Available online at www.sciencedirect.com
SCIENCE DIRECT®
International Journal of Mass Spectrometry 240 (2005) 249–259
International Journal of Mass Spectrometry
www.elsevier.com/locate/ijms

Probing protein dynamics and function under native and mildly denaturing conditions with hydrogen exchange and mass spectrometry

Igor A. Kaltashov*

Department of Chemistry, University of Massachusetts at Amherst, 710 North Pleasant Street, LGRT#701, Amherst, MA 01003, USA

Received 30 July 2004; accepted 27 September 2004
Available online 18 November 2004

Abstract

A combination of hydrogen exchange and mass spectrometry emerged in recent years as a powerful experimental tool capable of probing both structural and dynamic features of proteins. Although its concept is very simple, the interpretation of experimental data is not always straightforward, as a combination of chemical reactions (isotope exchange) and dynamic processes within protein molecules give rise to convoluted exchange patterns. This paper provides a historical background of this technique, candid assessment of its current state and limitations and a discussion of promising recent developments that can result in tremendous improvements and a dramatic expansion of the scope of its applications.

Keywords: Protein dynamics; Protein folding; Protein conformation; Hydrogen exchange; Mass spectrometry

1. Introduction

Hydrogen–deuterium exchange (HDX) is a technique that became in recent decades one of the major experimental tools to probe both structural and dynamic features of proteins. The analytical value of HDX as a tool for probing macromolecular structure was recognized almost immediately after the discovery of deuterium [1] and the subsequent development of efficient methods of heavy water production [2]. Initial studies of the exchange reactions between small organic molecules and ^2H_2O carried out by Bonhoeffer and Klar [3] established that hydrogen atoms attached to carbon atoms (e.g., $–CH_3$ groups) do not undergo facile exchange in solution, while the exchange rates for hetero-atoms (e.g., –OH groups) are generally high. Nevertheless, even such *labile* hydrogen atoms may exchange very slowly if they are not easily accessible by solvent, a situation that is rarely encountered among small organic molecules, but becomes increasingly common as the physical size of the molecule in question increases. Hvidt and Linderstrom-Lang [4,5] later used HDX to measure the solvent accessibility of labile hydrogen atoms as a probe of polypeptide structure. While this study was the first attempt to use HDX to probe protein structure in solution, the early studies of hydrogen exchange reactions between water and biopolymers preceded this work by more than 15 years [6].

Replacement of a labile hydrogen atom with deuterium (or vice versa) changes two fundamental parameters, namely nuclear spin and mass. The former can be identified using NMR spectroscopy, while the latter can be measured using mass spectrometry. A change in mass also results in a significant alteration of a vibrational frequency, enabling the use of IR spectroscopy to determine the isotope content of a macromolecule. It is probably fair to say that the tremendous popularity enjoyed by HDX as an experimental technique is in large part due to the emergence and spectacular progress of high-resolution NMR [7–10]. However, mass spectrometry (MS) is currently enjoying a dramatic surge in popularity in this field as well [11–14]. Interestingly, the idea to use mass measurements as a means to monitor the extent of HDX within macromolecules actually precedes the use of NMR for

* Tel.: +1 413 545 1460; fax: +1 413 545 4490.
E-mail address: kaltashov@chem.umass.edu.

1387-3806/$ – see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.ijms.2004.09.021

the same purpose. In the mid-1950s, Burley et al. [15] estimated the extent of ^{2}H incorporation into fibrous proteins by measuring the mass change of the protein sample as a result of HDX. In these experiments, quantitation of ^{2}H content was carried out gravimetrically, i.e., by measuring the mass of an entire protein sample prior to and following the completion of the exchange reactions using a quartz spring. Although this idea was adopted by several other groups in the late 1950s–early 1960s [16–18], a lack of accuracy afforded by such measurements limited its use. This shortcoming could have been addressed by employing mass spectrometry to directly measure mass changes of macromolecules, however it took several decades before the dramatic technological advances in the field of mass spectrometry enabled desorption and ionization of intact macromolecular species. In the early 1980s McCloskey and co-workers [19] demonstrated that ^{2}H incorporation into a peptide in solution can be accurately determined using fast atom bombardment (FAB) MS. Although initially this methodology was employed only as a means of providing an accurate count of labile hydrogen atoms within a peptide [19], its use has been later expanded to provide information on HDX kinetics in solution [20].

The advent of electrospray ionization (ESI) MS dramatically expanded the range of biopolymers for which the extent of ^{2}H incorporation can be measured directly under a variety of conditions [21]. As a result of these developments, HDX MS has now become a powerful experimental tool for probing protein higher order structure. The number of applications of the HDX MS methodology to probe both architecture and dynamics of biomolecules continues to expand, as it offers several important advantages over HDX NMR. These include tolerance to paramagnetic ligands and co-factors as well as much more forgiving molecular weight limitations, and superior sensitivity, which often allows the experiments to be carried out at concentrations close to or even below the endogenous levels.

2. Protein structure and dynamics reflected in HDX MS patterns

The concept of HDX experiments is very simple, although the interpretation of experimental data is not always straightforward. Indeed, a combination of chemical reactions (isotope exchange) and dynamic processes within a protein molecule often give rise to convoluted exchange patterns. Labile hydrogen atoms exchange slowly if they are shielded from solvent (e.g., reside in a hydrophobic core of a protein) or involved in a hydrogen-bonding network. In order for the exchange to occur, such hydrogen atoms must become *unprotected* (e.g., by exposing the protein interior to the solvent, breaking the hydrogen bond, etc.). Most protein HDX studies target backbone amide hydrogen atoms, since they are convenient intrinsic reporters of backbone dynamics that are evenly distributed throughout the entire polypeptide chain. The dynamic processes leading to the loss of protection under native

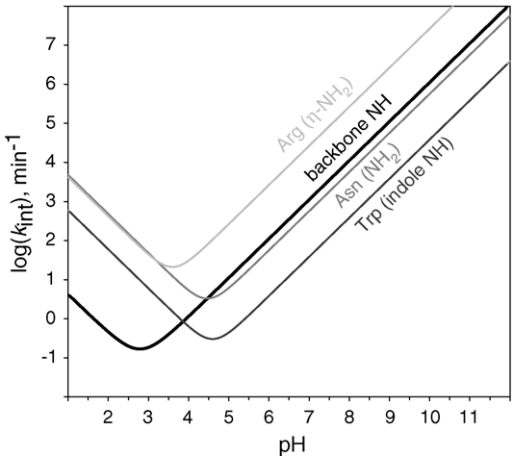

Fig. 1. The pH dependence of intrinsic exchange rates of several types of labile hydrogen atoms calculated based on data compiled by Dempsey [50].

or near-native conditions are usually transient, and it is the lifetime of such *exchange-competent states* that determines the extent of the exchange reactions occurring during a single opening event. A mathematical formalism that is often used to describe HDX kinetics was introduced several decades ago and is based upon a simple two-state kinetic model [22]:

$$\mathrm{N}^{i}\mathbf{H}(\text{incompetent}) \underset{k_{cl}}{\overset{k_{op}}{\rightleftarrows}} \mathrm{N}^{i}\mathbf{H}(\text{competent}) \xrightarrow{k_{int}} \mathrm{N}^{j}\mathbf{H}(\text{competent}) \rightleftarrows \mathrm{N}^{j}\mathbf{H}(\text{incompetent}), \quad (1)$$

where k_{op} and k_{cl} are the rate constants for the opening (unfolding) and closing (refolding) events that expose/protect a particular amide hydrogen to/from exchange with the solvent, and $^{i}\mathbf{H}$ and $^{j}\mathbf{H}$ represent a pair of hydrogen isotopes (typically, ^{2}H and ^{1}H, although radioactive ^{3}H was also used in the past). The rate constant of intrinsic exchange of an unprotected hydrogen atom (from the exchange-competent state) k_{int} depends on both solution pH and temperature and can be estimated using short unstructured peptides [23] (Fig. 1). The $\mathrm{N}^{i}\mathbf{H} \rightarrow \mathrm{N}^{j}\mathbf{H}$ transition is essentially irreversible, as HDX experiments are carried out in a significant excess (10–1000-fold) of exchange buffer. A general expression for an observed exchange rate constant for a single amide in this model is given by [22]:

$$k_{HDX} = \frac{k_{op} + k_{cl} + k_{int} - \sqrt{(k_{op} + k_{cl} + k_{int})^2 - 4k_{op}k_{int}}}{2}. \quad (2)$$

In most HDX studies, the exchange-incompetent state of the protein is in fact its native state. Since most HDX measurements are carried out under conditions that favor the native state (i.e., the effective unfolding equilibrium constant $K_{op} = k_{op}/k_{cl} \ll 1$), the expression for k_{HDX} can be simplified

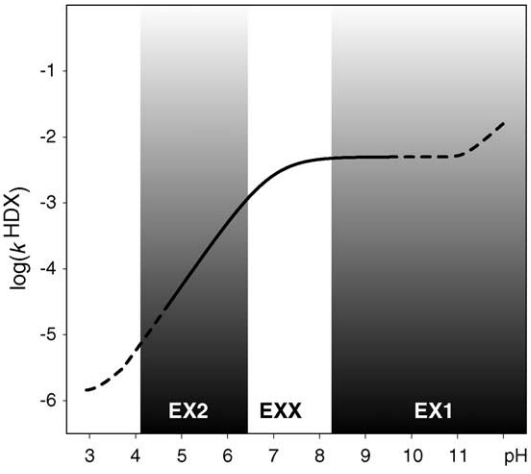

Fig. 2. Dependence of the rate of observed hydrogen exchange of backbone amide hydrogen atoms upon pH calculated with the assumption that the k_{op}/k_{cl} remains constant throughout the entire pH range. Dashed lines indicate expected deviations from the calculated dependence when the assumption of k_{op}/k_{cl} constancy is unreasonable.

to

$$k_{HDX} = \frac{k_{op}k_{int}}{k_{cl} + k_{int}}. \quad (3)$$

The latter expression is intuitive, as it presents the exchange rate as a product of the frequency of opening events (k_{op}) and the fraction of hydrogen atoms at a particular position that undergo exchange during a single opening event ($k_{int}/(k_{cl} + k_{int})$). Two extreme situations are usually considered, commonly referred to as EX1 and EX2 exchange mechanisms [24]. If the protein refolding rate constant k_{cl} is much higher than the intrinsic exchange rate k_{int}, only a small fraction of the exposed amide hydrogen atoms will be exchanged during each opening event. As a result, the exchange will proceed in small increments, following the kinetics with an apparent rate constant $k_{HDX} = k_{int} \cdot k_{op}/k_{cl}$ (a situation known as EX2 regime). HDX under native conditions almost always proceeds via EX2 mechanism. If, on the other hand, $k_{cl} \ll k_{int}$ (which can be achieved by carrying HDX reactions under denaturing conditions or at high pH), all exposed amides will be exchanged during a single opening even. The exchange rate in this case will actually be equal to the rate of transition from the closed to the open state of the protein, i.e., $k_{HDX} = k_{op}$ (the so-called EX1 exchange regime).

The ability to clearly identify the exchange regime is important, since such knowledge is crucial for correct quantitative interpretation of the experimental data (EX1 measurements provide kinetic information, while measurements under EX2 conditions yield thermodynamic information). In some cases, the exchange regime can be identified by determining the pH dependence of the rate constant k_{HDX} (Fig. 2), assuming the k_{op}/k_{cl} ratio stays nearly constant within a certain pH range. HDX MS measurements can often identify the EX1 exchange regime directly, as it often yields a *correlated* exchange pattern, while the EX2 regime always results in *uncorrelated* exchange. In the case of a two-state protein, correlated exchange results in characteristic bimodal isotopic distributions, a feature that is notably absent from uncorrelated exchange patterns (Fig. 3). It must be noted, however, that the uncorrelated exchange component is almost always present even under EX1 exchange conditions. For example, it is seen in Fig. 3B as a gradual shift of the higher-*m*/*z* isotopic cluster, and is often attributed to uncooperative dynamic events leading to a loss of protection by one amide (such as *local structural fluctuations* [25]). Furthermore, the EX1 limit represents only an extreme situation, which is not always achieved if HDX MS measurements carried out under mildly denaturing conditions. As a result of that, convoluted *semi-correlated* exchange patterns can be observed, signaling that $k_{cl} \sim k_{int}$ and, therefore, each opening even leads to a considerable, but not complete exchange of amides that become unprotected (Fig. 3C).

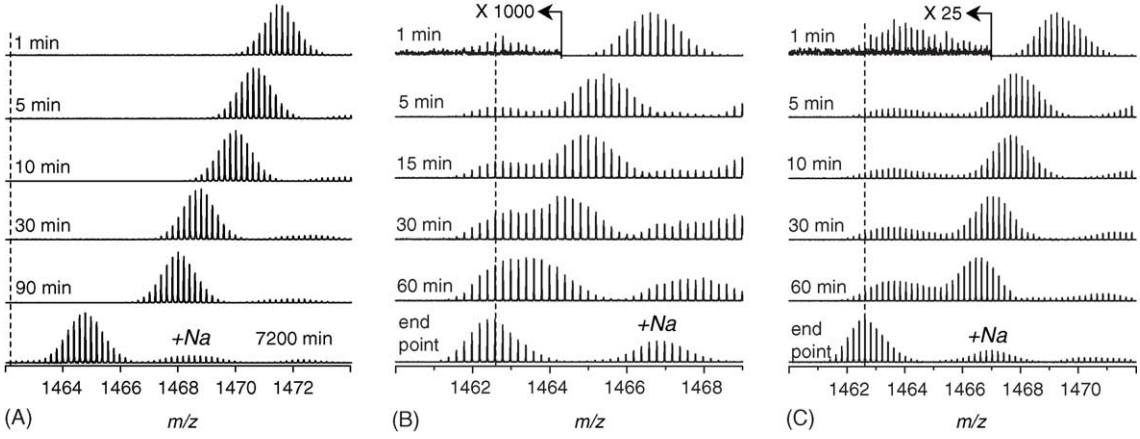

Fig. 3. HDX MS patterns of a small two-state protein chymotrypsin inhibitor 2 acquired under near-native conditions favoring fully uncorrelated exchange (A), pH 11, 60% methanol at 8 °C favoring correlated exchange (B) and pH 10, 70% methanol at 8 °C favoring semi-correlated exchange (C). A dotted line in each panel indicates the position of the centroid of an isotopic cluster of protein ions whose ^{2}H content is identical to that of solvent.

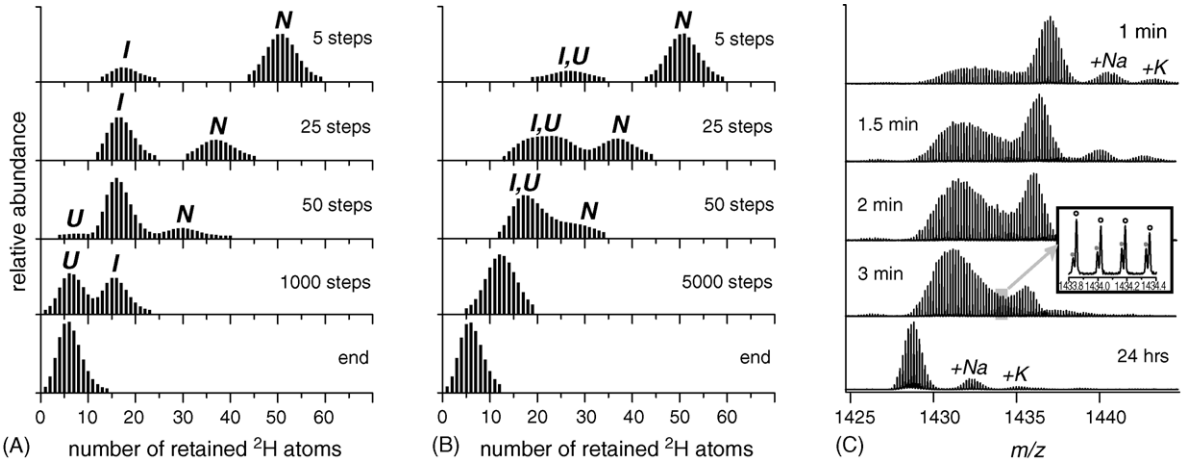

Fig. 4. Computer-simulated HDX patterns of a model three-state protein, in which transitions among all states ($N \rightarrow I$, $I \rightarrow U$ and $N \rightarrow U$) proceed under conditions favoring EX1 exchange regime (A) and only one transition ($N \rightarrow I$) favors Ex1 regime, producing semi-correlated exchange pattern (B). (C) HDX MS profiles of a three-state protein ubiquitin acquired under conditions when all three states are populated in solution (pH 7, 60% methanol). The inset illustrates resolution of isobaric protein ions: ^{2}H-rich, ^{23}Na-free (◯) and ^{2}H-depleted ^{23}Na-adducts (●).

Although HDX MS measurements of two-state proteins considered above are instructive and often provide valuable information on protein dynamics, it is the studies of multi-state proteins that benefit most from HDX MS experiments carried out under EX1 conditions. In an ideal case, such measurements may provide a means to map and characterize distinct protein conformers differing by their exposure to solvent (Fig. 4A). The ability of HDX MS to reveal the presence and characterize the behavior of distinct intermediate states based on their ^{2}H content is quite unique, as HDX NMR generates exchange data averaged across the entire ensemble of states, thus complicating the detection of distinct conformations. Although in some instances certain information about the intermediates may be obtained by grouping amides with similar exchange kinetics into cooperative unfolding-refolding units, or *foldons* [26], this approach may result in artifacts [27,28]. It must be mentioned, however, that an idealized correlated exchange pattern similar to that presented in Fig. 4A is rarely observed in practice, particularly under mildly denaturing conditions. The exchange patterns are often complicated due to the fact that while some transitions (e.g., from the native to the intermediate states) favor the EX1 exchange mechanism, others (e.g., from the intermediate to the fully unfolded states) might give rise to exchange occurring in the EX2 regime (Fig. 5). As a result, some protein states may escape a straightforward detection or else their protection levels may be misread. Two examples of such behavior are presented in Fig. 4B and C. Nonetheless, multiple protein conformers co-existing at equilibrium under mildly denaturing conditions can often be directly observed in carefully executed HDX MS experiments [29].

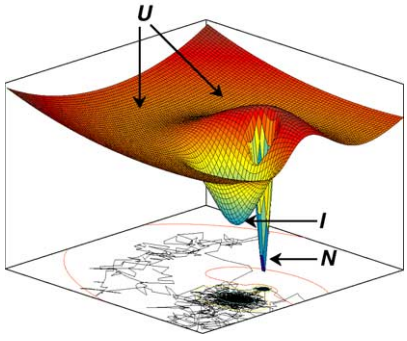

Fig. 5. A hypothetical energy diagram of a three-state protein that is expected to produce HDX pattern shown in Figs. 4B, C. Separation between two energy minima (***N***, global, and ***I***, local) is sufficient to trap protein molecules in the ***I***-state long enough to result in correlated or semi-correlated exchange. Lack of energy barrier between the ***I***-state and ***U***-state does not allow protein molecules to visit the latter long enough, so that only uncorrelated exchange is afforded in the ***U***-state. Sample trajectory of a simulated Brownian motion along this energy surface is shown at the bottom of the diagram.

3. Protein dynamics and function under native conditions: quantitative assessment of protein–ligand binding

The ability of HDX measurements to detect changes in the solvent exposure of polypeptide chains has been used in numerous studies of protein binding processes [30–32]. The basic premise of such analyses is that protein–protein (or, more generally, protein-large ligand) binding inevitably leads to solvent exclusion from the interface region, resulting in significant reduction of HDX rates for all amides located at the binding interface. In many cases, however, the ligand is too small to cause any significant solvent exclusion from the interface area or else does not form hydrogen bonds with the protein at all (e.g., lipophilic ligands, metal ions, etc.). Furthermore, many of these ligands do not even induce significant changes in the protein secondary structure detectable by HDX. Nevertheless, even in such situations HDX measurements often reveal intimate details of protein–ligand interaction processes. We will illustrate this point by

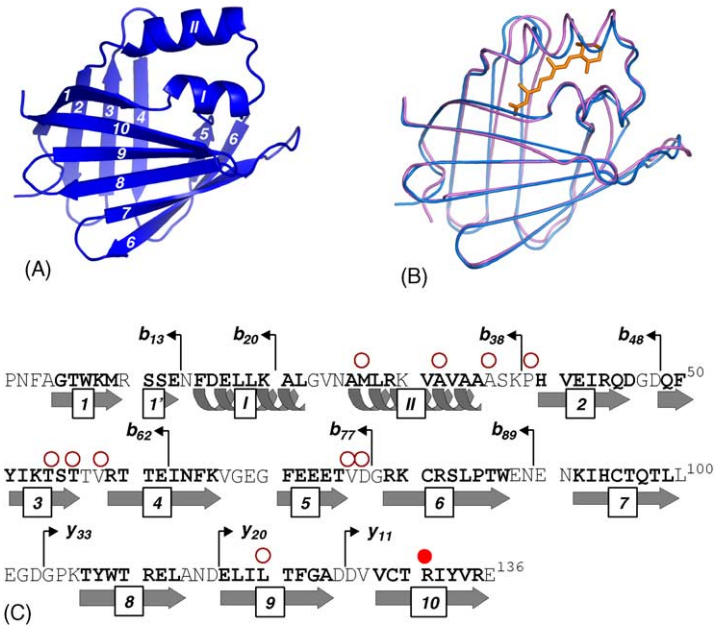

Fig. 6. Primary, secondary and tertiary structure of CRABP I. Crystal structures of the apo (blue) and holo (purple) forms of the protein are overlaid in panel B.

considering the binding of *all-trans retinoic acid* (RA) to its carrier, *cellular retinoic acid binding protein I* (CRABP I).

CRABP I is a member of a family of small soluble intracellular proteins that bind hydrophobic ligands such as fatty acids, lipids, and retinoids. The *raison d'etre* for CRABP I is binding RA, a major physiologically active metabolite of vitamin A, although it still is unclear whether the protein acts as RA transporter to the nuclear receptor site or simply restricts its availability in the cell by mere sequestration. The protein contains 136 residues forming two five-stranded β-sheets (Fig. 6A). The first two strands are connected by a helix-turn-helix motif, and all others by reverse turns. The two β-sheets are packed orthogonally to form a solvent-filled β-barrel. The ligand-binding pocket, which physiologically accommodates RA is located inside the barrel. RA binding to CRABP I introduces only minimal changes to the protein structure (Fig. 6B), yet it is clear that there must be transient dynamic events that allow RA to enter the cavity inside the protein. A *portal model* has been postulated, which invokes the notion of a highly dynamic segment within the protein that serves as an opening, which provides ligand access to the cavity [33]. Recent NMR measurements suggests that in CRABP I this portal region constitutes the helix-turn-helix motif and two flanking β-hairpins [33].

It is now generally accepted that protein structure under native conditions is not a rigid crystalline state, but undergoes a range of movements, involving anything from side chain rotation to rearrangement of secondary structure elements relative to each other. Although the existence of such motions within the native state of a protein can be detected with a variety of experimental techniques, their exact nature remains a subject of discussion in the literature. The commonly accepted models of the local dynamics within natively folded proteins invoke the notions of a structural fluctuation (already mentioned in the preceding section) or a *mobile defect* (which considers not only the emergence and dissipation of local disorder, but also the possibility of its propagation through the protein structure). Alternatively, the local dynamics can be described using a *solvent penetration* model (slow diffusion of the solvent molecules into and out of the protein interior) [34]. Such description is actually very similar to the mobile defect model, as applied to the integral solute–solvent system, instead of the protein molecule alone.

Above and beyond local structural fluctuations, protein dynamics under native conditions is exemplified by transiently sampling non-native (higher-energy or *activated*) conformations. Such protein states are often functionally important despite their low Boltzmann weight [35,36]. Transient loss of structure by a certain segment of CRABP I mentioned earlier as a requisite event for ligand entry and binding is consistent with the notion of sampling an activated protein state (Fig. 7). The second local minimum on the free energy diagram representing the apo-form of the protein corresponds to an activated state, whose structure allows unimpeded RA entry into the cavity. While the native conformation is highly favored by the protein (and is represented by the global energy minimum), the free energy difference between the two states is not too high, so that the protein samples the activated state frequently in the absence of the ligand. However, once the ligand is introduced into the system and binds to the protein in its activated state, the energy landscape changes quite dramatically. Multiple favorable interactions between the ligand and the protein in its native state provide significant "reinforcement" of the latter, making its free energy even more negative. As a result of that, the protein will be

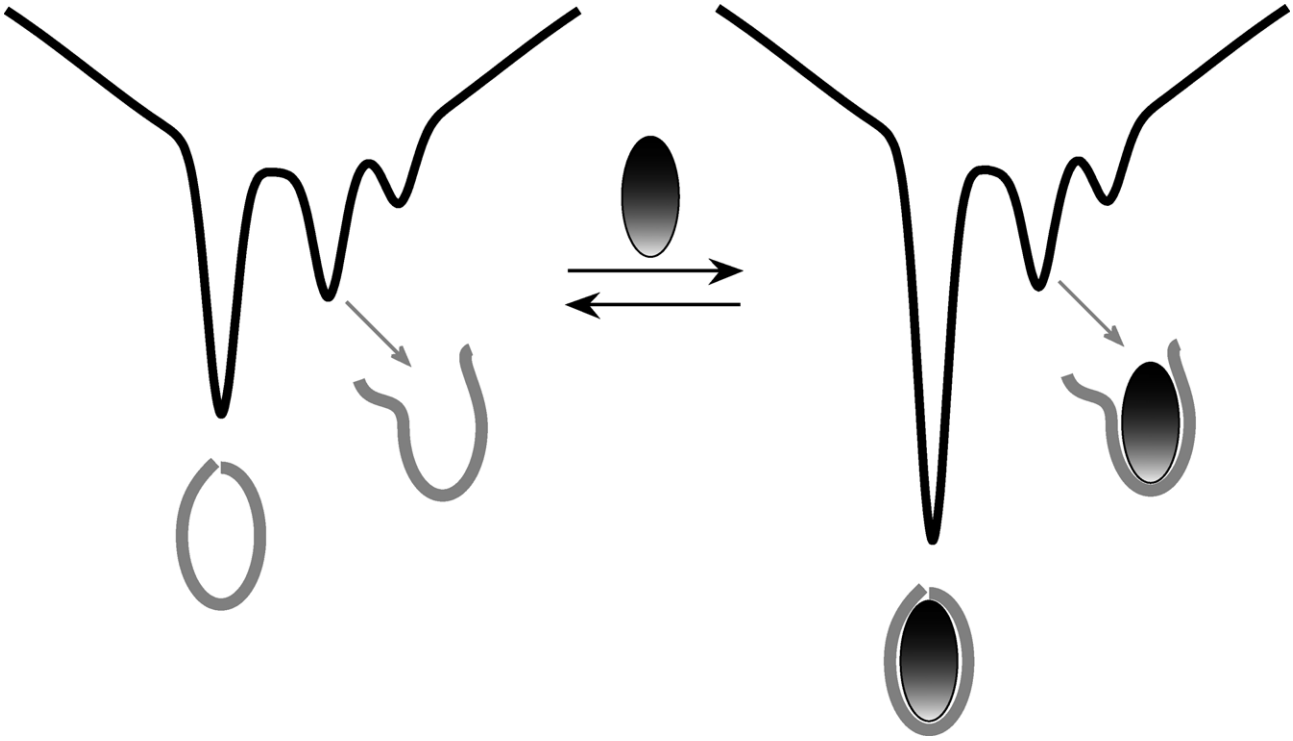

Fig. 7. Hypothetical free energy landscapes of CRABP I in the absence (lest) and in the presence of RA (right).

now much less likely to sample the alternative conformation (whose structure provides the ligand with an escape route).

The initial confirmation of this view of RA binding to CRABP I is provided by HDX MS experiments carried out under native conditions [37]. Hydrogen exchange becomes significantly slower in the presence of RA (Fig. 8), and such an efficient deceleration cannot be attributed to the solvent-shielding action of the ligand. Indeed, RA does not form any hydrogen bonds to the protein and the reduced rates of exchange are fully attributable to the diminished dynamics of the polypeptide chain. Since the exchange follows the EX2 mechanism under these conditions, we can use measured kinetic parameters of HDX reactions to calculate equilibrium constants of various unfolding processes (vide supra):

$$k_{\mathrm{HDX}} = \frac{k_{\mathrm{op}} k_{\mathrm{int}}}{k_{\mathrm{cl}}} = k_{\mathrm{int}} K_{\mathrm{op}}. \tag{4}$$

Since the slowest phase of the exchange corresponds to global unfolding, one can use (4) to calculate free energy of the native conformation as

$$\Delta G = -RT \ln(K_{\mathrm{unfolding}}) = -RT \ln\left(\frac{k_{\mathrm{HDX}}}{k_{\mathrm{int}}}\right) \tag{5}$$

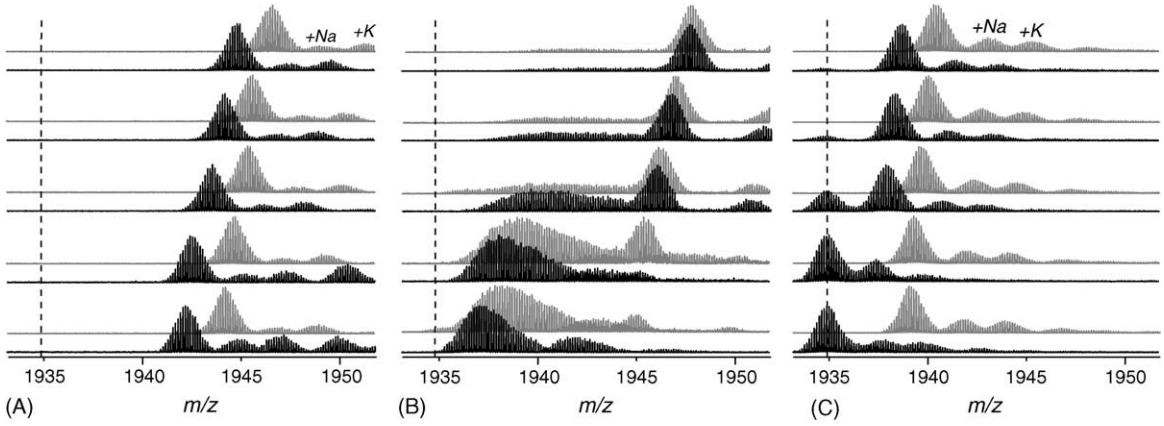

Fig. 8. Isotopic distributions of CRABP I ions (charge state +8). The protein undergoes HDX in solution at pH 6.8 (A), 3.4 (B) and 10 (C) in the absence (black traces) and in the presence of RA (gray traces). Dotted line in each panel indicates the position of the centroid of an isotopic cluster of a protein ion whose ^{2}H content is identical to that of the exchange buffer ($^{1}H/^{2}H = 96{:}4$).

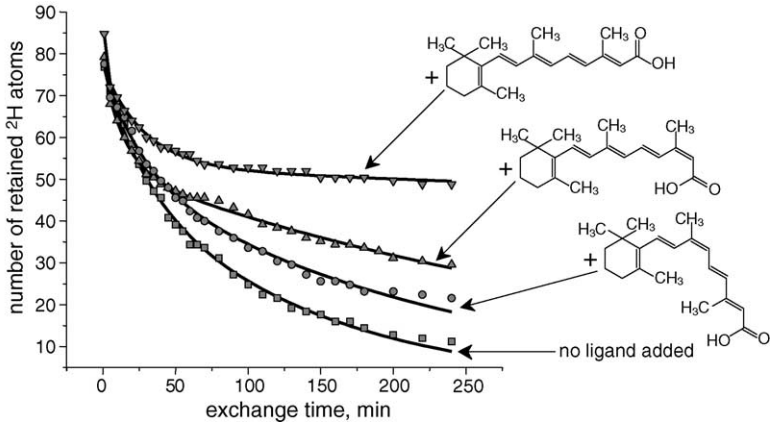

Fig. 9. HDX kinetics of CRABP I measured under native conditions in the absence and in the presence of natural retinoids. Adapted with permission from Xiao et al. [37].

using the fully unfolded state of the protein as a reference zero-energy conformation. Since the addition of the ligand is unlikely to change the energy of the fully unstructured state of the protein, it is possible to obtain an estimate of the protein–ligand interaction simply as

$$\Delta\Delta G = RT \ln\left(\frac{K^{\text{apo}}_{\text{unfolding}}}{K^{\text{holo}}_{\text{unfolding}}}\right) = RT \ln\left(\frac{k^{\text{apo}}_{\text{HDX}}}{k^{\text{holo}}_{\text{HDX}}}\right), \quad (6)$$

where $k^{\text{holo}}_{\text{HDX}}$ and $k^{\text{apo}}_{\text{HDX}}$ are the measured rate constants for the slowest HDX phases of CRABP I in the presence and in the absence of RA, respectively. This allows a fairly straightforward quantitation of RA binding to CRABP I to be carried out, yielding a value that is reasonably close to estimates provided by other biophysical techniques [37].

While RA is a cognate ligand for CRABP I, this protein also binds a range of other retinoids, particularly RA isomers, such as 9-*cis*-RA and 13-*cis*-RA. The hydrophobic skeletons of the latter two are altered sufficiently to introduce significant disruption to the network of native hydrophobic contacts, thus reducing the stability of the protein–ligand complex. HDX measurements of CRABP I carried out in the presence of 9-*cis*-RA and 13-*cis*-RA reveal a notable decrease of protein dynamics as compared to the apo-form of the protein, although not as significant as in the case of the cognate ligand (Fig. 9). CRABP I affinity to non-cognate ligands can also be estimated based on the measured HDX rates [37].

4. Protein dynamics and function under mildly denaturing conditions: a glimpse at the mechanism

HDX MS measurements carried out under native conditions allow the thermodynamic parameters of the RA-CRABP I interaction to be quantitatively assessed, however, they do not provide direct information on the putative intermediate states of the protein and their involvement in ligand binding. Indeed, the exchange under these conditions proceeds via the EX2 mechanism and, therefore, is uncorrelated. As a result, HDX MS produces a picture of protein dynamics averaged across the entire population. Various protein states can only be distinctly detected if the exchange is correlated or semi-correlated, which only becomes possible in the EX1 or EXX exchange regime, respectively.

Lowering solution pH to 3.4 results in a significant decrease of k_{int} and, therefore, is often expected to keep the exchange in the EX2 regime (see Fig. 2). Very often, however, low pH results in a significant destabilization of the protein, which affects refolding rates k_{cl} even more significantly than the intrinsic exchange rate k_{int}. In many instances, it is possible to find conditions under which $k_{\text{int}} \geq k_{\text{cl}}$, a situation leading to semi-correlated exchange (vide supra). An example of such behavior can be seen in Fig. 8B, which shows HDX MS profiles of CRABP I acquired at pH 3.4 in the absence and in the presence of RA. The isotopic distribution of CRABP I ions is clearly bimodal, with the high *m*/*z* cluster corresponding to a highly protected form of the protein. The lower *m*/*z* cluster is likely to contain contributions from several states of the protein (including a fully unfolded one, as well as one or more intermediates, which cannot be resolved under conditions favoring semi-correlated exchange). The observed exchange profiles clearly show that partial or full unfolding of the protein is significantly slowed down when RA is present in the exchange buffer. However, these data provide very little information on transitions between the non-native states of the protein.

In order to distinctly see the non-native states, the solution pH must be sufficiently high, so that the EX1 criterion $k_{\text{int}} \gg k_{\text{cl}}$ is met. At pH 10, the apo-form of CRABP I exhibits a clearly bimodal character, where the lower *m*/*z* cluster corresponds to the fully unfolded state of the protein (as soon as it becomes visible in the spectra, the ^{2}H content of this state is identical to that of solvent). The higher *m*/*z* cluster represents a partially structured intermediate state, which initially retains 30 deutrons, although this number slowly diminishes due to local structural fluctuations. Addition of RA to the exchange buffer results in the disappearance of the lower *m*/*z*

cluster, indicating a dramatic increase of the protein refolding rate caused by the presence of the ligand. When coupled with the analysis of protein ion charge state distributions [37,38], these studies provide conclusive evidence that partially unfolded states of CRABP I play a crucial role in binding its cognate ligand.

5. Understanding how proteins work: site-specific HDX MS measurements and protein ion fragmentation

While HDX MS unequivocally establishes the functional importance of partially unstructured protein states and even affords their detection, it provides little information on their structure. Indeed, site-specific assignment of ^{2}H incorporation is a challenging task due to the reversible character of HDX. Still, there are conditions (pH 2.5–3, $T = 0\,^\circ$C) under which the exchange of backbone amide hydrogen atoms is relatively slow (Fig. 1). Coupling of HDX carried out under native conditions with proteolysis under such *slow exchange conditions* can provide valuable information on the spatial distribution of ^{2}H along the polypeptide chain [11]. The proteolytic step (with pepsin) is followed by fast chromatographic separation and MS identification of the fragment peptides, and measurement of their ^{2}H content. Noticeable *back-exchange* occurs even under such conditions, limiting the resolution afforded by these site-specific *bottom-up* HDX MS measurements.

Fragmentation of protein ions in the gas phase in many cases offers an attractive alternative to site-specific measurements of ^{2}H content using a traditional bottom-up approach [29,39,40]. Protein ion fragmentation is typically induced by its collisional activation in the ESI interface (CAD), although alternative methods (such as infrared multi-photon dissociation, IRMPD) are occasionally utilized as well [41]. HDX CAD MS does not require that the measurements be carried out in a slow exchange mode. Therefore, such *top-down* experiments can be carried out on-line, allowing the back-exchange to be avoided. A major experimental concern in HDX CAD MS is the possibility of *hydrogen scrambling*, i.e., redistribution of ^{2}H atoms within the protein ion prior to its dissociation. Indeed, an average lifetime of protein ions undergoing collisional activation can be quite significant, exceeding in some cases hundreds of milliseconds. Such metastable ions may potentially exhibit significant intra-molecular hydrogen exchange in the gas phase (which we term long-distance internal exchange), which would render the results of top-down HDX CAD MS measurements meaningless. In addition, charge-directed cleavages of peptide bonds in the gas phase often involve local proton transfer and may cause selective alteration of ^{2}H content of some types of fragment ions without affecting others [42–44].

Our recent study demonstrated that the extent of hydrogen scrambling by internal long-distance proton exchange within multiply charged protein ions is controlled largely by collisional energy [45]. HDX CAD MS measurements carried out under rapid collisional heating conditions provide correct information on the backbone protection of protein segments (structural elements spanning several amino acid residues). While the rate of ion activation appears to be a major determinant of the extent of scrambling, diminished flexibility of the polypeptide chain in the gas phase limits the extent of scrambling even if it is favored energetically (if a slow heating method of ion activation is employed, e.g., sustained off-resonance irradiation, SORI) [45]. At the same time, local hydrogen transfer reactions associated with peptide bond cleavages do take place regardless of the ion activation rate, making it impossible to accurately measure the protection of each individual amide group and, therefore, providing realistic limits as far as the spatial resolution that can be achieved with this method [45].

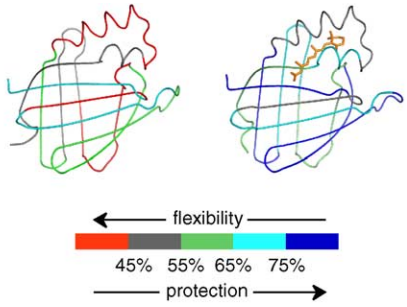

Fig. 10. Local conformational stability maps of CRABP I measured with HDX CAD MS under native conditions in the absence (left) and in the presence of RA (right).

We have recently used the top-down HDX CAD MS strategy to produce conformational stability maps of CRABP I under native and mildly denaturing conditions. A series of abundant fragment ions are observed in CAD mass spectra of the protein, providing relatively even coverage of its entire sequence (Fig. 6C). Time evolution of ^{2}H content of each of these fragment ions allows the amide protection of the key structural element of CRABP I to be calculated as a function of exchange time. Fig. 10 presents a color-coded protection map of the protein following 60 min of exchange (when the fast phase of HDX due to local structural fluctuations has been completed and the major contributors to the exchange are transitions to partially unstructured intermediate states of the protein). The map clearly identifies three high-mobility regions within the protein, namely α-helix ***II*** (α**II**), β-strands ***4***/***5*** (β**4**/**5**) and β-strand ***9*** (β**9**). The two least flexible structural elements of the protein are strands β**8** and β**10**. It is these two segments that contain amino acid residues Arg^{111} and Arg^{131}, which make direct contacts to RA via a salt-bridge [46]. The remarkable stability of these two strands even in the absence of the ligand suggests that they form a binding template, or a scaffold, which is likely to be an important part of the putative ligand-binding intermediate state of the protein. It has the ability to provide unobstructed ligand entry to the binding site (by lacking a structured α**II** segment, which

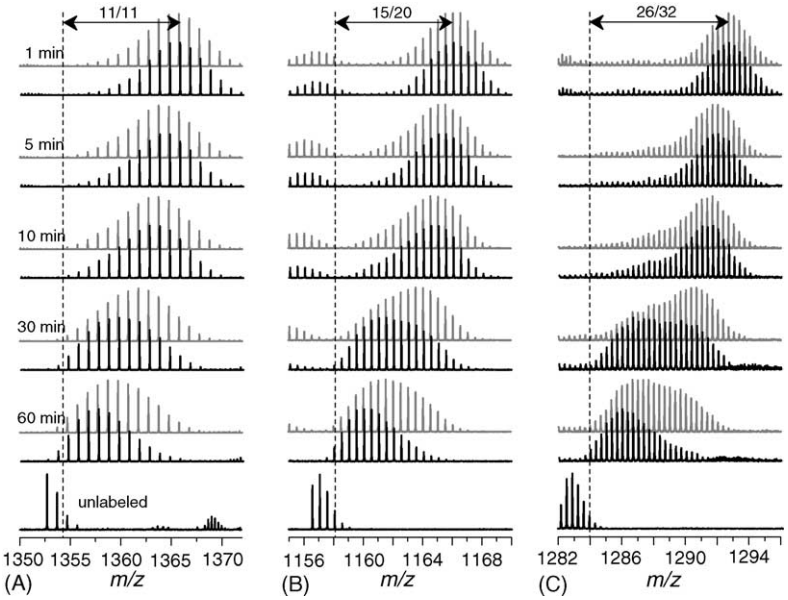

Fig. 11. Isotopic distributions of fragment ions y_{11}^{+} (A), y_{20}^{2+} (B) and y_{33}^{3+} (C) derived from CRABP I ions (CAD in the trapping hexapole region of a 4.7 T FT ICR MS). The protein undergoes HDX in solution at pH 3.4 in the absence (gray traces) and in the presence of RA (black traces). The bottom trace in each panel represents isotopic distribution of an unlabeled (essentially ^{2}H-free) fragment ion. Dotted line in each panel indicates the position of the centroid of an isotopic cluster of a fragment ion whose ^{2}H content is identical to that of the exchange buffer (^{1}H/^{2}H = 96:4).

normally restricts access to the cavity) and is also capable of guiding the ligand into the binding site and retaining it there efficiently (due to the presence of structurally stable β**8** and β**10** segments).

The presence of RA in the exchange buffer reduces local dynamics within every single segment of the protein, most dramatically in α**II** and β**4/5**. Reduced flexibility of these segments essentially eliminates the "escape routes" of RA from the protein cavity, in full agreement with the binding scenario presented earlier in the paper (see Fig. 7). These results are in good agreement with a recent study that utilized ^{1}H NMR to evaluate local dynamics of CRABP I [33]. However, neither HDX CAD MS carried out under native conditions, nor HDX NMR can characterize the individual structural scaffolds whose presence is clearly detected in such experiments. For example, the observed remarkable stability of β**8** and β**10** tells us nothing on whether these two segments form a single core state or represent two different intermediate states of the protein. An unequivocal conclusion can be reached only if HDX CAD MS experiments are carried out under conditions favoring correlated or semi-correlated exchange. Fig. 11 shows the time evolution of the isotopic distributions of several fragment ions (y_{11}^{+}, y_{20}^{2+} and y_{33}^{3+}, from which ^{2}H content of β**8**, β**9** and β**10** segments can be calculated) acquired under conditions favoring semi-correlated exchange. All distributions are clearly bimodal, although not well resolved in the case of smaller ions. Calculations of local ^{2}H content within the highly protected conformer indicate almost complete amide protection within β**8** and β**10** following 1 min of exchange (both retain ALL of their amide deutrons), while β**9** is highly flexible (retaining only half of its amide deutrons).

Similar results (high protection within β**8** and β**10** and poor protection within β**9**) were obtained under conditions favoring correlated exchange, providing strong evidence that stable β**8** and β**10** form a single binding template within a partially unstructured state of CRABP I. These experiments allow us to begin to reconstruct the architecture of the partially unfolded states of CRABP I that are so crucial for ligand binding. Understanding what makes these flexible ligand traps so efficient in catching their specific prey will undoubtedly contribute to our ability to engineer "smart" drug delivery systems for precisely targeted therapeutic uses in the future.

6. Conclusions

Functional importance of transient non-native protein structures can be seen in processes as diverse as recognition, signaling and transport. However, characterization and even distinct detection of such states is a challenging experimental task. Selected examples presented in this paper clearly demonstrate that HDX MS and HDX CAD MS carried out under native or mildly denaturing conditions often provide a unique opportunity to detect and characterize these elusive species in great detail. Despite being a relative newcomer to the field of biophysics, HDX MS has already made a plethora of important contributions to the field, and this number continues to grow. We are witnessing a very exciting time when continuous improvements in both MS hardware and new radical developments in experimental methodology dramatically expand the scope of problems that are amenable to study with this technique.

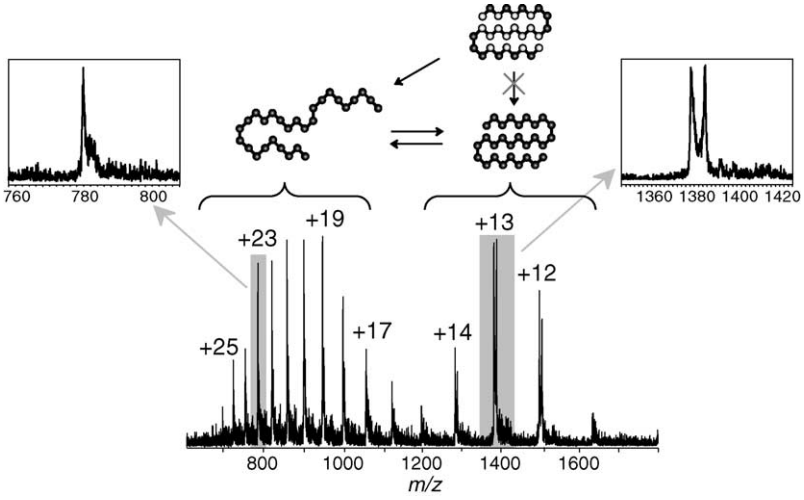

Fig. 12. ESI mass spectrum of a 17 kDa CRABP I variant acquired 10 min after a 1:50 dilution of the deuterated protein in $^{1}H_2O/C_2{}^{1}H_5O^{1}H$ solution. Higher charge states (+17 through +25) represent less structured protein conformations and lower charge states (+12 through +14) correspond to compact (tightly folded) protein states.

Despite the early success of HDX MS as a biophysical tool, there is certainly great room for both improving the experimental methodology and extending the scope of its applications. One of the challenges facing HDX MS is limited time resolution afforded in most experiments reported to date. Measurements carried out in real time do not allow the exchange reactions to be studied in the sub-second time scale. However, the temporal resolution can be improved quite dramatically by using a rapid mixing apparatus interfaced with the ESI source [14]. Another challenge that often complicates interpretation of HDX MS data is the presence of a large number of natural isotopes in protein molecules. The isotopic distributions of both intact protein ions and their fragments have significant width even if no exchange reactions are taking place. The finite widths of isotopic clusters pose a serious problem if HDX occurs in the EX1 regime, since the closely spaced peaks can overlap and make a distinct detection of respective intermediate states rather difficult (see, for example, Fig. 11). This problem may be solved by removing "irrelevant" isotopes from the protein, e.g., by expressing the protein in ^{13}C-, ^{15}N-depleted media [47]. HDX CAD MS methodology can also benefit greatly from utilizing charge-selected fragmentation of protein ions. Indeed, co-existence of several protein states in solution often leads to the appearance of bimodal charge state distributions in the mass spectra. The low charge density ions represent tightly folded conformers, while those at the higher charge states correspond to less compact protein states [38,48]. In many cases, the exchange behavior exhibited by these two groups is quite different, particularly if the exchange occurs in the EX1 regime [29]. It follows then that selective dissociation of protein ions from these two groups will allow the interpretation of the HDX CAD MS data to be greatly simplified. Indeed, the ^{2}H content of fragment ions derived from the highly charged protein ions will be characteristic of the less structured conformations, while isotopic signatures of the low charge density protein ions will contain contributions from all protein states (Fig. 12). Finally, the spectacular expansion of the range of proteins, from which abundant structurally diagnostic fragments ions can be generated using various gas-phase ion dissociation techniques [49], continues to broaden the scope of potential applications of HDX CAD MS.

Both hydrogen exchange and mass spectrometry have long histories, and their happy union seems to be a very recent event on this time scale. This relationship is obviously a synergistic one and is clearly poised to bring about a wealth of new exciting discoveries in biophysics and greatly advance our understanding of how biological macromolecules carry out their duties in living systems.

Acknowledgements

I wish to thank my group members and colleagues who have been directly involved with various stages of the work presented in this paper: Hui Xiao, Andras Dobo, Stephen J. Eyles and Joshua Hoerner. I am also very grateful to David L. Smith (University of Nebraska), Max Deinzer (Oregon State University), Lars Konermann (University of Western Ontario) and John Engen (University of New Mexico) for very helpful and informative discussions we have had in the past several years. The work presented in the paper was supported by a grant from the National Institutes of Health R01 GM61666.

References

[1] H.C. Urey, F.G. Brickedde, G.M. Murphy, Phys. Rev. 39 (1932) 164.
[2] G.N. Lewis, R.T. Macdonald, J. Am. Chem. Soc. 55 (1933) 3057.
[3] K.F. Bonhoeffer, R. Klar, Naturwissenschaften 22 (1934) 45.

[4] A. Hvidt, K. Linderstrom-Lang, Biochim. Biophys. Acta 14 (1954) 574.
[5] A. Hvidt, K. Linderstrom-Lang, Biochim. Biophys. Acta 16 (1955) 168.
[6] G. Champetier, R. Viallard, B Soc. Chim. 33 (1938) 1042.
[7] S.W. Englander, L. Mayne, Annu. Rev. Biophys. Biomol. Struct. 21 (1992) 243.
[8] H.J. Dyson, P.E. Wright, Annu. Rev. Phys. Chem. 47 (1996) 369.
[9] T.M. Raschke, S. Marqusee, Curr. Opin. Biotechnol. 9 (1998) 80.
[10] J.G. Kempf, J.P. Loria, Cell. Biochem. Biophys. 37 (2003) 187.
[11] J.R. Engen, D.L. Smith, Anal. Chem. 73 (2001) 256A.
[12] I.A. Kaltashov, S.J. Eyles, Mass Spectrom. Rev. 21 (2002) 37.
[13] A.N. Hoofnagle, K.A. Resing, N.G. Ahn, Annu. Rev. Biophys. Biomol. Struct. 32 (2003) 1.
[14] L. Konermann, D.A. Simmons, Mass Spectrom. Rev. 22 (2003) 1.
[15] R.W. Burley, C.H. Nicholls, J.B. Speakman, J. Text. Inst. 46 (1955) T427.
[16] J.L. Morrison, Nature 185 (1960) 160.
[17] J.L. Morrison, Biochim. Biophys. Acta 47 (1961) 606.
[18] O. Sepall, S.G. Mason, Can. J. Chem. 39 (1961) 1934.
[19] S.K. Sethi, D.L. Smith, J.A. McCloskey, Biochem. Biophys. Res. Commun. 112 (1983) 126.
[20] Z. Zhang, D.L. Smith, Protein Sci. 2 (1993) 522.
[21] V. Katta, B.T. Chait, Rapid Commun. Mass Spectrom. 5 (1991) 214.
[22] A. Hvidt, S.O. Nielsen, Adv. Protein Chem. 21 (1966) 287.
[23] Y. Bai, J.S. Milne, L. Mayne, S.W. Englander, Proteins 20 (1994) 4.
[24] S.W. Englander, Annu. Rev. Biophys. Biomol. Struct. 29 (2000) 213.
[25] H. Maity, W.K. Lim, J.N. Rumbley, S.W. Englander, Protein Sci. 12 (2003) 153.
[26] S.W. Englander, L. Mayne, J.N. Rumbley, Biophys. Chem. 101–102 (2002) 57.
[27] C.B. Arrington, L.M. Teesch, A.D. Robertson, J. Mol. Biol. 285 (1999) 1265.
[28] C.B. Arrington, A.D. Robertson, J. Mol. Biol. 300 (2000) 221.
[29] S.J. Eyles, T. Dresch, L.M. Gierasch, I.A. Kaltashov, J. Mass Spectrom. 34 (1999) 1289.
[30] J.R. Engen, Analyst 128 (2003) 623.
[31] G.S. Anand, D. Law, J.G. Mandell, A.N. Snead, I. Tsigelny, S.S. Taylor, L.F.T. Eyck, E.A. Komives, Proc. Natl. Acad. Sci. U.S.A. 100 (2003) 13264.
[32] J. Lanman, P.E. Prevelige Jr., Curr. Opin. Struct. Biol. 14 (2004) 181.
[33] V.V. Krishnan, M. Sukumar, L.M. Gierasch, M. Cosman, Biochemistry 39 (2000) 9119.
[34] D.W. Miller, K.A. Dill, Protein Sci. 4 (1995) 1860.
[35] C.D. Tsai, B. Ma, S. Kumar, H. Wolfson, R. Nussinov, Crit. Rev. Biochem. Mol. Biol. 36 (2001) 399.
[36] G.A. Papoian, P.G. Wolynes, Biopolymers 68 (2003) 333.
[37] H. Xiao, I.A. Kaltashov, S.J. Eyles, J. Am. Soc. Mass Spectrom. 14 (2003) 506.
[38] A. Dobo, I.A. Kaltashov, Anal. Chem. 73 (2001) 4763.
[39] S.J. Eyles, P. Speir, G. Kruppa, L.M. Gierasch, I.A. Kaltashov, J. Am. Chem. Soc. 122 (2000) 495.
[40] I.A. Kaltashov, S.J. Eyles, J. Mass Spectrom. 37 (2002) 557.
[41] N. Yamada, E. Suzuki, K. Hirayama, Rapid Commun. Mass Spectrom. 16 (2002) 293.
[42] Y.Z. Deng, H. Pan, D.L. Smith, J. Am. Chem. Soc. 121 (1966).
[43] M.Y. Kim, C.S. Maier, D.J. Reed, M.L. Deinzer, J. Am. Chem. Soc. 123 (2001) 9860.
[44] X. Yan, J. Watson, P.S. Ho, M.L. Deinzer, Mol. Cell. Proteomics 3 (2004) 10.
[45] J.K. Hoerner, H. Xiao, A. Dobo, I.A. Kaltashov, J. Am. Chem. Soc. 126 (2004) 7709.
[46] G.J. Kleywegt, T. Bergfors, H. Senn, P. Le Motte, B. Gsell, K. Shudo, T.A. Jones, Structure 2 (1994) 1241.
[47] A.G. Marshall, M.W. Senko, W. Li, M. Li, S. Dillon, S. Guan, T.M. Logan, J. Am. Chem. Soc. 119 (1997) 433.
[48] A. Mohimen, A. Dobo, J.K. Hoerner, I.A. Kaltashov, Anal. Chem. 75 (2003) 4139.
[49] Y. Ge, B.G. Lawhorn, M. ElNaggar, E. Strauss, J.H. Park, T.P. Begley, F.W. McLafferty, J. Am. Chem. Soc. 124 (2002) 672.
[50] C.E. Dempsey, Progr. Nucl. Magn. Res. Spectrosc. 39 (2001) 135.