Biochimica et Biophysica Acta 1834 (2013) 1202–1209

Contents lists available at SciVerse ScienceDirect

Biochimica et Biophysica Acta

journal homepage: www.elsevier.com/locate/bbapap

Review

Hydrogen-exchange mass spectrometry for the study of intrinsic disorder in proteins ☆

Deepa Balasubramaniam, Elizabeth A. Komives *

Department of Chemistry and Biochemistry, University of California, San Diego, La Jolla, CA 92093-0378, USA

ARTICLE INFO

Article history:
Received 20 August 2012
Received in revised form 11 October 2012
Accepted 15 October 2012
Available online 22 October 2012

Keywords:
Amide exchange
Intrinsically disordered proteins
Flow quench
Amyloid
Coupled folding and binding
Biopharmaceuticals

ABSTRACT

Amide hydrogen/deuterium exchange detected by mass spectrometry (HXMS) is seeing wider use for the identification of intrinsically disordered parts of proteins. In this review, we discuss examples of how discovery of intrinsically disordered regions and their removal can aid in structure determination, biopharmaceutical quality control, the characterization of how post-translational modifications affect weak structuring of disordered regions, the study of coupled folding and binding, and the characterization of amyloid formation. This article is part of a Special Issue entitled: Mass spectrometry in structural biology.

1. Introduction

1.1. Intrinsically disordered proteins in health and disease

At the time the complete human genome was sequenced, Dunker's group, attempting to predict the structures of all the new protein coding sequences, noted that the published genomes contained sequences that did not look like they would code for structured proteins [1]. Remarkably, some 30% of bacterial proteins, nearly 40% of archaeal proteins and perhaps as many as 50% of eukaryotic proteins were predicted to contain stretches of at least 40 consecutive disordered amino acids [2]. Around the same time, protein biophysicists discovered several proteins that appeared to not have a stable, unique structure, at least until they interacted with their physiological partners [3–6]. The first review describing these early experimental results was published in 1999 [7]. Many algorithms for prediction of disordered sequences in proteins have since been developed [8–15].

Intrinsically disordered proteins (IDPs) are found in functional niches that include transcriptional and translational control, cell cycle control and signaling. IDPs were found to be enriched among cancer related and signaling proteins [16], all processes that are in some way involved in physiological control. IDPs are frequently associated with diseases. Increases in the number of disordered repeat sequences increases susceptibility to the neurodegenerative diseases Huntington's and Parkinson's diseases [17–19]. Mutations in disordered regions can also result in the loss of important post-translational modification sites, leading to disease [20]. Diseases such as cancer, cardiovascular disease, amyloidosis, neurodegenerative diseases and diabetes often involve disordered proteins [17,21] probably because these often involve failures in protein signaling, structure, posttranslational modification and inability to interact correctly with physiological partners [21]. The more virulent strains of disease-causing viruses may be characterized by their increased levels of intrinsic disorder [22], and viruses in general appear to have a high proportion of genes potentially coding for disordered proteins [23].

1.2. Experimental characterization of intrinsic disorder

Intrinsically disordered regions of proteins can range from surface loops in otherwise structured proteins that are not observed in x-ray crystal structures to long stretches of polypeptide with low sequence complexity and no predicted structure. It is probably most helpful to think of intrinsic disorder as a continuum from slightly less than uniquely structured all the way to completely unfolded. Given this viewpoint, it is easy to see how different experimental measures will weigh-in differently on whether or not a particular region of a protein qualifies as intrinsically disordered. NMR spectroscopy is often a good experimental approach for detecting intrinsic disorder, as it provides a highly sensitive read-out of resonances that change their chemical environment and this is very useful for monitoring changes in an IDP upon interaction with a binding partner [24,25]. NMR is also very useful for detecting protein dynamics, however

Abbreviations: HXMS, hydrogen/deuterium exchange coupled to mass spectrometry; IDP, intrinsically disordered protein; LRP1 CT, the cytoplasmic tail of the low density lipoprotein receptor-related protein, LRP1; AR, ankyrin repeat

☆ This article is part of a Special Issue entitled: Mass spectrometry in structural biology.

* Corresponding author at: Department of Chemistry and Biochemistry, U. C. San Diego, 9500 Gilman Dr., La Jolla, CA 92093-0378, USA. Tel.: +1 858 534 3058; fax: +1 858 534 6174.

E-mail address: ekomives@ucsd.edu (E.A. Komives).

1570-9639/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.bbapap.2012.10.009

there is a blind spot for dynamic regions that move on the NMR timescale because resonances that exchange between different chemical environments will have broadened signals that may disappear completely. If the disordered region is so dynamic that the atoms move very quickly between different environments, the resonances will have averaged chemical shifts and will appear as strong signals. Other more global indicators of secondary structure, such as circular dichroism (CD) and/or packing of the hydrophobic cores, 1-anilino 8-naphthalene sulfonate (ANS) binding and fluorescence changes are also useful for detecting large regions of disorder and changes in disorder upon binding.

1.3. Amide exchange

The use of mass spectrometry to detect amide hydrogen/deuterium exchange (HDX) has found increasing applicability since its inception in the early 1990s [26]. Both electrospray ionization (ESI) [26] and MALDI-TOF [27] are suitable approaches for analyzing whether an amide group has exchanged with solvent deuterium when the sample is incubated in ^2H_2O (D_2O). H/D exchange analysis can be performed on intact proteins [28,29], however it is most often useful to perform the exchange reaction and then digest the protein with an acid-tolerant protease to localize the sites of exchange [30].

Mass spectrometry-detected amide exchange (HXMS) experiments were traditionally performed to study protein folding and unfolding and the measured amide proton exchange rates were used to interpret the amount of hydrogen bonding [31], however, HXMS can also be a useful indicator of solvent accessibility changes at protein surfaces due to ligand binding [32–34] and for epitope mapping of antibodies [35]. The rates of amide exchange in small peptides have been accurately measured by NMR, and these reveal subtle differences in the intrinsic rates of exchange due to sequence context [36]. NMR has the advantage that the rate of exchange of single amides is measured, but the disadvantage is that the protein must be of a size that is tractable by NMR, and it must be able to be expressed and isotopically labeled with ^{15}N so that the resonance assignments for the amide protons can be made. Mass spectrometric measurements avoid the need for isotopic labeling of the protein and can be performed on proteins of any size [37]. The mass spectrometry measurements rely on proteolytic digestion of the protein and analysis of the peptide fragments. Thus, the resolution is limited by the sizes of the peptides although, if many overlapping peptides are obtained, this can improve the resolution. Top-down approaches as well as various alternative ionization schemes have recently been demonstrated to yield accurate single amide resolution as well [38,39].

In folded proteins, the rates of exchange vary widely and depend on exposure to solvent as well as folding and/or dynamics. If the rate of exchange of a single amide can be measured, it is possible to compute a protection factor, which is simply the ratio of the intrinsic exchange rate to the observed rate. Protection factors vary by 10–12 orders of magnitude from amides in completely unfolded proteins to amides in very well-folded regions of proteins. Teasing apart the contributions of dynamics, solvent accessibility, and foldedness remains an unsolved problem. Computer simulations can reproduce the HDX results [40,41] and theoretical analyses of amide exchange in proteins are now allowing insights into the mechanisms of exchange [42].

This review will discuss what can be learned about IDPs and intrinsically disordered regions of proteins from HXMS experiments. In the first part, we will discuss the utility of HXMS for detecting intrinsically disordered regions in proteins. In the second part, we will discuss how HXMS can be used to monitor coupled folding and binding, which is a phenomenon that is extremely common among IDP sequences, as they often are only disordered prior to finding their binding partner. In the third part, we will discuss the use of HXMS to characterize aggregates and oligomers. Another property of IDPs is that they are prone to aggregation, and this often leads to diseases such as Parkinson's Alzheimers' and Creutzfeld–Jacob disease among others. HXMS is extremely useful in characterizing the oligomers that form from IDPs. With evidence mounting that the disease-causing forms of these aggregation-prone proteins may not be the final fibrillar form but an intermediate oligomer, HXMS is becoming critical for characterizing such intermediates. In the final part, we will discuss where developments could still be made to expand the information that can be gained from analyzing IDPs by HXMS.

1.4. Identification of disordered regions in proteins by HXMS

1.4.1. Identification and removal of intrinsically disordered regions for structure determination

Unstructured or disordered regions often pose practical problems when it comes to structure determination by X-ray crystallography and NMR. In X-ray crystallography, flexible or disordered regions often interfere with crystal formation and result in poor quality crystals. Even if crystals can be obtained, the electron density from disordered regions often cannot be resolved and appropriate structural models of these regions cannot be built. NMR resonances for disordered regions usually are highly overlapped making them hard to assign and sometimes mask the signal from the structured residues. Disordered regions often also weakly associate, and at the high protein concentrations necessary for structure determination, weak association can cause line broadening. It is preferable, therefore, to remove these unstructured regions prior to structure determination, as long as the structural integrity is not lost.

HXMS is being increasingly used to discover problematic disordered regions for construct optimization in structure determination. As outlined schematically in Fig. 1, regions of intrinsic disorder exchange within a few minutes, and are readily identified by a protocol in which the protein is incubated for short periods of time in deuterated buffer to exchange the exposed amides, then the reaction is quenched by decreasing the pH to 2.5 and the temperature to 0 °C, the sample is digested with pepsin, and finally mass spectrometry is used to determine the deuterium content of the resulting peptides (Fig. 1).

Pantazatos et al. have outlined an improved HXMS method to allow rapid identification of unstructured regions in proteins and have demonstrated the use of this technique in improving crystallographic success [43]. As part of a large scale structural genomics project, they screened 24 proteins from *T. maritima* by HXMS. The proteins were recombinantly expressed and purified from *E. coli* and protein fragmentation maps after pepsin digestion and ESI LC MS analysis before and after exchange with D_2O for 10 s were analyzed to determine the amides that readily exchanged and identify unstructured and highly dynamic parts of the proteins. The entire analysis was completed rapidly (within 2 weeks) and data were obtained for 21 of the 24 proteins screened. Complete coverage was obtained for 16 of the 21 proteins. Comparison of the HXMS data to the known crystal structure of the *T. maritima* thy1 protein TM0449 showed that they were able to accurately localize disordered regions within the protein. They identified two regions in the protein that exchanged with D_2O rapidly that corresponded to parts in the crystal structure that were disordered. One of the highly exchanging regions corresponded to the substrate binding region in the protein demonstrating how HXMS can provide information on binding sites within proteins [32]. They also compared the exchange results for the *T. maritima* GroES heat shock protein TM0505, which produced poorly diffracting crystals to the homologous *M. tuberculosis* structure and showed that the highly dynamic regions of the protein mapped mostly to the disordered residues in GroES that are involved in binding to GroEL. From the results of the HXMS analysis the authors were able to design constructs that would have better crystallographic success. Indeed, truncation of highly disordered C-termini detected by HXMS in both TM0160 and TM171 led to improved crystals leading to higher resolution structures of 1.7 Å and 2.3 Å,