2004

Optimization of Conditions for Studies of Protein Unfolding by Hydrogen Exchange/Mass Spectrometry

A.S. Raza
University of Nebraska - Lincoln

David L. Smith
University of Nebraska - Lincoln, dsmith7@unl.edu

Optimization of conditions for studies of protein unfolding by hydrogen exchange/mass spectrometry

A.S. Raza and D.L. Smith
Department of Chemistry, University of Nebraska, Lincoln, NE 68588-0304, USA. E-mail: dsmith7@unl.edu

Understanding the forces driving protein folding and aggregation is an essential step in developing means for controlling these important processes. Amide hydrogen exchange, coupled with mass spectrometry, has become an important method for studying protein unfolding and refolding. To extend procedures developed to study unfolding of relatively soluble proteins to less soluble, aggregation-prone proteins requires special considerations. This publication describes a general strategy developed using yeast transaldolase, which aggregates easily under conditions required to study its unfolding. Results presented here show that reducing the protein concentration to the nanomolar range is essential for managing aggregation of transaldolase. In addition, the present results point to use of relatively high concentrations of denaturants and short incubation times to minimize aggregation. These results also show how amide hydrogen exchange, coupled with mass spectrometry, can be used to study soluble aggregates.

Keywords: protein folding, hydrogen exchange, mass spectrometry, transaldolase

Introduction

Most of the several thousand atoms making up the individual molecules of a folded protein are located in small, highly specific regions of the protein. If only a few atoms are misplaced, the protein may not exhibit its normal properties. Proteins with misplaced atoms may be regarded as "misfolded". Understanding the forces that direct protein folding, both to give the native structure as well as misfolded forms, is important for our efforts to control certain processes in biotechnology and several diseases. From basic thermodynamics we know that the unfolded and folded states are connected by a free energy surface. Both theory and experiment have shown that folding energy surfaces for some proteins are smooth, whereas folding energy surfaces for other proteins have valleys and ridges. To begin to understand protein folding, we need to characterize these surfaces, which begins by characterizing the important valleys and ridges.

A variety of different theoretical and experimental methods have been used to study protein-folding free energy surfaces. The general goal of most experimental approaches is to detect and characterize partially folded intermediates whose populations are a result of specific combinations of valleys and ridges in the free energy surfaces. These intermediates may be highly populated and, therefore, detected when unfolded proteins are allowed to refold under certain conditions. In addition, there is increasing evidence that specific features of the protein-folding surface can be revealed by the reverse process, protein unfolding.[1,2]

Amide hydrogen exchange/mass spectrometry (HX-MS) has become an important tool for studying the intermediates populated during unfolding and refolding of several proteins.[3–7] The general approach used to study partially folded intermediates populated in unfolding studies involves incubation of a protein in denaturant to increase the populations of intermediates, labeling the protein in D_2O and analyzing by mass spectrometry.[7] If the incubation time is sufficiently long to establish equilibrium populations of the folded, partially folded and unfolded forms, their populations may be used to determine the free energy change in each step. However, aggregation and precipitation may occur before equilibrium is established. Yeast transaldolase is an example of a protein whose unfolding is complicated by aggregation of the partially unfolded protein. We present here results that illustrate the competition between unfolding and aggregation, and discuss experimental approaches we have used to minimize the confounding effects of aggregation.

Experimental

Sample preparation

Yeast transaldolase, phosphate buffer, urea and guanidine hydrochloride were purchased from Sigma Chemical

DOI: 10.1255/ejms.631 ISSN 1356-1049

Co. Deuterium oxide (99.9 atom %D) and formic acid were purchased from Aldrich Chemical Co. Aggregates and impurities present initially in the commercial transaldolase were separated from the native protein by size-exclusion chromatography (SEC). The SEC Superose 12 column was equilibrated in 5 mM phosphate buffer pH 6.8 prior to protein elution using 5 mM phosphate buffer pH 7. For a flow rate of 500 μL min^{-1}, the protein eluted within 20 min. This procedure also served to remove any salt present in the protein and to dilute the protein into the phosphate buffer used in the next step.

The general procedures used to study protein unfolding by HX-MS have been described previously.[3,7,8] Sample preparation was initiated by diluting 2 μM transaldolase into different concentrations of either guanidine hydrochloride (GdHCl) or urea. Following various incubation times, the samples were diluted 10–20 fold with D_2O (same concentration of denaturant and pH) to label unfolded regions. Both incubation and labeling solutions were buffered to pH 6.8 or 7.3 with 5 mM phosphate. Labeling times of 5 s and 10 s were used when the pH was 7.3 or 6.8, respectively. In either case, the labeling time was sufficient for deuterating all peptide linkages in unfolded regions of the protein. At the end of the labeling time, isotope exchange was quenched by decreasing the pH to 2.5 (via addition of 0.2 M HCl) and the temperature to 0°C.

Several reference samples were used in this study. The completely deuterated 100% reference was prepared by incubating the protein in 4 M urea/D_2O or 4 M GdHCl/D_2O at pD 6.8 or 7.3 until all of the amide hydrogens in the protein had exchanged with deuterium. This reference was used to determine the number of deuteriums lost by the protein during analysis. The unfolded reference was prepared by unfolding the transaldolase in 4 M urea or GdHCl for 1 h followed by pulse labeling with D_2O buffer for various times. Results of these experiments showed that exchange of amide hydrogens in the unfolded protein was complete after 10 s of labeling at pH 6.8 or 5 s labeling at pH 7.3. The native reference, which was prepared by incubating the protein in H_2O for 2 h and pulse labeling in D_2O, was used to determine the deuterium uptake by the native conformation. The 0% reference was prepared by the same procedure, except that H_2O was used in place of D_2O for labeling. The final H_2O:D_2O ratio after the quench step was similar in each sample, thereby facilitating direct comparison of the deuterium levels found for different samples.

Analyses by HPLC ESI-MS

Isotope analyses were performed by passing 1000 μL of sample through a 2 × 20 mm loop filled with perfusion material (POROS 10 R2, PerSeptive Biosystems). Most of the denaturant and salt was removed from the protein by passing 1000 μL of 0.25% formic acid through the loop at 0°C. Following this step of protein concentration and purification, the protein was eluted from the loop and through a similar column by a mobile phase composed of H_2O (0.25% formic acid, 0°C) and acetonitrile (0.25% formic acid, 0°C). The flow rate was 50 μL min^{-1}. A gradient of 30 to 85% acetonitrile in 5 min led to elution of the protein in approximately 5 min. The column, injector and solvents were submerged in an ice bath to minimize isotope exchange at the peptide amide linkages. For the conditions used in these experiments, deuterium loss from these positions was less than 10%. However, deuterium located in the sidechains, as well as at the N- and C-termini, was replaced with protium during HPLC.[9] Because the protein was unfolded under quench conditions, similar deuterium levels and distributions were found in all charge states. The deuterium levels and intermolecular distributions of the eluting proteins were determined by electrospray ionization mass spectrometry (ESIMS) using either a Platform II or QToF instrument (Micromass Inc.). The protein concentration during incubation in denaturant was 0.5 μM for analyses by the less sensitive Platform II and 0.1 μM for analyses by the more sensitive QToF. Commercially-available software (Peak Fit version 4, ASIN Software Inc.) was used to deconvolute some broad peaks and to determine the centroids of these deconvoluted peaks.

Results and discussion

Human transaldolase is a single domain, $(\beta\alpha)_8$-barrel protein that exists as a homodimer under native conditions. The X-ray crystal structure of yeast transaldolase, the protein used in this study, has not been reported. However, the amino acid sequence of yeast transaldolase has 60% identity with human aldolase.[10] Nucleic acid sequencing indicates that yeast transaldolase has 334 amino acid residues and a molecular weight of 36,950.[11] However, analysis of the commercially-available protein indicated a molecular weight of 35,460. This discrepancy was attributed to truncation of one residue at the N-terminus and 12 residues at the C-terminus. Lachais *et al.* reported finding multiple transaldolase isoforms using 2D-electrophoresis analysis.[12]

The experimental procedure used to study unfolding of transaldolase by HX-MS is illustrated in Figure 1. Aqueous, buffered solutions of the protein were diluted with aqueous, buffered solutions containing high concentrations of denaturant, either GdHCl or urea. Following various incubation times, these solutions were diluted with deuterated labeling solutions that also contained denaturant. After a brief time, the isotope exchange reaction was quenched and the samples were analyzed by HPLC-ESI MS. The labeling time was just sufficient to replace protium at peptide amide linkages in unfolded regions of the protein with deuterium. Chromatography, prior to mass spectrometry analysis, was essential for removing buffers and the high concentrations of denaturants required to destabilize the protein. In addition, the protic solvents used for the mobile phase lead to back exchange of deuterium located in rapidly exchanging

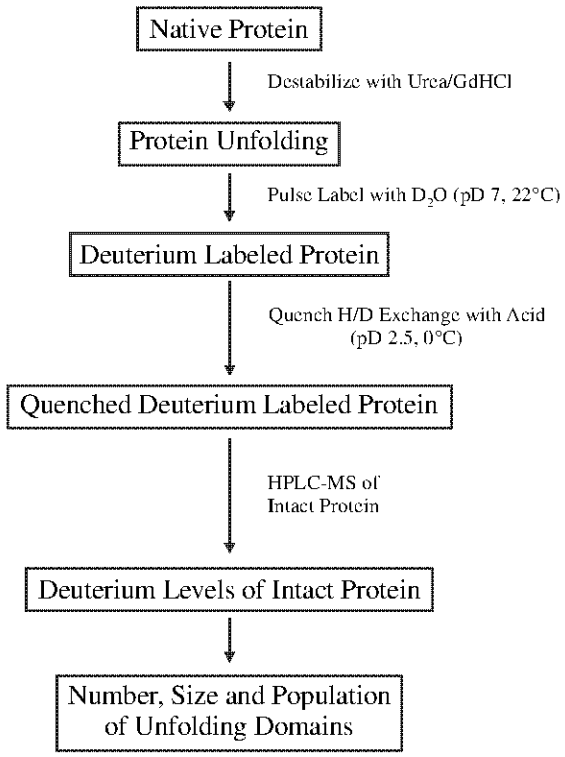

Figure 1. Experimental procedure used to study unfolding of yeast transaldolase by HX-MS.

positions in the side chains, as well as those at the N- and C-termini.[9,13,14] As a result, the mass increases directly reflect the number of deuteriums located at peptide amide linkages.

Incubation of proteins in denaturants may loosen the structure of the folded protein. In addition, denaturants may lead to population of structurally different forms (for example, partially unfolded and completely unfolded forms). The time required for the protein to complete these structural adjustments may span a range from a few seconds to many hours, depending on the protein and the concentration of the denaturant. Because pulse labeling gives an instantaneous view of the structural forms, the results may depend on the incubation time. When the incubation time is less than the time required to reach equilibrium, the populations of the different structural forms are controlled by the activation energies of each rate-limiting unfolding step. However, when the incubation time is longer than the time required to reach equilibrium, the populations are controlled by the free energy of each structural form. Whether unfolding intermediates are substantially populated under kinetic or thermodynamic control depends on the protein and the experimental conditions. Kinetic and thermodynamic control led to population of the same unfolding intermediates in aldolase[15] and the α-subunit of tryptophan synthase,[16] whereas only kinetic control led to population of unfolding intermediates in GroEL.[17]

Most previous studies of protein unfolding have focused on a small number of highly soluble proteins. As interest shifts to less soluble proteins, aggregation becomes an important issue. Yeast transaldolase is an example of a protein that undergoes aggregation under some conditions. In our initial studies of transaldolase unfolding, the protein concentration during incubation was 200 μM. Extensive aggregation and precipitation was observed, prohibiting any useful measurements. Decreasing the protein concentration to 0.1–0.5 μM substantially reduced the aggregation. However, the low concentration of protein led to sensitivity problems, which were resolved by using a packed injection loop that permitted rapid concentration of the protein. Details of this approach have been discussed elsewhere.[18]

Finding denaturant concentrations and incubation times that lead to measurable populations of all transaldolase structural forms was an essential step in this study. The effects of aggregation on finding these conditions, even when the protein concentration was only 0.5 μM, are illustrated in Figure 2. Mass spectra of the intact protein following incubation in low concentrations of GdHCl (< 0.5 M) for 6 h, 1 h or 1 min have one peak with *m/z* 1076.5 (+33 charge state) corresponding to the folded protein. Incubation in high concentrations of GdHCl (> 2.0 M) also leads to mass spectra with only one peak, but the *m/z* has increased to approximately 1082.8, which corresponds to that of the completely unfolded protein. Results presented in Figures 2(a) and 2(b)

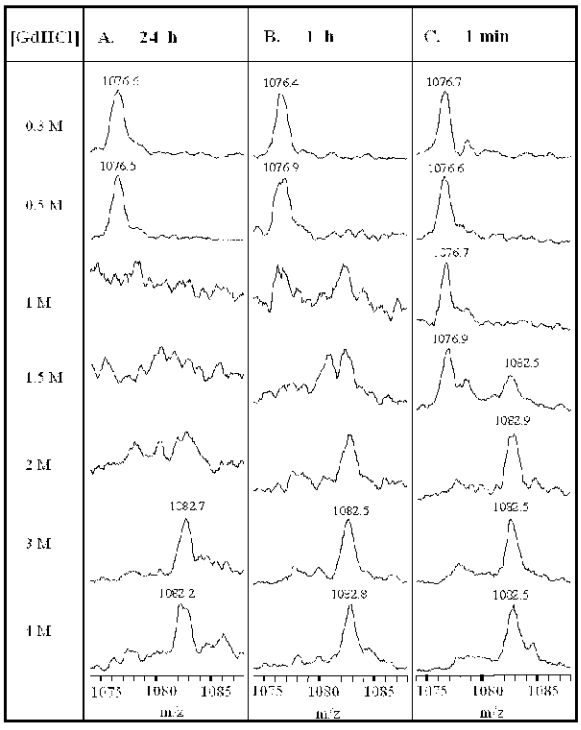

Figure 2. ESI mass spectra of the +33 charge state of aldolase incubated in various concentrations of GdHCl for: (a) 6 h, (b) 1 h and (c) 1 min prior to pulse labeling with D_2O.

show that incubation in intermediate concentrations of GdHCl leads to low quality, useless mass spectra. Although the low concentration of protein used in these experiments led to high noise levels in all spectra, the particularly low quality found for intermediate concentrations of GdHCl is due to protein aggregation under these conditions. Unfortunately, these are the conditions that are most likely to give measurable populations of different structural forms. It is noted that the mass spectrum of transaldolase incubated in 1.5 M GdHCl for only 1 min [Figure 2(c)] appears to have three peaks corresponding to three different structural forms. These results suggest that measurable levels of folded, partially unfolded intermediate(s) and completely unfolded protein were present in 1.5 M GdHCl.

Collectively, results presented in Figure 2 point to a general strategy of using relatively high concentrations of denaturant and short incubation times to avoid aggregation. Optimization of denaturant concentration for kinetically-controlled unfolding experiments requires consideration of three points: (a) the denaturant concentration must be sufficient to unfold the protein, (b) higher concentrations of denaturant lead to faster unfolding rates and (c) the minimum labeling time at pH 7 is approximately 5 s. For the case of transaldolase, results presented in Figure 2 show that GdHCl concentrations in the range of 1.5–2.0 are sufficient to unfold the protein. The time actually required to unfold transaldolase in 2.0 M GdHCl can be determined from results presented in Figure 3(a). Finding multiple peaks in the mass spectrum after incubation for only 10 s indicates population of several structural forms. If higher concentrations of denaturant were used, the unfolding rate would be too fast for the obligatory labeling time of approximately 5 s and there would be only one peak, which corresponds to the unfolded protein.

The multiple peaks in the mass spectra of transaldolase incubated in 2.0 M GdHCl for 10–30 s [Figure 3(a)] illustrate the conditions required to detect unfolding intermediates. Mass spectra of transaldolase in the native and unfolded reference states are presented at the top and bottom of Figure 3(a), respectively. The *m/z* values of the principal peaks in each of these reference spectra were used to identify structural forms represented by the multiple peaks in the mass spectra of transaldolase incubated in GdHCl for various times. Application of this approach will be illustrated using the 10 s spectrum. The *m/z* of the largest peak, 960.4, is only slightly greater than that of the native reference, indicating that the structure of this population is only slightly more unfolded than that of the native structure. The peak with *m/z* 965.7 represents another population, which is completely unfolded. The rather broad peak with *m/z* between these two extremes represents additional populations that are partially unfolded.

Although the actual number of structural forms responsible for the broad peak is unknown, this peak could be fitted by two peaks with widths similar to those of the two references. The *m/z* values of these two peaks, 961.9 and 962.7, may be used to determine the number of residues that are unfolded (i.e. lost protection towards hydrogen exchange) in each structural form. Similarly, the relative intensities of each of the peaks in the mass spectra may be used to determine the populations of each structural form as a function of incubation time. These intensities have been used with kinetic modeling to determine rate- and equilibrium-constants for unfolding.

Similar fast-unfolding experiments were performed using 4.0 M urea in place of 2.0 M GdHCl. In addition, the protein concentration during incubation was reduced from 0.5 μM to 0.1 μM. This reduction in protein concentration was possible because the analyses were performed with a QToF mass spectrometer. Results for rapid unfolding in urea are presented in Figure 3(b). It is important to note that spectra taken after incubation in GdHCl or urea for 10 s are very similar in terms of the number of intermediates and their populations. Furthermore, reducing the protein concentration five-fold had little effect on the general features of the mass spectra. It may also be of interest to note that the resolution in the QToF spectra is only slightly better than the resolution in the spectra taken using the Platform. Although the instrument resolution of the QToF is nearly 10-fold

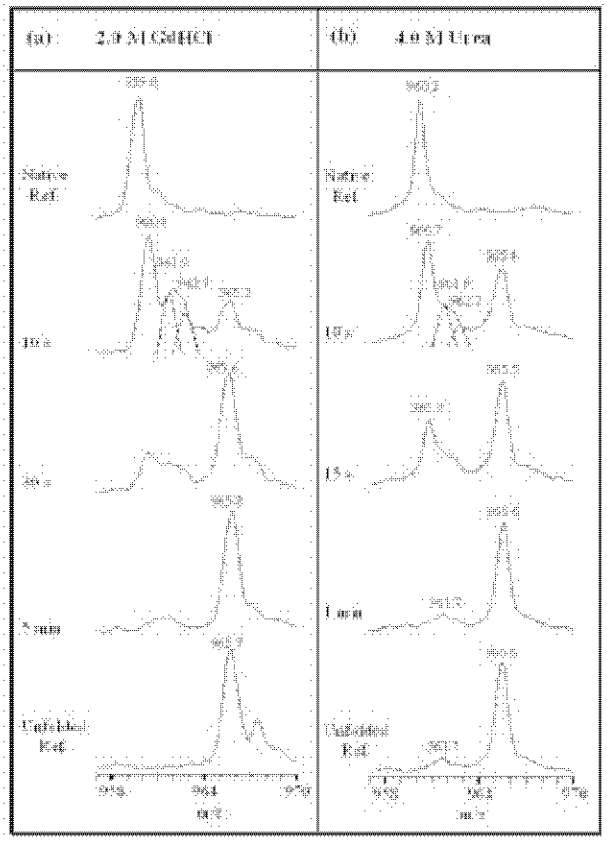

Figure 3. ESI mass spectra of the +37 charge state of transaldolase incubated for various times in: (a) 2.0 M GdHCl or (b) 4.0 M urea prior to pulse labeling with D_2O.

greater than that of the Platform, there is only a small gain in effective resolution because the intermolecular distribution of deuterium is broad.

For the experimental conditions used to obtain results presented in Figure 3, i.e. relatively high concentrations of denaturant and short incubation times, there was no evidence of aggregation. It is important to note that aggregation may have occurred even when the solution remains visually clear. The ability to study structural features of misfolded or aggregated proteins that remain in solution is illustrated by mass spectra presented in Figure 4. In these experiments, transaldolase (0.1 μM) was destabilized in 2.4 M urea and pulse labeled as described above. Comparing results in Figures 3(b) and 4 shows how decreasing the urea concentration from 4.0 to 2.4 M substantially reduces the unfolding rate. Although the protein unfolds much more slowly in the lower concentration of urea, the mass spectrum of transaldolase incubated in 2.4 M urea for 3 min shows population

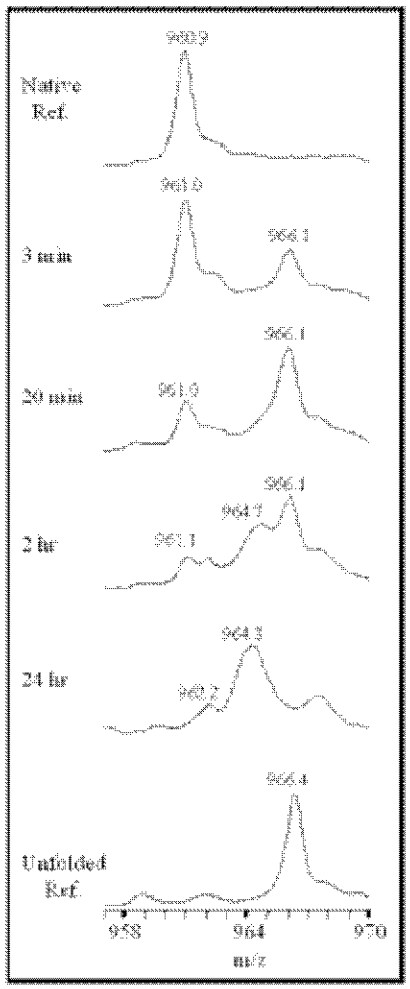

Figure 4 ESI mass spectra of the +37 charge state of transaldolase incubated for various times in 2.4 M urea prior to pulse labeling with D_2O.

of the same unfolding intermediates as were found for unfolding in 4.0 M urea after incubation for only 10 s. However, as the incubation time increased beyond 20 min, the number of peaks and their relative intensities changed. Incubation for 2 h gave four peaks, indicating population of at least four structural forms of the protein. Incubation for 24 h gave a major peak at *m/z* 964.4, which indicates that much of the backbone of molecules comprising this form has substantial protection against hydrogen exchange. This protection is due to misfolding or aggregation.

Conclusions

Most previous HX-MS studies of protein unfolding were performed under conditions where aggregation was minimal. These studies could be performed using relatively high concentrations of protein and a wide range of experimental conditions because the proteins were highly soluble. As interest shifts to unfolding and to fundamental aggregation properties of less soluble proteins, experimental procedures for controlling aggregation are required. These procedures will be helpful for studying unfolding in the absence of aggregation, as well as for studying the fundamental processes that initiate aggregation.

The present study showed that decreasing the protein concentration to the nanomolar range is helpful, but may not be sufficient for reducing aggregation. In addition, short incubation times lead to less aggregation. However, short incubation times usually mean that unfolding experiments must be performed under kinetic rather than equilibrium control. In either case, mass spectra of labeled protein can be used to determine the number of populated intermediates, as well as their population and extent of protection toward hydrogen exchange.

Acknowledgements

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

References

1. Y. Bai, T.R. Sosnick, L. Mayne and S.W Englander, "Protein folding intermediates: Native-state hydrogen exchange", *Science* **269,** 192 (1995).
2. H. Pan and D.L. Smith, "Quaternary structure of aldolase leads to differences in its folding and unfolding intermediates", *Biochemistry* **42,** 5713 (2003).
3. Y. Deng and D.L. Smith, "Rate and equilibrium constants for protein unfolding and refolding determined by hydrogen exchange-mass spectrometry", *Anal. Biochem.* **276,** 150 (1999).
4. X.R. Babu, A. Moradian and D.J. Douglas, "The methanol-induced conformational transitions of beta-lactoglobulin,

cytochrome *c* and ubiquitin at low pH: a study by electrospray Ionization mass spectrometry", *J. Am. Soc. Mass Spectrom.* **12,** 317 (2001).
5. D. Canet, A.M. Last, P. Tito, M. Sunde, A. Spencer, D.B. Archer, C. Redfield, C.V. Robinson and C.M. Dobson, "Local cooperativity in the unfolding of an amyloldogenic valiant of human lysozyme", *Nat. Struct. Biol.* **9,** 308 (2002).
6. D.A. Simmons and L. Konermann, "Characterization of transient protein folding intermediates during myoglobin reconstitution by time-resolved electrospray mass spectrometry with on-line isotopic pulse labeling", *Biochemistry* **41,** 1906 (2002).
7. A. Raza, H. Pan and D.L. Smith, "FFD exchange to study protein folding", *Encyclopedia of Mass Spectrometry* **2,** (2003).
8. Y. Deng, Z. Zhang and D.L. Smith, "Comparison of continuous and pulsed labeling amide hydrogen exchange/mass spectrometry for studies of protein dynamics", *J. Am. Soc. Mass Spectrom.* **10,** 675 (1999).
9. J.J. Englander, J.R. Rogero and S.W. Englander, "Protein hydrogen exchange studied by the fragment separation method", *Anal. Biochem.* **147,** 234 (1985).
10. S. Thorell, P. Gergely, Jr, K. Banki, A. Perl and G. Schneider, "The three-dimensional structure of human transaldolase", *FEBS Lett.* **475,** 205 (2000).
11. I. Schaaff, S. Hohmann and F.K. Zimmermann, "Molecular analysis of the structural gene for yeast transaldolase", *Eur. J. Biochem.* **188,** 597 (1990).
12. F. Lachaise, G. Martin, C. Drougard, A. Perl, M. Vuillaume, M. Wegnez, A. Sarasin and L. Daya-Grosjean, "Relationship between posttranslational modification of transaldolase and catalase deficiency in UV-sensitive repair-deficient xeroderma pigmentosum fibroblasts and SV40-transformed human cells", *Free Radic. Biol. Med.* **30,** 1365 (2001).
13. Y. Bai, J.S. Milne, L. Mayne and S.W. Englander, "Primary structure effects on peptide group hydrogen exchange", *Proteins: Struct. Funct. Genet.* **17,** 75 (1993).
14. G. Thévenon-Emeric, J. Kozlowski, Z. Zhang and D.L. Smith, "Determination of amide hydrogen exchange rates in peptides by mass spectrometry", *Anal. Chem.* **64,** 2456 (1992).
15. Y. Deng and D.L. Smith, "Identification of unfolding domains in large proteins by their unfolding rates", *Biochemistry* **37,** 6256 (1998).
16. T. Rojsajjakul, "Unfolding of alpha subunit of tryptophan synthase by hydrogen exchange/mass spectrometry", PhD thesis. University of Nebraska, Nebraska, USA (2002).
17. J. Chen and D.L. Smith, "Unfolding and disassembly of the chaperonin GroEL occurs via a tetradecameric intermediate with a folded equatorial domain", *Biochemistry* **39,** 4250 (2000).
18. L. Wang, H. Pan and D.L. Smith, "Hydrogen exchange-mass spectrometry: optimization of digestion conditions", *Mol. Cell Proteomics* **1,** 132 (2002).

Received: 30 May 2003
Revised: 19 September 2003
Accepted: 22 September 2003
Web Publication: 4 March 2004