entry 1ADO). (A) Contacts among the segments included in the Most Stable folding domain. (B) Contacts among the Most Stable, Intermediate, and Least Stable domains. Connected segments, indicated with red lines, have interatomic contacts of ≤4 Å (*38*).

have considerable contact with adjacent subunits. Direct measurements of interatomic distances show that residues 116, 117, 125, 127, 172, 175, and 179 are only 4 Å from residues in another subunit. These contacts are expected to stabilize the folded forms of these regions (see Table 2). The other segment exhibiting unexpected unfolding behavior, which includes residues 144−151, does not appear to have any contact with other subunits.

Proposed Folding Model for Aldolase. Folding and unfolding of the α/β-barrel, rabbit muscle aldolase, have been studied extensively by hydrogen exchange, CD, and activity. Results of these studies show that protection against hydrogen exchange, formation of α-helices, and enzymatic activity occur on the same time scale, when averaged over the entire population of molecules. Aldolase folding is envisioned to start with collapse to a compact state where four widely separated regions of the backbone, including a total of approximately 110 residues, become closely packed in an apparent highly cooperative action with a duration that is considerably shorter than the labeling time, 10 s. Following formation of this folding nucleus, two additional domains fold sequentially and cooperatively. Each of these domains also has approximately 110 residues. Finding that approximately 110 residues fold in each step is consistent with predictions that large proteins are likely to fold in steps involving approximately 100 (*36*).

Coalescence of four segments to form a highly compact folding nucleus might occur by assembly of local structural units [i.e., hierarchic model (*30*, *31*)]. However, finding approximately half of two α-helices in this domain while their other halves are in a different folding domain is not consistent with this model. Likewise, finding no apparent preference for particular types of secondary structure is also inconsistent with the hierarchic model. Coalescence of these segments might also be driven directly by hydrophobic forces. Although highly hydrophobic patches occur in all three folding domains, three of the four regions comprising the Most Stable folding domain are highly hydrophobic. Contact among these four regions (see Figure 7A) has been identified from the X-ray crystal structure of aldolase. In Figure 7, Most Stable, Intermediate, and Least Stable folding domains are colored red, green, and blue, respectively. The red lines indicate the number of contacts between these segments that are within approximately 4 Å of each other. These measurements show that the four regions making up the Most Stable folding domain have many points of mutual contact in folded aldolase.

Following formation of the Most Stable folding domain (red in Figure 5A), approximately 110 more residues achieve protection in each of the next two stages of folding. Both reactions appear to be highly cooperative, occurring in less than 10 s. As in the first step, each of these folding domains includes residues from widely different regions of the backbone with no preferences for particular types of secondary structure (Figure 6). This behavior is not consistent with many small steps, each involving another unit of local structure, though such behavior might be detected using shorter labeling times. Given the locations of these segments in folded aldolase, and given that aldolase has only one apparent structural domain, it also seems unlikely that these regions fold into a unit before binding to the initially formed domain. It seems more likely that the second and third domains fold using the domain folded in the preceding step as a template (*37*). The pattern of inter-residue contacts in folded aldolase (see Figure 7B) supports this view. The red lines in the top panel of Figure 7B indicate the number of direct contacts between the segments of the Most Stable folding domain (folding nucleus) and segments of the Intermediate (green) and Least Stable (blue) folding domains. This presentation indicates numerous contacts between the

Most Stable and Intermediate domains, but very few contacts between the Most Stable and Least Stable domains. However, there are many direct contacts between the Intermediate and Least Stable domains (see the bottom of Figure 7B). These results suggest that a folded region may function as a template for folding of another domain. The general validity of this model awaits similar studies of other $\alpha\beta$-barrel proteins.

ACKNOWLEDGMENT

Part of this manuscript was written while D.L.S. was on sabbatical leave at the Institut de Biologie Structurale in Grenoble, France.

REFERENCES

1. Raschke, T. M., and Marqusee, S. (1997) *Nat. Struct. Biol. 4*, 298−304.
2. Fuentes, E. J., and Wand, A. J. (1998) *Biochemistry 37*, 3687−3698.
3. Xu, Y., Mayne, L., and Englander, S. W. (1998) *Nat. Struct. Biol. 5*, 774−778.
4. Clarke, J., Itzhaki, L. S., and Fersht, A. R. (1997) *Trends Biochem. Sci. 22*, 284−287.
5. Li, R., and Woodward, C. (1999) *Protein Sci. 8*, 1571−1590.
6. Sygusch, J., Beaudry, D., and Allaire, M. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 7846−7850.
7. Gamblin, S. J., Cooper, B., Millar, J. R., Davies, G. J., Littlechild, J. A., and Watson, H. C. (1990) *FEBS Lett. 262*, 282−286.
8. Zhang, Z., Post, C. B., and Smith, D. L. (1996) *Biochemistry 35*, 779−791.
9. Resing, K. A., and Ahn, N. G. (1998) *Biochemistry 37*, 463−475.
10. Halgand, F., Dumas, R., Biou, V., Andrieu, J. P., Thomazeau, K., Gagnon, J., Douce, R., and Forest, E. (1999) *Biochemistry 38*, 6025−6034.
11. Chen, J., Walter, S., Horwich, A. L., and Smith, D. L. (2001) *Nat. Struct. Biol. 8*, 721−728.
12. Deng, Y., and Smith, D. L. (1998) *Biochemistry 37*, 6256−6262.
13. Deng, Y., and Smith, D. L. (1999) *J. Mol. Biol. 294*, 247−258.
14. Deal, W. C., Rutter, W. J., and Van Holde, K. E. (1963) *Biochemistry 2*, 246−251.
15. Tanford, C., Kawahara, K., and Lapanje, S. (1966) *J. Biol. Chem. 241*, 1921−1923.
16. Castellino, F. J., and Barker, R. (1968) *Biochemistry 7*, 4135−4138.
17. Zhang, Z., and Smith, D. L. (1993) *Protein Sci. 2*, 522−531.
18. Engen, J. R., and Smith, D. L. (2001) *Anal. Chem. 73*, 256A−265A.
19. Bai, Y., Milne, J. S., Mayne, L., and Englander, S. W. (1993) *Proteins: Struct., Funct., Genet. 17*, 75−86.
20. Englander, J. J., Rogero, J. R., and Englander, S. W. (1985) *Anal. Biochem. 147*, 234−244.
21. Deng, Y., Pan, H., and Smith, D. L. (1999) *J. Am. Chem. Soc. 121*, 1966−1967.
22. Zitzewitz, J. A., Gualfetti, P. J., Perkons, I. A., Wasta, S. A., and Matthews, C. R. (1999) *Protein Sci. 8*, 1200−1209.
23. Zitzewitz, J. A., and Matthews, C. R. (1999) *Biochemistry 38*, 10205−10214.
24. Bilsel, O., Zitzewitz, J. A., Bowers, K. E., and Matthews, C. R. (1999) *Biochemistry 38*, 1018−1029.
25. Godzik, A., Skolnick, J., and Kolinski, A. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 2629−2633.
26. Eder, J., and Kirschner, K. (1992) *Biochemistry 31*, 3617−3625.
27. Jasanoff, A., Davis, B., and Fersht, A. R. (1994) *Biochemistry 33*, 6350−6355.
28. Noland, B. W., Dangott, L. J., and Baldwin, T. O. (1999) *Biochemistry 38*, 16136−16145.
29. Hocker, B., Beismann-Driemeyer, S., Hettwer, S., Lustig, A., and Sterner, R. (2001) *Nat. Struct. Biol. 8*, 32−36.
30. Baldwin, R. L., and Rose, G. D. (1999) *Trends Biochem. Sci. 24*, 77−83.
31. Baldwin, R. L., and Rose, G. D. (1999) *Trends Biochem. Sci. 24*, 26−33.
32. Jennings, P. A., and Wright, P. E. (1993) *Science 262*, 892−896.
33. Deng, Y., and Smith, D. L. (1999) *Anal. Biochem. 276*, 150−160.
34. Teipel, J. W. (1972) *Biochemistry 11*, 4100−4107.
35. Chan, W. W.-C., Mort, J. S., Chong, D. K. K., and MacDonald, P. D. M. (1973) *J. Biol. Chem. 248*, 2778−2784.
36. Bryngelson, J. D., and Wolynes, P. G. (1990) *Biopolymers 30*, 177−188.
37. Rumbley, J., Hoang, L., Mayne, L., and Englander, S. W. (2001) *Proc. Natl. Acad. Sci. U.S.A. 98*, 105−112.
38. Vriend, G., and Krause, R. (2002) *WHAT IF*, http://cmbi1.cmbi.kun.nl:1100/WIWWWI/.

BI027388Q