



## RESEARCH ARTICLE

# Many Overlapping Peptides for Protein Hydrogen Exchange Experiments by the Fragment Separation-Mass Spectrometry Method

Leland Mayne,<sup>1,2</sup> Zhong-Yuan Kan,<sup>1</sup> Palaniappan Sevugan Chetty,<sup>3</sup> Alec Ricciuti,<sup>1,2</sup> Benjamin T. Walters,<sup>1,2</sup> S. Walter Englander<sup>1,2</sup>

<sup>1</sup>Johnson Research Foundation, Department of Biochemistry and Biophysics, Perelman School of Medicine, University of Pennsylvania, 1006 Stellar-Chance Labs, 422 Curie Boulevard, Philadelphia, PA 19104, USA

<sup>2</sup>Graduate Group in Biochemistry and Molecular Biophysics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA

<sup>3</sup>Department of GI, Nutrition, and Hepatology, Children's Hospital of Philadelphia, Philadelphia, PA 19104, USA

### Abstract

Measurement of the naturally occurring hydrogen exchange (HX) behavior of proteins can in principle provide highly resolved thermodynamic and kinetic information on protein structure, dynamics, and interactions. The HX fragment separation-mass spectrometry method (HX-MS) is able to measure hydrogen exchange in biologically important protein systems that are not accessible to NMR methods. In order to achieve high structural resolution in HX-MS experiments, it will be necessary to obtain many sequentially overlapping peptide fragments and be able to identify and analyze them efficiently and accurately by mass spectrometry. This paper describes operations which, when applied to four different proteins ranging in size from 140 to 908 residues, routinely provides hundreds of useful unique peptides, covering the entire protein length many times over. Coverage in terms of the average number of peptide fragments that span each amino acid exceeds 10. The ability to achieve these results required the integrated application of experimental methods that are described here and a computer analysis program, called ExMS, described in a following paper.

**Key words:** Hydrogen exchange, Protein, HX-MS, HDX

## Introduction

The main chain amide NH hydrogen, one on every amino acid (except proline) in every protein molecule, engages

in continual exchange with the hydrogens in water. Hydrogen exchange (HX) rates depend on pH, temperature, immediately neighboring side chains, and isotope effects in ways that are well understood [1–3], so that in the absence of protecting structure, HX rates can be accurately predicted. In structured proteins, the blocking of normal HX chemistry (H-bonding, sterics) can slow HX rates over five to 10 orders of magnitude. The exchange of structurally protected hydrogens is governed by dynamic structural fluctuations that transiently remove the structural barriers and expose the hydrogens to solvent where normal HX chemistry can proceed [4–6]. An understanding of these relationships

Palaniappan Sevugan Chetty, Alec Ricciuti and Benjamin T. Walters contributed equally to this work.

**Electronic supplementary material** The online version of this article (doi:10.1007/s13361-011-0235-4) contains supplementary material, which is available to authorized users.

Correspondence to: Leland Mayne; e-mail: mayne@mail.med.upenn.edu

Received: 20 May 2011  
Revised: 11 August 2011  
Accepted: 12 August 2011  
Published online: 14 September 2011

together with well-designed HX measurements can provide a powerful experimental approach for the study of protein structure, structure change, stability, dynamics, and interactions, resolved to the level of individual amino acids [7–19].

Early HX methods were designed to measure the unresolved H-D exchange kinetics of whole proteins, initially by a mass/density method [20], then IR spectroscopy [21], and then by an accurate and efficient tritium labeling/gel filtration method [22, 23]. In efforts to achieve greater structural resolution, the tritium labeling method was extended to include a proteolytic fragmentation—HPLC separation stage [24–26] that could provide resolution to the level of protein subfragments, typically ranging from five to perhaps 30 residues in length. This method was limited by the inability to rapidly separate many proteolytic fragments by HPLC alone. The rise of two-dimensional (2D) NMR capabilities [27–29] led to the ability to measure HX directly at amino acid resolution [30, 31] and, for a time, eclipsed the further development of the fragmentation separation method. However, routine NMR analysis is limited to relatively small (< ~250 residues) highly soluble proteins that are available in quantity. It has been shown that HX investigations of larger and more demanding protein systems, reaching up even to massively aggregated systems like amyloid, can be achieved by adding a stage of mass spectrometric (MS) separation and analysis to the fragmentation method [10, 32–39]. As in the earlier H-T exchange method [26], proteins are exposed to H-D exchange labeling, samples are taken in time and quenched into a low pH, low temperature condition where exchange is largely halted. The protein is proteolyzed into fragments, they are partially resolved by fast HPLC, and the LC eluant is injected into a mass spectrometer for a second dimension of peptide separation and mass analysis.

A number of problems continue to limit the HX-MS method. Complete online analysis systems have been developed but they have proven to be surprisingly problematic. The measurement provides resolution only to the level of individual fragments, which may wind over some sizeable region in the native protein. The peptide fragments obtained have generally represented something less than the whole length of the protein and, within any given peptide fragment, one generally measures the exchange of only a fraction of the residues because the different amides typically span many orders of magnitude on the HX time axis [40]. Given these problems, one cannot distinguish where in the fragment or precisely where in the protein the hydrogens that are measured are placed. When a change in HX rate is detected, one cannot distinguish which residues in any given fragment change and which do not, or whether the observation detects some small change shared by many sites or a larger change due to some specific subset. Therefore, one is unable to infer the size, identity, energetics, or exact placement of the structure change.

Recent progress shows that the sequence resolution problem can be overcome by a top-down analysis in which the exchange-labeled protein is fragmented while in flight in

the mass spectrometer. This approach has been plagued by the scrambling problem in which H and D atoms are redistributed due to the extreme bond vibrational energies that accompany collision induced dissociation (CID). The application of non-ergodic electron capture dissociation and electron transfer dissociation methods now accomplish near amino acid resolution [41–44] for relatively small proteins.

Alternatively, high structural coverage and resolution might be obtained by a bottom up fragmentation-separation approach if one could compare HX data for many sequentially overlapping fragments [45]. However, the ability to obtain very large numbers of peptide fragments and to efficiently analyze them in HX-MS experimentation has not been achieved. Here, we describe a system that can routinely provide many useful peptides, covering the protein length many times over. Also, we have written a computer program called ExMS that can analyze these many peptide fragments under conditions that are useful for HX-MS experiments.

This paper describes the methods used to obtain and validate large numbers of peptide fragments with four different proteins. A following paper describes the ExMS data analysis program [46]. These advances extend the capabilities of modern HX-MS experimentation and move toward ultimate amino acid resolution.

## Experimental

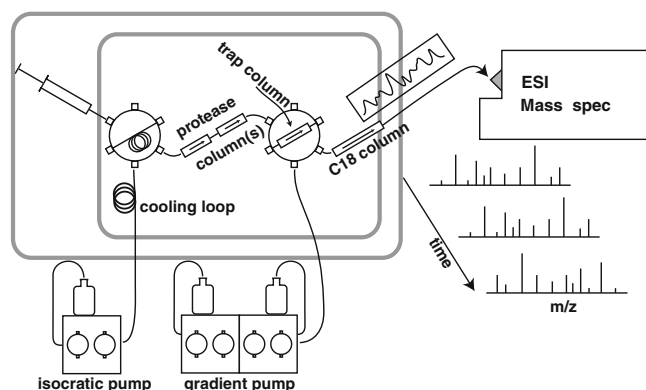
### *Proteins*

Three proteins— $\alpha$ -synuclein (140 residues), maltose binding protein (370 residues), and heat shock protein 104 (908 residues)—were expressed by recombinant methods in *E. coli*. Apolipoprotein A-1 (243 residues) was isolated from human blood. Purification followed previously described methods [47–50]. Proteases used for proteolytic fragmentation were porcine pepsin, fungal protease XIII, and the two in tandem (Sigma-Aldrich, St. Louis, MO, USA).

### *Hardware and System Details*

The online flow system used and system details and parameters are illustrated in Figure 1 and described below.

Experimental samples (~50  $\mu$ L of ~1  $\mu$ M protein in 0.1% formic acid at pH 2.5) are injected into the on-line flow system diagrammed in Figure 1. The flow system carries the protein solution at 50–120  $\mu$ L/min through an immobilized acid protease column for proteolysis (15 or 60  $\mu$ L total volume each). Digested peptides are directed through a second valve to a small C8 or C18 trap column (1  $\times$  5 mm, 5  $\mu$ m beads). After sufficient flow (~3 min at 50  $\mu$ L/min) to transport the peptides into the trap column and wash away buffer salts, the second valve is switched, placing the trap column in the flow of a low volume HPLC pump. A water/acetonitrile gradient (6  $\mu$ L/min, 10%–50% AcCN over 10–15 min) elutes peptides from the trap column through an



**Figure 1.** On-line HX-MS analysis system. The entire flow system is contained within a Peltier cooled chamber ( $21 \times 15 \times 25$  cm; from an automobile accessory supplier) that maintains  $0 \pm 1$  °C, monitored with a thermocouple thermometer. An internal fan circulates air across the Peltier element and through the chamber. Liquid flow is precooled in a large loop positioned in the airflow. Switching valves are mounted on a ridged board backed with foam insulation with valve handles outside easily accessible for manual operation. The wetted parts of the valves, the injection loop, and columns are contained in the cold chamber. Liquid lines and the thermometer leads are threaded through small holes in the insulation and mounting board. A short length of outflow tubing from the C18 analytical column to the MS electrospray source is packed with ice-filled plastic bags

analytical C18 column ( $0.3 \times 50$  mm,  $3 \mu\text{m}$  beads) for rough peptide separation, and then to the electrospray needle for further separation of the peptides by mass. Narrow tubing ( $25$  or  $65 \mu\text{m}$  i.d.) is used in the slow flow chromatography stage. To minimize eluant peptide overlap, we use an acetonitrile gradient shaped to elute constant numbers of peptides per unit time. Typical peptide elution peak widths are  $20$  sec wide at baseline. We have not found peak widths to be limiting in the MS data even for hundreds of peptides, due especially to the ability of the ExMS analysis program [46] to recognize peptides in MS data even in the presence of significant spectral overlap.

Cleaning steps between serial experiments (several up-down gradients) elute very large peptides not useful for HX-MS experiments, helping to avoid peptide carryover from earlier MS runs and to maintain low column back pressure. If some peptides present a particularly difficult carryover problem [51], the availability of many other peptides makes it feasible to simply remove them from the experimental list. Overall, each experimental cycle takes  $20$ – $25$  min followed by a  $10$ – $15$  min cleaning cycle, resulting in a  $40$  min total time for each run. This allows for as many as  $10$  experimental HX runs each day in addition to an all-H run to calibrate chromatographic retention time for each peptide and an all-D run to calibrate their back exchange.

The flow rates and column sizes noted were chosen as a compromise between the competing demands of transit time, resolution, and back pressure. Particular care must be taken

with flow system connections to assure against leaks under significant back pressure and to minimize unswept dead volume, which can lead to peak trailing and problems with peptide identification, although the ExMS analysis program operates to minimize this problem. Tubing is carefully cut, inspected, and connected, and ports are sprayed out with clean pressurized air on assembly. Common problems include slippage of connectors due to under-tightening, blockage due to over-tightening, and the trapping of particles at connections.

### Proteolysis

The quenched experimental protein ( $\text{pH} \sim 2.5$ ,  $0$  °C) can be fragmented by adding an acid protease in a batch mode but passage through an immobilized protease column is more effective by far [52, 53]. When necessary, protein unfolding can be promoted and proteolysis improved by the inclusion of low concentrations of denaturant (urea, GdmCl, GdmSCN;  $\sim 0.5$  M) or TCEP to promote disulfide reduction [54]. The protease columns used here are  $1$  or  $2$  mm  $\times$   $20$  mm guard columns packed with POROS AL to which either pepsin or fungal protease XIII was coupled using  $\text{Na}_2\text{SO}_4$  for salting out the protein during the coupling reaction at room temperature. It may be useful to gel-filter protease preparations to remove any contaminating amines such as Tris buffer, which will compete for coupling sites on the beads. Protease columns can be stably stored at  $4$  °C in  $\text{pH} 2.5$  formic acid for months.

To maximize the number of overlapping peptides obtained, we use these two proteases, with broad but different specificities, separately and in tandem. Additional peptides can be produced by additional acid proteases with somewhat different specificity [18, 26]. At the flow rates noted, the protease column transit time, from  $3$  to  $35$  s, does not produce a final limit digest. Therefore, larger and smaller columns and slower and faster flow rates can be used to bias toward populations of shorter and longer peptides, respectively. These and other operations can increase the peptide numbers listed here.

### Mass Spectrometer

The mass spectrometer used was a ThermoScientific LTQ Orbitrap XL (Thermo Fisher Scientific, Waltham, MA, USA) operated at a  $60,000$  resolution ( $\sim 1$  s/scan). The MS fragment separation capability is illustrated in Figure S1. Data were collected in profile mode with an AGC target of  $10^6$ . The mass calibration was checked daily and recalibrated when necessary ( $< 2$  ppm rms deviation over  $9$  masses/ $100$  scans in the calibration mix). Source parameters were: spray voltage  $3.5$  kV; capillary voltage  $40$  V; tube lens  $170$  V; capillary temperature  $150$  °C. MS/MS CID fragment ions were usually detected in the LTQ stage of the instrument (see main text) in centroid mode at normal scan rate with an AGC target value of  $10^4$ . CID fragmentation was at  $35\%$  energy for  $30$  ms at  $Q$  of  $0.25$ . HCD fragmentation energy was  $35$ .

## Results and Discussion

The proteolytic fragments described here were identified initially by data-dependent MS/MS experiments. Peptides in the MS/MS list were evaluated by a quality score ( $P_{\text{pep}}$ ) that we independently calibrated to reduce the probability of false assignments to 1/1000. Large numbers of high scoring peptide fragments were found. We tested the ability of the ExMS analysis program [46] to find and definitively validate these many peptides in MS spectra under conditions that mimic MS-HX experiments with partially deuterated samples. The accuracy of the peptide fragment assignments was validated in a variety of ways. These steps are detailed below.

### Peptide Identification

Protein fragments were prepared for MS/MS analysis by passage through the same online system designed for HX-MS experiments (Figure 1). Samples unfolded at the low pH quench condition used in HX-MS experiments were manually injected into the cold online flow system, where the protein was proteolyzed in immobilized acid protease columns, the peptide fragments were caught on a trap column, washed, and then gradient eluted through a small HPLC column directly into the mass spectrometer by electrospray for further mass analysis.

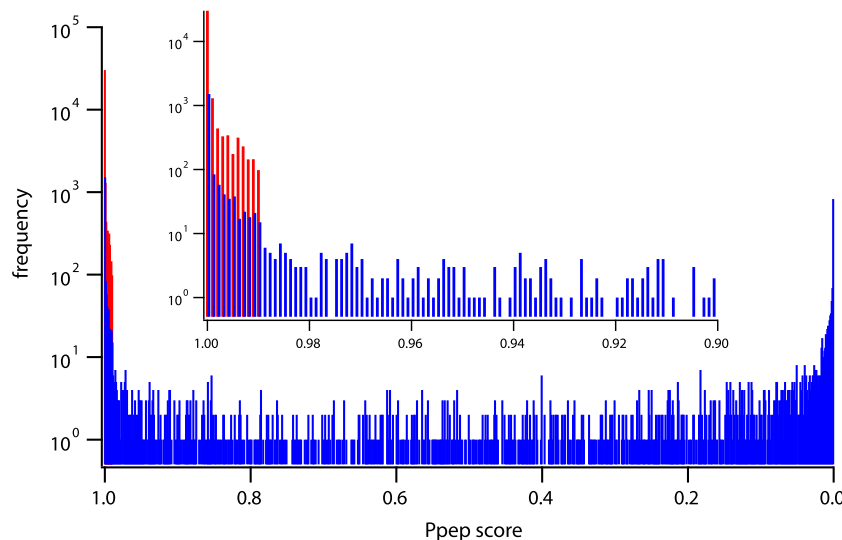
Starting with all-H protein samples, high resolution parent ion data were collected in the Orbitrap detector, selected peptide ions were CID-fragmented, and then detected in either the LTQ or Orbitrap stage. The four most intense peptide ions in each scan were selected for MS/MS

analysis using a dynamic exclusion list to deselect already identified peptides. Three separate MS/MS runs were done for each proteolysis condition (pepsin, fungal protease XIII, the two in tandem), using an exclusion list for the second and third runs to avoid previously processed peptides with excellent  $P_{\text{pep}}$  scores (see below).

SEQUEST (ThermoScientific Bioworks 3.3.1) was used to analyze the MS/MS results against a large database of protein sequences including the four experimental proteins studied here and all potential contaminants, namely the acid proteases used, the list of proteins used in this lab, human and dog keratin, and the entire *E. coli* proteome because the experimental proteins were expressed in *E. coli*. An enlarged decoy set containing also the human proteome was used with apoA-1 because it was isolated from human blood. Search tolerance was 4 ppm for parent ions and 0.1 u for MS/MS fragment ions.

### $P_{\text{pep}}$ Calibration and Use

We used the Bioworks  $P_{\text{pep}}$  score to grade the quality of peptide MS/MS assignments.  $P_{\text{pep}}$  is a proprietary, relatively poorly documented statistical goodness of fit parameter provided in Bioworks. Although this quality parameter was designed for a different purpose, it proved to be useful here. We calibrated the ability of the  $P_{\text{pep}}$  score to distinguish false peptide identifications in the present system. MS/MS data were analyzed against a decoy sequence database with the amino acid sequence of all proteins reversed [55]. In this case, matches obtained clearly represent false identifications. Figure 2 shows in red the



**Figure 2.** Distributions of Bioworks  $P_{\text{pep}}$  scores. Known false identifications, in red, represent SEQUEST hits for MS/MS data run against a large database with reversed sequences. Data in blue represent hits on the experimental protein for MS/MS data run against the large database with all proteins in their forward sequences. Results for the four proteins studied are merged to provide a large statistical sampling. The inset focuses on the poorest scoring matches. These results show that the cutoff  $P_{\text{pep}}$  value necessary to reduce false identifications to <0.1% is 0.990. (Other measures of assignment quality could be calibrated in the same way.)

number distribution of false identifications plotted against their  $P_{\text{pep}}$  score. (Higher  $P_{\text{pep}}$  score indicates lower identification quality.) The red distribution in Figure 2 shows that in order to reduce the false identification probability to  $<1/1000$ , the  $P_{\text{pep}}$  cutoff score would have to be set at 0.990. We adopted this cutoff value.

The MS/MS data were also analyzed with the experimental protein and all other proteins in the large sequence database in their correct sequence. Figure 2 shows in blue the distribution of hits on the experimental protein versus  $P_{\text{pep}}$  score. A list of potentially useful peptides was assembled from the MS/MS results by rejecting apparent peptide matches with  $P_{\text{pep}}$  scores  $>0.990$ , which amounted to 40% of hits on the experimental proteins.

Contaminant peptides with scores below (better than) the cutoff score were found to originate only from the pepsin used (two overlapping C-terminal peptides). No *E. coli* protein had any peptide with a good score whether the experimental protein itself was included or omitted in the search database. In the case of apo A-I, hits were found on some other lipoproteins included in the human data base and seen at very low levels in PAGE analysis.

To confirm the usefulness of these many peptides for HX-MS experiments, another requirement is that the peptides must be found by the ExMS program [46] in MS spectra under HX-MS conditions. For each experimental protein, we did MS runs with fully protonated samples and with samples that had been equilibrated in 50%  $D_2O$  solvent, just as for HX-MS experimentation. The ExMS analysis searches through the MS spectra for each peptide in turn and subjects each found peptide to a series of validation tests. The all-H MS runs provided an additional test for the correctness of the MS/MS identification of the peptides since their exact theoretical isotopic peak positions and envelope shape are known. ExMS directly found close to 100% of the listed peptides. Peptide identification is most challenging in the 50% D condition where isotopic envelope widths and therefore spectral overlap are maximized and  $m/z$  values to be expected for individual peaks are less well defined. These experiments provided the opportunity to test and optimize the various settable parameters in the ExMS program under the most difficult conditions. The optimized version of ExMS automatically found 50% to 75% of the peptide list, and most of the rest were easily determined by using the manual check function of the program. Peptides not definitively found ( $\sim 20\%$ ) were removed from the peptide list.

### Peptide Numbers

For the four proteins studied here, Figure 3 shows the unique peptides (many with more than one charge state, not shown) that were initially identified by MS/MS with  $P_{\text{pep}} < 0.990$  and were also found by ExMS in both all-H and 50% D MS runs. The number of useful peptides approach or exceed the number of amino acids in each protein, namely 222 from

$\alpha$ -synuclein (140 amino acids), 239 for human apolipoprotein A-I (243 amino acids), 443 for maltose binding protein (370 amino acids), and 664 for heat shock protein 104 (908 amino acids). Figure 3 also shows the coverage per amino acid that is provided by the multiple overlapping peptides. The amino acid coverage obtained is variable, sometimes even falling to zero when high probability cut sites are closely spaced [56]. More importantly, the usual coverage per residue is large, averaging over 10 peptides per amino acid position.

Table S1 reports the survival history of peptides for each protein through the stages of identification and validation. The record indicates that these numbers could be substantially increased by the use of additional acid proteases and MS/MS runs, as well as manipulation of the degree of proteolysis (flow rate, protease column size, pH).

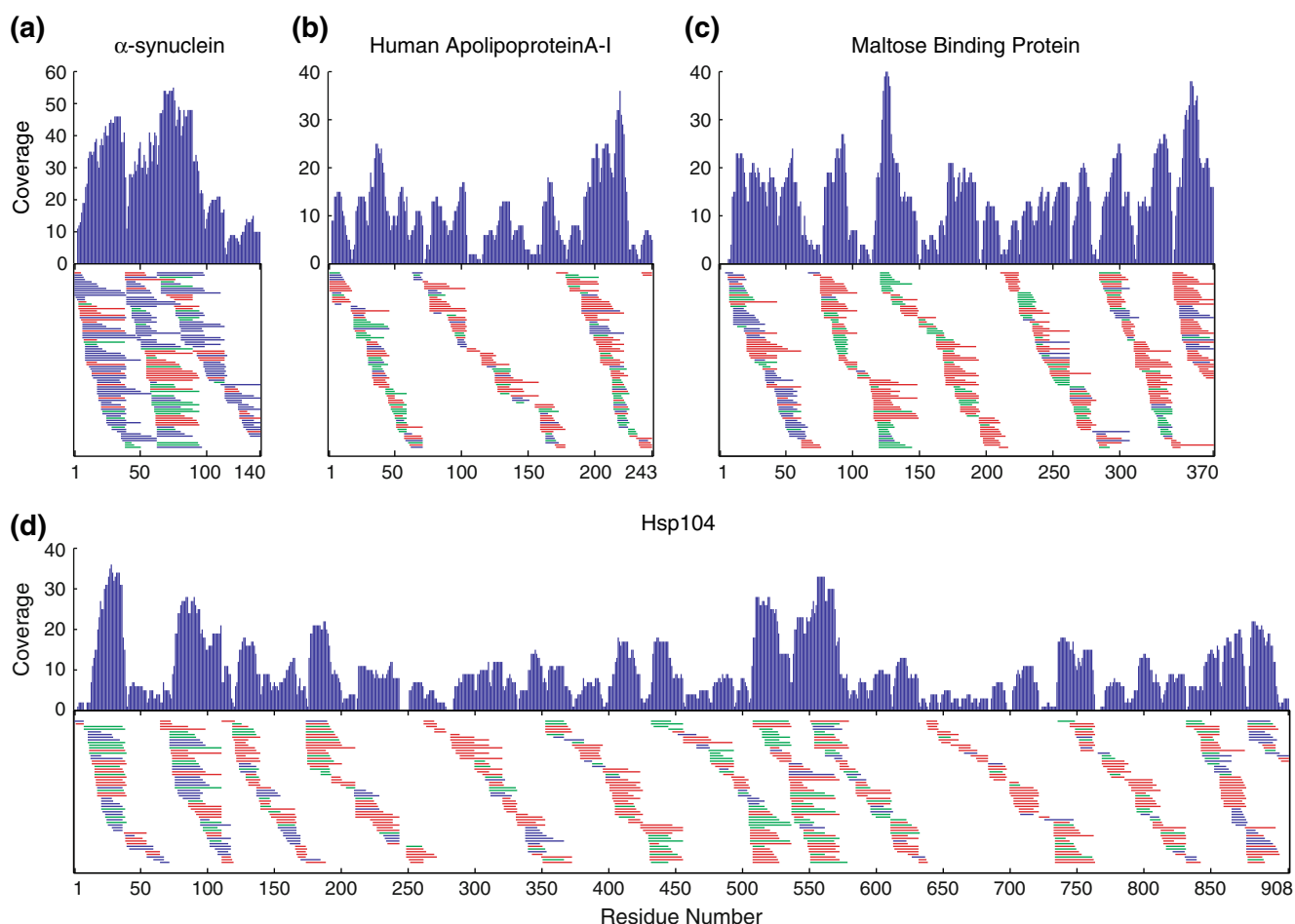
### The Validity of MS/MS Peptide Identification

We empirically calibrated in our system a quality score ( $P_{\text{pep}}$ ) against known incorrect MS/MS identifications and accepted only those assignments with a score that reduces the false positive probability to  $<1/1000$ . These peptides were further tested by ExMS to ensure that they exhibit their theoretically expected isotopic peaks and envelope shapes. The unusually large number of useful peptide fragments found raises the possibility that they may include incorrectly identified peptides or correctly identified peptides that in fact represent contaminant proteins. Accordingly, further tests were done.

To control for peptide fragments that might come from contaminants in our protein preparations, the SEQUEST analysis of MS/MS data was run against a large data base containing all possible contaminant proteins. Individual hits with good  $P_{\text{pep}}$  scores on any contaminant protein were rare, as expected from the analysis in Figure 2, with the exceptions noted before (two C-terminal pepsin fragments; known low level apolipoprotein contaminants). This was true whether the experimental protein was included in the comparison data base or not. Another indicator is that the number of peptides found for our different proteins increases in proportion to the size of the experimental protein (Figure 3), whereas the same contaminant decoy set was used for the different proteins except for the greatly expanded apoA-I set (the human proteome), which did not produce more hits.

The possibility of peptide production by in-source fragmentation was checked by increasing the source fragmentation voltage, and in other experiments the capillary and tube lens voltage. The number of peptides found did not change except at high source fragmentation voltage ( $>60$  V) where the number fell precipitously.

Tests were done to check the quality of the MS/MS identifications. Figure S2 shows the number of b and y daughter ions found in MS/MS analyses of  $\alpha$ -synuclein peptides with good  $P_{\text{pep}}$  score. The number averages about



**Figure 3.** Unique peptides and amino acid coverage obtained for four proteins by the methods described here. Peptides are represented by horizontal bars. The different colors distinguish peptides obtained by proteolysis with porcine pepsin (red), fungal protease XIII (blue), and the two proteases in tandem (green). The coverage graphs show the number of peptides that span each amino acid residue position. For each protein, the list of peptides identified by MS/MS was culled to reduce potentially false identifications to one per thousand, and to eliminate those not definitively found by the ExMS program in MS spectra from runs with all-H protein and 50% D-labeled protein

1.5 times the number of residues in each peptide. Of the 543 peptides found (all three protease combinations), only one has a suspiciously low number of observed daughter ions.

In the MS/MS experiments described before, parent peptides selected based on Orbitrap scans were fragmented by CID and daughter ions were analyzed in the LTQ instrument. In tests done with  $\alpha$ -synuclein, we fragmented by both HCD and CID and detected daughter ions in the Orbitrap. Of 129 previously identified peptic peptides derived from CID/LTQ analysis, 107 were validated by CID/Orbitrap analysis and 103 by HCD/Orbitrap analysis.

A large fraction (~70%) of our peptides were found in more than one charge state. The MS/MS analysis reached the same identification from these independent CID analyses. Almost all of the peptides paired in this way show very similar chromatographic retention time ( $\Delta RT < 5$  s). For the very few retention time violations found, one peptide of the pair had poor  $P_{\text{pep}}$  score and appears to be due to low level chromatographic peak tailing.

MS/MS analysis is more difficult if two or more different peptides are co-fragmented by CID. The situation was checked for our largest protein where this problem will be maximized. For heat shock protein 104 cleaved by pepsin + fungal protease, CID/SEQUEST identified 564 peptides. Of these, 70 had co-eluting ions that fell within  $\Delta m/z = \pm 0.5$  of the maximum intensity peak and so were co-selected for fragmentation in the LTQ. Of these, 66 had excellent scores in the MS/MS analysis and four were simply good.

The fact that many groupings of peptides are found with common termini (Figure 3), due to high probability cut points, speaks for their accurate identification. We checked our pepsin cut sites against the list of prohibited sites found by Hamuro et al. [56] (no Arg, His, Lys, or Pro at the P1 site). One violation was found, at an Arg Phe site.

A severe test is offered by checking for consistency in HX-MS experiments. In H-D exchange experiments so far done, identical but differently charged peptides yielded results in excellent quantitative agreement, indicating that

they were correctly identified. Among sequentially overlapping peptides, no inconsistencies have been found.

In summary, many independent measures document the reality and accurate identification of the large numbers of peptide fragments found in this work.

### Peptide Inventory

It is commonly assumed that the number of useful peptide fragments in HX-MS experiments is limited by the small number of fragments produced by proteolysis. We inventoried the peptide fragments in a pepsin plus fungal protease digest of  $\alpha$ -synuclein. A home-written program was used to search the MS spectra for “features” that appear to represent peptide fragments. There are thousands, spread over orders of magnitude in intensity. Analogous observations have been made before. We culled the list down to about 700 distinctly peptide-like features with non-redundant mono-isotopic mass. All were found to match some theoretical  $\alpha$ -synuclein peptide.

The large number of peptides produced and observed that appear to represent true  $\alpha$ -synuclein peptide fragments compares with the much smaller number identified in this work, namely 97 unique peptides in three MS/MS runs for  $\alpha$ -synuclein (pepsin + fungal protease, Table S1). One assumes that a similar analysis would find similar numbers for other proteins. These results tend to change one’s perspective. The question is not how the many peptides that we find could possibly be present but rather why so few are commonly found of the many that are there. A partial answer can be found in the record of peptide identification in Table S1, which suggests that the numbers found might be easily increased.

### Conclusions

To gain access to the large amount of structural, biophysical, and functional information that is in principle available from HX-MS experimentation, and ultimately to extend experimental HX-MS resolution to the amino acid level, it is most important to maximize the number of peptide fragments studied. Analysis shows that acid proteolysis produces many more peptides than have usually been identified (see inventory above). This paper describes methods that access and validate a number of peptide fragments that approximate the number of amino acids in each of four different proteins ranging in size from 140 to 908 residues. At the level of initial MS/MS identification, important determinants are the use of multiple independent proteases, multiple non-overlapping MS/MS runs, the reliability of the physical set up, and the resolution and sensitivity of the mass spectrometer. At the level of MS analysis, this work also helped to develop the ExMS computer analysis program and used it to efficiently and definitively find, validate, and characterize these many peptides under conditions pertinent for HX-MS experiments. These same approaches should be widely

applicable in other laboratories. The ExMS program is described in a following paper [46].

### Acknowledgments

The authors thank John Engen, Yoshitomu Hamuro, and Virgil Woods Jr. for exceptionally helpful comments on the work and on the manuscript. They gratefully acknowledge support by research grants from the NIH (GM031847), NSF (MCB1020649), and the Mathers Foundation.

### References

- Bai, Y., Milne, J.S., Mayne, L., Englander, S.W.: Primary Structure Effects on Peptide Group Hydrogen Exchange. *Proteins* **17**, 75–86 (1993)
- Connelly, G.P., Bai, Y., Jeng, M.F., Englander, S.W.: Isotope Effects in Peptide Group Hydrogen Exchange. *Proteins* **17**, 87–92 (1993)
- Molday, R.S., Englander, S.W., Kallen, R.G.: Primary Structure Effects on Peptide Group Hydrogen Exchange. *Biochemistry* **11**, 150–158 (1972)
- Hvidt, A., Nielsen, S.O.: Hydrogen Exchange in Proteins. *Adv. Protein Chem.* **21**, 287–386 (1966)
- Linderström-Lang, K. Deuterium Exchange and Protein structure. *Symposium on Protein Structure*; Methuen: London, **1958**.
- Englander, S.W., Kallenbach, N.R.: Hydrogen Exchange and Structural Dynamics of Proteins and Nucleic-acids. *Q. Rev. Biophys.* **16**, 521–655 (1983)
- Kang, S., Prevelige, P.: Hydrogen/deuterium exchange studies in viruses. *Mass Spectrometry Analysis for Protein-Protein Interactions and Dynamics*, pp. 105–122. John Wiley and Sons, Hoboken, NJ (2008)
- Gross, M.L., Zhou, M., Hambly, D.: Two Approaches to Mass Spectrometric Protein Footprinting: PLIMSTEX and FPOP. *Mass Spectrometry analysis for Protein-Protein Interactions and Dynamics*, pp. 243–270. John Wiley and Sons, Hoboken, NJ (2008)
- Hamuro, Y., Coales, S., Hamuro, L., Woods Jr., V.L.: Use of Enhanced Peptide Amide Hydrogen/Deuterium Exchange-Mass Spectrometry (DXMS) in the Examination of Protein-Protein Interactions. *Mass Spectrometry Analysis for Protein-Protein Interactions and Dynamics*, pp. 123–156. John Wiley and Sons, Hoboken, NJ (2008)
- Komives, E.A.: Deuterium Exchange Approaches for Examining Protein Interactions: Case Studies of Complex Formation. *Mass Spectrometry Analysis for Protein-Protein Interactions and Dynamics*, pp. 91–104. John Wiley and Sons, Hoboken, NJ (2008)
- Tang, L., Hopper, E., Tong, Y., Sadowsky, J., Peterson, K., Gellman, S., Fitzgerald, M.: H/D Exchange- and Mass Spectrometry-Based Strategy for the Thermodynamic Analysis of Protein-Ligand Interactions. *Anal. Chem.* **79**, 5869–5877 (2007)
- Dai, S., Fitzgerald, M.: A mass spectrometry-based probe of equilibrium intermediates in protein folding reactions. *Biochemistry* **45**, 12890–12897 (2006)
- Engen, J.R.: Analysis of Protein Conformation and Dynamics by Hydrogen/Deuterium Exchange MS. *Anal. Chem.* **81**, 7870–7875 (2009)
- Kaveti, S., Engen, J.R.: Protein Interactions Probed with Mass Spectrometry. *Methods Mol. Biol.* **316**, 179–197 (2006)
- Engen, J.R.: Analysis of Protein Complexes with Hydrogen Exchange and Mass Spectrometry. *Analyst* **128**, 623–628 (2003)
- Garcia, R.A., Pantazatos, D., Villarreal, F.J.: Hydrogen/Deuterium Exchange Mass Spectrometry for Investigating Protein-Ligand Interactions. *Assay Drug Devel. Tech.* **2**, 81–91 (2004)
- Tsutsui, Y., Wintrod, P.L.: Hydrogen/deuterium, exchange-mass spectrometry: A powerful tool for probing protein structure, dynamics and interactions. *Current Med. Chem.* **14**, 2344–2358 (2007)
- Wales, T.E., Engen, J.R.: Hydrogen Exchange Mass Spectrometry for the Analysis of Protein Dynamics. *Mass Spectrom. Rev.* **25**, 158–170 (2006)
- Kaltashov, I.A., Eyles, S.J.: Studies of Biomolecular Conformations and Conformational Dynamics by Mass Spectrometry. *Mass Spectrom. Rev.* **21**, 37–71 (2002)

20. Linderstrom-Lang, K.U.: Deuterium Exchange Between Peptides and Water. *Chem. Soc. Spec. Publ.* **2**, 1–20 (1955)
21. Hvidt, A., Kagi, J.H.R., Ottesen, M.: Effect of Oxidized Nicotinamide-Adenine Dinucleotide on Hydrogen-Deuterium Exchange of Yeast Alcohol Dehydrogenase as Measured by Infrared Spectrophotometry. *Biochim. Biophys. Acta* **75**, 290–296 (1963)
22. Englander, S.W.A.: Hydrogen Exchange Method Using Tritium and Sephadex. Application to Ribonuclease. *Biochemistry* **2**, 798–807 (1963)
23. Englander, S.W., Downer, N.W., Teitelbaum, H.: Hydrogen Exchange. *Ann. Rev. Biochem.* **41**, 903–924 (1972)
24. Rosa, J.J., Richards, F.M.: An Experimental Procedure for Increasing the Structural Resolution of Chemical Hydrogen Exchange Measurements on Proteins: Application to Ribonuclease S Peptide. *J. Mol. Biol.* **133**, 399–416 (1979)
25. Englander, S. W.; Calhoun, D. B.; Englander, J. J.; Kallenbach, N. R.; Liem, R. K. H.; Malin, E. L.; Mandal, C.; Rogero, J. R. Individual Breathing Reactions Measured in Hemoglobin by Hydrogen Exchange Methods. *Biophys. J.* **32**, 577–590 (1980)
26. Englander, J.J., Rogero, J.R., Englander, S.W.: Protein Hydrogen Exchange Studied by the Fragment Separation Method. *Anal. Biochem.* **147**, 234–244 (1985)
27. Nagayama, K., Kumar, A., Wüthrich, K., Ernst, R.R.: Experimental Techniques of Two-Dimensional Correlated Spectroscopy. *J. Magn. Res.* **40**, 321–334 (1980)
28. Sorensen, O. W.; Eich, G. W.; Levitt, M. H.; Bodenhausen, G.; Ernst, R. R. Product Operator Formalism for the Description of NMR Pulse Experiments. *Prog. NMR Spectrom.* **16**, 163–192 (1984)
29. Bax, A.; Lerner, L. Two-Dimensional NMR Spectroscopy. *Science* **232**, 960–967 (1986)
30. Wagner, G., Wüthrich, K.: Amide Proton Exchange and Surface Conformation of BPTI in Solution: Studies with 2D NMR. *J. Mol. Biol.* **160**, 343–361 (1982)
31. Wand, A. J.; Englander, S. W. Two-Dimensional Proton NMR Studies of Cytochrome c. *Biochemistry* **24**, 5290–5294 (1985)
32. Zhang, Z., Smith, D.L.: Determination of Amide Hydrogen Exchange by Mass Spectrometry: A New Tool for Protein Structure Elucidation. *Protein Sci.* **2**, 522–531 (1993)
33. Hoofnagle, A.N., Resing, K.A., Ahn, N.G.: Protein Analysis by Hydrogen Exchange Mass Spectrometry. *Ann. Rev. Biophys. Biomol. Struct.* **18**, 18 (2003)
34. Kaltashov, I.A., Eyles, S.J.: *Mass Spectrometry in Biophysics: Conformation and Dynamics of Biomolecules*, p. 458. John Wiley, Hoboken, NJ (2005)
35. Kaltashov, I.A., Engen, J.R., Gross, M.L.: Hydrogen Exchange and Covalent Modification: Focus on Biomolecular Structure, Dynamics, and Function. *J. Am. Soc. Mass Spectrom.* **17**, I–II (2006)
36. Konermann, L., Tong, X., Pan, Y.: Protein Structure and Dynamics Studied by Mass Spectrometry: H/D Exchange, Hydroxyl Radical Labeling, and Related Approaches. *J. Mass Spectrom.* **43**, 1021–1036 (2008)
37. Yan, X., Maier, C.S.: Hydrogen/Deuterium Exchange Mass Spectrometry. *Methods Mol. Biol.* **492**, 255–271 (2009)
38. Marcisisin, S.R., Engen, J.R.: Hydrogen Exchange Mass Spectrometry: What Is It and What Can It Tell Us? *Analyt. Bioanalyt. Chem.* **397**, 967–972 (2010)
39. Del Mar, C., Greenbaum, E.A., Mayne, L., Englander, S.W., Woods, V. L.: Structure and Properties of  $\alpha$ -Synuclein and Other Amyloids Determined at the Amino Acid Level. *Proc. Natl. Acad. Sci. USA* **102**, 15477–15482 (2005)
40. Coales, S.J., Sook Yen, E., Lee, J.E., Ma, A., Morrow, J.A., Hamuro, Y.: Expansion of Time Window for Mass Spectrometric Measurement of Amide Hydrogen/Deuterium Exchange Reactions. *Rapid Commun. Mass Spectrom.* **24**, 3585–3592 (2010)
41. Pan, J.X., Han, J., Borchers, C.H., Konermann, L.: Hydrogen/Deuterium Exchange Mass Spectrometry with Top-Down Electron Capture Dissociation for characterizing Structural Transitions of a 17 kDa Protein. *J. Am. Chem. Soc.* **131**, 12801–12808 (2009)
42. Abzalimov, R.R., Kaplan, D.A., Easterling, M.L., Kaltashov, I.A.: Protein Conformations Can be Probed in Top-Down HDX MS Experiments Utilizing Electron Transfer Dissociation of Protein Ions without Hydrogen Scrambling. *J. Am. Soc. Mass Spectrom.* **20**, 1514–1517 (2009)
43. Rand, K.D., Zehl, M., Jensen, O.N., Jorgensen, T.J.D.: Protein Hydrogen Exchange Measured at Single-Residue Resolution by Electron Transfer Dissociation Mass Spectrometry. *Anal. Chem.* **81**, 5577–5584 (2009)
44. Zehl, M., Rand, K.D., Jensen, O.N., Jorgensen, T.J.D.: Electron Transfer Dissociation Facilitates the Measurement of Deuterium Incorporation into Selectively Labeled Peptide with Single Residue Resolution. *J. Am. Chem. Soc.* **130**, 17453–17459 (2008)
45. Althaus, E., Canzar, S., Ehrler, C., Emmett, M., Karrenbauer, A., Marshall, A., Meyer-Base, A., Tipton, J., Zhang, H.: Computing H-D Exchange Rates of Single Residues from Data of Proteolytic Fragments. *BMC Bioinformatics* **11**, 424 (2010)
46. Kan, Z.Y., Mayne, L., Chetty, P.S., Englander, S.W.: ExMS: Data Analysis for HX-MS Experiments. *J. Am. Soc. Mass Spectrom.* (2011). doi:10.1007/s13361-011-0236-3
47. Chun, S., Strobel, S., Bassford, P.J., Randall, L.: Folding of Maltose Binding Protein. Evidence for the Identity of the Rate-Determining Step In Vivo and In Vitro. *J. Biol. Chem.* **268**, 20855–20862 (1993)
48. Weisweiler, P.: Isolation and Quantitation of Apolipoproteins A-I and A-II from Human High-Density Lipoproteins by Fast-Protein Liquid Chromatography. *Clin. Chim. Acta* **169**, 249–254 (1987)
49. Giasson, B.I., Murray, I.V., Trojanowski, J.Q., Lee, V.M.: A Hydrophobic Stretch of 12 Amino Acid Residues in the Middle of  $\alpha$ -Synuclein is Essential for Filament Assembly. *J. Biol. Chem.* **276**, 2380–2386 (2001)
50. Shorter, J., Lindquist, S.: Hsp104 Catalyzes Formation and Elimination of Self-Relicating Sup35 Prion Conformers. *Science* **304**, 1793–1797 (2004)
51. Fang, J., Rand, K.D., Beuning, P.J., Engen, J.R.: False EX1 signatures caused by sample carryover during HXMS analyses. *Int. J. Mass Spectrom.* **302**, 19–25 (2011)
52. Wu, Y., Kaveti, S., Engen, J.R.: Extensive Deuterium Back-Exchange in Certain Immobilized Pepsin Columns Used for H/D Exchange Mass Spectrometry. *Anal. Chem.* **78**, 1719–1723 (2006)
53. Wang, L.T., Pan, H., Smith, D.L.: Hydrogen Exchange-Mass Spectrometry—Optimization of Digestion Conditions. *Mol. Cell. Proteom.* **1**, 132–138 (2002)
54. Hamuro, Y., Burns, L., Canaves, J., Hoffman, R., Taylor, S., Woods, V.: Domain Organization of d-akap2 Revealed by Enhanced Deuterium Exchange-Mass Spectrometry (DXMS). *J. Mol. Biol.* **321**, 703–714 (2002)
55. Hather, G., Higdon, R., Bauman, A., von Haller, P.D., Kolker, E.: Estimating False Discovery Rates for Peptide and Protein Identification Using Randomized Databases. *Proteomics* **10**, 2369–2376 (2010)
56. Hamuro, Y., Coales, S.J., Molnar, K.S., Tuske, S.J., Morrow, J.A.: Specificity of Immobilized Porcine Pepsin in H/D Exchange Compatible Conditions. *Rapid Commun. Mass Spectrom.* **22**, 1041–1046 (2008)