Protein Science (1997), 6:1316–1324. Cambridge University Press. Printed in the USA.

Hydrogen exchange properties of proteins in native and denatured states monitored by mass spectrometry and NMR

EVONNE W. CHUNG,[1] EWAN J. NETTLETON,[1] CHARLES J. MORGAN, MICHAEL GROß, ANDREW MIRANKER, SHEENA E. RADFORD,[2] CHRISTOPHER M. DOBSON, AND CAROL V. ROBINSON

Oxford Centre for Molecular Sciences, New Chemistry Laboratory, University of Oxford, South Parks Road, Oxford OX1 3QT, United Kingdom

(RECEIVED December 16, 1996; ACCEPTED March 11, 1997)

Abstract

The extent of deuterium labeling of hen lysozyme, its three-disulfide derivative, and the homologous α-lactalbumins, has been measured by both mass spectrometry and NMR. Different conformational states of the proteins were produced by varying the solution conditions. Alternate protein conformers were found to contain different numbers of ^2H atoms. Furthermore, measurement in the gas phase of the mass spectrometer or directly in solution by NMR gave consistent results. The unique ability of mass spectrometry to distinguish distributions of ^2H atoms in protein molecules is exemplified using samples prepared to contain different populations of ^2H-labeled protein. A comparison of the peak widths of bovine α-lactalbumin in alternate solution conformations but containing the same average number of ^2H atoms showed dramatic differences due to different ^2H distributions in the two protein conformers. Measurement of ^2H distributions by ESI-MS enabled characterization of conformational averaging and structural heterogeneity. In addition, a time course for hydrogen exchange was examined and the variation in distributions of ^2H atom compared with simulations for different hydrogen exchange models. The results clearly show that exchange from the native state of bovine α-lactalbumin at 15 °C is dominated by local unfolding events.

Keywords: hydrogen exchange, mass spectrometry, NMR, protein dynamics, protein structure

The introduction of electrospray ionization (ESI) to mass spectrometry (MS) has had a major impact on protein structural studies (Fenn et al., 1989). An important breakthrough in the development of ESI MS was the ability to study proteins from conditions in which the protein remains in its native state during introduction to the ESI source allowing structural studies of proteins using hydrogen exchange labeling techniques (Katta & Chait 1991; Miranker et al., 1993; Robinson et al., 1994; Yi & Baker 1996). The labeling process exploits the fact that proteins contain exchangeable protons at sites that include amides and certain side chains. Those sites involved in secondary structure and/or buried from solvent exchange slowly due to H-bonding interactions and sequestering from solvent. Thus, such labeling techniques are widely used in the study of both kinetic and thermodynamic properties of proteins (Hvidt & Nielsen 1966; Woodward et al., 1982; Baldwin, 1993; Bai et al., 1995).

Hydrogen-exchange monitored by ESI MS (HX-MS) has shown that a number of ^2H are protected from exchange with solvent protons in the native state of cytochrome C (Katta & Chait, 1991). HX-MS has also been used to study lysozyme during protein folding experiments (Miranker et al., 1993) and to compare the relative stabilities of ferrocytochrome c_2 and site-directed mutants (Jaquinod et al., 1996). Different exchange rates are observed for myoglobin in its *apo* and *holo* forms (Johnson & Walsh, 1994) and the solution dynamics of β-sheet and α-helical peptides have also been measured by HX-MS (Wagner et al., 1994). Furthermore, a two-state unfolding transition has been clearly demonstrated by the combination of ESI MS and NMR (Yi & Baker, 1996). All of the above studies involve intact proteins. However, proteolysis of labeled proteins has been shown to retain deuterons protected against exchange with solvent protons (Zhang & Smith, 1993; Zhang et al., 1996) allowing investigation of some site-specific properties.

The ability to study proteins and retain their hydrogen exchange labeling pattern in the native state enables protein folding studies by MS (Robinson & Radford, 1995; Miranker et al., 1996). When

Reprint requests to: Carol V. Robinson, Oxford Centre for Molecular Sciences, New Chemistry Laboratory, University of Oxford, South Parks Road, Oxford OX1 3QT, United Kingdom; e-mail: carolr@bioch.ox.ac.uk.

[1] E.J.N. and E.W.C. contributed equally to this work.

[2] Present address: Department of Biochemistry and Molecular Biology, University of Leeds, Leeds LS2 9JT.

combined with a time resolved technique, such as pulse labeling (Baldwin et al., 1993), in vitro protein folding pathways can be examined by ESI MS (Miranker et al., 1993; Eyles et al., 1994; Kragelund et al., 1995). These studies have provided important insight into the nature of folding intermediates and have highlighted the unique attributes of MS to study populations of molecules on protein folding pathways, and hence, to elucidate the cooperativity of folding events (Hooke et al., 1995). In addition to the study of native states of proteins, we have extended the capabilities of MS to examine less stable non-native states, for example, partially folded molten globule and denatured states (Kuwajima, 1989). We have shown, by comparison with proteins in non-native states, that the conformation of a protein ligand bound within the cavity of an 800 kDa molecular chaperone complex resembles a partially folded molten globule state (Robinson et al., 1994). In addition we have shown that hydrogen exchange protection is preserved in dihydrofolate reductase during iterative rounds of binding to GroEL (Groß et al., 1996). The success of these studies arises from the unique ability of MS to dissociate large weakly bound protein ligands in the gas phase where intermolecular hydrogen exchange can no longer take place.

The aims of the present study are twofold. First, to identify the experimental procedures necessary to ensure that the hydrogen exchange measured by MS and 1D ^{1}H-NMR under identical solution conditions are coincident within the errors of the two techniques. Second, to demonstrate the unique ability of HX-MS to measure directly collective hydrogen exchange properties of many labile sites. By comparison of peak widths of proteins in ESI mass spectra with those simulated for different hydrogen exchange models, information about conformational states of proteins is obtained and hydrogen exchange mechanisms are delineated.

Results and discussion

Hydrogen exchange in native states

The hydrogen exchange properties of the model globular protein hen lysozyme were examined. Hen lysozyme is a good model protein for comparative studies by MS and NMR because extensive studies by NMR have lead to a precise picture of the hydrogen exchange behavior of this protein (Pedersen et al., 1991; Radford et al., 1992). This protein possesses a total of 255 exchange labile protons including 126 from backbone amides, 126 from side chains, and 3 from the termini. The majority of the side-chain labile sites are on the surface of the protein, and NMR has confirmed that in solution these sites, as well as a subset of the backbone amides, exchange rapidly ($t^{1/2} = 10^2$ s) at pH*3.8 and 20 °C (Pedersen et al., 1991). By contrast, the remaining amides are protected from solvent in the native state resulting in exchange over a much longer time scale ($t^{1/2} = 10^4$–10^8 s) (Pedersen et al., 1991).

Typical conditions for operating ESI mass spectrometers involve introducing the protein solution into a delivery solvent, which is continuously infused into the ESI probe for analysis. In general mixtures of water with organic solvents, for example, 50% acetonitrile and 50% water containing 1% formic acid, are used as a delivery solvent and source temperature is set to 70 °C to aid desolvation of the electrospray droplets. Under these conditions, an average of only 18 amides remain protected against hydrogen exchange in hen lysozyme. Because most proteins are unable to tolerate the presence of organic solvent and extremes of pH and temperature, HX-MS of native states is carried out at more neutral pH, room temperature, and in the absence of organic solvents. The HX-MS kinetics of hen lysozyme from a 100% aqueous solution without heating are shown in Figure 1. Under these conditions the protein was in its native state up until, and possibly including, the time it enters the gas phase of the mass spectrometer. These experiments were carried out in ^{1}H solvent (pH5 and 22 °C) with protein in which all labile sites are pre-exchanged for ^{2}H and with protein containing natural abundance isotopes prior to exchange in ^{2}H solvent (pH*5 and 22 °C). Although the two data sets are similar, the rate of exchange monitored by ESI MS from ^{2}H-protein into H_2O is consistently slower than for ^{1}H-protein in D_2O under these solution conditions (pH*5 and 22 °C). Observed differences in rates can arise from a number of complex factors including changes in effective pH (Connelly et al., 1993), viscosity and diffusion rate (Itzhaki & Evans, 1996), and changes in protein stability (Makhatadze et al., 1995). Because the stability of proteins at this pH appears to be largely unchanged in D_2O and H_2O (Makhatadze et al., 1995), it seems likely that changes in the effective pH, and therefore, OH^- and OD^- concentrations, are responsible for the observed rate decrease in H_2O, from pH 5.4 in D_2O to pH 5.0 in H_2O (Bai et al., 1993). The NMR and MS data were found to be coincident within the estimated errors of the two techniques (±5 and ±2, respectively). The error in the NMR measurement is systematic and arises principally from solvent-induced baseline error, while the error in the MS measurements arises from the standard deviation of the charge states and represents the worst-case scenario in the data reported.

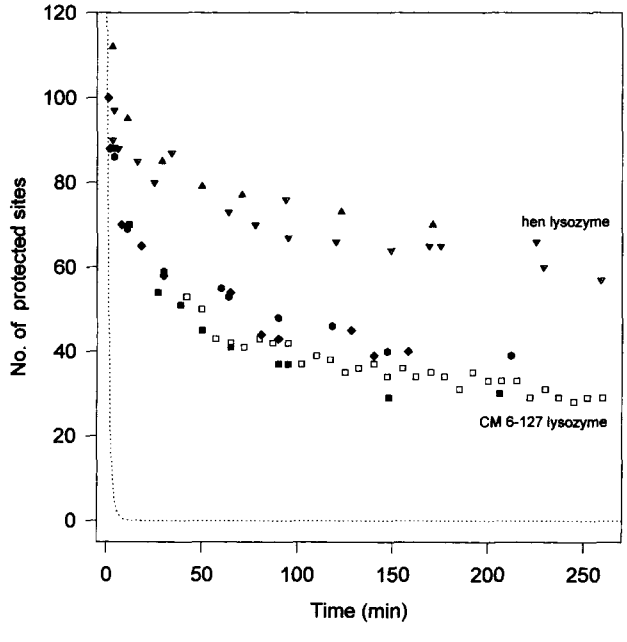

Fig. 1. Kinetic profiles of hydrogen exchange at pH5 and 22 °C for hen and CM^{6-127} lysozyme monitored by MS and 1D ^{1}H-NMR. Filled up triangle—^{2}H-hen lysozyme in H_2O and filled down triangle ^{1}H-hen lysozyme in D_2O both measured by MS and open down triangle ^{1}H-hen lysozyme in D_2O by 1D ^{1}H-N. MR. ■ ^{1}H-CM^{6-127} lysozyme in D_2O at pH*3.8 and 22 °C measured by MS and □ ^{1}H-CM^{6-127} lysozyme at pH*3.8 in D_2O and 22 °C measured by 1D ^{1}H-NMR. Filled diamond, ^{1}H-CM^{6-127} lysozyme in D_2O at pH*5 and room temperature measured by MS and filled hexagon ^{1}H-CM^{6-127} lysozyme in D_2O at pH*5 22 °C for exchange followed by pH jump to pH*2.9 and 4 °C prior to obtaining mass spectra. Data for the exchange profile of completely unstructured hen lysozyme and CM^{6-127} lysozyme at pH 5 and 22 °C calculated from near-neighbor inductive effects on acid base and water catalysis (Bai et al., 1993).

The native state of hen lysozyme contains a highly protected core with a significant number of backbone amides remaining after 6 h of exchange as observed by both 1D ^{1}H-NMR and ESI MS. Less stable proteins can also be measured by HX-MS. For example, reduction and carboxymethylation of the terminal disulfide bond (CM^{6-127} lysozyme), leads to a substantial reduction in stability. At pH 3.8, the native protein has a T_m of 77 °C (Wedin et al., 1982), while CM^{6-127} lysozyme has a T_m of 53 °C (Radford et al., 1991). In a previous NMR study (Eyles et al., 1994) individual rate constants were measured for 63 amide probes for CM^{6-127} lysozyme at pH*3.8 and 22 °C. A bulk exchange profile was simulated from the individual rate constants measured for the 63 individual amide probes (data not shown) and compared with MS measurements made at the same pH. While excellent agreement with the MS data was noted after the first 60 min of hydrogen exchange, the number of protected sites measured by MS during the first hour was consistently higher than that calculated from the simulated bulk exchange. This apparent discrepancy during the first hour arises because HX-MS measures the entire population of exchanging sites, which may include some slowly exchanging side chains, 2D NMR experiments monitored selected amides, in this case 63 of the 127 backbone amides present. HX-MS from solution in which CM^{6-127} lysozyme is in its native state (pH*3.8 and 22 °C) are shown in Figure 1. When compared with bulk exchange measurements by 1D ^{1}H NMR of CM^{6-127} lysozyme, under identical solution conditions (Fig. 1), the data sets are in agreement. In this comparison all labile sites are studied by both NMR and MS. HX-MS of CM^{6-127} lysozyme (pH*5 and 22 °C; Fig. 1) enables direct comparison with native lysozyme data under the same conditions. HX-MS reveals a reduction to about a half the number of protected amides observed in native lysozyme for CM^{6-127} lysozyme in accord with the reduced stability reported in previous studies (Radford et al., 1991; Eyles et al., 1994).

Because the majority of proteins exist in their native state only at or near pH 7.0 and the quality of ESI spectra is not optimal at this pH, the ability to measure hydrogen exchange from the native state and to lower the pH immediately prior to obtaining the ESI mass spectrum is particularly attractive. Under acidic conditions, below pH 3, CM^{6-127} has been shown to adopt a highly unstructured state (Radford et al., 1991) (see denatured states below). This reduction in stability of CM^{6-127} at low pH allows investigation of a pH jump experiment in which exchange is allowed to proceed from the native state at pH*5 then, immediately prior to obtaining the ESI mass spectrum, the pH of the protein solution is lowered. Ice-cooled acidified water is used to minimize exchange. The reduction in pH from pH*5 to pH*2.8 (meter reading after 10-fold dilution of pH*5 in pH*2) enhances the ESI mass spectra; in general, an increase in signal intensity is observed coupled with a reduction in metal adducts. The data from this pH jump experiment, when compared to data from native state exchange without the pH jump, reveal closely similar hydrogen exchange behavior (Fig. 1) and an increase in signal intensity by a factor of 10 in the ESI mass spectrum.

Bovine α-lactalbumin (BLA) shares the same overall fold as hen lysozyme and binds Ca^{++} in its native state with a high binding affinity (K_a = 2.7 × 10^6 M^{-1}) (Aramini et al., 1992). Other metal ions have also been shown to bind to BLA, and the existence of a second lower affinity Ca^{++} binding site (K_a = 3.1 × 10^4 M^{-1}) has been reported (Aramini et al., 1992). HX-MS of this metal binding protein in its native state is complicated, owing to the difficulty in distinguishing metal binding and deuterons present because both will contribute to the measured mass (Fig. 2a). Interestingly, the number of deuterons calculated for each peak in the mass spectrum was found to be identical irrespective of metal binding in the gas phase. When the protein is washed thoroughly and equilibrated at pH 5.0 in the presence of equimolar concentrations of Ca^{++}, the *apo* form is generally the predominant species observed (Fig. 2b). NMR studies in solution (V. Forge, pers. comm.) and HX-MS from protein in solution free from metal ions (Fig. 2c) have demonstrated that in the absence of Ca^{++} protection against exchange is decreased. Thus, while absence of Ca^{++} in solution has a dramatic effect, loss of Ca^{++} or addition of Na^+ in the gas phase has no effect on HX-MS. This observation suggests that loss of Ca^{++} or metal ion addition occurs late in the ESI process where hydrogen exchange can no longer take place due to the absence of solvent in gas phase protein molecule. This reinforces the observation that HX-MS reflects the solution conformation and is not affected by changes in binding observed in the gas phase of the mass spectrometer (Robinson et al., 1996). A stable core of sites, protected against hydrogen exchange, is observed for the native state of the protein, in the presence of Ca^{++}, after 6 h of exchange by both NMR and MS (Fig. 3).

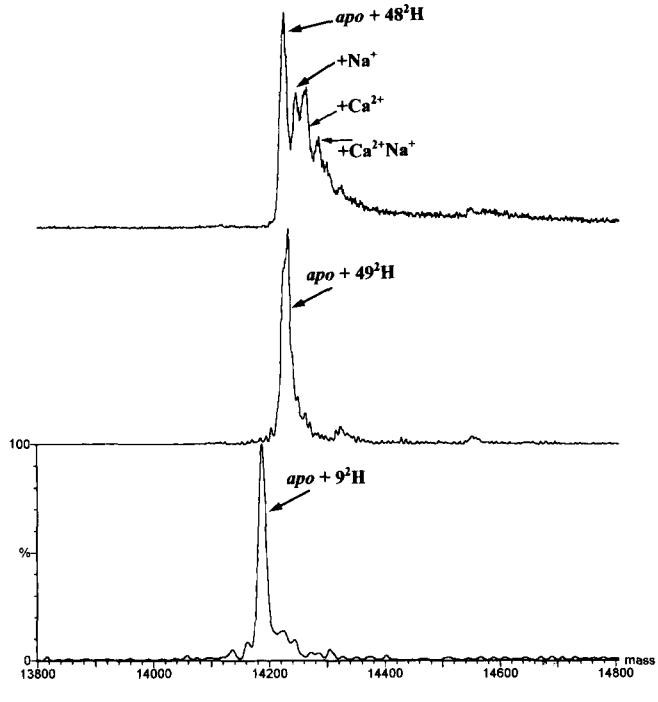

Fig. 2. Regions of the ESI mass spectra of ^{2}H-BLA in H_2O under a variety of different solution conditions at 22 °C. At pH5, without washing, 120 min after exchange was initiated, (**a**), after extensive washing of the protein on Centricon-10 (Amicon) and addition of an equimolar amount of Ca^{++}, 122 min after exchange was initiated, (**b**) and after removal of all traces of Ca^{++} at pH2, and 50 min after exchange was initiated (**c**). The multiplicity of the peak in (**a**) arises from metal binding to the molecular ion of the protein. The calculated masses for ^{1}H-BLA: 14,178, ^{1}H-BLA +1 Na^+: 14,200, ^{1}H-BLA +1 Ca^{++}: 14,216 and ^{1}H-BLA + 1 Ca^{++}, and 1 Na^+ 14,238 Da, measured masses were 14,226.5 ± 1, 14,249.1 ± 1.5, 14,264.0 ± 0.7, 14,285.3 ± 2 Da, respectively. The differences in mass observed in the ESI mass spectrum for each species arises from protection from exchange of 48 ± 1 sites occupied by ^{2}H under these solution conditions. The mass spectrum of BLA in the presence of equimolar Ca^{++} in solution, (**b**), shows essentially *apo* protein in the gas phase but has similar hydrogen exchange protection to that observed in (**a**) above. Under acidic conditions at pH2.0 in the absence of Ca^{++} only nine protected sites remain.

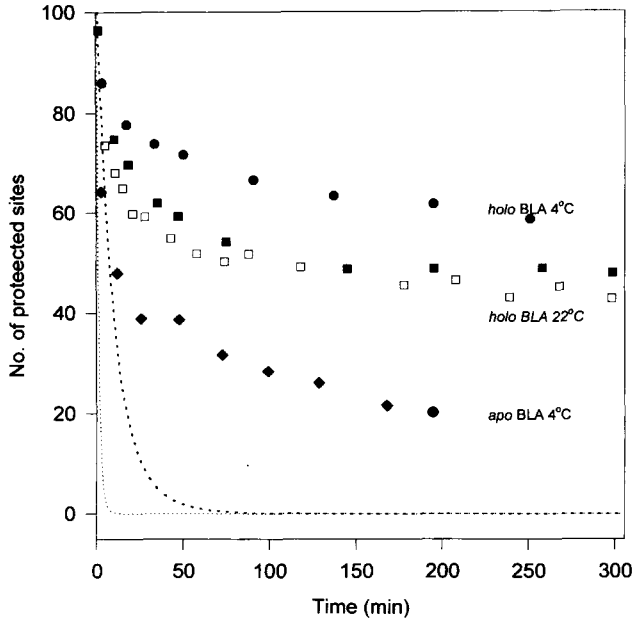

Fig. 3. Kinetic profiles of hydrogen exchange for ^{1}H-BLA diluted into D_2O at pH*3.8 under a variety of solution conditions and measured by MS and NMR. ● BLA at 4 °C with 1.7 molar excess of Ca^{++} and ■ BLA at 22 °C with 1.7 molar excess of Ca^{++} both measured by MS. Open square, BLA at 22 °C with 1.1 molar equivalent of Ca^{++} monitored by NMR. Filled diamond, apo BLA at pH 3.8 at 4 °C measured by MS. Dotted lines, data for the exchange profile of completely unstructured BLA at pH 3.8 at 22 °C (...) and 4 °C (—) calculated from near-neighbor inductive effects on acid base and water catalysis (Bai et al., 1993).

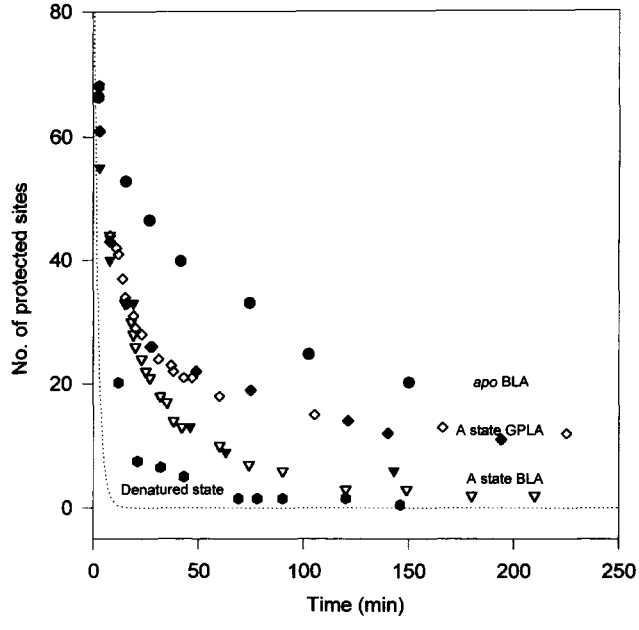

Fig. 4. GPLA and BLA molten globule states measured by NMR and ESI MS and compared with denatured states. ■ ^{2}H-BLA at pH2 on ice, monitored by MS. Filled down triangle, and open down triangle, ^{2}H-BLA in H_2O at pH2 and 20 °C, measured by MS and NMR, respectively. Filled diamond and open diamond, ^{2}H-GPLA in H_2O pH2 at 20 °C measured by MS and NMR, respectively. Filled hexagon, HX-MS of the denatured state derived from ^{2}H-$CM^{6\text{-}127}$ lysozyme in H_2O at pH2 and 20 °C and comparison with random coil data calculated from nearest neighbor inductive effects (Bai et al., 1993) for BLA at pH 2 and 4 °C.

Hydrogen exchange in denatured states

HX-MS of denatured states is complicated by their labile nature and conformational heterogeneity. Any increase in internal energy of the protein conformer during introduction into the ESI interface could lead to unfolding of denatured states and, if solvent is still present, loss of hydrogen exchange label. Previously, we have reported HX-MS data for BLA in its acid induced molten globule state (A state) at pH2 and room temperature (Robinson et al., 1994). In this study we report HX-MS data from a second molten globule state formed by removal of Ca^{++} to give *apo* BLA at pH3.8. Because the protein can be studied in the presence and absence of Ca^{++} at pH3.8, in its native and molten globule states respectively, conditions can be matched to enable direct comparison of these different protein conformers. The difference between these two states is clearly seen in Figure 3, the apo protein being much less stable than its *holo* counterpart.

In addition to states of BLA, we have also studied the A state formed from guinea pig α-lactalbumin (GPLA) at pH 2, which has been shown by NMR to exhibit higher levels of hydrogen exchange protection than the bovine protein (Chyan et al., 1993). In order to compare HX-MS data with NMR data already available for the A state of GPLA, hydrogen exchange was carried out at room temperature. This requires not only aqueous solvents and sample introduction without heating, but reduction of the temperature in the ESI interface by cooling nebulizer gas, delivery solvent, and protein solution before sample introduction. These precautions effectively quench exchange during introduction to the electrospray interface. The HX-MS data are shown in Figure 4, together with those for the bovine protein reported previously (Robinson et al., 1994) and demonstrate the increased protection of the GPLA A state compared with that derived from the bovine protein. Further comparison with existing NMR data, for both GPLA and BLA A states (Wormwald, 1991) (Fig. 4) confirm the extent of HX-MS.

Hydrogen exchange in an unstructured state is dictated by the intrinsic rate of exchange and can be simulated for the amino acid sequence of the protein using near-neighbor inductive effects on acid catalysis (Bai et al., 1993). Using $CM^{6\text{-}127}$ lysozyme at pH2 as a model for a highly unstructured state (Radford et al., 1992) and the conditions developed for HX-MS in partially folded states, the exchange rate for $CM^{6\text{-}127}$ lysozyme has been measured (Fig. 4). A simulation of the exchange rate profile for $CM^{6\text{-}127}$ lysozyme under these solution conditions, assuming no residual secondary structure, is also shown in Figure 4. Interestingly, HX-MS from the unfolded $CM^{6\text{-}127}$ lysozyme at pH2 is consistently slower than that predicted for a totally unstructured state, suggesting some residual structure remains protected against hydrogen exchange. This is in accord with recent studies on unstructured polypeptides and denatured proteins, which suggest that these states adopt an ensemble of conformers weakly protected against hydrogen exchange rather than an entirely random coil structure (Smith et al., 1996.)

Analysis of mass distribution

The peak width of a molecular ion in the ESI mass spectrum of a protein arises from the contribution of the natural abundance isotopes of the individual charge states, principally the contribution

from 1% of ^{13}C, and the resolution capabilities of the mass spectrometer. Hen lysozyme and *apo* BLA have similar molecular formulae ($C_{629}H_{956}N_{162}O_{195}S_9$, and $C_{613}H_{951}N_{193}O_{185}S_{10}$, respectively), and hence, show similar peak widths. A binomial distribution of these natural abundance isotopes is compared to these MS peaks (Fig. 5). Measurements of the MS peakwidths at half-height, 9.95 Da ± 0.10 Da for hen lysozyme and 10.26 Da ± 0.15 Da for *apo* BLA, and comparison with the theoretical peak width at half-height obtained from the binomial distribution, 7.55 Da, provides a measure of the instrument contribution to the peak width. Although the overall shape the peak should take is given by a binomial distribution, a Gaussian distribution of isotopes is compared with the raw data (Fig. 5) (Winkler et al., 1975). Although the base of the Gaussian distribution does not agree well with the raw data, the peak width at half-height is a close approximation, allowing convenient interpretation of peak widths (see below).

The peak shape of a protein containing ^2H provides information about the distribution of ^2H in the molecules. For example, lysozyme samples were prepared to contain 50% ^2H protein and 50% ^1H protein. While one sample was heated prior to analysis to produce a random distribution of ^2H throughout the protein, a second sample was not heated prior to analysis and retains distinct populations of protein molecules containing ^2H and ^1H (Fig. 6). These two samples are clearly distinct by ESI MS and would be difficult to distinguish by NMR methods because the average proton occupancy in the different samples is identical.

The heated sample containing 50% ^2H-protein and 50% ^1H-protein, has a probability of 0.5 of being occupied by ^2H at all 255 labile sites and results in significant broadening of the peak width in the ESI mass spectrum. The additional broadening of a mass spectrum, as a result of incorporation of deuterium, is a convolution of the natural abundance distribution of isotopes, shown in Figure 5, with the distribution of deuterium in the protein. A Gaussian approximation to these peaks is convenient as convolutions of Gaussians are themselves Gaussians; i.e., a Gaussian with mean m and width at half-height w, convoluted with a second Gaussian, mean m' and width at half-height w', yields a Gaussian with mean $m + m'$ and width $\sqrt{(w^2 + w'^2)}$. Furthermore, a set of n labile sites having a probability p of being occupied by ^2H, for large n, can be approximated by a Gaussian with mean $n \times p$ and with width at half-height $\sim 5np(1 - p)$. Thus, the above sample should have a mass shift of $255 \times 0.5 = 122.5$ and a width of $\sqrt{[9.95^2 + (5 \times 255 \times 0.5 \times 0.5)]} = 20.4$ Da, which compares well with the observed results (mass shift 120.5 and peak width 19.8 ± 0.20 Da).

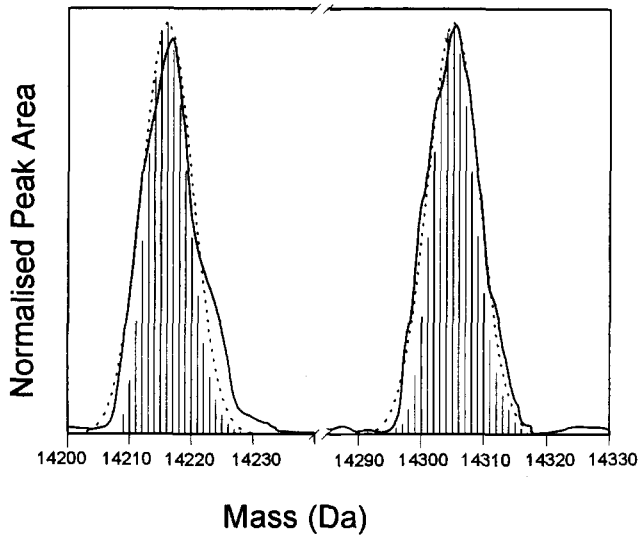

Fig. 5. A comparison of the molecular ions of *holo* BLA and hen lysozyme (continuous line) with a binomial distribution (vertical lines) and Gaussian distribution (dotted line) for protein containing natural abundance isotopes. The peak widths are calculated from the +11 charge states of both proteins, are multiplied by their charge and superimposed on a mass scale. ESI mass spectra were obtained from 100% aqueous solution, at pH3.8 and 22 °C, and under high resolution mass spectrometry conditions.

Prior to heating the sample containing 50% ^2H-lysozyme the populations are distinct because 50% of the molecules protect a core of ^2H while the remaining 50% protect a core of ^1H. Assuming that all exposed labile sites are labeled with 50% ^2H, the number of unprotected sites measured from the mass shift is 170 ($n \times 0.5 = 85$, $n = 170$). The peak widths of the two populations, however, are broader than predicted for the discrete populations containing a protected core (calculated 17.6 Da, observed 21.2 Da ± 0.20 Da). This broadening arises because the samples are mixed and analyzed immediately and thus continue to undergo hydrogen exchange, resulting in a shift in mass distribution, during analysis (~2 min). Similarly, mixtures of different proportions of ^2H-lysozyme in ^2H-solvent and ^1H-lysozyme in ^1H-solvent, either with or without heating, are presented in Figure 6b and c, and compared to simulations. These samples illustrate the potential for HX-MS to analyze ^2H and ^1H populations from masses and peak shapes in ESI mass spectra.

Changes in mass distribution can be clearly seen as a result of HX-MS of alternative protein conformers. For example, *apo* and *holo* BLA, in molten globule and native state conformations, respectively, containing an average of 158 ± sites occupied by ^2H are shown in Figure 7. The peak width for these protein conformers are remarkably different, 8.9 ± 0.2 Da for *holo* and 16.2 ± 0.2 Da for *apo*. In the native state, the protein's labile hydrogens are either solvent exposed or are within a protected core, while fluctuations in the molten globule state expose many sites for exchange leading to a wider distribution of masses. The *holo* BLA data are in agreement with the protected core model for $n = 158$ and with the hydrogen exchange kinetics under these conditions (data not shown). The *apo* BLA data, however, are not in agreement with either the protected core ($n = 158$) or random exchange model ($n = 220$). The number of sites remaining, obtained from the plateau of hydrogen exchange kinetics under these conditions (Fig. 3), shows 200 labile sites (20 protected sites). The peak width remains consistently broader than that predicted for 200 labile sites throughout the exchange time course. This broad distribution of mass suggests, therefore, that the molten globule structure is heterogeneous and involves an ensemble of denatured states each exposing an average of 200 labile sites for exchange.

The Gaussian approximation makes it straightforward to extract mechanistic information from experimentally determined peak widths. Consider, for example, the experimental data for the time course for exchange of the native state of BLA at 15 °C and pH 3.8 (Fig. 8). The measured peak widths in the HX-MS data were compared with a random exchange model in which all 220 labile sites are exposed for exchange via global unfolding and with a core protection model in which 158 labile sites are exposed via local fluctuations (Fig. 8). The peak width of the molecular ion during the first 16 min of hydrogen exchange is significantly broader than

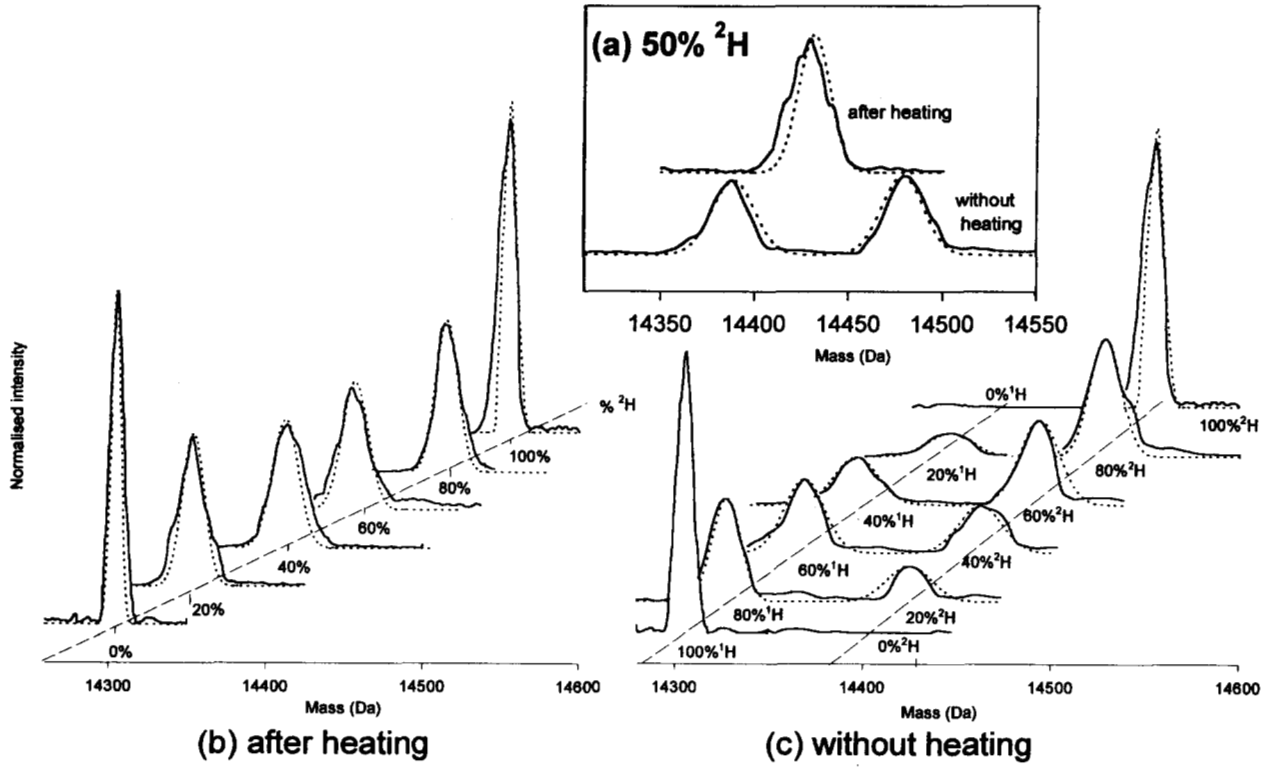

Fig. 6. Samples of hen lysozyme prepared to contain different proportions of ^2H and ^1H protein either in distinct populations or heated to produce a random distribution of isotopes. (**A**) Fifty percent ^2H-protein in D_2O and ^1H-protein in H_2O with and without heating. (**B**) From 0–100% ^2H at intervals of 20% after heating, and (**C**) the same isotopic composition as in (**B**) but heated after mixing to give a random distribution of ^2H. All samples are normalized and compared with a Gaussian simulation (dotted line) of the protein containing natural abundance isotopes and the same average number of ^2H atoms as the experimental data, arranged in a random distribution throughout the protein molecules. The increase in mass with increasing percentage of D_2O arises because molecules undergo increasing levels of hydrogen exchange of ^1H for ^2H.

predicted for $n = 158$. This broadening could arise from a number of different effects: first, because exchange is rapid in this time period and thus continues during analysis; second, because the rates of exchange of labile sites in different protein molecules may differ; and third, because exchange may occur via different mechanisms. However, for longer exchange times, in excess of 16 min, close agreement is found with the core protection model (Fig. 8). The analysis of peak widths highlights the potential for distinguishing between global unfolding, involving ^1H in the core of the protein, and local unfolding, in which a protected core of ^1H remain unaffected by exchange. Thus, although the underlying populations of ^2H are not resolved in this analysis, the data demonstrate the potential of HX-MS for delineating global versus local unfolding during hydrogen exchange events.

Concluding remarks

The utility of the hydrogen exchange process resulting from its sensitivity to changes in solution conditions and protein conformation is exemplified in this study. Comparison of the MS data of native and denatured states with the hydrogen exchange measured by 1D ^1H-NMR under the same solution conditions clearly show agreement within the experimental errors of the two techniques. Furthermore, the results reinforce the supposition that the levels of hydrogen exchange protection observed in the gas phase of the mass spectrometer are those present in solution, irrespective of ligand binding or dissociation (Robinson et al., 1996). We have shown that both variation of the average mass as well as the distribution of mass of a protein with time provide valuable insight about the structure and dynamics of proteins in solution. The ability to distinguish between global versus local unfolding mechanisms, using the measured peak width in an ESI mass spectrum, provides a new perspective for populations of molecules in hydrogen exchange environments. Furthermore, analysis of peak widths has shown that the arrangement of ^2H in protein molecules is different for alternative protein conformers. The methods described here for assessing protein conformation from HX-MS measurements are likely to provide valuable insight not only in assessing the small model proteins described in this study but also for the study of much larger protein–protein and protein–ligand interactions.

Materials and methods

Hen lysozyme and BLA were purchased from Sigma Chemical Company (St. Louis, MO) and washed on Centricon 10 concentrators (Amicon) to remove salts. CM^{6-127} lysozyme was prepared by selective reduction of Cys6–Cys127 and carboxymethylated as described previously (Eyles et al., 1994). Guinea pig α-lactalbumin was prepared from whole milk as described previously (Chyan et al., 1993). In all experiments ultrapure water (ELGA maxima

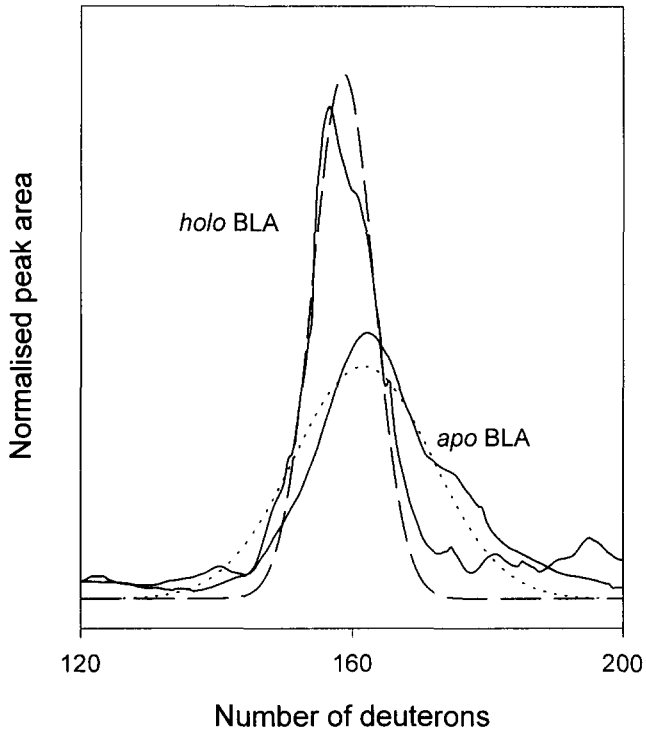

Fig. 7. Comparison of the peak width for *apo* and *holo* BLA containing 164 ± 1 ^{2}H with models for random and protected core exchange. Spectra were obtained at pH*3.8 after 1.5 min (*apo* BLA) and 180 min (*holo* BLA) of exchange in D_2O. The experimental peak widths are compared with a random exchange model (dotted line) and a core protection model (dashed line) in which 220 and 158 sites are available for exchange, respectively.

system) and deuterium oxide (Aldrich Chemical Company, Milwaukee, WI) were adjusted to the required pH by the addition of formic acid (Fisons) or ammonia (AnalaR). No attempt was made to correct for differences in pH in H_2O and D_2O. All pH readings in D_2O are represented by pH*.

Hydrogen exchange from native states

^{2}H-Lysozyme was prepared by incubation of the protein in D_2O at pH*3.8, heating to 75 °C for 1 h, followed by lyophilization. This process was repeated three times. Labile sites in CM^{6-127} lysozyme were exchanged for ^{2}H by incubation of the protein in D_2O at pH*4 for 1 h at 60 °C, lyophilization and the process repeated three times. ^{2}H-BLA was prepared by overnight incubation of the *apo* protein in D_2O at pH*5.0 and washing five times on a Centricon-10 (Amicon) in D_2O pH*5.0.

Hydrogen exchange from molten globule states

Labile sites in the acid induced molten globule states were exchanged for ^{2}H by incubation of protein overnight in D_2O at pH*2 and washed five times on a Centricon-10 (Amicon) with D_2O at pH*2. The *apo* state of BLA was prepared by washing five times in H_2O at pH2 on Centricon-10 (Amicon) and the final wash was adjusted to pH3.8. Exchange was initiated by dilution into water at pH3.8.

Hydrogen exchange from denatured states

The denatured state of CM^{6-127} lysozyme was prepared by dissolving lyophilized ^{2}H-protein, prepared as described above, in

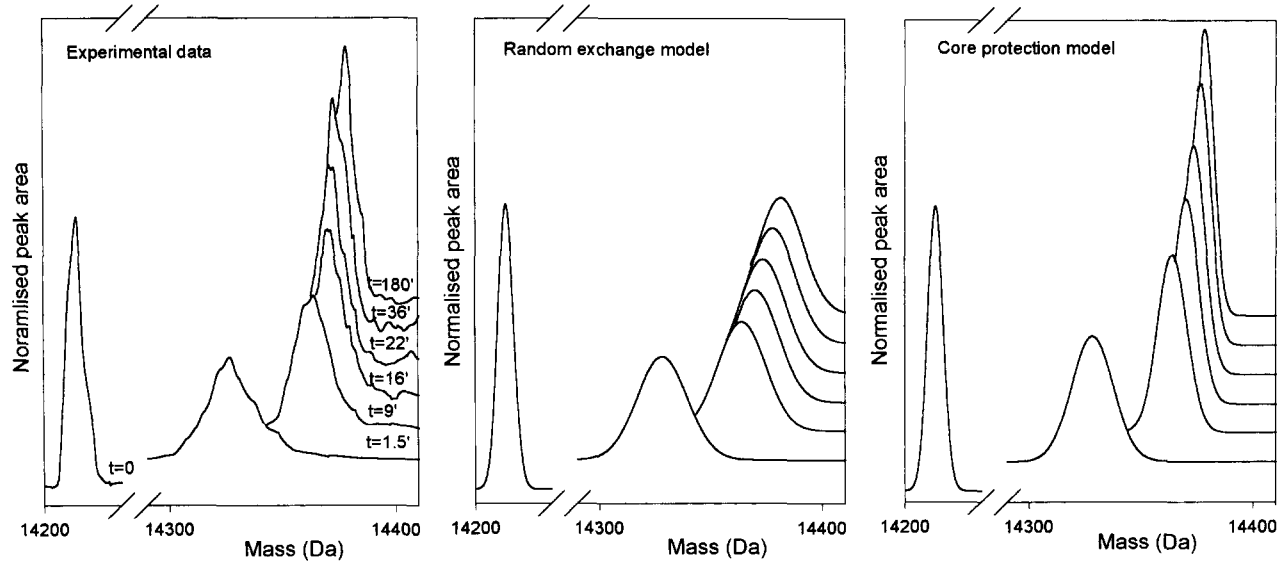

Fig. 8. Variation of the peak width during hydrogen exchange of ^{1}H-*holo* BLA at 15 °C and from solution at pH*3.8 containing a 1.7 molar excess of Ca^{++}. The experimental data are compared with two simulations: a random exchange model, in which exchange takes place via global unfolding and exchange from any of the 220 sites within the protein, and a core protection model, where 158 sites are able to exchange. The similarity between the core protection model and the experimental data suggests that exchange takes place via local fluctuations rather than global unfolding.

pH*2 and washing with D_2O at pH*2 five times on a Centricon-10 concentrator (Amicon).

Exchange from ^{2}H-protein

Lyophilized protein was dissolved in 10 μL D_2O at the pH* for exchange to give a concentration of 2 mM. Hydrogen exchange was initiated by 100-fold dilution in H_2O at the same pH to give a final concentration of 20 μM.

Exchange from ^{1}H-protein

HX-MS of Ca^{++}-bound native state of BLA, and pH*5 CM^{6-127} lysozyme, and native lysozyme was followed at pH*3.8. The *holo* BLA sample contained a 1.7 molar excess of Ca^{++}. Exchange was initiated by 100-fold dilution in D_2O at pH*3.8 and the mass measured as a function of time.

Mass spectrometry

All mass spectra were recorded on either a Platform I or II mass spectrometer (Micromass). The ESI source was operated without heating and in the absence of organic co-solvents. Samples were introduced via a Rheodyne injector, and solvent delivery was by means of a fluid delivery module (Michrom Bioresources). The background solvent was at the same pH and was matched isotopically to the protein sample solution. For HX-MS of labile states at room temperature (including denatured, molten globule, and A states), samples were pre-cooled on ice before introduction and the solvent delivery system was cooled using an ice salt bath. The temperature of the nebulizer gas was also reduced by passing the gas through a copper coil immersed in an ice salt bath. These measures effectively quench hydrogen exchange during sample introduction and, thus, the time recorded for exposure to the exchange environment was taken as the time at which samples were plunged into the ice salt bath. While it is possible to initiate exchange of proteins from lyophilized samples, either containing ^{2}H-protein or ^{1}H-protein, more consistent results were obtained when exchange was initiated by dilution into the appropriate solution. This method of sample dilution to initiate exchange gives rise to low levels of the original solvent in the exchanging solution, and it is necessary to adjust measurements to account for the isotopic composition of the solvent. The precise ratio of D_2O to H_2O is measured by heating the sample to 50–90 °C, depending on the T_m of the protein, to accelerate the exchange and remove any residual protection from the 3D structure of the protein. This measurement provides an accurate measure of the isotopic content of the solvent.

The average masses of protein samples were calculated from at least three charge states, and all mass spectra were calibrated using hen lysozyme. The estimated error in these HX-MS measurements, calculated from the standard deviation of the charge states, is ±2 Da. The number of protected sites was calculated by subtraction of the mass of the fully protonated protein from that of the mass measured at each time point (^{2}H-protein into H_2O). Alternatively, the measured mass is subtracted from the mass of the protein plus the number of labile sites calculated from the amino acid sequence (^{1}H-protein into D_2O). The masses measured are the centroid values, which take into account the asymmetry of the peak arising from the natural isotopic distribution. The HX-MS kinetic data, normalization of peak areas and simulations of HX-MS data, were carried out using Sigma Plot (Jandell Scientific), and all mass spectra presented represent the raw data, transformed onto a mass scale, with minimal smoothing and with no resolution enhancement. The binomial distribution of isotopes was produced using MassLynx software (Micromass).

^{1}D ^{1}H-NMR Measurements

Bulk hydrogen exchange measurements by 1D ^{1}H-NMR were performed on a GE electronics Omega 500 MHz instrument. Hydrogen exchange was initiated by direct dissolution of protein in D_2O at the appropriate pH to give a protein concentration of 1 mM in 600 μL and followed at 22 °C. A total of 32 scans of 2 K complex data points were collected at time intervals, varying from five minutes during the early period of exchange to 2 h toward the end of exchange. The number of protected sites was determined by integration of the amide and aromatic region of the frequency domain spectrum, followed by normalization to the intensity of upfield shifted methyls between 0 and 1 ppm. The assignments of the 1D ^{1}H-NMR spectra (Redfield & Dobson, 1988) were used to subtract the contribution of non-exchanging aromatic protons. The estimated error for the NMR data sets is of the order of ±5 sites and arises principally from solvent induced baseline error.

Acknowledgments

We thank Dennis Benjamin, Steve Eyles, and Brian Green (Micromass) for helpful discussions and R.G. Procter for expert technical assistance. This is a contribution from the Oxford Centre for Molecular Sciences, which is funded by the BBSRC, EPSRC, and MRC. E.W.C. has a Hong Kong Studentship and an Overseas Research Studentship. E.J.N. is grateful for financial support from Glaxo-Wellcome, and C.J.M. is a Harris Scholar at St. Hugh's College, Oxford. The research of M.G. is funded by a BBSRC Technology Foresight Junior Research Fellowship, and A.M. is a Junior Research Fellow of Christ Church, Oxford. The research of C.M.D. is supported in part by an International Research Scholars award from the Howard Hughes Medical Institute. S.E.R. and C.V.R. thank the Royal Society for support.

References

Aramini JM, Drakenberg T, Hiraoki T, Ke Y, Nitta K, Vogel HJ. 1992. Calcium-43 NMR studies of calcium-binding lysozymes and α-lactalbumins. *Biochemistry 31*:6761–6768.

Bai YW, Milne JS, Mayne L, Englander SW. 1993. Primary structure effects on peptide group hydrogen exchange. *Proteins Struct Funct Genet 17*:75–86.

Bai YW, Englander JJ, Mayne L, Englander SW. 1995. Thermodynamic parameters from hydrogen-exchange measurements. *Methods Enzymol 259*:344–356.

Baldwin RL. 1993. Pulsed H/D-exchange studies of folding intermediates. *Curr Opin Struct Biol 3*:84–91.

Chyan C-L, Wormwald C, Dobson CM, Evans PA, Baum J. 1993. Structure and stability of the molten globule state guinea pig α-lactalbumin. *Biochemistry 32*:5681–5691.

Connelly GP, Bai Y, Jeng M-F, Englander SW. 1993. Isotope effects in peptide group hydrogen exchange. *Protein Struct Funct Genet 17*: 87–92.

Eyles SJ, Radford SE, Robinson CV, Dobson CM. 1994. Kinetic consequences of the removal of a disulfide bond from hen egg white lysozyme. *Biochemistry 33*:13038–13048.

Fenn JB, Mann M, Meng CK, Wong SF, Whitehouse CM. 1989. Electrospray ionization for mass spectrometry of large biomolecules. *Science 246*:64–71.

Groß M, Robinson CV, Mayhew M, Hartl FU, Radford SE. 1996. Significant hydrogen exchange protection in GroEL-bound DHFR is maintained during iterative rounds of substrate cycling. *Protein Sci 5*:2506–2513.

Hooke SD, Eyles SJ, Miranker A, Radford SE, Robinson CV, Dobson CM. 1995. Cooperative elements in protein folding monitored by electrospray mass spectrometry. *J Am Chem Soc 117*:7548–7549.

Hvidt A, Nielsen SO. 1966. Hydrogen exchange in proteins. *Adv Protein Chem 21*:287–385.

Itzhaki LS, Evans PA. 1996. Solvent isotope effects on the refolding kinetics of hen egg white lysozyme. *Protein Sci 5*:140–146.

Jaquinod M, Guy P, Halgand F, Caffrey M, Fitch J, Cusanovich M, Forest E. 1996. Stability study of *Rhodobacter capsulatus* ferrocytochrome c_2 wild-type and site directed mutants using hydrogen/deuterium exchange monitored by electrospray ionization mass spectrometry. *FEBS Lett 380*:44–48.

Johnson RS, Walsh KA. 1994. Mass spectrometric measurement of protein amide hydrogen exchange rates of *apo-* and *holo*-myoglobin. *Protein Sci 3*:2411–2418.

Katta V, Chait BT. 1991. Conformational changes in proteins probed by hydrogen exchange and electrospray mass spectrometry. *Rapid Commun Mass Spectrom 5*:214–217.

Kragelund BB, Robinson CV, Knudsen J, Dobson CM, Poulsen FM. 1995. Folding of a four helix bundle: Studies of Acyl CoA binding protein. *Biochemistry 34*:7217–7224.

Kuwajima K. 1989. The molten globule state as a clue for understanding the folding and cooperativity of globular protein structure. *Proteins 6*:87–103.

Makhatadze GI, Clore GM, Gronenborn AM. 1995. Solvent isotope effect and protein stability. *Nat Struct Biol 2*:852–855.

Miranker A, Robinson CV, Radford SE, Aplin RT, Dobson CM. 1993. Detection of transient folding intermediates by mass spectrometry. *Science 262*:896–899.

Miranker A, Robinson CV, Radford SE, Dobson CM. 1996. Investigation of protein folding by mass spectrometry. *FASEB 10*:93–101.

Pedersen TG, Sigurskjold BW, Andersen KV, Kjaer M, Poulsen FM, Dobson CM, Redfield C. 1991. *J Mol Biol 218*:413–426.

Radford SE, Wolfson DN, Martin SR, Lowe G, Dobson CM. 1991. A three disulfide derivative of hen-lysozyme structure dynamics and stability. *Biochem J 273*:211–217.

Radford SE, Buck M, Topping KD, Dobson CM, Evan PA. 1992. Hydrogen exchange in native and denatured states of hen egg-white lysozyme. *Proteins Struct Funct Genet 14*:237–248.

Robinson CV, Groß M, Eyles SJ, Ewbank JJ, Mayhew M, Hartl FU, Dobson CM, Radford SE. 1994. Conformation of GroEL bound α-lactalbumin probed by mass spectrometry. *Nature 372*:646–651.

Robinson CV, Radford SE. 1995. Weighing the evidence for structure: Electrospray ionization mass spectrometry of proteins. *Structure 3*:861–865.

Robinson CV, Chung EW, Kragelund BB, Aplin RT, Poulsen FM, Dobson CM. 1996. Probing the nature of non-covalent interactions by mass spectrometry. *J Am Chem Soc 118*:8646–8653.

Smith LJ, Fiebig KM, Schwalbe H, Dobson CM. 1996. The concept of a random coil. Residual structure in peptides and denatured proteins. *Folding Design 1*:95–96.

Wagner DS, Melton LG, Yan Y, Erickson BW, Anderegg RJ. 1994. Deuterium exchange of α-helices and β-sheets as monitored by electrospray ionization mass spectrometry. *Protein Sci 3*:1305–1314.

Wedin RE, Delepierre M, Dobson CM, Poulsen FM. 1982. Mechanisms of hydrogen exchange in proteins from nuclear magnetic resonance studies of individual tryptophan indole NH hydrogens in lysozyme. *Biochemistry 21*:1098–1103.

Winkler RL, Hays WL. 1975. Special probability distributions. *Statistics, probability, inference and decision*. New York: Holt, Rinehart & Winston. pp 203–270.

Woodward C, Simon I, Tuchsen E. 1982. Hydrogen exchange and the dynamic structures of proteins. *Mol Cell Biochem 48*:135–160.

Wormwald C. 1991. NMR studies of α-lactalbumins [thesis]. Oxford University D. Philos. Thesis.

Yi Q, Baker D. 1996. Direct evidence for a two-state protein unfolding transition from hydrogen-deuterium exchange and mass spectrometry, and NMR. *Protein Sci 5*:1060–1066.

Zhang Z, Smith DL. 1993. Determination of amide hydrogen exchange by mass spectrometry: A new tool for protein structure elucidation. *Protein Sci 2*:522–531.

Zhang Z, Post CB, Smith DL. 1996. Amide hydrogen exchange determined by mass spectrometry: Application to rabbit muscle aldolase. *Biochemistry 35*:779–791.

Appendix

A Gaussian approximation to a binomial distribution simulating a peak in a mass spectrum may be constructed using Equation 1. The equation relates the intensity $\mathbf{I}(M)$ at a given mass, M, for a peak with measured maximum $\mathbf{I}_{max}$ whose isotopic content yields an average mass of μ_{tot} and standard deviation, σ_{tot}.

$$\mathbf{I}(M) = \mathbf{I}_{max} \times \exp\left[\frac{-0.5(M - \mu_{tot})^2}{\sigma_{tot}^2}\right] \quad (1)$$

where

$$\mu_{tot} = \mu_{^{12}C} + \mu_{^{13}C} + \mu_{^{1}H} + \mu_{^{2}H} + \mu_{^{3}H} + \mu_{^{16}O} + \mu_{^{18}O} \ldots$$

and

$$\sigma_{tot} = \sqrt{\sigma^2_{^{12}C} + \sigma^2_{^{13}C} + \sigma^2_{^{1}H} + \sigma^2_{^{2}H} + \sigma^2_{^{3}H} + \sigma^2_{^{16}O} + \sigma^2_{^{18}O} \ldots}$$

and

$$\mu_{A'} = n_{\mathbf{A}} \times m_{A'} \times F_{A'}$$

and

$$\sigma_{A'} = \sqrt{n_{\mathbf{A}} \times f_{A'} \times (1 - f_{A'})}$$

for each isotope A' of $\mathbf{A}$ and $n_{\mathbf{A}}$ = No. of atoms of type $\mathbf{A}$, $m_{A'}$ = exact mass of isotope A', $f_{A'}$ = fractional abundance of A'.

The binomial distribution of isotopes was calculated using Micromass MassLynx software.