Characterization of Transient Protein Folding Intermediates during Myoglobin Reconstitution by Time-Resolved Electrospray Mass Spectrometry with On-Line Isotopic Pulse Labeling†

Douglas A. Simmons and Lars Konermann*

Department of Chemistry, The University of Western Ontario, London, Ontario N6A 5B7, Canada

Received August 21, 2001; Revised Manuscript Received November 12, 2001

ABSTRACT: A novel technique for studying protein folding kinetics is presented. It is based on a continuous-flow setup that is coupled to an electrospray (ESI) mass spectrometer and allows initiation of a folding reaction, followed by isotopic pulse labeling. The protein is electrosprayed "quasi-instantaneously" after exposure to the deuterated solvent. This approach yields structural information from the ESI charge state distribution and from the H/D exchange levels of individual protein states, while at the same time noncovalent interactions can be monitored. This technique is used to study the reconstitution of holomyoglobin (hMb) from unfolded apomyoglobin (aMb) and free heme. MS/MS is used to establish that a short-lived folding intermediate with two heme groups attached represents a protein-bound heme dimer. This state appears to have a compactness close to that of native hMb; however, isotopic labeling indicates a significantly perturbed structure. Another intermediate is bound to a single heme group and shows a charge state distribution similar to that of unfolded aMb. Exchange levels exhibited by this state are lower than for unfolded aMb, indicating that fewer hydrogens are exposed to the solvent and/or that more of them are involved in hydrogen bonding. Native hMb leads to the formation of low charge state ions (hMb^{9+}, hMb^{8+}) and shows low exchange levels. However, early during reconstitution, a slightly unfolded form of the heme−protein complex contributes to the observed hMb^{9+} ions. A peak width analysis reveals that the structural heterogeneity of some of the observed protein species decreases as reconstitution proceeds.

The occurrence of partially folded transient intermediates during protein folding is well documented (*1*−*3*). Important insights into the folding mechanism can be gained from a structural characterization of these short-lived states (*1*, *4*, *5*). Many kinetic intermediates have lifetimes in the millisecond (*6*, *7*) or even microsecond (*8*, *9*) time range. Furthermore, the folding of single-domain proteins is often highly cooperative so that intermediate states do not become strongly populated (*10*). Therefore, it is often difficult to obtain structural information on these transient states. A variety of experimental approaches have been developed for monitoring the kinetics of protein folding reactions and for the structural characterization of short-lived intermediates. Techniques that use optical triggers for initiating protein folding can provide a time resolution on the order of nanoseconds to microseconds (*11*−*14*). Continuous-flow mixing experiments allow the examination of folding events in the range from 50 μs to several milliseconds (*8*, *15*). This provides an effective complement to stopped-flow methods which are commonly used for studies in the millisecond to second time range (*9*, *16*, *17*). All of these techniques have significantly advanced our understanding of early folding events. However, the optical detection methods used in many of these experiments provide only limited information on the structure of transient folding intermediates.

Hydrogen/deuterium (H/D) exchange techniques are among the most powerful tools available to protein scientists. Labile hydrogens located on the protein backbone and on amino acid side chains can be exchanged with deuterium from the solvent (or vice versa). The rate of exchange is sensitive to protein structure; labile hydrogens that are engaged in hydrogen bonding or that are sterically protected from the solvent exchange very slowly (*18*, *19*). Thus, the degree of hydrogen exchange incurred by a protein during exposure to a deuterated solvent can be used as a probe of protein conformation. The application of H/D exchange in conjunction with quench-flow mixing and multidimensional NMR spectroscopy has been particularly fruitful for studying folding kinetics (*6*, *7*, *20*, *21*). These experiments allow structural information to be obtained in a residue-specific manner and with submillisecond time resolution. Quench-flow pulse-labeling strategies have also been used in conjunction with electrospray ionization mass spectrometry (ESI MS)[1] (*19*, *22*−*24*). Whereas NMR monitors average exchange levels at individual sites across the entire protein population, ESI MS is able to distinguish coexisting subpopulations that exhibit different H/D exchange levels (*22*, *25*).

† Supported by the Canada Foundation for Innovation, the Ontario Research and Development Challenge Fund, The Natural Sciences and Engineering Research Council of Canada, and The University of Western Ontario.

* Corresponding author. Phone: (519) 661-2111 ext 86313. Fax: (519) 661-3022. E-mail: konerman@uwo.ca. WWW: http://publish.uwo.ca/~konerman/.

10.1021/bi011697j CCC: $22.00
Published on Web 01/16/2002

ESI MS (*26*, *27*) has emerged as a tool for the investigation of protein structure only within the past decade or so. ESI produces intact and multiply charged (usually protonated) gas-phase ions from proteins in solution. It is well documented that unfolded proteins form higher protonation states than those in folded conformations (*28*−*38*). The physical basis of this relationship is still a matter of debate (*39*). There is substantial evidence that gas-phase protein ions retain at least some of their solution structure (*40*−*42*); i.e., unfolded proteins in solution will generally produce more extended conformations in the gas phase. These extended structures have a greater capacity to stabilize protons that become attached to basic sites on the protein surface. Most likely, this occurs through intramolecular solvation and through the greater spatial separation of individual charges, which reduces the Coulombic energy of the gas-phase ions. In this sense, the ESI charge state distribution can be considered to be a probe of the overall "compactness" of a protein in solution (*38*). Because ESI is a very gentle ionization technique, it even allows the mass spectrometric analysis of noncovalent complexes (*43*−*47*). Thus it is possible to monitor changes in the conformation of a protein while at the same time studying the interactions with metal ions, prosthetic groups, peptides, or other proteins (*29*, *48*, *49*). Recent work has shown that ESI MS in combination with on-line mixing devices can be used for monitoring protein folding kinetics (*50*, *51*). This approach, termed "time-resolved" ESI MS, allows production of ESI mass spectra at various times (from several milliseconds to seconds) following initiation of a reaction. It enables tracking of protein charge state distribution and noncovalent interactions as a function of reaction time, such that kinetic and mechanistic data can be generated (*50*−*52*). Similar techniques were previously used to study the kinetics of enzymatic (*53*, *54*) and bioorganic reactions (*55*, *56*).

Holomyoglobin (hMb) is an oxygen storage protein consisting of 153 amino acid residues. The native protein comprises eight α-helices. It adopts a largely spherical tertiary structure and forms a hydrophobic binding pocket into which heme (iron protoporphyrin IX) is noncovalently bound. Heme binding occurs through van der Waals interactions, coordination of the central heme iron with the proximal histidine (His^{93}), and hydrogen bonds involving the two heme propionate groups (*57*). Folding and unfolding processes of myoglobin have been the subject of numerous studies (see, e.g., refs *24* and *58*−*63* and references cited therein). However, surprisingly little is known about the reconstitution mechanism of this heme−protein complex, during which native hMb is formed from free heme and unfolded apomyoglobin (aMb). Initial studies led to the conclusion that reconstitution is a simple second-order reaction (*64*, *65*). However, more recent work indicates the occurrence of several kinetic intermediates (*66*). The utility of ESI MS for monitoring reconstitution under equilibrium conditions was demonstrated by Feng and Konishi (*67*). These authors reported that a transition from acidic to slightly basic pH results in a two-step reconstitution process with unfolded aMb first refolding to a compact conformation, followed by binding of the apoprotein to the heme group.

In a recent study, our group has used time-resolved ESI MS for studying the reconstitution kinetics of hMb (*52*). Native hMb was briefly exposed to acidic conditions, resulting in unfolding of the protein and disruption of the heme−protein complex. Subsequently, the solution was neutralized, and the reconstitution process was monitored as a function of time. The time-resolved ESI MS data revealed a surprisingly complex reaction mechanism that could be described in a multipathway model. For the majority of the proteins, reconstitution proceeds through two consecutive steps, i.e., refolding of aMb into a relatively compact state, followed by binding to a heme group. The order of these two steps is reversed for a significant subpopulation of proteins. In addition, some highly unusual complexes were observed early during the reaction, where myoglobin was bound to two and even three heme groups.

The aim of the current study is to obtain additional structural information on the various intermediates that are involved in myoglobin reconstitution and to gain a better understanding of the overall reaction mechanism. A novel ESI MS-based experimental approach is employed that allows initiation of protein folding and aging of the reaction mixture on the millisecond time scale, followed by on-line isotopic pulse labeling. The reaction mixture is electrosprayed for mass analysis immediately following the labeling pulse, thus obviating the need for a separate quenching step. This technique provides structural information based on the ESI charge state distribution, the heme binding state of the protein, and its H/D exchange behavior.

EXPERIMENTAL PROCEDURES

Materials. Horse skeletal myoglobin, equine hemin chloride, butanone, and bradykinin (Sigma, St. Louis, MO), deuterium oxide (Cambridge Isotope Laboratories, Inc., Andover, MA), acetic acid and ammonium hydroxide (Fisher Scientific, Nepean, ON), and hydrochloric acid (Caledon Laboratories, Georgetown, ON) were used without further purification. aMb was prepared by butanone extraction (*32*, *68*). pH and pD values were measured with an AB15 pH meter (Fisher Scientific, Nepean, ON). Reported measurements are as read directly from the pH meter, uncorrected for isotopic effects (*69*, *70*). Readings for D_2O-containing solvents are referred to as pD.

Time-Resolved ESI MS with On-Line H/D Exchange. A schematic of the apparatus used for this work is shown in Figure 1. Syringes 1 and 2 are connected via fused silica capillaries (TSP075150, 75 μm i.d.; Polymicro Technologies, Phoenix, AZ) to a mixing tee (M1, dead volume 29 nL; Upchurch Scientific, Oak Harbor, WA) which empties into a "reaction capillary" of the same i.d. At a second mixer (M2, constructed in-house, dead volume ≈3 nL), the reaction capillary joins another capillary that is connected to syringe 3. The outlet of mixer M2 is a final "labeling capillary" that leads to the ESI source. Reconstitution time was controlled by altering the length of the reaction capillary. All three syringes were advanced simultaneously by syringe pumps (Model 22; Harvard Apparatus, South Natick, MA). The contents of syringe 1 (200 μM aMb in 1% acetic acid at 14

[1] Abbreviations: aMb, apomyoglobin; aMb^{n+}, gas-phase protein ion $[aMb + nH]^{n+}$; E_{HX}, relative hydrogen exchange level of protein ions; ESI, electrospray ionization; fwhm, full width at half-maximum; hMb, holomyoglobin; MS, mass spectrometry; i.d., inner diameter; MS/MS, tandem mass spectrometry; 2hMb, myoglobin state with two heme groups attached; *t*, time elapsed after initiation of reconstitution.

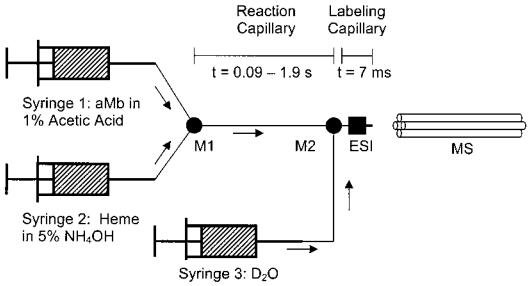

FIGURE 1: Schematic of the continuous-flow capillary apparatus used for this study. Heme in a solution of 5% (v/v) concentrated NH_4OH and acid denatured aMb in 1% (v/v) acetic acid are mixed at mixer M1, where reconstitution is initiated. At mixer M2, the reaction mixture is diluted with D_2O for isotopic pulse labeling. On-line analysis by ESI MS occurs immediately after the second mixing step.

μL/min) and syringe 2 (55 μM heme in 5% NH_4OH at 56 μL/min) were mixed at M1 for a flow rate of 70 μL/min and pH 10.4 and a stoichiometric heme:protein ratio of 1.1. Previous work has shown that a pH in this range does not have detrimental effects on the reconstitution of the heme−protein complex (*52*). Efficient mixing at M1 has been demonstrated by comparing time-resolved ESI MS data with results from stopped-flow spectroscopy (*50*, *51*, *54*). At M2 the reaction mixture was diluted with D_2O from syringe 3, resulting in a total flow rate of 350 μL/min and pD of 10.2. The final solution was comprised of 79.9% D_2O (taking into account the protons contributed by H_2O, acetic acid, and ammonium hydroxide). A small peptide, bradykinin, was chosen as a test compound to determine the extent to which unprotected labile hydrogens were exchanged during the labeling pulse. All labile hydrogens in bradykinin are accessible to the solvent and should be readily exchangeable (*71*). After labeling, the peptide showed an exchange level of 0.778, slightly lower than the expected 0.799 (data not shown). This small discrepancy may be due to rapid back-exchange of deuterated sites with residual water vapor in the ion source region (*71*). This control experiment confirms that the labeling conditions are adequate for unprotected hydrogens to exchange with an efficiency close to unity (0.778:0.799 = 0.974).

Reconstitution times from 90 ms to 1.9 s, corresponding to reaction capillary lengths between 2.3 and 50.0 cm, were studied by using the continuous-flow setup described above. Reaction capillaries longer than 50 cm could not be used due to excessive back-pressure that caused leaking at the capillary connections. Data for reconstitution times longer than 1.9 s were obtained from manual mixing experiments. H/D exchange was carried out at basic pD (10.2) to ensure that unprotected hydrogens on the protein were exchanged on the millisecond time scale (*6*, *19*). The exchange was allowed to proceed for roughly 7 ms during transfer through the labeling capillary. Subsequently, the reaction mixture was electrosprayed for mass analysis. To accommodate a short (0.85 cm) labeling capillary, a custom ESI source was constructed similar to that described in ref *30*. Ions were generated by pneumatically assisted ESI in the positive ion mode at a voltage of 10 kV. Dry air was used as the nebulizer gas. An additional air flow of 10 L/min at room temperature ("turbo gas") was applied to aid desolvation; this also served to reduce the water vapor which could be present in the ion source region. All data were recorded on an API 365 triple quadrupole mass spectrometer (SCIEX, Concord, ON). Relatively soft desolvation conditions had to be used (orifice 15 V, ring electrode 150 V) to prevent fragmentation of hMb ions in the gas phase (*72*). At least 15 scans were averaged for mass spectra of nonlabeled myoglobin by using a dwell time of 1 ms and a step size of 0.1 *m*/*z* unit. Pulse-labeling data were recorded by scanning a window of 12 *m*/*z* units centered around each peak maximum, with a step size of 0.2 *m*/*z* unit and a dwell time of 10 ms. Typically 100 scans were averaged for each peak. For experiments without isotopic labeling, syringe S3 contained water instead of D_2O. Collision potentials for MS/MS were in the range of 10−120 V. Nitrogen was used as the collision gas.

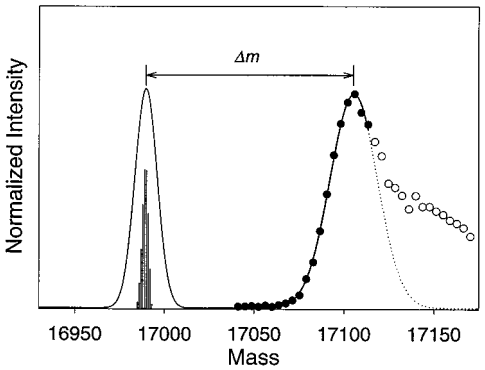

FIGURE 2: Calculation of the mass shift due to isotopic pulse labeling, illustrated for aMb^{19+} at t = 90 ms. Experimental data used for determination of the Gaussian fit (solid line, right peak in the figure) are shown as solid circles. Data shown as open circles were not taken into account for the analysis. The left peak in the figure has been calculated by convoluting a Gaussian template peak with the expected binomial distribution of charge carriers (i.e., protons and deuterons, shown as bar diagram). The mass shift, Δm, due to isotopic pulse labeling corresponds to the difference between the two peak maxima.

Data Analysis. The observed mass shift of protein ions is due to a combination of two effects: (i) isotopic exchange of hydrogens located along the protein backbone and in amino acid side chains and (ii) the fact that deuterons, as well as protons, can act as charging agents during ESI. Structural information is provided only by the former contribution. The latter is a purely statistical consequence of the isotopically impure nature of the solvent; it leads to the formation of $[M + kD + (n - k)H]^{n+}$ ions, where k = 0, ..., n. The probability of acquiring k deuterons and $n - k$ protons is given by $p_k{}^n = \binom{n}{k} p^k(1 - p)^{n-k}$, where $\binom{n}{k}$ is the binomial coefficient and p = 0.799. The procedure used to determine the exchange levels of individual protein species is illustrated in Figure 2. Initially, the most intense peaks in the mass spectrum of unlabeled aMb were averaged. The resulting average peak was fit to a Gaussian curve with a maximum of 16955 Da and an fwhm of 13.9 Da. This curve was convoluted with the expected binomial distribution of H^+ and D^+, resulting in simulated peaks that are broadened and shifted to higher mass due to the attachment of a combination of protons and deuterons. Protein peaks measured after isotopic pulse labeling were converted to a mass scale by multiplying each *m*/*z* value by the corresponding charge number. Subsequently, the peaks were fitted to

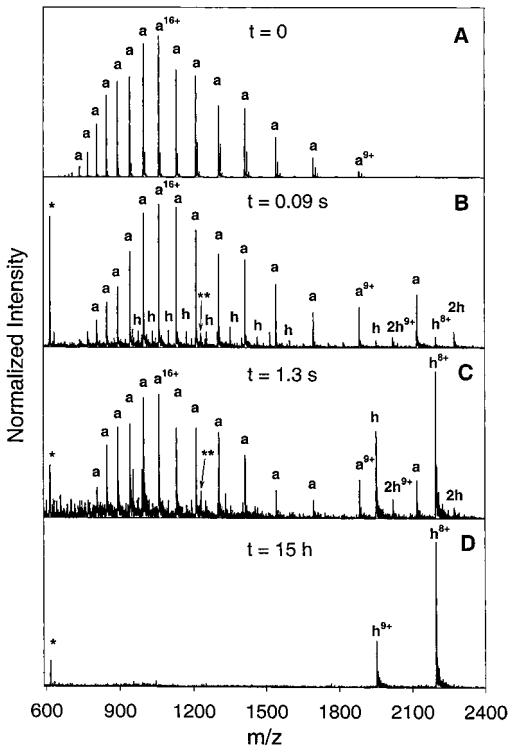

FIGURE 3: Time-resolved mass spectra recorded at selected times during myoglobin reconstitution, following a pH jump from 2.7 to 10.4. (A) Spectrum recorded prior to initiation of reconstitution. (B−D) Spectra recorded 90 ms, 1.3 s, and 1 h after initiation of reconstitution, respectively. The data were recorded without isotopic labeling. Notation: a, apomyoglobin (aMb); h, holomyoglobin (hMb); 2h, 2-holomyoglobin (2hMb); *, heme; **, heme dimer. The charge states of some selected protein ions are as indicated.

Gaussians. It is common during ESI MS to observe some peak tailing, caused by incomplete desolvation. Because of this effect, a number of data points on the high mass side of the measured peaks had to be discarded for the fitting procedure. The mass shift due to H/D exchange is given by the difference between the fitted Gaussian and the simulated protein peak.

aMb contains 153 amino acid residues with 262 exchangeable hydrogens. Of these, 148 are amide backbone hydrogens, 3 on the termini and 111 on the amino acid side chains (*25*). H/D exchange levels are reported as $E_{HX} = \Delta m/(262 \times 0.799)$, where Δm is the mass shift due to isotopic labeling. To simplify the graphic presentation of exchange levels, the same expression was used for all protein states, thus neglecting the two exchangeable hydrogens located on each heme group.

RESULTS AND DISCUSSION

Myoglobin Reconstitution Monitored by Time-Resolved ESI MS. Figure 3 shows ESI mass spectra recorded at different times after mixing heme and acid-unfolded aMb at pH 10.4. This initial set of experiments was performed without isotopic labeling. The data are very similar to those reported in our previous reconstitution study (*52*); they are shown here primarily to facilitate the discussion of the pulse-labeling experiments presented below. The spectrum depicted in Figure 3A represents a largely unfolded state of aMb under acidic conditions, prior to exposure to heme (*29*, *67*, *73*). It is characterized by a broad charge state distribution centered around aMb^{16+}. Following 90 ms of reconstitution (Figure 3B) this distribution remains a dominant feature of the spectrum; however, a significant portion of aMb now appears in low charge states (aMb^{8+}, aMb^{9+}), indicating the presence of a subpopulation of apoprotein that has adopted a more compact structure. Figure 3B also shows a significant contribution of hMb ions in high charge states (hMb^{10+}–hMb^{19+}), marking the presence of proteins that are bound to a heme group while maintaining a largely unfolded conformation (*50*, *51*). At the same time, the spectrum also shows hMb ions in low charge states (hMb^{8+}, hMb^{9+}) which are usually attributed to native hMb (*25*, *50*, *67*, *74*). In addition to the protein peaks, free heme (marked as *) and a heme dimer (marked as **) are observed in the spectrum. Heme in basic solution is known to exist at least partly in the form of noncovalent dimers (*75*, *76*); the observation of heme dimers in the ESI mass spectrum is therefore not unexpected. Reconstitution also results in the transient formation of a myoglobin subpopulation which is bound to two heme groups ($2hMb^{8+}$, $2hMb^{9+}$). The low charge states formed for this "2hMb state" imply a relatively compact conformation. Figure 3C shows the spectrum recorded 1.3 s after initiation of reconstitution. By this time, the relative contribution from highly charged aMb has become noticeably weaker as more of the protein has folded to form hMb in a compact conformation (represented by hMb^{8+} and hMb^{9+}). This trend continues until, for long reconstitution times, only hMb^{8+} and hMb^{9+} ions are observed. As an example, Figure 3D shows the spectrum recorded for $t = 15$ h. It is virtually identical to that of nonreconstituted native hMb (data not shown). Similar data for native hMb have been reported previously (*25*, *29*, *50*, *67*, *74*). Also, the H/D exchange levels measured for the $t = 15$ h sample and for nonreconstituted native protein are almost identical (see next section). Therefore, it appears that the reconstituted protein at $t = 15$ h and nonreconstituted native hMb have a similar structure; they may even be indistinguishable. Overall, the ESI MS data in Figure 3 show that hMb reconstitution is a complex process involving at least three transient intermediates, namely, an unfolded form of hMb, a relatively compact form of aMb, and a state where the protein is bound to two heme groups.

Studying the Structure of 2hMb by Tandem Mass Spectrometry (MS/MS). Previous work (*52*) strongly indicates that myoglobin ions with two heme groups attached correspond to a "real" reconstitution intermediate that exists in solution; it is unlikely that these ions are formed as an artifact of the ESI process. Given the known tendency of heme to form noncovalent dimers under the solvent conditions used here (*75*, *76*), it seems feasible that the protein might be able to accommodate such a dimer in a distorted heme binding pocket. Alternatively, the heme groups could be present as unassociated monomers, possibly seated in separate locations on a misfolded protein. To elucidate the heme group configuration, tandem mass spectrometry was performed. Figure 4 shows an MS/MS spectrum that was obtained after the $2hMb^{8+}$ parent ion was fragmented at a collision potential of 60 V. These data were recorded at a reconstitution time of 280 ms. The dominant peak in the spectrum corresponds to monomeric heme at *m*/*z* 616 (the isotopic distribution confirms this to be a singly charged ion, thus ruling out the possibility of having a doubly charged dimer). The spectrum

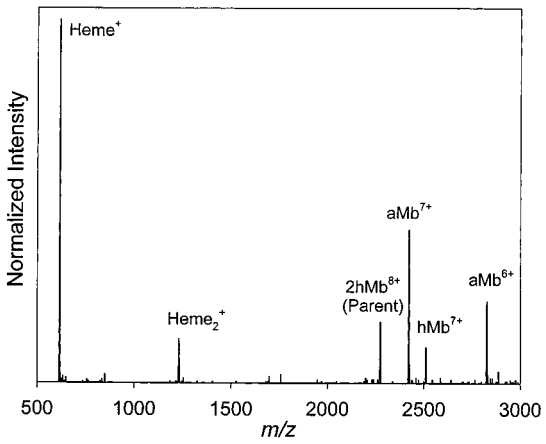

FIGURE 4: Fragment spectrum of $2hMb^{8+}$ (parent ion) at a collision potential of 60 V and a reconstitution time of 280 ms.

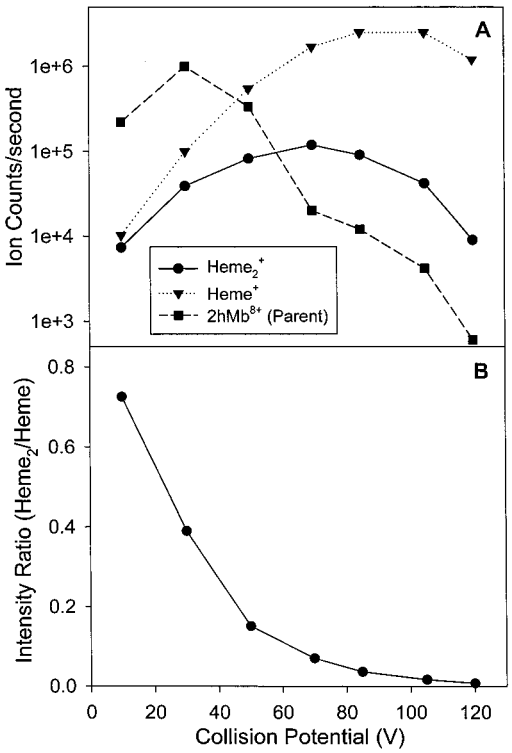

FIGURE 5: (A) Intensities of heme, heme dimer, and $2hMb^{8+}$ (parent ion) as a function of collision potential. (B) Plot of the heme dimer: heme monomer intensity ratio.

also shows a singly charged heme dimer and a number of protein ions (aMb^{7+}, hMb^{7+}, aMb^{6+}) as products of the fragmentation process. The information provided in this spectrum is somewhat inconclusive. The observation of a $heme_2^+$ ion shows that at least some of the 2hMb complexes exist in the form of a heme dimer that is bound to the protein. Also, monomeric heme could be a fragmentation product of such a complex; however, it could just as well be formed from protein ions that are bound to two separate heme groups.

Ramping of the collision potential while observing the intensities of $heme^+$ and $heme_2^+$ fragment ions (Figure 5A,B) provides some insight into the origin of the large $heme^+$ signal observed in the tandem mass spectrum. At high collision energy (120 V), the signal observed for $heme^+$ is roughly 2 orders of magnitude higher than that of $heme^{2+}$.

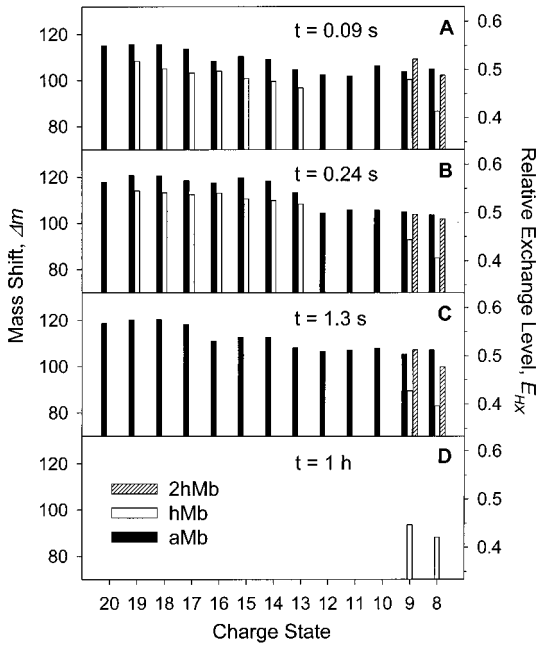

FIGURE 6: Deuterium uptake after isotopic pulse labeling observed for different charge and heme binding states. The reconstitution times for panels A−D are as indicated. Error bars have been omitted to simplify the graphic presentation. The errors calculated from the fitting program are generally less than 1 mass unit.

The relative contribution of the heme dimer increases as the collision potential is reduced. Intensities of comparable magnitude were observed for $heme^+$ and $heme^{2+}$ at a potential of 10 V. Unfortunately, experiments at even lower collision potential could not be carried out because the ion transmission was reduced to background levels. However, extrapolation of the observed data implies that more gentle fragmentation conditions would yield an even greater relative intensity for the heme dimer. The observed effects can be explained by assuming that, at high collision energies, dissociation of a heme dimer−protein complex generates thermally activated $heme_2^+$, which subsequently fragments to generate singly charged monomeric heme (and neutral heme which cannot be detected). At lower collision energy a smaller amount of excess energy is imparted to the $heme_2^+$ ions, thus increasing their lifetime in the mass analyzer and facilitating their detection. Some variations of this model are possible; e.g., collisional activation of the parent ion at high energy might first lead to fragmentation of a protein-bound heme dimer, followed by disruption of the heme−protein interactions. However, all of these scenarios require that most, if not all, of the 2hMb gas-phase ions represent proteins that are bound to heme dimers.

H/D Exchange. Figure 6 summarizes the hydrogen exchange levels (E_{HX}) for different reconstitution times. Selected peaks for $t = 240$ ms, together with the corresponding Gaussian fits, are shown in Figure 7. The extremely wide bimodal charge state distribution observed for aMb (see Figure 5B,C) represents a structurally heterogeneous protein population, ranging from relatively compact conformers all the way to substantially unfolded polypeptide chains (*50*). The most unfolded proteins in this population generate charge states around 20+. Exchange levels for these ions are just below 0.6 (Figure 6A,B), thus indicating that even for these "unfolded" states a significant portion of the labile sites is

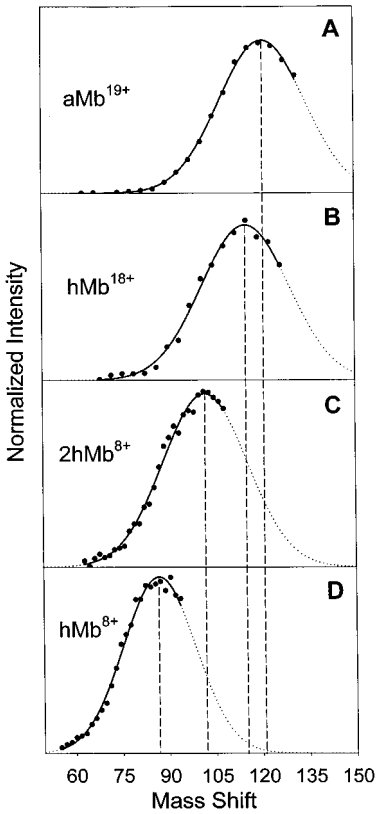

FIGURE 7: Mass shifts due to isotopic pulse labeling observed for protein ions of different charge states and heme binding states at t = 240 ms: (A) aMb^{19+}; (B) hMb^{18+}; (C) $2hMb^{8+}$; (D) hMb^{8+}. Solid lines are Gaussian fits to the experimental data.

protected from exchange. The exchange levels gradually drop off for aMb ions of lower charge states, to around 0.5 for aMb^{9+} and aMb^{8+}. These low charge state ions represent a subpopulation of aMb that has a compactness close to that of native hMb (*25*, *52*, *67*). However, the observed exchange levels are significantly higher than those observed for the corresponding hMb ions. Much of this effect may be due to the stabilizing effect that the heme group has on the overall structure of the holoprotein (*62*). The steric protection of labile sites by the heme is also likely to play a role.

Panels A and B of Figure 6 show that the exchange levels of ions representing the unfolded hMb intermediate (hMb^{12+}–hMb^{19+}) are around 0.05 lower than those of the corresponding aMb ions. Again, this effect can be attributed to the steric protection and/or stabilization of the protein by the heme group. However, the E_{HX} values are higher than for native hMb (Figure 6D). The observation of unfolded holoprotein during reconstitution has interesting implications. For azurin and α-lactalbumin it has been shown that binding of a cofactor to a largely unfolded state can be a key step of the folding process, resulting in the formation of an early intermediate with a nativelike cofactor binding site (*77*, *78*). In a similar fashion, some proteins act as "folding templates" for other proteins (*79*). The observation of an unfolded holoprotein intermediate in the current study suggests that the heme group acts as a folding template during myoglobin reconstitution.

The MS/MS data presented in this work indicate that 2hMb represents a protein-bound heme dimer. This complex leads to the formation of a charge state distribution similar to that

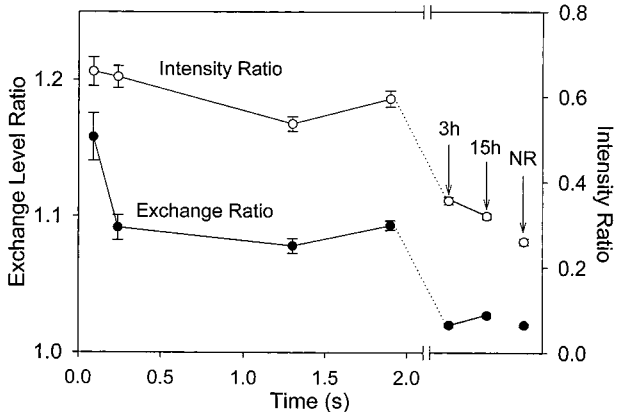

FIGURE 8: hMb^{9+}/hMb^{8+} exchange level ratio and hMb^{9+}/hMb^{8+} intensity ratio as a function of reconstitution time. Note that the time axis after the break is not linear. NR, data obtained for nonreconstituted native protein.

of native hMb (Figure 3). However, 2hMb exhibits an exchange level of roughly 0.5, which is substantially higher than that of the corresponding hMb ions. This finding is unexpected, given the apparent compactness of 2hMb and the likely steric protection of a significant part of the protein surface by the heme dimer. The exchange behavior exhibited by the 2hMb complex therefore indicates a significantly perturbed protein structure. Considering the likely role of monomeric heme as a folding template, we propose that heme dimers can play an analogous role. They may induce the formation of a misfolded protein structure, possibly by prompting a hydrophobic collapse of nonpolar side chains around the π-electron system of the porphyrin dimer.

hMb^{8+} consistently exhibits the lowest values of E_{HX} (around 0.4; Figure 6). The same E_{HX} was observed for hMb^{8+} ions generated from nonreconstituted native hMb (data not shown). The exchange level for hMb^{9+} early during reconstitution is significantly higher; e.g., $E_{HX} = 0.48$ for t = 90 ms (Figure 6A). As reconstitution proceeds, it approaches that of hMb^{8+} (around 0.4). This process is displayed in Figure 8, where the hMb^{9+}/hMb^{8+} exchange level ratio is plotted as a function of time. It drops from 1.16 for $t = 90$ ms to 1.03 for $t = 15$ h. The $t = 15$ h time point is almost identical to that obtained for nonreconstituted native hMb. Also shown in Figure 8 is the ratio of the hMb^{9+}/hMb^{8+} peak intensities as a function of time. Initially, the intensities of hMb^{9+} and hMb^{8+} are of similar magnitude, corresponding to a ratio of about 0.7. For $t = 15$ h, this ratio drops to 0.32, which is close to the value of 0.26 that was measured for nonreconstituted native hMb. The data depicted in Figure 8 clearly show that early during reconstitution there is a significant contribution of slightly unfolded hMb to the hMb^{9+} peak. It has a higher value of E_{HX} and leads to the formation of higher charge states than the proteins corresponding to hMb^{8+}. As the reaction proceeds, the relative concentration of this species decreases, as seen from the diminishing hMb^{9+}/hMb^{8+} peak intensity ratio.

Conformational Heterogeneity. After pulse labeling, a conformationally heterogeneous protein population will show a wider distribution of E_{HX} values than a homogeneous population (*24*, *80*). Consequently, while the heterogeneity of the overall protein compactness can be monitored through its ESI charge state distribution (*49*, *50*), peak widths of

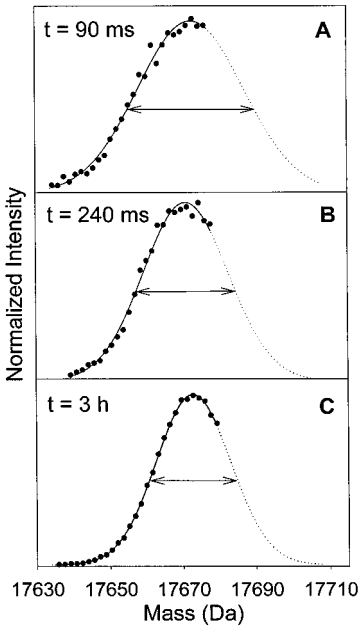

FIGURE 9: Mass spectral data and Gaussian fits (solid lines) measured after pulse labeling for hMb^{8+} after 90 ms (A), 240 ms (B), and 3 h (C) of reconstitution. Horizontal arrows represent the fwhms of the fitted peaks.

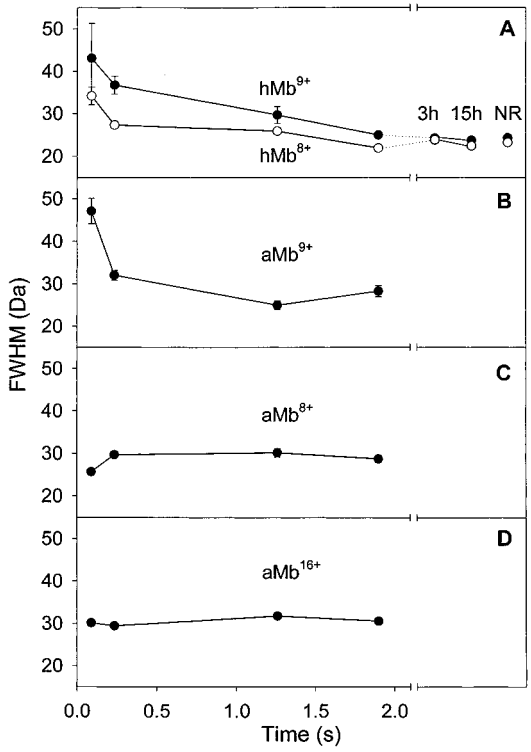

FIGURE 10: Peak widths (fwhm) for selected charge and heme binding states as a function of reconstitution time. Error bars represent uncertainties of the measured fwhms as determined by the fitting program. Note that the time axis after the break is not linear. NR, data obtained for nonreconstituted native protein.

isotopically labeled proteins can be used as a probe of conformational heterogeneity with respect to their exchange accessibility. Figure 9 shows ESI MS profiles for hMb^{8+}, obtained after pulse labeling, for reconstitution times of 90 ms, 240 ms, and 3 h. The widths (fwhm) of these three peaks are 34.1, 27.4, and 23.8 Da, respectively. A plot of the hMb^{8+} peak width as a function of time is shown in Figure 10A. Also included is a reference point measured for nonreconstituted native protein. It is almost identical to the value measured for $t = 15$ h. These data imply that hMb^{8+} ions observed early during reconstitution represent a nativelike, but structurally heterogeneous ensemble of heme−protein complexes in solution. The protein becomes structurally more homogeneous as reconstitution proceeds.

An even more pronounced decrease of the measured peak width is observed for hMb^{9+} (Figure 10A). This is consistent with the conclusions reached above, namely, that early during reconstitution a mix of nativelike and slightly unfolded heme−protein complexes contribute to this ion. The conformational heterogeneity decreases as the relative concentration of this slightly unfolded state is diminished (see the discussion of Figure 8). Previous studies have shown that freshly reconstituted hMb represents a mixture of proteins with their heme groups in a non-native orientation (rotated by 180° about the heme $\alpha-\gamma$-meso axis) and in the correct binding geometry (*81*−*84*). Subsequent reorientation of the inverted heme groups occurs on the time scale of minutes to hours (*82*). It is possible that this heme reorientation contributes to the effects depicted in Figure 10A. However, a decreasing peak width is also observed for the ions aMb^{9+}−aMb^{12+}, as exemplified in Figure 10B. Therefore, the conformational heterogeneity observed early during reconstitution cannot be solely attributed to factors that are related to heme binding.

Figure 10C shows that aMb^{8+} has a peak width that is relatively constant (Figure 10B). This is also the case for the higher charge states, aMb^{13+}−aMb^{20+}, as exemplified for aMb^{16+} in Figure 10D. Thus the structural heterogeneity of the corresponding protein states in solution undergoes only very little change. Reliable isotopic labeling data for the unfolded hMb intermediate could only be obtained for the two earliest time points. During this time interval, the observed peak widths of the corresponding peaks showed a notable decrease, e.g., from 50.1 Da for hMb^{19+} at 90 ms to 39.5 Da at 240 ms. Relatively constant widths around 33 Da were observed for the 2hMb peaks in the spectrum (data not shown). The most narrow peaks are observed for reconstituted hMb at $t = 15$ h and for nonreconstituted native protein, indicating that these heme−protein complexes represent a state with minimum structural heterogeneity.

CONCLUSIONS

The current study uses a novel approach where the reaction mixture is electrosprayed on-line and quasi-instantaneously after isotopic pulse labeling, without using chemical quenching. The ESI charge state distribution generated under these conditions reflects the various protein conformations that are populated at the time of the labeling pulse. The charge state distribution represents a structural probe that is complementary to the measured H/D exchange levels (*25*, *35*, *73*, *85*). In addition, information on the ligand binding state of the protein is obtained. H/D exchange levels of multiple coexisting species can be determined in a straightforward fashion, and information on the structural heterogeneity of individual protein subpopulations is obtained.

One limitation of this approach is that high concentrations of salts and buffers in the reaction mixture will interfere with

the ESI process (*86*). This problem is not encountered in traditional quench-flow experiments (*22*−*24*, *87*), because unwanted solvent additives can be removed prior to the off-line analysis of the reaction mixture. It is not entirely clear whether the technique used in this work is compatible with the use of chemical denaturants such as urea or guanidine hydrochloride. Any denaturant in syringe S1 will be diluted substantially by the time it reaches the ESI source. High-quality mass spectra of proteins have been reported for urea concentrations up to 0.7 M (*88*). For the conditions used in this work, an initial denaturant concentration of 8 M in S1 would be reduced to 0.32 M at the ion source. Even higher dilution factors could be obtained by reducing the flow rate from S1, relative to that from S2 and S3. Therefore, it appears that the approach used in this study could be compatible with the use of chemcial denaturants. However, it remains to be seen whether the quality of the ESI MS data obtained would be high enough for an accurate determination of H/D exchange levels under these conditions.

This work has shown that continuous-flow techniques can be used for studies on selected folding intermediates by tandem mass spectrometry. It is hoped that this MS/MS approach can be extended to obtain time-resolved *and* spatially resolved information on the H/D exchange pattern of selected folding intermediates, in a fashion that is comparable to 2D NMR experiments (*6*, *7*) and proteolytic digestion/fragment separation techniques (*19*, *87*). Data from a number of research groups indicate the feasibility of this approach (*89*−*94*). Efforts in this direction are currently under way in our laboratory.

REFERENCES

1. Rumbley, J., Hoang, L., Mayne, L., and Englander, S. W. (2001) *Proc. Natl. Acad. Sci. U.S.A. 98*, 105−112.
2. Baldwin, R. L., and Rose, G. D. (1999) *Trends Biol. Sci. 24*, 77−83.
3. Privalov, P. L. (1996) *J. Mol. Biol. 258*, 707−725.
4. Dill, K. A., and Chan, H. S. (1997) *Nat. Struct. Biol. 4*, 10−19.
5. Nymeyer, H., Socci, N. D., and Onuchic, J. N. (2000) *Proc. Natl. Acad. Sci. U.S.A. 97*, 634−639.
6. Roder, H., Elöve, G. A., and Englander, S. W. (1988) *Nature 335*, 700−704.
7. Udgaonkar, J. B., and Baldwin, R. L. (1988) *Nature 335*, 694−699.
8. Shastry, M. C. R., and Roder, H. (1998) *Nat. Struct. Biol. 5*, 385−392.
9. Roder, H., and Shastry, M. C. R. (1999) *Curr. Opin. Struct. Biol. 9*, 620−626.
10. Creighton, T. E. (1990) *Biochem. J. 270*, 1−16.
11. Ballew, R. M., Sabelko, J., and Gruebele, M. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 5759−5764.
12. Telford, J. R., Wittung-Stafshede, P., Gray, H. B., and Winkler, J. R. (1998) *Acc. Chem. Res. 31*, 755−763.
13. Chen, E., Wittung-Stafshede, P., and Kliger, D. S. (1999) *J. Am. Chem. Soc. 121*, 3811−3817.
14. Abbruzzetti, S., Viappiani, C., Small, J. R., Libertini, L. J., and Small, E. W. (2001) *J. Am. Chem. Soc. 123*, 6649−6653.
15. Shastry, M. C. R., Luck, S. D., and Roder, H. (1998) *Biophys. J. 74*, 2714−2721.
16. Englander, S. W., Sosnick, T. R., Mayne, L. C., Shtilterman, M., Qi, P. X., and Bai, Y. (1998) *Acc. Chem. Res. 31*, 737−744.
17. Aronsson, G., Brorsson, A.-C., Sahlman, L., and Jonson, B.-H. (1997) *FEBS Lett. 411*, 359−364.
18. Englander, S. W., Bai, L. M., and Sosnick, T. R. (1997) *Protein Sci. 6*, 1101−1109.
19. Smith, D. L., Deng, Y., and Zhang, Z. (1997) *J. Mass Spectrom. 32*, 135−146.
20. Raschke, T. M., and Marqusee, S. (1997) *Nat. Struct. Biol. 4*, 298−304.
21. Kuwata, K., Shastry, R., Cheng, H., Hoshino, M., Batt, C. A., Goto, Y., and Roder, H. (2001) *Nat. Struct. Biol. 8*, 151−155.
22. Miranker, A., Robinson, C. V., Radford, S. E., Aplin, R., and Dobson, C. M. (1993) *Science 262*, 896−900.
23. Hooke, S. D., Eyles, S. J., Miranker, A., Radford, S. E., Robinson, C. V., and Dobson, C. M. (1995) *J. Am. Chem. Soc. 117*, 7546−7547.
24. Tsui, V., Garcia, C., Cavagnero, S., Siuzdak, G., Dyson, H. J., and Wright, P. E. (1999) *Protein Sci. 8*, 45−49.
25. Wang, F., and Tang, X. (1996) *Biochemistry 35*, 4069−4078.
26. Fenn, J. B., Mann, M., Meng, C. K., Wong, S. F., and Whitehouse, C. M. (1989) *Science 246*, 64−71.
27. Loo, J. A. (1995) *Bioconjugate Chem. 6*, 644−665.
28. Chowdhury, S. K., Katta, V., and Chait, B. T. (1990) *J. Am. Chem. Soc. 112*, 9012−9013.
29. Katta, V., and Chait, B. T. (1991) *J. Am. Chem. Soc. 113*, 8534−8535.
30. Konermann, L., Collings, B. A., and Douglas, D. J. (1997) *Biochemistry 36*, 5554−5559.
31. Konermann, L. (1998) *Sci. Prog. 81*, 123−140.
32. Konermann, L., and Douglas, D. J. (1998) *Rapid Commun. Mass Spectrom. 12*, 435−442.
33. Konermann, L., and Douglas, D. J. (1998) *J. Am. Soc. Mass Spectrom. 9*, 1248−1254.
34. Mirza, U. A., Cohen, S. L., and Chait, B. T. (1993) *Anal. Chem. 65*, 1−6.
35. Wagner, D. S., and Anderegg, R. J. (1994) *Anal. Chem. 66*, 706−711.
36. Przybylski, M., and Glocker, M. O. (1996) *Angew. Chem., Int. Ed. Engl. 35*, 806−826.
37. Loo, J. A., Edmonds, C. G., Udseh, H. R., and Smith, R. D. (1990) *Anal. Chem. 62*, 693−698.
38. Miranker, A. D. (2000) *Proc. Natl. Acad. Sci. U.S.A. 97*, 14025−14027.
39. Fenn, J. B. (1993) *J. Am. Soc. Mass Spectrom. 4*, 524−535.
40. Jarrold, M. F. (1999) *Acc. Chem. Res. 32*, 360−367.
41. Sullivan, P. A., Axelsson, J., Altmann, S., Quist, A. P., Sunqvist, B. U. R., and Reimann, C. T. (1996) *J. Am. Soc. Mass Spectrom. 7*, 329−341.
42. Hunter, C. H., Mauk, A. G., and Douglas, D. J. (1997) *Biochemistry 36*, 1018−1025.
43. Loo, J. A. (1997) *Mass Spectrom. Rev. 16*, 1−23.
44. Loo, J. A. (2000) *Int. J. Mass Spectrom. 200*, 175−186.
45. van Berkel, W. J. H., van den Heuvel, R. H. H., Versluis, C., and Heck, A. J. R. (2000) *Protein Sci. 9*, 435−439.
46. Fändrich, M., Tito, M. A., Leroux, M. R., Rostom, A. A., Hartl, F. U., Dobson, C. M., and Robinson, C. V. (2000) *Proc. Natl. Acad. Sci. U.S.A. 97*, 14151−14155.
47. Bruce, J. E., Smith, V. F., Liu, C., Randall, L. L., and Smith, R. D. (1998) *Protein Sci. 7*, 1180−1185.
48. Nemirovskiy, O. V., Ramanathan, R., and Gross, M. L. (1997) *J. Am. Soc. Mass Spectrom. 8*, 809−812.
49. Vis, H., Heinemann, U., Dobson, C. M., and Robinson, C. V. (1998) *J. Am. Chem. Soc. 120*, 6427−6428.
50. Sogbein, O. O., Simmons, D. A., and Konermann, L. (2000) *J. Am. Soc. Mass Spectrom. 11*, 312−319.
51. Konermann, L., Rosell, F. I., Mauk, A. G., and Douglas, D. J. (1997) *Biochemistry 36*, 6448−6454.
52. Lee, V. W. S., Chen, Y.-L., and Konermann, L. (1999) *Anal. Chem. 71*, 4154−4159.
53. Paiva, A. A., Tilton, R. F., Crooks, G. P., Huang, L. Q., and Anderson, K. S. (1997) *Biochemistry 36*, 15472−15476.
54. Zechel, D. L., Konermann, L., Withers, S. G., and Douglas, D. J. (1998) *Biochemistry 37*, 7664−7669.
55. Sam, J. W., Tang, X. J., and Peisach, J. (1994) *J. Am. Chem. Soc. 116*, 5250−5256.
56. Sam, J. W., Tang, X. J., Magliozzo, R. S., and Peisach, J. (1995) *J. Am. Chem. Soc. 117*, 1012−1018.

57. Evans, S. V., and Brayer, G. D. (1990) *J. Mol. Biol. 213*, 885−897.
58. Gilmanshin, R., Dyer, R. B., and Callender, R. H. (1997) *Protein Sci. 6*, 2134−2142.
59. Jennings, P. A., and Wright, P. E. (1993) *Science 262*, 892−896.
60. Eliezer, D., Jennings, P. A., Dyson, H. J., and Wright, P. E. (1997) *FEBS Lett. 417*, 92−96.
61. Hughson, F. M., Wright, P. E., and Baldwin, R. L. (1990) *Science 249*, 1544−1548.
62. Johnson, R. S., and Walsh, K. A. (1994) *Protein Sci. 3*, 2411−2418.
63. Piro, M. C., Militello, V., Leone, M., Gryczynski, Z., Smith, S. V., Brinigar, W. S., Cupane, A., Friedman, F. K., and Fronticelli, C. (2001) *Biochemistry 40*, 11841−11850.
64. Adams, P. A. (1976) *Biochem. J. 159*, 371−376.
65. Adams, P. A. (1977) *Biochem. J. 163*, 153−158.
66. Kawamura-Konishi, Y., Kihara, H., and Susuki, H. (1988) *Eur. J. Biochem. 170*, 589−595.
67. Feng, R., and Konishi, Y. (1993) *J. Am. Soc. Mass Spectrom. 4*, 638−645.
68. Teale, F. W. J. (1959) *Biochim. Biophys. Acta 35*, 543.
69. Zhang, Z., Post, C. B., and Smith, D. L. (1996) *Biochemistry 35*, 779−791.
70. Englander, J. J., Rogero, J. R., and Englander, S. W. (1985) *Anal. Biochem. 147*, 234−244.
71. Katta, V., and Chait, B. T. (1993) *J. Am. Chem. Soc. 115*, 6317−6321.
72. Collings, B. A., and Douglas, D. J. (1996) *J. Am. Chem. Soc. 118*, 4488−4489.
73. Babu, K. R., Moradian, A., and Douglas, D. J. (2001) *J. Am. Soc. Mass Spectrom. 12*, 317−328.
74. Li, Y.-T., Hsieh, Y.-L., Henion, J. D., and Ganem, B. (1993) *J. Am. Soc. Mass Spectrom. 4*, 631−637.
75. Gallagher, W. A., and Elliott, W. B. (1968) *Biochem. J. 108*, 131−136.
76. Kuzelova, K., Mrhalova, M., and Hrkal, Z. (1997) *Biochim. Biophys. Acta 1336*, 497−501.
77. Pozdnyakova, I., and Wittung-Stafshede, P. (2001) *J. Am. Chem. Soc. 123*, 10135−10136.
78. Kuwajima, K., Mitani, M., and Sugai, S. (1989) *J. Mol. Biol. 206*, 547−561.
79. Li, M., Phylip, L. H., Lees, W. E., Winther, J. R., Dunn, B. M., Wlodawer, A., Kay, J., and Gustchina, A. (2000) *Nat. Struct. Biol. 7*, 113−117.
80. Chung, E. W., Nettleton, E. J., Morgan, C. J., Gross, M., Miranker, A., Radford, S. E., Dobson, C. M., and Robinson, C. V. (1997) *Protein Sci. 6*, 1316−1324.
81. La Mar, G. N., Davis, N. L., Parish, D. W., and Smith, K. M. (1983) *J. Mol. Biol. 168*, 887−896.
82. La Mar, G. N., and Krishnamoorthi (1984) *J. Am. Chem. Soc. 106*, 6395−6401.
83. Aojula, H. S., Wilson, M. T., and Drake, A. (1986) *Biochem. J. 237*, 613−616.
84. Light, W. R., Rohlfs, R. J., Palmer, G., and Olson, J. (1987) *J. Biol. Chem. 262*, 46−52.
85. Babu, K. R., and Douglas, D. J. (2000) *Biochemistry 39*, 14702−14710.
86. Xu, N., Lin, Y., Hofstadler, S. A., Matson, D., Call, C. J., and Smith, R. D. (1998) *Anal. Chem. 70*, 3553−3556.
87. Yang, H., and Smith, D. L. (1997) *Biochemistry 36*, 14992−14999.
88. Le Blanc, J. C. Y., Guevremont, R., and Siu, K. W. M. (1993) *Int. J. Ion. Mass Spectrom. Ion Proc. 125*, 145−153.
89. Anderegg, R. J., Wagner, D. S., Stevenson, C. L., and Borchardt, R. T. (1994) *J. Am. Soc. Mass Spectrom. 5*, 425−433.
90. Miranker, A., Kruppa, G. H., Robinson, C. V., Aplin, R. T., and Dobson, C. M. (1996) *J. Am. Chem. Soc. 118*, 7402−7403.
91. Deng, Y., Pan, H., and Smith, D. L. (1999) *J. Am. Chem. Soc. 121*, 1966−1967.
92. Eyles, S. J., Speir, J. P., Kruppa, G. H., Gierasch, L. M., and Kaltashov, I. A. (2000) *J. Am. Chem. Soc. 122*, 495−500.
93. Akashi, S., and Takio, K. (2001) *J. Am. Soc. Mass Spectrom. 12*, 1247−1253.
94. Kim, M.-Y., Maier, C. S., Reed, D. J., and Deinzer, M. L. (2001) *J. Am. Chem. Soc. 123*, 9860−9866.

BI011697J