



APPLICATION NOTE

HDXFinder: Automated Analysis and Data Reporting of Deuterium/Hydrogen Exchange Mass Spectrometry

Danny E. Miller, Charulata B. Prasannan, Maria T. Villar, Aron W. Fenton, Antonio Artigues

Biochemistry and Molecular Biology, The University of Kansas Medical Center, MS 3030, 3901 Rainbow Blvd., Kansas City, KS 66160, USA

Abstract

Hydrogen/deuterium exchange in combination with mass spectrometry (H/D MS) is a sensitive technique for detection of changes in protein conformation and dynamics. However, wide application of H/D MS has been hindered, in part, by the lack of computational tools necessary for efficient analysis of the large data sets associated with this technique. We report a novel web-based application for automatic analysis of H/D MS experimental data. This application relies on the high resolution of mass spectrometers to extract all isotopic envelopes before correlating these envelopes with individual peptides. Although a fully automatic analysis is possible, a variety of graphical tools are included to aid in the verification of correlations and rankings of the isotopic peptide envelopes. As a demonstration, the rate constants for H/D exchange of peptides from rabbit muscle pyruvate kinase are mapped onto the structure of this protein.

Key words: Hydrogen/deuterium exchange, Mass spectrometry, Protein dynamics, Fourier transform ion cyclotron mass spectrometry

Introduction

There is a growing appreciation that hydrogen/deuterium exchange mass spectrometry (H/D MS) can provide insights into different aspects of protein conformation and dynamics. Thus, there is a current need for computational tools that can aid in the analysis of the very large data sets produced by these techniques. Efforts in program

development to-date focus on isolated aspects of data analysis; examples include centroid calculation [1, 2], calculation of deuterium content in proteins or peptic peptides [3], and determining envelope data [4]. These tools are useful for the statistical analysis, visualization, and plotting of deuterium content.

There are several problems associated with complex multi-dimensional H/D MS experiments that have not been fully addressed by the programs listed above. One is the detection of peptide ion isotopic envelopes, complicated by the shift in mass and complexity as a consequence of varying deuterium content. Another is the isolation of individual isotope envelopes from overlapping peptide ions. Programs such as Hydra [5], HeXicon [6], the Deuterator [7], and its successor, HD Desktop [8] attempt to address these problems using statistical methods for fitting experimental data to theoretical isotope envelopes. Unfortunately, these programs require,

Electronic supplementary material The online version of this article (doi:10.1007/s13361-011-0234-5) contains supplementary material, which is available to authorized users.

Correspondence to: Aron W. Fenton; e-mail: afenton@kumc.edu, Antonio Artigues; e-mail: aartigues@kumc.edu

Received: 24 March 2011
Revised: 3 August 2011
Accepted: 12 August 2011
Published Online: 15 November 2011

as initial input, the HPLC retention times of the peptide ions of interest. Thus, even with these aids, user input is needed at all phases of data interpretation, making this process time-consuming and prone to human error. There are also recent developments in the analysis of H/D MS data acquired on a 14.5 Tesla FT-ICR instrument [9], but these instruments are not commonly available.

Here we report HDXFinder, an application for peptide detection and for automatic peak assignments, which does not depend on HPLC retention times. Detection and isolation of isotope peptide envelopes, including extraction of overlapping peptide ions, are fully automatic functions performed by HDXFinder. The set of tools included in HDXFinder has been tested at a resolution power as low as 12,500 for an ion with $m/z=400$.

Program Description and Results

Programming Background

HDXFinder is a Linux-based web application accessible from any web browser. The program is a combination of Java, JavaScript, and Perl. The program utilizes a MySQL database for storage and retrieval of all data. Most software used in the development and operation of HDXFinder is freely available as open-source software. The only exception is the Fusion Charts v. 3.0 package used to develop most of the graphics for data analysis.

HDXFinder is freely accessible from the KU Medical Center Mass Spectrometry Core Laboratory (<https://HDXAnalyzer.kumc.edu/>) to all not-for-profit users. Due to the recent use of the name “HDX-Analyzer” by another group, we have modified the name of the program described here from “HDXAnalyzer” to “HDXFinder.” Therefore, the address of the program (<https://HDXAnalyzer.kumc.edu/>) does not match the program name (HDXFinder).

Data Input and Analysis

Three sets of data are required as input for the program: (1) a data file(s) containing the mass spectrometric information, (2) the sequence of the protein of interest, and (3) a list of peptides generated from the protein of interest. Additionally, a pdb datafile can be uploaded for representation of the output data on the three-dimensional structure of the protein. HDXFinder can be used for data analysis of files generated by a variety of instruments, independent of manufacturer. Thus, it accepts data in the mzXML [10] or mzML [11] format, both developed to facilitate exchange of mass spectrometric information across multiple platforms. There are freely available converters for the specific data files generated by most manufacturers of mass spectrometers (<http://tools.proteomecenter.org>) [12].

The list of peptides of interest can be generated from any of the search engines broadly used for protein identification (a topic beyond the scope of this article). For each peptide included in this list, the following information must be included in a comma-separated (csv) format: the peptide sequence, the peptide mass ($M + H^+$), and the peptide charge. For the experiments described here, the list of peptides of interest from the peptic digest of rabbit muscle pyruvate kinase was generated from Sequest (see [Supplementary information](#)).

The basic components and steps of the analysis of a full H/D MS experiment are outlined in Figure 1. The chart shows a schematic of the workflow of HDXFinder. Upon input of mzXML or mzML data, the program automatically begins a series of consecutive algorithms to detect and isolate isotopic envelopes, measure the charges of these envelopes, and calculate the average mass of the represented ions. All data are stored in a MySQL database and can be obtained in a csv format. HDXFinder will match (i.e., identify envelopes that could correspond to peptide ions with

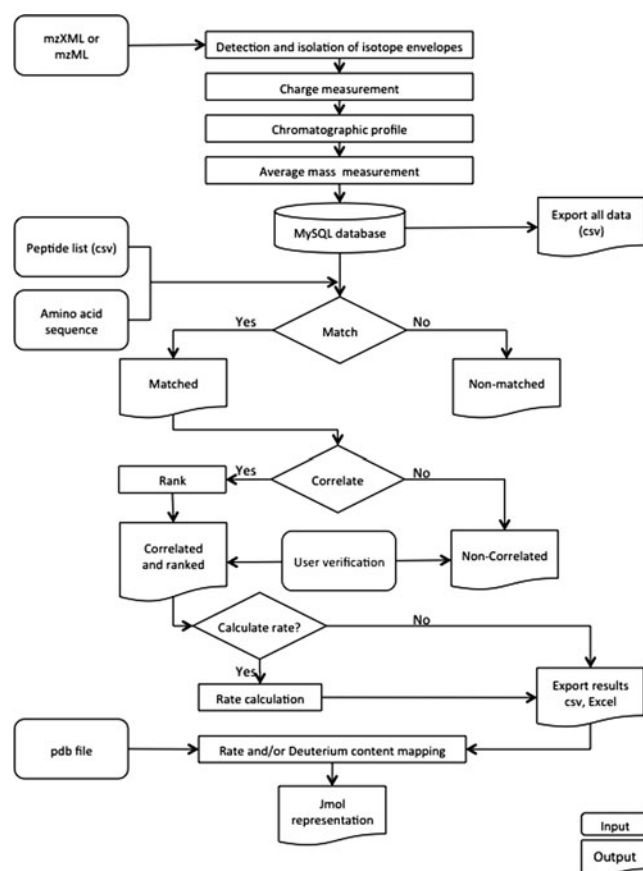


Figure 1. HDXFinder workflow. The final output is a table, which may be exported as a csv file containing, for each peptide and experimental time (or condition), the peptide’s HPLC retention time, average mass, the number of deuterium acquired and overall rate of exchange. Finally, if a pdb file is associated with the project, the deuterium content or rates can be displayed on a three-dimensional image of the protein

a range of deuteration content, depending on the number of exchangeable amide protons), correlate (i.e., assign isotopic envelopes to individual peptides), and rank this data (i.e., associates isotopic envelopes across different H/D reaction times) (see [Supplementary material](#) for details). The user has the option to validate these correlations and rankings and/or change the ranking. If desired, rate constants for H/D exchange can be calculated and mapped onto the user provided pdb structure.

Determining the Rate Constants of Exchange

Deuterium content is measured as the difference of average mass between the deuterated and the theoretical mass of the undeuterated peptide. Once data have been correlated and ranked, rate constants of exchange can be calculated for each peptide. Rate constants of exchange are determined by following the mass of the peptide as a function of time of the exchange reaction, using a single exponential fit. Users may also choose to export data to enable determination of the rate constants of exchange with other fit models.

Results and Output

The basic output of HDXFinder is an html table that can be downloaded as a csv file. This table lists all the peptides of interest, ordered according to their relative position on the protein amino acid sequence, and includes amino acid sequence, retention time, average mass, delta mass for each experimental time/condition, and rate constant of exchange (if that analysis was requested). Additionally, if a pdb file was provided, HDXFinder can produce a view with the rate constants of exchange mapped onto the protein structure. This structure with peptide mass information can be visualized with Jmol in the web browser. See Figure 2 using pyruvate kinase as an example of this output.

Discussion

HDXFinder works by first extracting all isotopic envelopes from an MS experiment. This process is automatic and is independent of any consideration for the nature of isotopic envelopes (i.e., chemical composition) or for their deuterium content. The only limiting variables considered are that the isotopic peptide envelopes should have a measurable charge ($z=1$ to 6) and a chromatographic elution profile consistent with an eluting ion (see [Supplementary material](#)). There are several advantages to extracting all isotopic envelopes before correlating them to individual peptides: (1) the extraction of isotopic envelopes is performed only once, resulting in a significant increase in the overall speed of the analytical process. (2) The extraction of all isotopic envelopes present in a given experiment, independent of the identification of the corresponding amino acid sequence, allows for the reporting of the isotopic envelopes of

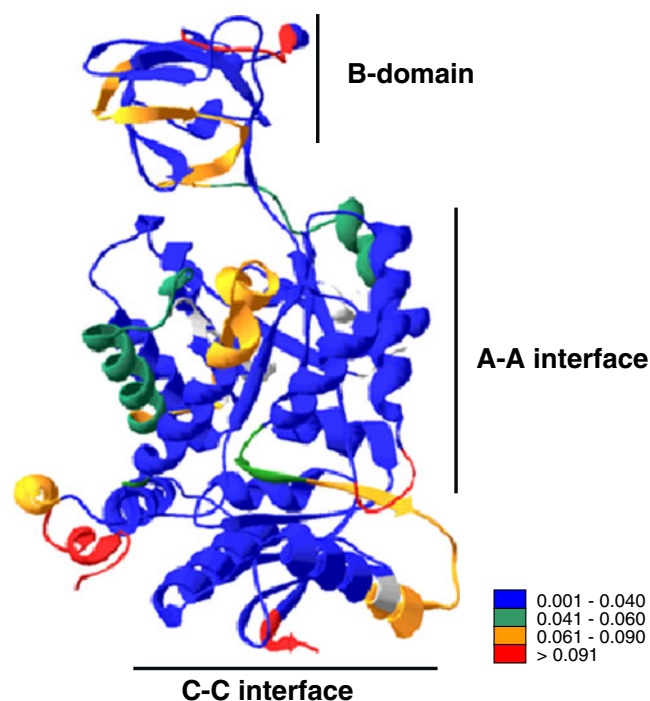


Figure 2. Rate constants of H/D exchange mapped to the monomer of pyruvate kinase from mammal brain/muscle (M₁-PYK) (PDB code 2G50). The two types of subunit interfaces within the homotetramer are labeled A-A and C-C. The rate constants of H/D exchange have been grouped in four categories: Blue (0.001–0.040), green (0.041–0.060), orange (0.061–0.090), red (>0.091). All rate constants of exchange are from single exponential fits—independently of the goodness of the fit—and are expressed in min⁻¹.

unknown primary sequence. This may allow the user to increase protein coverage by re-attempting peptide identification using an inclusion list for these particular ions. (3) The application of the algorithms used for this process results in the extraction of overlapping isotopic envelopes and the removal of spurious noise.

To help prevent back exchange and maintain deuterium label, peptides are eluted quickly (i.e., steep gradients and high flow rates) from HPLC reversed-phase chromatographic columns. However, despite considerable effort in instrument and method development [13–16], steep gradients often result in the co-elution of peptide ions with overlapping isotopic envelopes [9]. This makes it difficult to determine accurately the experimental average mass of the peptides. As mentioned above, programs such HD Desktop [8] and HeXicon [6] determine average masses by performing curve-fitting of the isotopic envelopes to a theoretical envelope.

In contrast to other programs that incorporate an evaluation of HPLC retention times as a step for extracting average masses of peptides, HDXFinder does not require, as an input, the retention times for the detection of the eluting peptides. Instead, all isotopic peptide envelopes are detected automatically and sub-

sequently matched and correlated to the peptides included in the list. The retention time along with the intensity of the ion, as measured by HDXFinder, are utilized for the ranking of ions (see [Supplementary material](#) for details).

The main advantage of the set of web-based tools included in HDXFinder is the ability to extract, match,

correlate and rank isotopic envelopes with the peptides provided, at all H/D reaction times in an unattended, automatic mode. At the same time, HDXFinder gives the user an option to re-correlate and re-rank all matched peptide data as needed. To aid manual steps, a variety of graphical tools are included, which will allow the user to (Figure 3 a–d): (1) examine individual averaged isotopic envelopes in

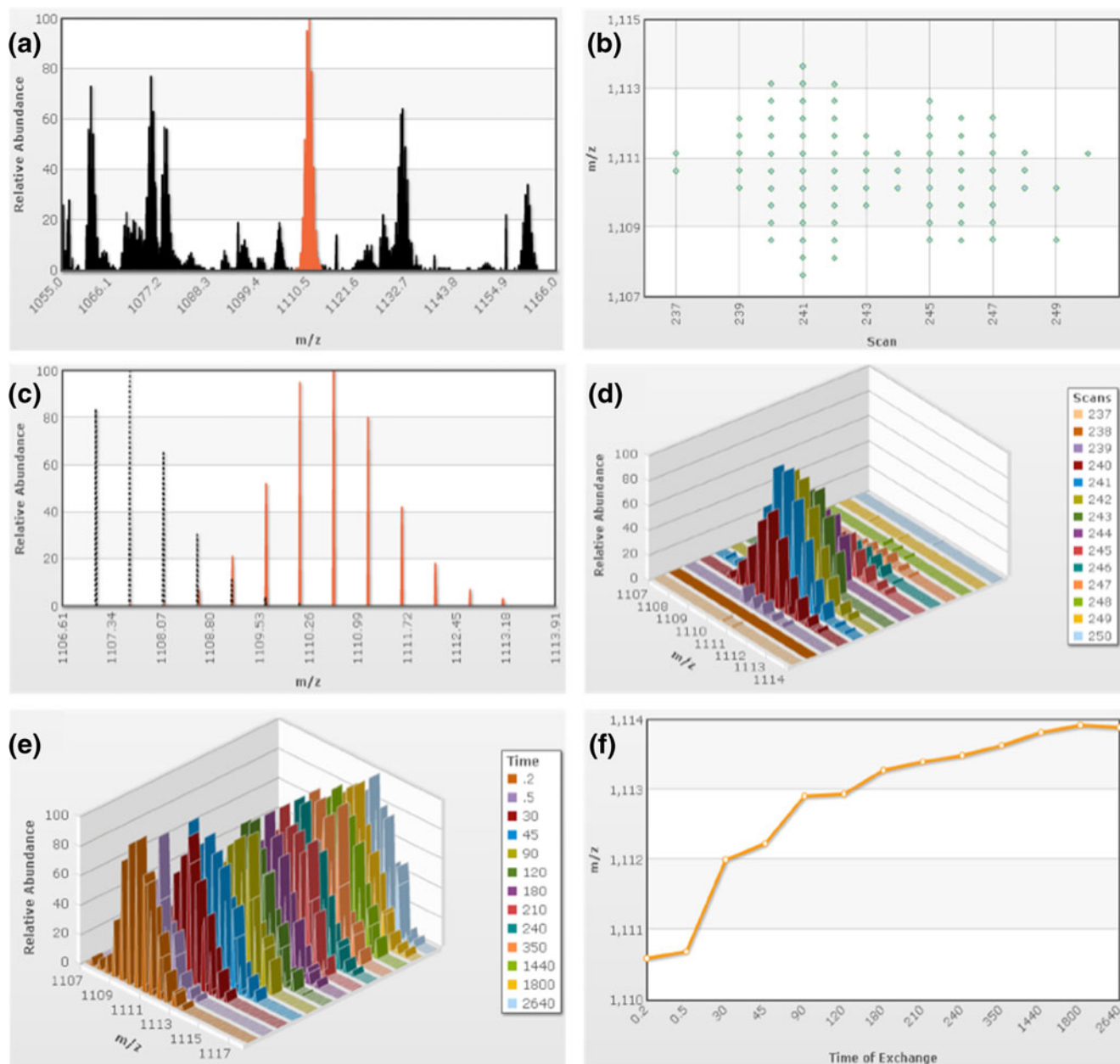


Figure 3. Verification tools. To aid in the process of verification several graphic tools have been built into HDXFinder. At the individual reaction point, the user can view: **(a)** a plot of the eluting peptide compared to other peptides. In this example the view is zoomed out 10% with the peptide of interest outlined in orange. **(b)** A scatter plot of all individual scans, indicating the position of isotopic peaks in the envelope under consideration. **(c)** The averaged isotopic envelope (orange) in comparison with the theoretical isotopic envelope for the peptide under consideration (black). **(d)** A three-dimensional plot of the scans showing the peptide eluting over time. At the overall range of the exchange reaction [panels **(e)** and **(f)**], the user can plot the change in m/z as a function of time either **(e)** as a three-dimensional plot or **(f)** average mass change for a peptide ion over time

comparison with the theoretical envelope of the undeuterated peptide; (2) display the elution profile of an isotopic envelope on a three-dimensional graph (i.e., m/z versus retention time versus intensity); and (3) make use of scatter plots of m/z versus retention of isotopic envelopes to establish chromatographic elution profiles. For analysis across the range of the exchange reaction (Figure 3e, f), the user can: (1) display a two-dimensional graph of the average mass for a peptide ion over the time of the exchange reaction; (2) display the deuterium exchange profiles for a peptide ion over the time of the exchange reaction as a three-dimensional profile of averaged isotopic envelopes (m/z versus abundance versus time of exchange).

In summary, HDXFinder provides a simple, fast, and automated routine for the identification of peptides in H/D MS. With these advantages, this program should find broad application in the mass spectrometry community.

Acknowledgments

The authors acknowledge support in part for this work by NIH grant DK78076 awarded to A.W.F. and by ITS grant from the University of Kansas to A.A.S. The authors thank Dr. David Nualart for valuable discussions.

References

- Hotchko, M., Anand, G.S., Komives, E.A., Ten Eyck, L.F.: Automated extraction of backbone deuteration levels from amide H/2H mass spectrometry experiments. *Protein Sci* **15**, 583–601 (2006)
- Nikamanon, P., Pun, E., Chou, W., Koter, M.D., Gershon, P.D.: “TOF2H”: a precision toolbox for rapid, high density/high coverage hydrogen-deuterium exchange mass spectrometry via an LC-MALDI approach, covering the data pipeline from spectral acquisition to HDX rate analysis. *BMC Bioinformatics* **9**, 387 (2008)
- Weis, D.D., Engen, J.R., Kass, I.J.: Semi-automated data processing of hydrogen exchange mass spectra using HX-express. *J Am Soc Mass Spectrom* **17**, 1700–1703 (2006)
- Althaus, E., Canzar, S., Ehrler, C., Emmett, M.R., Karrenbauer, A., Marshall, A.G., Meyer-Bäse, A., Tipton, J.D., Zhang, H.M., Eggertson, M., Stapels, M., Ahn, J., Fadgen, K.: Computing H/D-exchange rates of single residues from data of proteolytic fragments. *BMC Bioinformatics* **11**, 424 (2010)
- Slysz, G.W., Baker, C.A., Bozsa, B.M., Dang, A., Percy, A.J., Bennett, M., Schriemer, D.C.: Hydra: software for tailored processing of H/D exchange data from MS or tandem MS analyses. *BMC Bioinformatics* **10**, 162 (2009)
- Lou, X., Kirchner, M., Renard, B.Y., Kothe, U., Boppel, S., Graf, C., Lee, C.T., Steen, J.A., Steen, H., Mayer, M.P., Hamprecht, F.A.: Deuteration distribution estimation with improved sequence coverage for HX/MS experiments. *Bioinformatics* **26**, 1535–1541 (2010)
- Pascal, B.D., Chalmers, M.J., Busby, S.A., Mader, C.C., Southern, M. R., Tsinoremas, N.F., Griffin, P.R.: The Deuterator: software for the determination of backbone amide deuterium levels from H/D exchange MS data. *BMC Bioinformatics* **8**, 156 (2007)
- Pascal, B.D., Chalmers, M.J., Busby, S.A., Griffin, P.R.: HD desktop: an integrated platform for the analysis and visualization of H/D exchange data. *J Am Soc Mass Spectrom* **20**, 601–610 (2009)
- Kazazic, S., Zhang, H.M., Schaub, T.M., Emmett, M.R., Hendrickson, C.L., Blakney, G.T., Marshall, A.G.: Automated data reduction for hydrogen/deuterium exchange experiments, enabled by high-resolution fourier transform ion cyclotron resonance mass spectrometry. *J Am Soc Mass Spectrom* **21**, 550–558 (2010)
- Pedrioli, P.G., Eng, J.K., Hubley, R., Vogelzang, M., Deutsch, E.W., Raught, B., Pratt, B., Nilsson, E., Angeletti, R.H., Apweiler, R., Cheung, K., Costello, C.E., Hermjakob, H., Huang, S., Julian, R.K., Kapp, E., McComb, M.E., Oliver, S.G., Omenn, G., Paton, N.W., Simpson, R., Smith, R., Taylor, C.F., Zhu, W., Aebersold, R.: A common open representation of mass spectrometry data and its application to proteomics research. *Nat Biotechnol* **22**, 1459–1466 (2004)
- Deutsch, E.W.: Mass spectrometer output file format mzML. *Methods Mol Biol* **604**, 319–331 (2010)
- Deutsch, E.W., Mendoza, L., Shteynberg, D., Farrah, T., Lam, H., Tasman, N., Sun, Z., Nilsson, E., Pratt, B., Prazhen, B., Eng, J.K., Martin, D.B., Nesvizhskii, A.I., Aebersold, R.: A guided tour of the trans-proteomic pipeline. *Proteomics* **10**, 1150–1159 (2010)
- Wang, L., Smith, D.L.: Downsizing improves sensitivity 100-fold for hydrogen exchange-mass spectrometry. *Anal Biochem* **314**, 46–53 (2003)
- Chalmers, M.J., Busby, S.A., Pascal, B.D., He, Y., Hendrickson, C.L., Marshall, A.G., Griffin, P.R.: Probing Protein–ligand interactions by automated hydrogen/deuterium exchange mass spectrometry. *Anal Chem* **78**, 1005–1014 (2006)
- Chalmers, M.J., Busby, S.A., Pascal, B.D., Southern, M.R., Griffin, P. R.: A two-stage differential hydrogen/deuterium exchange method for the rapid characterization of protein/Ligand interactions. *J Biomol Tech* **18**, 194–204 (2007)
- Zhang, H.M., Bou-Assaf, G.M., Emmett, M.R., Marshall, A.G.: Fast Reversed-Phase Liquid Chromatography to Reduce Back Exchange and Increase Throughput in H/D Exchange Monitored by FT-ICR Mass Spectrometry. *J Am Soc Mass Spectrom* **20**, 520–524 (2008)