COMPUTATIONAL METHODS AND CHALLENGES IN HYDROGEN/DEUTERIUM EXCHANGE MASS SPECTROMETRY

Jürgen Claesen[1]* and Tomasz Burzykowski[1,2]
[1]*I-BioStat, Hasselt University, Campus Diepenbeek, Agoralaan Gebouw D, Diepenbeek 3590, Belgium*
[2]*Statistics and Medical informatics Unit, Medical University of Bialystok, Białystok, Poland*

Received 26 May 2015; revised 8 May 2016; accepted 18 August 2016

Published online 7 September 2016 in Wiley Online Library (wileyonlinelibrary.com). DOI 10.1002/mas.21519

Hydrogen/Deuterium exchange (HDX) has been applied, since the 1930s, as an analytical tool to study the structure and dynamics of (small) biomolecules. The popularity of using HDX to study proteins increased drastically in the last two decades due to the successful combination with mass spectrometry (MS). Together with this growth in popularity, several technological advances have been made, such as improved quenching and fragmentation. As a consequence of these experimental improvements and the increased use of protein-HDXMS, large amounts of complex data are generated, which require appropriate analysis. Computational analysis of HDXMS requires several steps. A typical workflow for proteins consists of identification of (non-)deuterated peptides or fragments of the protein under study (local analysis), or identification of the deuterated protein as a whole (global analysis); determination of the deuteration level; estimation of the protection extent or exchange rates of the labile backbone amide hydrogen atoms; and a statistically sound interpretation of the estimated protection extent or exchange rates. Several algorithms, specifically designed for HDX analysis, have been proposed. They range from procedures that focus on one specific step in the analysis of HDX data to complete HDX workflow analysis tools. In this review, we provide an overview of the computational methods and discuss outstanding challenges. © 2016 Wiley Periodicals, Inc. Mass Spec Rev 36:649–667, 2017

Keywords: *hydrogen/deuterium exchange; mass spectrometry; analysis*

I. INTRODUCTION

Proteins are a diverse and abundant class of biomolecules. They display a huge variety of processes involved in virtually all aspects of cell structure and function such as, for example, inhibition/catalysis of reactions, structural support, cellular signaling, immunologic responses, and many more. The structure of a protein is the main determinant of its biological function. However, the protein structure is not rigid. It undergoes changes before and/or after binding, affecting only a few atoms, the secondary structure elements (e.g., partial rearrangement of the protein to allow interaction between the binding site and the ligand), protein domains (e.g., hemoglobin after binding with oxygen), or even the complete protein (e.g., intrinsically disordered proteins). Therefore, when the protein function is of interest, one should not solely focus on the relation between protein structure and protein function, but also on the link with protein dynamics (Kaltashov & Eyles, 2005; Wales & Engen, 2006).

Several techniques exist to study the structure and/or dynamics of proteins. One of these methods is hydrogen/deuterium exchange (HDX) (Linderstrøm-Lang, 1958; Englander et al., 1997; Englander, 2006). Hydrogen atoms from OH-, NH-, and SH-groups of polar amino-acid-side-chains and from the NH-group of peptide bonds can be replaced by deuterium or tritium. Monitoring the exchange behavior of these labile hydrogen atoms should, in principle, provide detailed information about protein structure and protein dynamics. However, obtaining highly resolved structural information is only possible with sensitive recorders such as neutron crystallography (Kossiakoff, 1982; Mason, Bentley & McInTyre, 1983) and multi-dimensional nuclear magnetic resonance (NMR) (Wagner & Wuthrich, 1982; Ernst, Bodenhausen & Wokaun, 1988; Bax, 1994). Katta & Chait (1991) successfully combined HDX in solution with electrospray ionization (ESI) and mass spectrometry (MS). Winger et al. (1992) were the first to study the structure of proteins with gas-phase HDXMS. Although neutron crystallography and multi-dimensional NMR have single residue resolution, that is, atom-specific information, the number of applications of medium-resolution HDXMS (Zhang & Smith, 1993) increased dramatically in the last two decades. In comparison to NMR and crystallography, the use of mass spectrometers as *recorder* has several advantages. For instance, it needs a very low protein concentration. Additionally, in theory, there is no limit in protein mass, which allows studying large proteins and protein complexes.

The exchange of one lighter hydrogen isotope, ^{1}H, for a heavier deuterium isotope, ^{2}H, or *vice versa*, can be easily monitored with mass spectrometry as the mass of the protein under study changes with approximately one dalton. Kinetic information and the location of exchanged hydrogen atoms cannot be inferred directly from recording the mass shift. In order to draw conclusions regarding the protein conformation and dynamics, accurate data processing, analysis, and interpretation is needed. A typical computational workflow for digested and/or gas-phase fragmented proteins consists of four steps (Wales, Eggertson & Engen, 2013) (see Fig. 1):

- Step 1: Identification of peptides and/or fragmented ions. For intact proteins this step is not required.

Correspondence to: Jürgen Claesen, I-BioStat, Universiteit Hasselt, Campus Diepenbeek, Agoralaan Gebouw D, Diepenbeek 3590, Belgium. E-mail: jurgen.claesen@uhasselt.be

© 2016 by Wiley Periodicals, Inc.

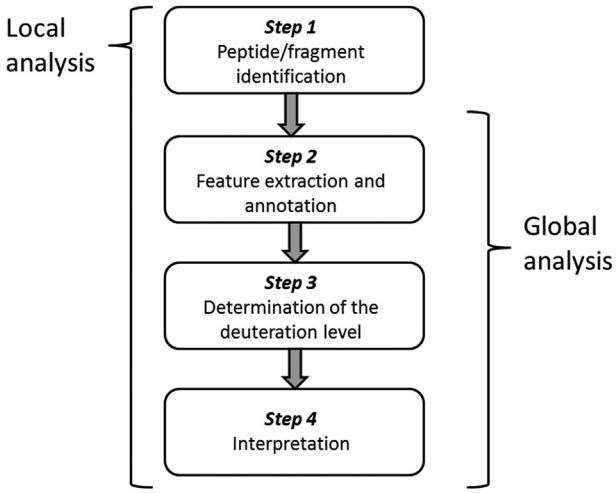

FIGURE 1. Computational workflow for HDXMS. In case of a local analysis, steps 2–4 have to be repeated for each identified peptide or fragment.

- Step 2: Feature extraction and annotation of the (partially) deuterated ions.
- Step 3: Determination of the deuteration level, that is, the number of incorporated deuterium atoms, of the annotated features.
- Step 4: Interpretation.

Technological advancements, such as automation of the experimental workflow, made it relatively easy to study protein structure and dynamics with MS. As a result of these advancements, large amounts of data are being generated, which makes it no longer possible to process and analyze HDXMS data by hand. Several (semi-)automated tools for the computational workflow have been proposed in the last 15 years (Table 1). To our knowledge, no comprehensive overview of these tools and of methods that focus on a specific computational step is available. In Sections III–VI, we provide a detailed overview of the available, non-commercial methods and tools for HDXMS with ESI and discuss remaining difficulties and challenges. Commercial tools, such as DynamX (Wei et al., 2012), or HDExaminer (Hamuro et al., 2003, 2006; Horn et al., 2006), are not included in this review.

II. HYDROGEN/DEUTERIUM EXCHANGE MASS SPECTROMETRY

Before discussing the computational workflow in detail, we provide a short overview of HDX kinetics and the experimental workflow used to obtain HDXMS data.

A. HDX Kinetics

The structure of proteins in solution changes not only during interaction with ligands or proteins. It also undergoes alterations ranging from small, local fluctuations, to global unfolding events. However, the latter are very rare under native and non-native conditions. Through these reversible conformational fluctuations, labile hydrogen atoms involved in hydrogen bonding and/or hydrogens buried in the protein core can become readily available for exchange with deuterium. As a consequence, for continuous labeling experiments, HDX is generally described by a two-state kinetic model (Hvidt, 1964; Hvidt & Nielsen, 1966):

$$A-H_{\text{folded}} \underset{k_{\text{close}}}{\overset{k_{\text{open}}}{\rightleftharpoons}} A-H_{\text{unfolded}} \underset{D_2O}{\overset{k_{\text{int}}}{\rightarrow}} A-D_{\text{unfolded}} \underset{k_{\text{open}}}{\overset{k_{\text{close}}}{\rightleftharpoons}} A-D_{\text{folded}}$$

The overall exchange-rate, k_{ex}, depends on the opening and closing rate constants (k_{open} and k_{close}) and on the intrinsic exchange-rate, k_{int}, of the unstructured protein/peptide A:

$$k_{\text{ex}} = \frac{k_{\text{open}} \times k_{\text{int}}}{k_{\text{close}} + k_{\text{open}} + k_{\text{int}}} \quad (1)$$

For the majority of proteins and peptides under physiological conditions, the exchange reaction is much slower than the refolding (Englander et al., 1996). Hence, the probability of exchange in a single unfolding-refolding sequence is small. As a consequence, the overall exchange rate can be written as

$$k_{\text{ex}} = \frac{k_{\text{open}}}{k_{\text{close}}} \times k_{\text{int}} \quad (2)$$

In such cases, the biomolecule is said to follow an EX2 kinetic exchange regime. Monitoring proteins that exchange under EX2 conditions allows thermodynamical interpretation of the structural fluctuations (Hoofnagle, Resing, & Ahn, 2003; Konermann, Tong, & Pan, 2008; Konermann, Pan, & Liu, 2011; Jaswal, 2013).

In non-native, denaturing conditions and environments with high pH, proteins can exchange deuterium during a single unfolding/refolding event. A limited number of proteins follow, under native conditions, such an exchange regime, often referred to as EX1. The overall exchange rate for this regime is equal to k_{open}.

Some proteins undergo a mixed EX1–EX2 regime. For such proteins, a number of hydrogens exhibit EX1 exchange characteristics, while simultaneously other hydrogens exchange according to the EX2 regime.

Next to investigating the structure and dynamics of a protein with in-solution-HDX, it is also possible to assess these features with gas-phase-HDX. In contrast to the numerous applications of solution-phase-HDXMS, gas-phase-HDXMS has received little attention. Several exchange mechanisms have been proposed for different deuterium sources (Campbell et al., 1995; Wyttenbach & Bowers, 1999; Schaaff, Stephenson, & McLuckey, 2000). D_2O, the least reactive and most discriminant reagent, triggers HDX according to a relay mechanism. This involves the relocation of a proton from the N-terminus to a less basic site. N_3D induces exchange according to the omnium-ion mechanism: proton transfer and the solvation of the resulting ammonium ion occur simultaneously.

For more details on solution- and gas-phase hydrogen exchange mechanisms, we refer to numerous excellent reviews (Englander et al., 1996, 1997; Smith, Deng, & Zhang, 1997; Hoofnagle, Resing, & Ahn, 2003; Kaltashov & Eyles, 2005; Chalmers et al., 2006, 2011a,b; Wales & Engen, 2006; Konermann, Tong, & Pan, 2008; Engen, 2009; Morgan &

10982787, 2017, 5, Downloaded from https://analyticalsciencejournals.onlinelibrary.wiley.com/doi/10.1002/mas.21519 by Universiteit Hasselt, Wiley Online Library on [01/06/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 1. Overview of non-commercial HDX computational workflow tools.

Tool	Step 1 Identification	Step 2 Feature extraction and annotation	Step 3 Determining the deuteration level	Step 4 Interpretation
AUTOHD (Palmblad, Buijs, & Håkansson, 2001)	In-silico	Annotation based on master list	Deconvolution – theoretical fit	None
CalcDeut (Chik, Vande Graaf, & Schriemer, 2006)	N/A	N/A	Deconvolution – mass shifts	None
Deuterator (Pascal et al., 2007)	MS/MS	Extraction and annotation based on master list	Deconvolution – theoretical fit centroid	Visualization
HD Desktop (Pascal et al., 2009)	MS/MS	Extraction and annotation based on master list	Deconvolution – theoretical fit centroid	Visualization
HDX Workbench (Pascal et al., 2012)	Combination	Extraction and annotation based on master list	Deconvolution – theoretical fit centroid	Visualization statistical analysis
DEX (Hotchko et al., 2006)	Unknown	Extraction and annotation based on master list	Deconvolution – mass shifts	None
ExMS (Kan et al., 2011)	MS/MS	Extraction and annotation based on master list	Centroid	Visualization
HDsite(Kan et al., 2013)	MS/MS	Extraction and annotation based on master list	Deconvolution – theoretical fit	Exchange rates
HDX-Analyzer (Liu et al., 2012)	MS/MS	Unknown	Centroid	Visualization statistical analysis
HDXFinder (Miller et al., 2012)	MS/MS	Annotation based on master list	Centroid	Visualization exchange rates
HeXicon (Lou et al., 2010)	Combination	Extraction and annotation based on master list	Alternative	Exchange rates
HeXicon (*bimodal)* (Kreshuk et al., 2011)	Combination	Extraction and annotation based on master list	Alternative	Exchange rates
HeXicon2 (Lindner et al., 2014)	Combination	Annotation based on master list	Deconvolution – mass shifts	Visualization
HX-Express (Weis, Engen, & Kass, 2006)	Unknown	Extraction and annotation based on master list	Centroid	Visualization
HX-Express2 (Guttman et al., 2013)	Unknown	Extraction and annotation based on master list	Deconvolution – theoretical fit	Visualization
Hydra (Slysz et al., 2009)	MS/MS	Extraction and annotation based on master list	Deconvolution – mass shifts centroid	Visualization statistical analysis
Mass Spec Studio (Rey et al., 2014)	MS/MS	Unknown	Centroid	Visualization
MSTools (Kavan & Man, 2011)	N/A	N/A	N/A	Visualization
QUDeX-MS (Salisbury, Liu, & Agar, 2014)	Combination	Unknown	Alternative	None
Fajer, Bou-Assaf, & Marshall (2012)	N/A	N/A	N/A	Exchange rates
Kazazic et al. (2010)	In-silico	Extraction and annotation based on master list	Centroid	None
Zhang (2012a)	In-silico	Independent from master list	Centroid	Visualization protection extent

Engen, 2009; Marcsisin & Engen, 2010; Konermann, Pan, & Liu, 2011; Winkler, Dzyuba, & Schalley, 2011; Percy et al., 2012; Wales, Eggertson, & Engen, 2013).

B. Experimental Workflow of HDXMS

The experimental workflow of HDXMS consists of several steps, and is dependent of the exchange phase (solution or gas) and the aim of the study. Typically, the experimental workflow is repeated a number of times. These replicate experiments can be used, for instance, to assess the variability and reproducibility of the measurements (Burkitt & O'Connor, 2008; Houde, Berkowitz, & Engen, 2011; Iacob & Engen, 2012; Zhang, 2012a).

1. Solution-Phase Exchange

There are two types of solution-phase-HDX experiments: continuous labeling and pulse labeling (Deng, Pan, & Smith, 1999; Konermann & Simmons, 2003). In continuous labeling, a protein is incubated in a medium containing an excess of deuterium. The exchange reaction proceeds for a specified period in time, typically ranging from a few seconds to several days. Subsequently, the reaction is quenched by lowering the pH to 2.5 and the temperature to 0˚C. Quenching is needed to retain the deuteration level. Monitoring the deuterium content as a function of the incubation time provides information on protein dynamics under equilibrium.

With pulse labeling experiments, the short-lived folding intermediates are of interest (Konermann & Simmons, 2003). The protein under study is first denatured, followed by folding which is triggered by transferring the protein in a folding-inducing buffer. In contrast to continuous labeling, the protein is shortly exposed to D_2O (often for a few milliseconds) at a specified point in time. Next, the proteins are mixed with quenching buffer.

For both of these labeling approaches, the deuterated protein as a whole is ionized with a soft ionization method (MALDI and especially ESI are commonly used) and analyzed by a (high resolution) mass spectrometer. This procedure of monitoring a complete protein is known as *global exchange analysis* (Katta & Chait, 1991). In contrast to neutron crystallography and multidimensional NMR, information regarding which H-atoms exchanged is not readily available. Increasing the spatial resolution of HDXMS is obtained by proteolytic digestion after quenching and/or fragmentation in the mass spectrometer (*local exchange analysis*, Zhang & Smith, 1993). In case of digestion, liquid chromatography (LC) has to be applied to reduce the number of overlapping isotopic clusters in a mass spectrum. In case of global exchange analysis or gas-phase-fragmentation-based local exchange analysis, an LC step is not required but often applied for desalting purposes.

10982787, 2017, 5, Downloaded from https://analyticalsciencejournals.onlinelibrary.wiley.com/doi/10.1002/mas.21519 by Universiteit Hasselt, Wiley Online Library on [01/06/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Generating reproducible results requires careful control of experimental factors such as pH and temperature. Automation of sample handling and LCMS analysis (Woods & Hamuro, 2001; Hamuro et al., 2003; Chalmers et al., 2006, 2007, 2011b) increases the reproducibility and precision of HDXMS experiments (Burkitt & O'Connor, 2008; Iacob & Engen, 2012).

2. Gas-Phase Exchange

In comparison to solution-phase exchange, the experimental procedure for hydrogen exchange of gaseous proteins is less complex. An undeuterated protein is ionized by ESI and transferred to a flow/drift tube (Geller & Lifshitz, 2005), an ion trap (Suckau et al., 1993; Wood et al., 1995; Freitas et al., 1999), or traveling-wave ion guides (Rand et al., 2009, 2011b, 2012). In these compartments, a gaseous deuterium source, such as D_2O or ND_3, rapidly deuterates the ionized protein with collisions. After deuteration, the protein can be separated by ion mobility (Khakinejad et al., 2014) and/or fragmented by electron-transfer dissociation (ETD) or electron-capture dissociation (ECD).

III. STEP 1: IDENTIFICATION OF PEPTIDES AND/OR FRAGMENTED IONS

In case of a local exchange analysis, the first step in the computational workflow (see Fig. 1) is the identification of proteolytic peptides or fragment ions. This step should be executed with as much care as possible, because the interpretation of an HDXMS experiment, that is, deuterium incorporation (where, at which rate, etc.), depends on the annotation of the peptides and/or fragments. These are commonly identified using MS analysis of the non-deuterated protein. In order to be able to use the identified undeuterated ions as a guide to find and extract their (partially) deuterated versions, the experimental conditions of the sample preprocessing and handling steps (such as pH, temperature, and the chromatographic flow rate) should be identical to the conditions of the experimental HDXMS workflow.

There are three approaches to identify peptides and/or fragment ions generated in an HDXMS experiment. The most commonly applied method is database-driven identification of tandem MS spectra for the undeuterated (digested) protein. Similarly to standard shotgun proteomics, the selection of peptides to be fragmented can be data-dependent or -independent. The latter requires prior knowledge regarding the mass-to-charge (m/z) value and retention-time of the proteolytic peptides. Most computational HDXMS workflow tools do not identify the fragments themselves. They rely on well-established algorithms such as *Sequest* (Eng, McCormack, & Yates, 1994) or *Mascot* (Perkins et al., 1999).

The second approach to identify proteolytic peptides or fragment ions is based upon *in silico*-generated data. One can compare the mass of isotopic clusters of the undeuterated peptides detected in MS spectra against a database with the theoretical masses of all possible amino acid sequences found in the protein sequence. Unambiguous identification, however, is not always possible. Discarding the isotopic cluster or data-independent tandem MS are two possible remedies in such situations. Alternatively, instead of creating a database with all possible amino-acid sequences, one can also predict the proteolytic cleavage of the protein. However, as shown by Hamuro et al. (2008), the specificity of pepsin is low, which complicates prediction and identification of the observed isotopic clusters. The HDXMS analysis program (Zhang, 2012a) identifies peptides after feature extraction (see Section IV) by comparing their observed masses and fragmentation spectra to the predicted MS/MS spectra (Zhang, 2004, 2005, 2010, 2011). One should be careful with *in silico*-driven identification based upon mass, as it does not consider fragment ion information. Additionally, the mass resolution might be too low to clearly separate peptides with similar masses.

Not every detected peptide and/or fragment ion can be identified with one of these two identification approaches. As a result, the protein sequence coverage might be incomplete. This can limit the interpretation of the deuterium incorporation. The third approach to peptide identification is to increase the protein sequence coverage by combining MS/MS identification with the *in silico*-driven approach. Such an approach is an integral part of *HeXicon* (Lou et al., 2010), *HeXicon2* (Lindner et al., 2014), and is possible in *HDX Workbench* (Pascal et al., 2012) and *QUDeXMS* (Salisbury, Liu, & Agar, 2014). After tandem-MS-based identification, the unidentified masses in the MS spectra can be matched with the theoretical amino acid sequences (*HeXicon*), or with the predicted proteolytic peptides (*HeXicon2* and *QUDeXMS*). With *HDX Workbench*, the user can choose to combine these two *in silico* identification methods.

Similarly to other proteomic MS-experiments, quality control is important and should not be ignored (Wales, Eggertson, & Engen, 2013). For HDXMS experiments, one can check the spectra of the undeuterated reference for digestion and/or fragmentation performance. For instance, *MSTools* (Kavan & Man, 2011) can be used to show the protein sequence coverage after digestion. The spectra of the deuterated samples can be checked for the presence of identified ions with low signal intensity or strongly overlapping isotope distributions. In order not to complicate analysis of the deuterated protein, successfully identified ions that did not pass the quality control may be discarded. As a result, a (peptide) master list is very often created. This list contains all identified peptides or fragments that passed the quality control and that are of interest. As a consequence, completed sequence coverage is not always obtained nor wanted.

In the analysis workflow of *HeXicon*, a random-forest classifier divides the identified peptides and their corresponding deuteration level in high and low quality results. In order to do so, quality features related to abundance, m/z-values, retention-time, D-distribution, and identification are used to score each reported identification. The identified peptides are subsequently ranked according to their quality score. For bimodal distributions, arising due to EX1 exchange kinetics, some quality features are not suited and removed. The ranking of the identified peptides is also replaced by filtering specifically aimed at identifying bimodal deuterium distributions.

IV. STEP 2: FEATURE EXTRACTION AND ANNOTATION OF (PARTIALLY) DEUTERATED IONS

This step is an integral part of local and global HDXMS analyses (see Fig. 1). For bottom-up HDXMS, several methods have been proposed to detect and identify relevant signals. The methods can be divided into three groups: methods in which

10982787, 2017, 5, Downloaded from https://analyticalsciencejournals.onlinelibrary.wiley.com/doi/10.1002/mas.21519 by Universiteit Hasselt, Wiley Online Library on [01/06/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

10982787, 2017, 5, Downloaded from https://analyticalsciencejournals.onlinelibrary.wiley.com/doi/10.1002/mas.21519 by Universiteit Hasselt, Wiley Online Library on [01/06/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

feature extraction and annotation are both dependent on the peptide master-list (Section IV.A); methods in which only annotation of the deuterated peptides is based on the pre-defined master-list (Section IV.B); and methods in which feature extraction and annotation are independent from a peptide master list (Section IV.C). Some of the proposed methods are also applicable to top-down HDXMS. One tool, *Hydra* (Slysz et al., 2009), explicitly provides the option to extract features from HDX-MS/MS spectra.

Applications of gas-phase HDXMS to analyze protein structure and/or dynamics are limited. The processing of these experiments can be done manually or, analogously to solution-phase top-down HDXMS, processed with tools designed for bottom-up HDXMS experiments.

Although the proposed methods can be split in three groups, they are different and often optimized for a particular type of mass spectrometers or resolution-level. For this reason, we discuss these methods separately in the following subsections. Additionally, we also discuss an alternative LCMS alignment method specifically designed for HDXMS data.

A. The Peptide Master-List Is Used for Feature Extraction and Feature Annotation

The methods in this category require information that is present in the master list, such as retention-time, mass, charge, and the amino-acid sequence. The information is essential to determine the retention-time range and the mass range, which are generally used to reduce the search space.

In case of the retention-time range, it is assumed that the chromatographic properties of deuterated peptides are identical to the properties of the undeuterated variant. As a consequence, their retention-time should be the same. To account for small differences in retention-time, a window around the retention-time of the undeuterated peptide is defined. In some cases, the spectra are aligned before the retention-time window is determined. For the mass range, the expected mass increase due to deuteration is taken into account. In case of high resolution data, the mass range can be divided in subranges that are centered around the *m*/*z*-values of the (expected) aggregated isotopic variants.

Extracting and annotating the features should be repeated for each peptide in the master list.

1. Deuterator, HD Desktop, and HDX Workbench

In *Deuterator* (Pascal et al., 2007), the retention-time window is calculated by adding and subtracting a user-defined number of minutes to the retention-time of the undeuterated peptide of interest. For the mass range, the lower bound corresponds to the theoretical mass of the non-deuterated peptide ($m/z_{0\%}$). This value is based upon the charge and amino-acid sequence of the peptide found in the master list. The upper mass bound is equal to the theoretical *m*/*z*-value of the fully deuterated ion ($m/z_{100\%}$). This value is calculated as follows:

$$m/z_{100\%} = m/z_{0\%} + (\text{nbAa} - \text{nbPro} - 2)/z, \qquad (3)$$

where nbAa is the number of amino-acid residues, nbPro is the number of Prolines, which do not have exchangeable backbone hydrogens, and two corresponds to the first two amino-acids in the sequence for which no deuterium incorporation can be recorded (Bai et al., 1993).

Based upon these retention-time and mass ranges, a small part of the limited number of mass spectra, which supposedly contain the signal of interest, is retained. *Deuterator* extracts and summarizes this signal, that is, the isotope distribution, by averaging the intensities of each *m*/*z*-value over the selected spectra. This process of co-adding isotope distributions increases the signal-to-noise ratio. Highly resolved data are further processed to extract the isotope distribution more precisely. For each isotopic variant, a sub-window is determined. The center of these windows are the *m*/*z*-values, calculated based upon the known mono-isotopic mass (m_0/z) and the spacing ($1/z$) between them. The width of the sub-windows is set to values that correspond to the instruments' resolving power. Detected ions outside the ranges are removed.

To assess the accuracy of the retention-time window, the signal in each spectrum is represented as an average mass value. Subsequently, an extracted ion-chromatogram (XIC) for this average mass is generated over all scans. The XIC allows the user to manually re(de)fine the retention-time window.

With *HD Desktop* (Pascal et al., 2009), which is a successor of *Deuterator*, the optimal retention-time window for a peptide is determined in an automated manner. Within the user-defined retention-time range, a narrow window slides from the lower to the upper bound of the wider range with steps of 3 sec. At each step, the averaged isotope distribution is calculated and compared to the theoretical distribution. The retention-time window, for which the observed and expected signal matches best, is chosen as the optimal retention-time range. This moving window procedure can also be executed without a user-defined retention-time range.

HDX Workbench (Pascal et al., 2012) is exclusively designed for high-resolution MS instruments. For this reason the moving window approach is no longer needed and is replaced by filtering based on mass accuracy, *m*/*z*-range, retention-time range, and intensity.

2. Hydra

The standard feature extraction and annotation module in *Hydra* (Slysz et al., 2009) generates XICs based upon the *m*/*z*-value of a peptide from the master list. Subsequently, a peak in the XIC and its corresponding mass spectra are selected within a user-defined retention-time range. Thereafter, an isotopic-profile detection algorithm is applied to the selected spectra. This algorithm matches the expected *m*/*z*-values of the aggregated isotopic variants within an *m*/*z*-window centered around observed *m*/*z*-values.

Hydra also provides two methods that enable MS/MS data processing. The first algorithm is an MS/MS generator which extracts tandem MS spectra for each peptide ion (precursor) and merges multiple MS/MS spectra guided by *m*/*z* mass tolerance and width of eluted chromatographic peaks. The second tool, the MS/MS fragment analyzer, combines the isotopic-profile detection procedure and an algorithm to determine the amount of deuterium (see Section V).

3. The Method Proposed by Kazazic et al. (2010)

For each peptide in the pre-defined master-list, a user-defined retention-time window is used to select a subset of mass spectra.

In these spectra, the isotope distribution of a peptide is extracted with a "*m*/*z*-sub-window" method, specifically designed for high-resolution Fourier-transform-ion-cyclotron-resonance (FTICR) MS data. Within the mass range of a peptide that stretches from the mono-isotopic mass of the undeuterated peptide ($m_0/z_{0\%}$) to the theoretical mass of the fifth aggregated isotopic variant of the fully deuterated peptide ($m_4/z_{100\%}$), the width and the position of sub-windows for each isotopic variant are calculated based on the elemental composition, charge state, and number of labile hydrogens. For every sub-window, the midpoint between the minimum and maximum number of exchangeable H-atoms is computed and compared with the masses of the observed spectral peaks. The observed spectral peak with a mass identical to or closest to the midpoint-mass is considered to be part of the isotope distribution of interest, and is thus extracted.

4. ExMS and HDsite

ExMS (Kan et al., 2011) and *HDsite* (Kan et al., 2013) use the same feature extraction and annotation method. In each spectrum within the retention-time range of a peptide, spectral peaks are detected and centroided with a 9-point algorithm. This method selects a peak if the central part is greater than four adjacent spectral peaks on both sides. The centroided *m*/*z*-values of these selected spectral peaks are compared with the theoretical masses of isotopic variants of the peptide of interest in the mass interval

$$[m_0/z_{0\%}, (m_0 + \text{maxH} + \text{maxD})/z],$$

where $m_0/z_{0\%}$ is the undeuterated mono-isotopic mass, maxH is the maximum number of hydrogen atoms, and maxD the maximum number of deuterium atoms.

A spectral peak is considered to match an expected *m*/*z*-value if it is within a given mass-accuracy tolerance. After the spectral peaks are extracted, several validation steps are performed to check if, for instance, there is no mis-assignment due to overlapping peptide envelopes.

A sliding-window procedure is applied to all adjacent spectra containing a validated feature to redefine the retention-time range and to obtain a single summed envelope that after validation and a manual check will be used in the remaining computational analysis steps.

5. DEX

The "deconvolution of exchange data"-algorithm (*DEX*, Hotchko et al., 2006) extracts the isotope distribution of known peptides based upon their mass. The observed masses in a spectrum are resampled into evenly spaced increments of 0.1 mass units. Based upon the amino-acid sequence and charge, the mass range is determined. The mass range should be minimally 1.5 times the sequence length. The lower bound is $1.5/z$ Da smaller than the undeuterated, theoretical mono-isotopic mass ($m_0/z_{0\%}$), while the upper bound is $1.5/z$ Da times the number of residues in the amino-acid sequence higher than $m_0/z_{0\%}$.

6. HX-Express and HX-Express2

HX-Express (Weis, Engen, & Kass, 2006; Weis, Wales, & Engen, 2006) and *HX-Express2* (Guttman et al., 2013) define the mass range for a peptide from the master list exactly in the same way as *DEX*. However, instead of using all data in this mass range, these two tools further reduce the number of spectral peaks by identifying the aggregated isotopic variants and by determining the relative width of the (deuterated) isotope distribution. The aggregated isotopic variants that form the peptide's isotope distribution are identified by selecting the spectral peaks that will be nominated as aggregated isotopic variant if following criteria are met: the *m*/*z*-distance between isotopic variants should be equal to $1/z \pm 0.03$ Da; and the intensity of an isotopic variant, *i*, should be higher than the user-defined noise threshold and should be between the intensities of its neighboring isotopic variants, that is, $I_{i\text{-}1} \leq I_i \leq I_{i+1}$ for the variants with a lower mass than the most abundant peak and $I_{i\text{-}1} \geq I_{\text{i}} \geq I_{i+1}$ for the heavier variants. After identifying the aggregated isotopic variants of an isotopic cluster, the width of the isotope distribution is determined as follows:

$$\text{width} = z \times \Delta m/z, \tag{4}$$

where $\Delta m/z$ is the *m*/*z*-difference between the two most extreme points of the isotope distribution with an intensity equal to 20% of the maximum intensity. If there is no point or isotopic variant with an intensity smaller than this percentage, the width is directly determined. Plotting the width against the labeling time can be used to detect EX1 kinetics. In contrast to EX2 exchange, the width of the isotope distribution of an ion undergoing EX1 kinetics will change noticeably (Weis, Wales, & Engen, 2006).

The centroid of the isotope distribution is calculated as the intensity-weighted average of the observed mass-to-charge values of the isotopic variants.

7. HeXicon

Within *HeXicon* (Lou et al., 2010), each peptide in the master list is represented by a number of basis functions. The functions combine the natural isotope distribution of a peptide with all possible levels of deuteration, and are used to extract features and derive the deuteration distribution. After defining the mass range and retention-time window (referred to as LCMS segmentation), spectral peaks that belong to the same peptide are grouped and integrated. Thereafter, the exact retention-time position of the peak group is estimated and the abundance (β) of the basis functions of the peptide of interest is determined with *NITPICK* (Renard et al., 2008). For each incubation time, the identified peak groups and their characteristics (m_0/z, β, z, retention-time) are listed. Subsequently, the correspondence between the undeuterated peak group and the peptide sequence of interest is determined by matching the observed peak masses to their theoretical *m*/*z*-value. Correspondence between the undeuterated and deuterated peak groups is determined by an Euclidean distance measure that weighs the contributions of different characteristics including the charge and retention-time.

In 2011, *HeXicon* was modified to detect and analyze bimodal isotopic-peak distribution due to EX1 exchange kinetics (Kreshuk et al., 2011). One of these modifications included the LCMS alignment of the total ion chromatograms of all samples to a reference sample. As a result, the corresponding estimation improves, especially when the estimated retention-times of the samples are too different. Additionally, the stopping

10982787, 2017, 5, Downloaded from https://analyticalsciencejournals.onlinelibrary.wiley.com/doi/10.1002/mas.21519 by Universiteit Hasselt, Wiley Online Library on [01/06/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

criterion of *NITPICK* was changed to prevent ignoring overlapping and correlated basis functions which are needed to detect the bimodal exchange-pattern of peptides.

B. The Peptide Master-List Is Only Used for Feature Annotation

Three workflow analysis tools include procedures that detect and extract features without any prior knowledge. The detected ions matched to entries of the peptide master-list will be retained. Feature extraction and annotation should be repeated for each scan with different labeling time points. After annotation, features with the same identification, but with different labeling times, are grouped together and passed on to deuteration-level-determination (see Section V).

1. AUTOHD

AUTOHD (Palmblad, Buijs, & Håkansson, 2001) relies on external programs such as *THRASH* (Horn, Zubarev, & McLafferty, 2000) for feature detection and extraction. The resulting isotopic clusters are identified by comparing the theoretical and experimental masses and by comparing the observed and expected intensities (see Section V). A peptide from the predefined master list is selected if its expected m/z-value with $x\%$ deuteration ($0 \leq x\% \leq 100$) corresponds to the observed m/z-value.

2. HDXFinder

HDXFinder (Miller et al., 2012) extracts and annotates spectral features in an automated manner. It gradually reduces the number of peaks detected in a spectrum by applying a series of filtering steps. After detecting, centroiding, and denoising of spectral peaks, clusters of peaks are created by grouping all peaks which are separated by maximally one dalton. Within each cluster, for each pair of peaks, the charge (z^*) is determined. If at least three adjacent peaks are separated by the same m/z-distance, that is, $(1 + \Delta)/z^*$ with Δ denoting the mass tolerance, these peaks are considered to be part of an isotopic cluster. The isotopic clusters are subsequently grouped over multiple spectra and retained if they are found in at least four consecutive spectra. The clusters are summarized, that is, their average mass is calculated and matched with the master list peptides. In order to be considered to be matching, the observed average mass should be between the boundaries of following mass range:

$$[m/z_{0\%}; m/z_{0\%} + (\mathrm{nbAa} - \mathrm{nbPro} - 1) \times m_{1\mathrm{H}}/z],$$

where $m_{1\mathrm{H}}$ denotes the mass of one proton (for the definition of other symbols, see Equation (3)).

3. HeXicon2

To improve the computational efficiency and accuracy of the analysis of high-resolution MS data, *HeXicon* was restructured and modified into *HeXicon2* (Lindner et al., 2014). *NITPICK* is no longer applied to the deuterated data, but only to the undeuterated reference spectra. As a result, feature detection is no longer based on the peptide master-list. For each detected peak in the undeuterated spectra, *NITPICK* generates an overcomplete set of candidate isotope distributions using an improved averagine model (Senko, Beu, & McLafferty, 1995; Renard et al., 2008). For each of the undeuterated spectra, a linear combination of those distributions is created. All spectral peaks with a candidate isotope distribution for which the estimated abundance is different from zero are selected. Each of these selected peaks are subsequently identified by matching their m/z-values to the masses of the peptides in the master list. Thereafter, the isotope distributions of the identified peptides are deuterated *in silico* and their isotope-variant masses are compared to the observed masses of the deuterated spectra. Matching signals are selected and their retention-times are predicted with the LCMS alignment. This matches the deuterated distributions to a reference, but also reduces the number of searched spectra. The intensities of the isotopic variants are extracted in predicted m/z-regions, similarly as in *ExMS*, *HDX Workbench* and the approach by Kazazic et al. (2010).

C. No Peptide Master-List Is Used for Feature Extraction and Feature Annotation

In contrast to the three methods presented in Section IV.B, one can annotate all detected features instead of matching them to the undeuterated ions from the master-list. The computational analysis tool for high-resolution data proposed by Zhang (2012a) is the only method at this moment that works in this manner. It uses *MassAnalyzer* (Zhang, 2009) in combination with an iterative divide-and-conquer LCMS alignment algorithm (Zhang, 2012a) to process bottom-up HDXMS experiments. After ion detection, based upon the isotope distribution, XICs are generated for each detected ion. A detectable peak in the XIC is a considered to be a confirmation that the detected ion is a peptide. This step ensures that spurious signals and background noise are not retained and further processed.

An XIC is typically created by summing the intensities of all ions within a m/z-window. *Mass Analyzer*, however, uses a weighting approach that considers the shape of the isotope distribution of the screened ion. For centroided, well-resolved data, narrow sub-windows are centered around the expected m/z-values of the isotopic variants. Intensities within these sub-windows are summed, while all other spectral peaks get weight zero and are ignored. This approach reduces interferences of neighboring ions and improves the accuracy of the XIC peak-area calculation.

From the list of confirmed ions, *anchor features* are selected. These features have a clearly observable mono-isotopic mass, well-resolved XIC peaks, and a chromatographic peak width that is not too wide. The iterative LCMS alignment method (Zhang, 2012a) uses the anchor features to determine the retention-time shifts of all detected ions. It aligns the retention-time of all anchor features against a reference. After adjusting the retention-times, the chromatogram is divided in two overlapping sub-parts and the alignment step is re-executed. This procedure is repeated until no more than ten features are left in one sub-section or if the range of a section is at most 1/4 of a typical XIC peak.

10982787, 2017, 5, Downloaded from https://analyticalsciencejournals.onlinelibrary.wiley.com/doi/10.1002/mas.21519 by Universiteit Hasselt, Wiley Online Library on [01/06/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Mass spectra of a detected feature are combined using a matched filter (Zhang & McElvain, 1999) and background subtraction. From the combined spectra, the average mass of a detected ion is calculated based upon the centroided isotopic-variant peaks. After adjusting the retention-time and combining the mass spectra, corresponding ions across different labeling times are grouped together. Peak detection is repeated to extract features that are missed in the first iteration.

The observed masses of the grouped features can be used to identify each peptide or fragment ion against predicted MS/MS spectra (see Section III). Extracted, post-translational modified proteins can be identified with this approach.

D. An Alternative LCMS Alignment Method for HDXMS Data

Venable, Scuba, & Brock (2013) introduced an alternative method to align spectral data with mass-shifted features. This method iteratively estimates a retention-time mapping function for the features. For HDXMS data, the detected deuterated peptides or fragmented protein ions are paired with undeuterated features based upon their first observable isotopic variant. This isotopic variant is the mono-isotopic variant for the undeuterated feature, but for deuterated features it is not always the case. To account for potentially unobservable mono-isotopic masses, the window, wherein deuterated features are matched with undeuterated ions, is very wide. A consequence is that a large number of pairs are incorrect. These erroneously matched features are eliminated by iteratively fitting a fourth-order polynomial and excluding matched pairs that are outside a retention-time-based band-constraint centered around the estimated retention-time mapping function.

V. STEP 3: DETERMINING THE DEUTERATION LEVEL OF THE ANNOTATED FEATURES

The next step in the computational workflow (see Fig. 1) is to determine—for each extracted peptide, fragment or protein, and at every labeling time point—how many hydrogens have been exchanged for deuterium (Section V.A). Next to quantifying the deuteration level, one can correct the obtained deuterium content (Section V.B). In case of overlapping protein-fragments, one can also try to increase the spatial resolution (Section V.C).

A. Calculating the Deuterium Content

There are two main approaches to retrieving the number of exchanged deuteriums from the isotope distribution: a centroid approach, which returns the relative deuterium content, and a deconvolution-based approach. Additionally, *QUDeXMS* and *HeXicon* offer two alternative methods.

1. Centroid Approach

The simplest way to calculate the deuterium content is to reduce the extracted isotope distributions of the protein, peptide, or a fragment ion to the intensity-weighted average mass (Zhang & Smith, 1993). At each labeling or folding time point, t, the relative deuterium content of a deuterated ion (D_t) can then be computed by subtracting the average mass of the undeuterated ion ($m_{\mathrm{avg},t0}$) from the calculated, deuterated average mass ($m_{\mathrm{avg},\,t}$):

$$D_t = m_{\mathrm{avg},t} - m_{\mathrm{avg},t0}. \tag{5}$$

This approach is commonly used in many HDXMS analysis workflow tools (Weis, Engen, & Kass, 2006; Weis, Wales, & Engen, 2006; Pascal et al., 2007, 2009, 2012; Slysz et al., 2009; Kazazic et al., 2010; Kan et al., 2011; Liu et al., 2011; Miller et al., 2012; Zhang, 2012b; Rey et al., 2014).

2. Deconvolution-Based Approach

Next to the centroid approach, deconvolution-based methods are also available. They "dismantle" the observed isotope distribution of a (partially) deuterated molecule in the natural isotope distribution and the deuterium content.

Theoretical fitting. "Theoretical fitting"—methods consider the observed, (partially) deuterated isotope distribution to be a convolution of the undeuterated, natural isotope distribution and the exchangeable-hydrogen distribution.

In *AUTOHD*, a Fourier-Transform-based convolution method (FT) is used to calculate the intensities of the expected isotope distribution of a (partially) deuterated protein:

$$I_{\mathrm{calc}} = F^{-1}\Big(F(I_C)^{\mathrm{nC}} \times F(I_H)^{\mathrm{nH}} \times F(I_O)^{\mathrm{nO}} \times F(I_N)^{\mathrm{nN}} \times F(I_S)^{\mathrm{nS}} \times F(I_{\mathrm{sol}})^{\mathrm{nsol}} \times F(I_D)\Big), \tag{6}$$

where $F(I_X)$ denotes the Fourier-transformed intensity distributions of X, I_{sol} denotes the intensity distribution of H-atoms that exchange too fast to be recorded, I_H is the intensity distribution of H-atoms that do not exchange, and I_D is the intensity distribution of the exchangeable H-atoms, which is assumed to be binomial:

$$I_D = \mathrm{Bin}(P_{\mathrm{exch}}, n_{\mathrm{labile}}), \tag{7}$$

where P_{exch} is the probability of exchange and n_{labile} the number of exchangeable H-atoms. Note that Equation (7) assumes that all labile H-atoms exchange with the same probability. The exchange probability, P_{exch}, is estimated by minimizing χ^2:

$$\chi^2 = \sum_{i=0}^{K-1} \frac{\left(I_{\mathrm{calc},i} - I_{\mathrm{obs},i}\right)^2}{\mathrm{cI}_{\mathrm{obs},i} + \sigma^2_{\mathrm{noise}}}, \tag{8}$$

where K denotes the number of observed isotopic variants, c is a constant term determined from peaks with similar intensities and m/z-values in control experiments, and $\sigma^2_{\mathrm{noise}}$ is the random noise variance observed in the analyzed spectra.

For each labeling time point, and for each candidate that has a mass corresponding to one of the peptides in the master list (see Section IV.B), one P_{exch} is estimated. The candidate with the smallest χ^2-value is saved, together with the estimated P_{exch}.

10982787, 2017, 5, Downloaded from https://analyticalsciencejournals.onlinelibrary.wiley.com/doi/10.1002/mas.21519 by Universiteit Hasselt, Wiley Online Library on [01/06/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

HX-Express2 uses an approach which is related to the method implemented in *AUTOHD*. After deconvoluting the deuterated isotope distribution with the natural isotope distribution, the average deuterium content is estimated with asymmetric least squares:

$$\text{ALS} - \text{score} = \sum_{D=0}^{n_{\text{labile}}} \lambda \times \left(I_D^{\text{calc}} - I_D^{\text{exp}}\right)^2, \tag{9}$$

where I_D^{calc} and I_D^{exp} are the peak intensities of, respectively, the calculated and deconvoluted deuteration distribution with D deuterons, and λ is the asymmetry-term which allows correcting for overlapping isotope distributions. Similar to Palmblad, Buijs, & Håkansson (2001), the deuterium-content distribution is considered to be binomially distributed:

$$I_D^{\text{calc}} = A \times \frac{n_{\text{labile}}!}{D!(n_{\text{labile}} - D)!} p^D (1-p)^{n_{\text{labile}} - D}, \tag{10}$$

where p denotes the average deuteration probability and Aa weighting term.

In case of bimodal deuterated isotope distributions (EX1 exchange kinetics), a combination of two binomial distributions is used, where the exchange probability of the highly deuterated species (p_2) is larger than the probability of the less deuterated species (p_1), and less than the fully deuterated labeling probability ($p_{100\%}$).

The deconvolution procedure used in *HX-Express2* adds zeros to the lower and higher mass end of the deuterated isotope distribution as suggested by Chik, VandeGraaf, & Schriemer (2006).

In *HDsite*, *Deuterator* and its successors, *HD Desktop*, and *HDX Workbench*, another approach, similar to the method implemented in *AUTOHD* (Palmblad, Buijs, & Håkansson, 2001), is used. For the selected feature, a hundred theoretical isotope distributions are calculated with *qmass* (Rockwood & Haimi, 2006). Each of these distributions corresponds to a fixed percentage of deuterium incorporation, ranging from 1% to 100% of all hydrogens. These theoretical distributions are compared to the observed distribution. The percentage of incorporated deuterium of the theoretical isotope distribution with the smallest "ordinary least squares"-score (*Deuterator*) or χ^2-score (*HD Desktop* and *HDX Workbench*) is saved.

"Mass-shifted-distribution" approach. Other computational workflow algorithms have implemented methods that consider the isotope distribution of a deuterated ion to be a linear combination of mass shifted natural isotope distributions (see Fig. 2):

$$\text{DI}_i = \sum_{j=0}^{i} \omega_j \times I_{i-j}, \tag{11}$$

where DI_i denotes the normalized intensity of the deuterated ith isotopic variant, ω_j is the weight or abundance of j exchanged deuteriums ($\Sigma\omega_j = 1$), and I_{i-j} is the intensity of the undeuterated $(i-j)$th isotopic variant.

In contrast to the "centroid" and "theoretical fitting" methods, these approaches return, at every labeling time point, a distribution of relative deuterium-content. As indicated in several articles (Kan et al., 2013; Zhang et al., 2013), the centroided deuterium-content does not always return enough information to draw correct conclusions regarding the structural properties and dynamics of the protein under study.

Zhang, Guan, & Marshall (1997) discussed three different approaches to determine ω_j. If the number of ω_j's is smaller than the number of observed isotopic variants (K), an exact solution can be retrieved by directly solving Equation (11) for $i = 0, \ldots, K\text{-}1$. However, this is rarely the case. An alternative method is to estimate ω_j with ordinary least-squares (*OLS*), that is, minimizing

$$C = \sum_{i=0}^{K-1} \left(\sum_{j=0}^{i} \omega_j \times I_{i-j} - I_{\text{obs},i} \right)^2, \tag{12}$$

where $I_{\text{obs},i}$ denotes the observed intensity of the ith isotopic variant.

OLS-fitting is relatively fast and tries to find the optimal ω_j-values. However, it is not desirable to have an exact fit to the experimental data as the intensities of the isotope distribution are often distorted by experimental measurement errors. Directly solving Equation (11) is even more prone to noise (Zhang, Guan, & Marshall, 1997).

The third method to determine the deuteration level, proposed by Zhang, Guan, & Marshall (1997), is the maximum-entropy method (*MEM*). It estimates ω_j-values that fit the observed data while accounting for measurement errors. *MEM* maximizes the entropy of $\omega_j(S)$ and minimizes, at the same time, a least-squares criterion:

$$Q = S - \lambda C^*, \tag{13}$$

where λ is a Lagrange multiplier, $C^* = C/\sigma_i^2$, and σ_i^2 the variance of the intensity of isotopic variant i. The combined maximization of S and minimization of C^* results in a smoothed deuteration distribution, but also in increased computation time.

As indicated by Chik, VandeGraaf, & Schriemer (2006), ω_{K-1} is only determined by the last observed isotopic variant. The abundance of this selected peak is generally low, and as such prone to measurement errors. As a consequence, the estimated deuteration distribution is perturbed. Chik, VandeGraaf, & Schriemer (2006) proposed to add a moderate number of zeroes (Z) to the higher mass end of the extracted isotope distribution:

$$\sum_{i=0}^{K-1} \text{DI}_i = \sum_{i=0}^{K-1} \left(\sum_{j=0}^{i} \omega_j \times I_{i-j} \right) + \sum_{i=K}^{K+Z-1} I_i^*, \tag{14}$$

with $I_i^*, \ldots, I_{K+Z-1}^* = 0$. Extending the extracted deuterated isotopic-cluster with zeroes to account for missing or truncated isotopic variants improves the fit and estimated deuteration

10982787, 2017, 5, Downloaded from https://analyticalsciencejournals.onlinelibrary.wiley.com/doi/10.1002/mas.21519 by Universiteit Hasselt, Wiley Online Library on [01/06/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

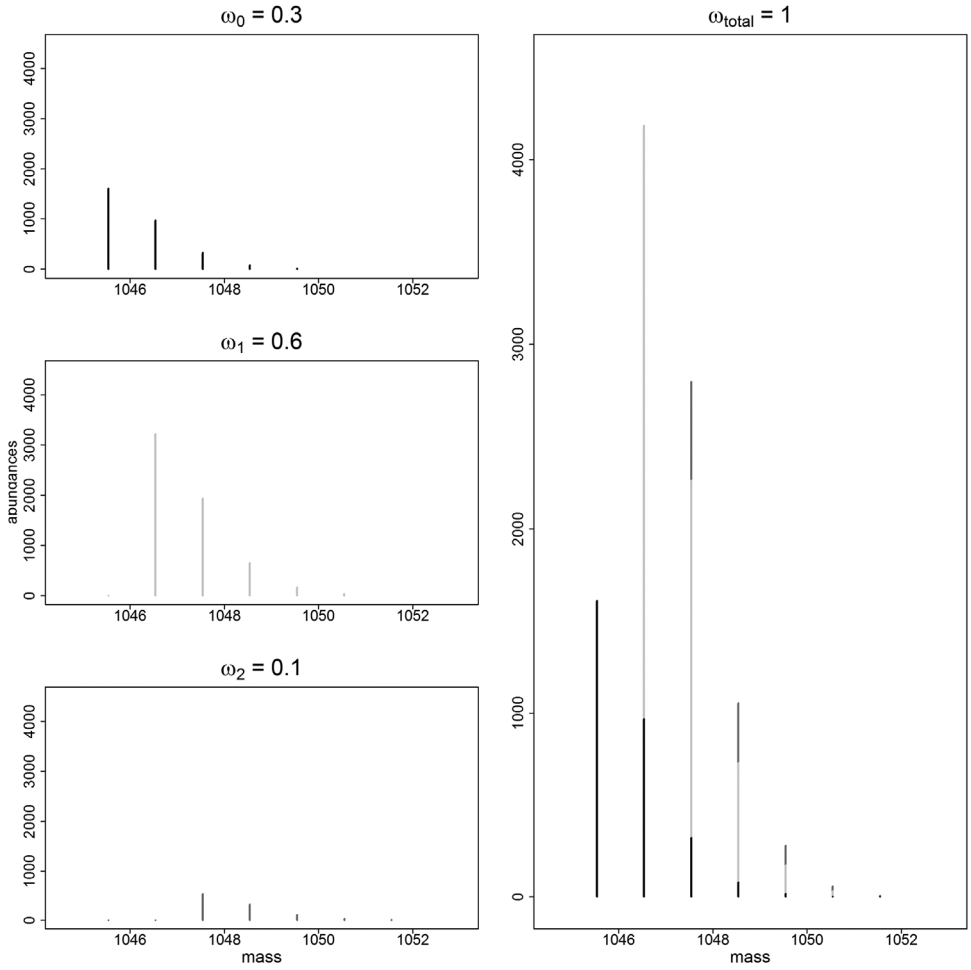

FIGURE 2. The partially deuterated isotope distribution (right-hand-side panel) is the weighted sum of three isotope distributions with different levels of deuteration, that is, no deuterium (top left panel), one deuterium atom (middle left panel), and two deuterium atoms (bottom left panel).

distribution (Chik, VandeGraaf, & Schriemer, 2006). This approach is implemented in *CalcDeut* (Chik, VandeGraaf, & Schriemer, 2006) and *Hydra* (Slysz et al., 2009).

Next to the *MEM*-procedure, *DEX* and *HeXicon2* also try to reduce the impact of the experimental noise of the estimated distribution. In *DEX*, ω_j is extracted from the observed isotope distribution with FT. FT deconvolution of the isotope distribution locally averages the random experimental noise, which results in a moderate reduction of the experimental noise (Hotchko et al., 2006).

HeXicon2 uses an implementation of Gold's iterative deconvolution algorithm (Morháč, 2006) to retrieve a smoothed estimate of the incorporated deuterium distribution for annotated features that have a required number of consecutive isotopic variants and clearly noticeable intensities.

3. Alternative Approaches

HeXicon. In *HeXicon*, an alternative procedure to obtain the deuteration distribution (ρ) of a feature g is proposed. In the feature annotation and extraction step (see Section IV), for each feature, the abundance (β) of a basis function corresponding to the deuteration level, D, is determined. The deuteration distribution of a feature is then calculated based upon the estimated abundance:

$$\rho(D) = \frac{\beta_{g,t}(D)}{\sum \beta_{g,t}}. \tag{15}$$

QuDeXMS. QUDeXMS (Salisbury, Liu, & Agar, 2014) is specifically designed for ultra-high-resolution HDXMS. As a result, the fine isotope distribution of a (partially) deuterated distribution is recorded. Instead of deconvoluting the whole isotope distribution, *QUDeXMS* extracts the mono-isotopic peak and the pseudo-mono-isotopic peaks (Liu, Easterling, & Agar, 2014). The pseudo-mono-isotopic peaks are all fine isotopic variants with m/z-values that are equal to

$$m_0/\mathrm{z} + [(\mathrm{m}_{2H} - \mathrm{m}_{1H})/z] \times i$$

with $i = 0, \ldots, n$, where n is the maximum number of exchangeable hydrogen atoms. These pseudo-mono-isotopic

10982787, 2017, 5, Downloaded from https://analyticalsciencejournals.onlinelibrary.wiley.com/doi/10.1002/mas.21519 by Universiteit Hasselt, Wiley Online Library on [01/06/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

peaks are used to calculate the average number of incorporated deuterium atoms:

$$D_t = \sum_{i=0}^{n} P_{F,i} \times i, \tag{16}$$

where $P_{F,i}$ is the relative contribution of the pseudo-mono-isotopic peak i to the total abundance/intensity of all pseudo-mono-isotopic peaks and the mono-isotopic peak. In cases where the mono-isotopic peak is not the most abundant peak of the undeuterated isotope distribution, this procedure is repeated, but the average deuterium-content is estimated based on the most-abundant ^{13}C peak and the corresponding pseudo-most-abundant peaks.

B. Correcting the Calculated Deuterium-Content

The relative deuterium-content determined by methods discussed in Section V.A cannot immediately be used due to imprecisions linked to in- and back-exchange and minor variations in experimental conditions and in back-exchange.

1. In-Exchange and Back-Exchange

In bottom-up HDXMS, during proteolysis and separation, the incorporated deuterium can be replaced by H-atoms even with fast separation and under quenching conditions, that is, lowered pH and temperature. On the other hand, in-exchange, that is, incorporation of deuterium during digestion, is also possible (Hoofnagle, Resing, & Ahn, 2003). These processes of in-exchange and back-exchange alter the deuterated isotope distribution. As a consequence, the determined deuterium-content and/or deuteration distribution do not reflect the true HDX process.

Zhang & Smith (1993) and Hoofnagle, Resing, & Ahn (2003) proposed similar methods to correct the determined deuterium content. The correction mechanism of Zhang & Smith (1993) is the most commonly applied:

$$D_{\mathrm{corr},t} = \frac{D_t}{m_{\mathrm{avg},100\%} - m_{\mathrm{avg},t_0}} \times n_{\mathrm{labile}}, \tag{17}$$

where $D_{\mathrm{corr},t}$ denotes the corrected deuterium-content at time t, n_{labile} is the number of exchangeable H-atoms, and $m_{\mathrm{avg},100\%}$ and m_{avg,t_0} are, respectively, the average mass of the fully deuterated peptide under study and the average mass of the peptide at time t_0.

There are two issues with this correction mechanism. First, fully deuterating the protein is not always possible. Second, Equation (17) is applicable to 92% of the peptides, where the correction error is less than 10%. For the other peptides, the error is significantly higher (Zhang & Smith, 1993; Wales, Eggertson, & Engen, 2013).

Hsu, Johnson, & Traugh (2008) calculate the back-exchange correction factor, $B\%$, as a ratio between the incorporated deuterium after 24 hr of exchange and the total number of exchangeable sites. The deuterium content is corrected as follows:

$$D_{\mathrm{corr},t} = \frac{m_t - m_0}{B\%} - D_{s,t}, \tag{18}$$

where $D_{s,t}$ denotes the residual deuterium present in the side-chains at time t.

The aforementioned methods assume that all deuteriums have the same back-exchange, which is rarely the case. In *HDsite* (Kan et al., 2013), the extracted deuterium-content of each amino-acid residue is corrected with different factors. First, the total number of lost deuterium atoms is determined at the peptide-level using a fully deuterated protein. Based upon the back-exchange rates reported for unstructured peptides (Bai et al., 1993; Connelly et al., 1993), the effective back-exchange time for each peptide is determined. These effective back-exchange times are subsequently used to calculate the loss of deuterium for each residue in the peptide.

An alternative method to correct for back-exchange is proposed by Zhang (2012a). Instead of correcting the deuterium content, the effective back-exchange time is modeled, and used to estimate the protection extent of each labile hydrogen (see Section VI.C.3).

For each experimental run r, based upon the deuterium content of a fully deuterated fragment j, the effective back-exchange time of fragment j is estimated by minimizing the difference between the observed deuterium content and the expected deuterium content $D_{j,r}$. $D_{j,r}$ is calculated as follows:

$$D_{j,r} = \sum_i e^{-k_i^B t^B_{\mathrm{eff},j,r}} \times e^{-k_i^A t^A_{\mathrm{eff},j,r}}, \tag{19}$$

where k_i^B and k_i^A, and $t^B_{\mathrm{eff},j,r}$ and $t^A_{\mathrm{eff},j,r}$ denote, respectively, the kinetic exchange rates, for a labile hydrogen i, and effective back-exchange times before and after digestion. The exchange rates are calculated theoretically, according to the method described by Bai et al. (1993). The effective back-exchange time accounts for small variations between different experimental runs and is calculated as follows:

$$t^B_{\mathrm{eff},j,r} = t_j^B + \Delta t_r^B, \tag{20}$$

$$t^A_{\mathrm{eff},j,r} = t_j^A + \Delta t_r^A. \tag{21}$$

The minimization routine used to determine the effective back-exchange time values requires the parameters (t_j^B, t_j^A, Δt_r^B, and Δt_r^A) to be constrainted to ensure meaningful outcomes (Zhang, 2012a).

According to Engen and co-workers (Wales & Engen, 2006; Morgan & Engen, 2009; Marcsisin & Engen, 2010; Wales, Eggertson, & Engen, 2013), there is little advantage in correcting for back-exchange. Back-exchange correction is necessary in cases where the calculated deuterium-content is compared between different proteins, fragments, or peptides, or when one is interested in the exchange rates or protection extent of the exchangeable H-atoms.

10982787, 2017, 5, Downloaded from https://analyticalsciencejournals.onlinelibrary.wiley.com/doi/10.1002/mas.21519 by Universiteit Hasselt, Wiley Online Library on [01/06/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

2. Variations in $\%D_2O$

The level of deuteration is influenced by experimental parameters such as the pH, temperature, concentration of deuterium, etc. Despite automation of the experimental workflow, minor fluctuations of the experimental conditions cannot be excluded. As such, reproducibility and especially interday reproducibility, remains a challenge (Iacob & Engen, 2012). Sheff & Schriemer (2014) proposed a method to correct for dispensing errors. Caffeine in two forms ("heavy" and "light") are used as internal standards to determine a correction factor:

$$X_{\text{corr}}(\%D_2O) = (\text{heavy}/(\text{heavy}+\text{light})), \tag{22}$$

where "heavy" and "light" represent the XIC intensities for the heavy and light form of caffeine, respectively. This correction factor should be used to divide the calculated relative deuterium-content.

3. Variations of Deuterium Back-Exchange

Zhang (2012b) proposed to use internal standards to correct for experimental run-to-run variations of deuterium back-exchange. Three short peptides are used as internal standards. Assuming that all backbone amide hydrogens of these peptides are exchanged during the shortest deuterium labeling period, a correction factor b is calculated and used to normalize the deuterium content of all features to the deuterium level of the fully deuterated internal standard ($D^0_{100\%}$):

$$b = \frac{1}{n_{\text{IS}}}\Sigma\frac{D^0_{100\%}}{D_{\text{IS},\,t}}, \tag{23}$$

$$D_{\text{corr},t} = D_t \times b, \tag{24}$$

where $D_{\text{IS},t}$ is the deuterium level of an internal standard at timepoint t and n_{IS} the total number of internal standards and D_t the deuterium level at time t.

C. Increasing the Spatial Resolution

As mentioned before, the spatial resolution of HDXMS is limited. Multiple experimental techniques to increase the spatial resolution have been proposed: proteolytic digestion (with one acidic enzyme (Rosa & Richards, 1979; Englander, Rogero, & Englander, 1985; Wang, Pan, & Smith, 2002) or a mixture of enzymes with different specificities (Woods & Hamuro, 2001; Cravello, Lascoux, & Forest, 2003)), gas-phase fragmentation (Anderegg et al., 1994; Eyles et al., 1999), or a combination of these methods (Deng, Pan, & Smith, 1999; Kim et al., 2001; Kaltashov, Bobst, & Abzalimov, 2009; Rand et al., 2009). These experimental procedures increase the spatial resolution, ideally at the individual labile H-atom level. However, most often a large amount of fragments with a limited number of backbone amides are generated. If some of these short stretches are overlapping, one can apply computational methods to artificially increase the obtained resolution.

For overlapping fragments that share the start- or end-residue, the deuteration level of the non-overlapping segment can be obtained by subtracting the deuterium content of the shortest fragment from the longest fragment (Fig. 3). This subtraction-approach is implemented in *HD Desktop* (Pascal et al., 2009) and commonly applied.

An alternative approach, proposed by Abzalimov & Kaltashov (2006), determines the deuterium distribution of a non-overlapping segment of two overlapping fragments by a MEM-based deconvolution after deconvolution approach. There are two options. One method deconvolutes the extracted deuterium distributions of the two overlapping fragments:

$$\omega_{\text{AB},j} = \sum_{j=0}^{i}\omega_{A,j} \times \omega_{B,i-j}, \tag{25}$$

with

$$\text{DI}_{\text{AB},i} = \sum_{j=0}^{i}\omega_{\text{AB},j} \times I_{\text{AB},i-j},$$

while the other one extracts the deuterium distribution from the deconvoluted isotope distributions of the deuterated and

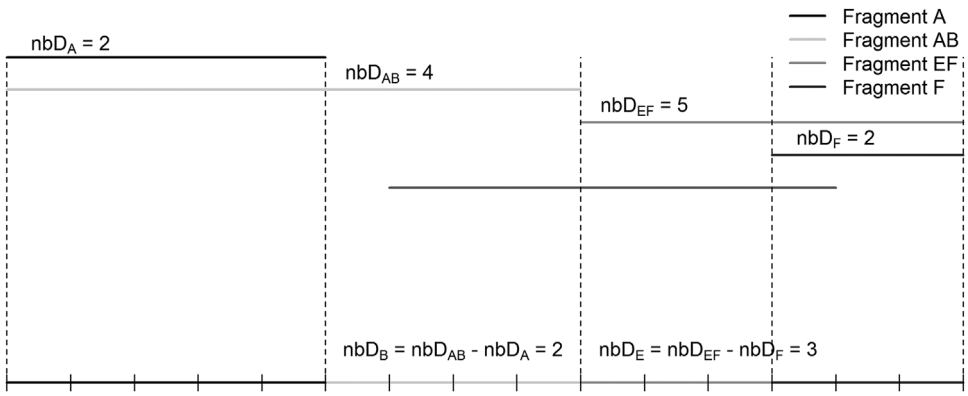

FIGURE 3. Illustration of the subtraction method to increase the spatial resolution. The deuterium contents of segments B and E are calculated based upon fragments A and AB, and fragments EF and F, respectively.

10982787, 2017, 5, Downloaded from https://analyticalsciencejournals.onlinelibrary.wiley.com/doi/10.1002/mas.21519 by Universiteit Hasselt, Wiley Online Library on [01/06/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

non-deuterated non-overlapping segment:

$$\mathrm{DI}_{B,i} = \sum_{j=0}^{i} \omega_{B,j} \times I_{B,i-j}, \tag{26}$$

with

$$\mathrm{DI}_{\mathrm{AB},i} = \sum_{j=0}^{i} \mathrm{DI}_{A,j} \times \mathrm{DI}_{B,i-j}$$

$$I_{\mathrm{AB},i} = \sum_{j=0}^{i} I_{A,j} \times I_{B,i-j}.$$

In the equations above, $I_{\mathrm{AB,i}}$ is the intensity of the undeuterated ith isotopic variant of the longest overlapping fragment (AB), $I_{A,i}$ is the intensity of the undeuterated ith isotopic variant of the shortest overlapping fragment (A), $I_{B,i}$ is the unobserved intensity of the undeuterated ith isotopic variant of the non-overlapping segment (B), ω_X the deuterium distribution of fragment X and $\mathrm{DI}_{X,i}$ the intensity of the deuterated ith isotopic variant of fragment X.

HDsite uses a theoretical-fitting approach to achieve the highest possible spatial resolution for the protein under study. Two modifications have been introduced to *AUTOHD* (Palmblad, Buijs, & Håkansson, 2001). Instead of using one exchange probability per peptide, the binomial exchange distribution for a peptide is calculated with as many exchange probabilities as there are amino acids in the peptide. This binomial exchange distribution is convoluted with the natural abundance distributions of the other elements. The resulting isotope distribution is compared against the observed isotope distribution. In contrast to *AUTOHD* this goodness-of-fit statistic combines the differences between the observed and calculated isotope distributions of all mutually overlapping peptides.

The subtractive method, the "MEM-based deconvolution after deconvolution" algorithm, and the modified theoretical-fitting approach implemented in *HDsite* (Kan et al., 2013) assume that the measured deuterium-content of a residue is independent from the fragments it is a part of. However, it has been shown that back-exchange can be influenced by the secondary structure (Rand et al., 2011a; Zhang, 2012b) and by interactions between the peptides and the LC column (Sheff, Rey, & Schriemer, 2013). As a consequence, the back-exchange rate can differ from fragment to fragment and different deuteration levels for a residue shared by several fragments can be observed (Sheff, Rey, & Schriemer, 2013). This could lead to questionable conclusions when multiple fragments, sharing one or more residues, are used to increase the spatial resolution.

Keppel & Weis (2015) proposed a residue averaging approach to increase the spatial resolution. Instead of the deuterium content of each peptide, the time required to reach 50% deuteration is used ($\bar{t}_{50\%}$). The residue-resolved time values ($t_{50\%}$) are calculated as the weighted average of the peptide 50% deuteration time:

$$t_{50\%,j} = \frac{\Sigma_{\mathrm{i}} E_{i,j} \times \bar{t}_{50\%} \times w_i}{\Sigma_i \omega_i}, \tag{27}$$

where i is a peptide, j a residue, E an exchangeability matrix. The exchangeability matrix indicates whether a residue is part of a peptide i and whether the residue is exchangeable or not. The latter is the case for proline residues, and for the first two N-terminal residues of peptide i. The weight for each peptide, w_i, is inversely proportional to the length of the peptide:

$$w_i = \left(\sum_i E_{i,j} \right)^{-2}.$$

VI. STEP 4: INTERPRETATION

Interpreting the level of deuteration of each annotated feature with respect to the protein structure and dynamics can be a challenging task. In order to facilitate this (final) step in the computational workflow (see Fig. 1), several auxiliary methods can be used, such as visualization and estimation of exchange rates and/or protection factors.

A. Visualization

Several tools exist to represent the deuteration level. *MSTools* (Kavan & Man, 2011) combines a number of scripts that can be used to visualize the detected deuteration level. In contrast to *MSTools*, most computational workflow tools have limited visualization options.

A popular representation is the deuterium incorporation plot (Fig. 4). The (relative) deuteration content of one extracted feature is plotted against the labeling time (for

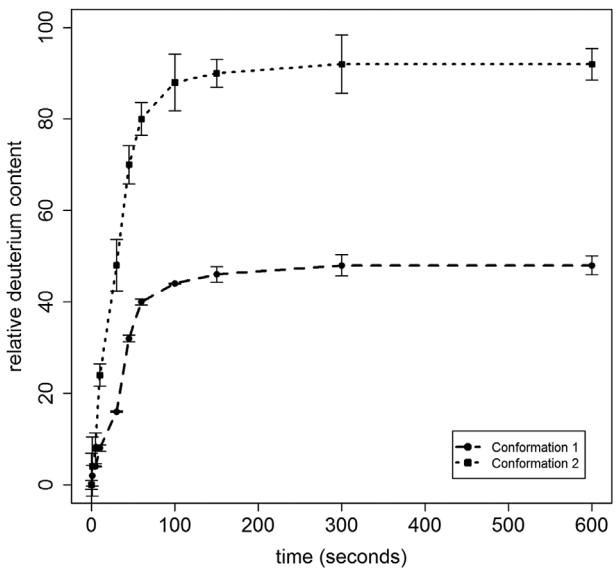

FIGURE 4. Deuterium incorporation plot of a fragment with two conformational states.

10982787, 2017, 5, Downloaded from https://analyticalsciencejournals.onlinelibrary.wiley.com/doi/10.1002/mas.21519 by Universiteit Hasselt, Wiley Online Library on [01/06/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

continuous labeling) or folding time (in case of pulsed labeling). The resulting deuteration curves can provide clues about the protein conformation (open or closed) and protein dynamics (flexible or rigid). When comparing two or more conformational states of the same protein, conformational differences can be made apparent by visualizations of this nature.

The deuterium incorporation plots do not contain information with respect to the (approximate) deuteration location in the global protein. Other types of plots provide this information. For example, the crystal structure of a protein, when available, can be combined with the level of deuteration of each fragments. The resulting three-dimensional plot is displayed in multiple colors according to the deuterium content at a given point in time. Based upon these 3D representations, conclusions about the protein conformation can be drawn. However, differences between the crystal structure and protein structure in solution or in gas-phase are possible.

A stacked barplot (Hamuro et al., 2003) or a heatmap (Kavan & Man, 2011) shows the deuterium levels of several fragments at multiple timepoints. Instead of plotting the deuterium content versus the timepoints, the changes in the deuterium level across time are illustrated with color gradients, and mapped against the amino acid sequence of the protein.

When comparing multiple conformational states, the figures already mentioned can be used. In addition to these plots, a mirror (butterfly) plot (Houde, Bergowitz, & Engen, 2011) or a difference plot can be created. The mirror plot depicts the deuterium content of all extracted features at once. The extracted features are summarized by unique numbers that indicate their order in the amino acid sequence of the protein. For each unique number, two deuteration levels per timepoint are available. These levels come, respectively, from a reference conformation and an experimental conformation. The deuteration content of the reference conformation is plotted on the positive end of the *y*-axis, while the data from the experimental conformation is plotted on the opposite end.

The difference plot has several versions (Chalmers et al., 2011a; Houde, Bergowitz, & Engen, 2011; Kavan & Man, 2011; Tiyanont et al., 2011; Street et al., 2012; Tsutsumi et al., 2012). All of them show the pairwise difference in deuterium content of each fragment for multiple states.

When interpreting the figures, the spatial resolution of the experiment should always be taken into account. HDX of a small number of H-atoms in a protein fragment is confined to a limited number of amino-acids, and it does not occur throughout the entire fragment (Morgan & Engen, 2009). One should also be careful not to overinterpret these plots, especially when multiple structures of the same protein are investigated. Because of these reasons, in our opinion, visualizing the deuterium content of a protein should not be considered as an endpoint, but merely as a starting point for the interpretation.

B. Statistical Analysis of Differential HDXMS

Three computational workflow tools, that is, *Hydra*, *HDX-Analyzer* (Liu et al., 2011), and *HDX Workbench* offer the possibility to statistically interpret HDX when comparing multiple protein structures based upon replicated experiments. In *Hydra* and *HDX Workbench*, Student's *t*-test is used to assess, at each timepoint, the differences in deuteration between two samples. When more than two conformations have to be compared, *HDX Workbench* uses a two-stage approach as proposed by Chalmers et al. (2007, 2011a). First, a single representative timepoint is selected. This timepoint should reflect the maximum deuteration difference between all samples. Second, the deuterium content is compared by using one-way ANOVA and Tukey's multiple comparison test to check if there are statistically significant differences and, if any, which conformations are differing.

In *HDX-Analyzer*, there are two options to test statistical significance of differences: a paired *t*-test and an ANCOVA model:

$$Y = \beta_t \times \text{time} + \beta_g \times \text{group} + \beta_{\text{tg}} \times (\text{time} \times \text{group}), \quad (28)$$

where Y denotes the m/z-values or deuterium-incorporation rate, "time" denotes the labeling or folding time point, and "group" is used as an indicator for the different protein structures.

Houde, Berkowitz, & Engen (2011) used a different approach to analyze replicate differential HDXMS experiments. For each fragment i, the difference in deuterium content between two conformations at timepoint t, $\Delta_{t,i}$, and the difference across multiple timepoints, $\Delta_s(i)$, are calculated:

$$\Delta_t(i) = D_{\text{ref},t}(i) - D_{\text{exp},t}(i), \quad (29)$$

$$\Delta_s(i) = \sum_t^T |\Delta_t(i)|, \quad (30)$$

where $D_{\text{ref},t}(i)$ and $D_{\text{exp},t}(i)$ denote the deuterium content of fragment i at timepoint t in, respectively, the reference and experimental protein conformation, and T is the number of time points.

The standard deviations of these differences are averaged across all fragments and subsequently used to calculate confidence limits. A fragment is considered to be different if the computed absolute values of $\Delta_t(i)$ and $\Delta_s(i)$ are greater than these confidence limits.

C. Estimating the Exchange Rates and/or Protection Extent

The deuteration incorporation plot (see Fig. 4) gives an intuitive idea about the kinetic exchange rate of a peptide, fragment, or protein. The exchange rates can be estimated based upon the back-exchange corrected deuterium content or the (partially) deuterated isotope distribution.

1. Exchange-Rate Estimation Based Upon Deuterium Content

In 1993, Zhang & Smith (1993) proposed a three-component model to estimate, based upon the deuteration level, kinetic exchange rates for three groups of hydrogens, that is, fast,

10982787, 2017, 5, Downloaded from https://analyticalsciencejournals.onlinelibrary.wiley.com/doi/10.1002/mas.21519 by Universiteit Hasselt, Wiley Online Library on [01/06/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

intermediate, and slow exchanging H-atoms:

$$D_t = N_1\left(1 - e^{-k_1 t}\right) + N_2\left(1 - e^{-k_2 t}\right) + N_3 k_3 t, \tag{31}$$

where N_x is the number of exchangeable hydrogens of group x and their respective exchange rates, k_x. These parameters are estimated using an OLS-optimization.

Smith, Deng, & Zhang (1997) extended this three-component model such that for each exchangeable H-atom a kinetic exchange rate can be determined:

$$D_t = N - \sum_{i=1}^{N} e^{-k_i t}, \tag{32}$$

where N is the total number of exchangeable H-atoms. Note that this extension can suffer from non-identifiability issues, that is, different combinations of k_i can give the same result. The optimization routine is also modified. Minimization of *OLS* is replaced by a *MEM*-algorithm (Zhang, Guan, & Marshall, 1997), similar to the method used to deconvolute the isotope distribution (see Section V.A.2).

The estimated exchange rates do not provide location-specific kinetic information, that is, which H-atom exchanges at which rate. Equations (31) and (32) require that there are more time points than unknown parameters, that is, five and N for Equations (31) and (32), respectively.

With *HDsite* it is possible to obtain location-specific exchange rates. As mentioned in Section V.C, the deuteration level of each labile hydrogen atom is retrieved with the help of mutually overlapping peptides.

Two computational methods have been proposed to increase the spatial resolution, based upon the deuterium content of overlapping fragments. The algorithm of Althaus et al. (2010) uses a combinatorial approach. This procedure determines, for each exchangeable H-atom, the membership to a distinct class of exchange rates, for instance, slow, intermediate, or fast. The determination is based upon the amino-acid sequence of the protein, the amino-acid sequence of the annotated fragments, and the number of labile hydrogens that are part of an exchange rate class. The numbers and the exchange rates are determined externally with the three-component model of Zhang & Smith (1993).

The method proposed by Fajer, Bou-Assaf, & Marshall (2012) estimates the exchange rate of amino-acid i based upon two equations:

$$D_t = N - \sum_{i=1}^{N} e^{-k_i t}, \tag{33}$$

$$D_t = \sum_{i=1}^{N} a_{ij}\left(1 - e^{-k_i t}\right), \tag{34}$$

where N is the total number of amino-acids with a labile hydrogen, and a_{ij} is an indicator for the presence of amino acid i in the jth fragment. If two or more amino-acids are not resolved, exchange rates are determined but cannot be assigned to specific labile H-atoms.

2. Exchange-Rate Estimation Based Upon the Isotope Distribution

Reuben et al. (2003) and Geller & Lifshitz (2004) proposed a method for gas-phase HDXMS that estimates the exchange rates directly from the isotope distribution. For each timepoint t, all exchangeable hydrogens, n_{labile}, are represented by $2^{n_{\text{labile}}}$ binary vectors, containing "1"s and "0"s, where "1" corresponds to exchanged, and "0" to not exchanged. The vectors represent all possible combinations of exchanging hydrogens in the analyzed protein or fragment, ranging from no exchange to all H-atoms are exchanged. The overall probability of each vector, $p_{\text{tot},t}$, is calculated as the product of individual exchange probabilities:

$$p(D_i)_t = 1 - e^{-k_i t}, \tag{35}$$

$$p(H_i)_t = 1 - p(D_i)_t = e^{-k_i t}, \tag{36}$$

where k_i denotes the exchange rate of labile hydrogen i. Binary vectors with the same number of exchanged hydrogens are combined, and their probabilities are summed. These summed overall probabilities are compared with the normalized intensities of the protein or fragment under study. The exchange rates of each exchangeable H-atom are estimated by minimizing the overall mean-square-deviance or the Kullback–Leibler divergence.

3. Protection-Extent Estimation

In case of an EX2 exchange regime, the protection factor of a labile hydrogen atom can be calculated based upon the (estimated) exchange rate of this H-atom:

$$P_f = \frac{k_{\text{int}}}{k_{\text{int}}}, \tag{37}$$

where k_{int} is the intrinsic exchange-rate and k_{ex} is the estimated exchange-rate. The intrinsic exchange-rate can be calculated by the formula described by Bai et al. (1993).

The workflow analysis tool proposed by Zhang (2012b) calculates the protection factor in a different way. First, the protection factor of a peptide internal standard, PPPI, is determined. The actual intrinsic exchange rate of each labile hydrogen is calculated by dividing the theoretical intrinsic exchange-rate by this protection factor. Together with the actual back-exchange times, $t^A_{\text{eff},j,r}$ and $t^B_{\text{eff},j,r}$, the protection factor of each exchangeable H-atom is estimated by minimizing:

$$\chi^2 = \frac{1}{n_{\text{meas}} - n_{\text{labile}} n_{\text{cond}}} \sum_j \sum_r \frac{\left(D_{\text{calc},j,r} - D_{\text{obs},j,r}\right)^2}{\sigma^2_{j,r}}, \tag{38}$$

where n_{meas} is the total number of deuterium measurements, n_{labile} is the total number of labile hydrogens, n_{cond} is the number

10982787, 2017, 5, Downloaded from https://analyticalsciencejournals.onlinelibrary.wiley.com/doi/10.1002/mas.21519 by Universiteit Hasselt, Wiley Online Library on [01/06/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

of experimental conditions or conformations, $D_{\text{obs},j,r}$ is the observed deuterium content for fragment j in run r, $\sigma^2_{j,r}$ is the variance of $D_{\text{obs},j,r}$. The expected deuterium content for fragment j in run r ($D_{\text{calc},j,r}$) is calculated as follows:

$$D_{\text{calc},j,r} = \sum_{i}^{n_{\text{labile}}} \left(1 - e^{-Kt_r}\right)\left(e^{-k_i^B t_{\text{eff},j,r}^B}\right)\left(e^{-k_i^A t_{\text{eff},j,r}^A}\right), \tag{39}$$

$$K = \frac{k_{\text{int},i}/P_f(PPPI)}{P_{f,i,c}}, \tag{40}$$

where $P_f(\text{PPPI})$ denotes the protection factor of PPPI, $k_{\text{int},i}$ is the theoretical intrinsic exchange rate of hydrogen i, and $P_{f,i,c}$ is the protection factor of hydrogen i in condition c.

Similarly to estimating the effective back-exchange times, two additional constraints to the minimization procedure are applied. The first constraint ensures that the protection factor of a hydrogen is similar to the factors of the neighboring hydrogens. The second constraint compels the protection factor of a hydrogen to be similar across different conditions. The estimation procedure of the individual protection factors, $P_{f,i,c}$, is repeated 4000 times and the average of the best 20 results is reported.

VII. CONCLUSION

Since the inception of HDXMS, the popularity of HDXMS applications to study protein conformations and dynamics grew almost exponentially as can been seen by the increase in number of published papers (Pirrone, Iacob, & Engen, 2015). This growth has been (partially) expedited by technological advancements. In the last decade, several methods have been proposed to process and analyze large amounts of mass-spectral HDXMS data. As shown in this review, many methods tackle the same issues in ways with no or a limited overlap. A consequence of this heterogeneity is that it is difficult to judge these methods with respect to their merits. Although a tremendous amount of effort has been put in the development of (semi)-automated computational methods, there are still outstanding issues and challenges. We list and briefly discuss some of these challenges.

A. Identification of Peptides and/or Fragmented Ions

Creation of a (peptide) master list is a commonly applied approach. However, features which are not in this list will not be analyzed, and useful information regarding protein structure or dynamics might be lost. Extending the databases used to identify features in the control run with, for instance, post-translational modifications can alleviate this issue.

B. Feature Extraction and Annotation

The occurrence of overlapping isotope distributions complicates feature extraction. There are two main causes for this issue: ions with small differences in *m/z*-values and ions with similar *m/z*-values but different exchange rates. Using (ultra)-high-resolution mass spectrometers such as FTICRs and orbitraps, increasing the chromatographic run times, or using ion mobility are potential technological solutions, but they can complicate the processing of the MS data. The application of mixture models and using alternate charge states have also been proposed (Abzalimov & Kaltashov, 2006; Kaltashov, Bobst, & Abzalimov, 2009), but mixture models are difficult to apply if it is unclear how many peptides are overlapping, and different charges could imply different protein conformations.

C. Determining the Deuteration Level of the Annotated Features

The occurrence of back-exchange is considered to be a major hurdle in the analysis of in-solution HDXMS experiments. Several approaches have been proposed to correct for the presence of back-exchange (see Section V.B.1). However, at this moment there is no commonly accepted method. Additionally, the proposed methods are not able to correct for back-exchange without simplifications or assumptions. It is possible that novel experimental technologies or improvements are required to further reduce or exclude the occurrence of back-exchange.

D. Interpretation

As mentioned by Liu et al. (2011), the vast majority of the effort to develop computational methods of HDXMS has been put in the processing of MS data. Substantially less effort has been put in the interpretation of the data, especially in the statistical analysis.

A very important but neglected aspect is the design of the HDXMS experiment. Commonly, an experiment is repeated a number of times. However, the number of independent runs is often chosen arbitrarily, irrespectively of the aim of the experiment and of the (expected) variability in the measurements. Moreover, when an experiment, with replicate runs, is carried out over multiple days, randomization, stratification, or blocking is required to avoid unwanted, complicating factors. Additionally, the number and the values of time points (e.g., after 5 sec, 30 sec, 1 min, etc.), are often chosen without any clear motivation. As a consequence, the information content of the HDXMS experiment can be too low. Carefully designing the experiment will be beneficial for the information content, and for the interpretation of the obtained data.

Another aspect is the uncertainty associated with the estimated parameters. The standard errors of the estimated deuterium-content, kinetic exchange-rates, and/or protection factors are rarely reported. Considering the estimated values with their uncertainty is needed, especially when comparing two or more conformational states. Moreover, when determining the exchange rates and/or protection extent based upon the estimated deuterium-content, the standard errors of the deuteration level should also be taken into account.

Lastly, when comparing different conformational states of a protein, the probability that one or more detected differences are false increases with the number of executed tests. In order to control this chance of false-positive findings, multiple-testing procedures (Dudoit & van der Laan, 2008; Benjamini, 2010) should be applied.

10982787, 2017, 5, Downloaded from https://analyticalsciencejournals.onlinelibrary.wiley.com/doi/10.1002/mas.21519 by Universiteit Hasselt, Wiley Online Library on [01/06/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

REFERENCES

Abzalimov RR, Kaltashov IA. 2006. Extraction of local hydrogen exchange data from HDX CAD MS measurements by deconvolution of isotopic distributions of fragment ions. J Am Soc Mass Spectrom 17:1543–1551.

Althaus E, Canzar S, Ehrler C, Emmett MR, Karrenbacher A, Marshall AG, Meyer-Bäse A, Tipton JD, Zhang HM. 2010. Computing H/D exchange rates of single residues from data of proteolytic fragments. BMC Bioinformatics 11:424.

Anderegg RJ, Wagner DS, Stevenson CL, Borchardt RT. 1994. The mass spectrometry of helical unfolding in peptides. J Am Soc Mass Spectrom 5:425–433.

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

Bax A. 1994. Multidimensional nuclear magnetic resonance methods for protein studies. Curr Opin Struct Biol 4:738–744.

Benjamini Y. 2010. Simultaneous and selective inference: Current successes and future challenges. Biom J 52:708–721.

Burkitt W, O'Connor G. 2008. Assessment of the repeatability and reproducibility of hydrogen/deuterium exchange mass spectrometry measurements. Rapid Commun Mass Spectrom 22:3893–3901.

Campbell S, Rodgers MT, Marzluff EM, Beauchamp JL. 1995. Deuterium exchange reactions as a probe of biomolecule structure. Fundamental studies of gas phase H/D exchange reaction of protonated Glycine oligomers with D_2O, CD_3OD, CD_3CO_2D, and ND_3. J Am Chem Soc 117:12840–12854.

Chalmers MJ, Busby SA, Pascal BD, He Y, Hendrickson CL, Marshall AG, Griffin PR. 2006. Probing protein ligand interactions by automated hydrogen/deuterium exchange spectrometry. Anal Chem 78:1005–1014.

Chalmers MJ, Busby SA, Pascal BD, Southern MR, Griffin PR. 2007. A two-stage differential hydrogen deuterium exchange method for the rapid characterization of protein/ligand interactions. J Biomol Tech 18: 194–204.

Chalmers MJ, Busby SA, Pascal BD, West GM, Griffin PR. 2011a. Differential hydrogen/deuterium exchange mass spectrometry analysis of protein-ligand interactions. Expert Rev Proteomics 8:43–59.

Chalmers MJ, Pascal BD, Willis S, Zhang J, Iturria SJ, Dodge JA, Griffin PR. 2011b. Methods for the analysis of high precision differential hydrogen deuterium exchange data. Int J Mass Spectrom 302:59–68.

Chik JK, Vande Graaf JL, Schriemer DC. 2006. Quantitating the statistical distribution of deuterium incorporation to extend the utility of H/D exchange MS data. Anal Chem 78:207–214.

Connelly GP, Bai Y, Jeng MF, Englander SW. 1993. Isotope effects in peptide group hydrogen exchange. Proteins 17:87–92.

Cravello L, Lascoux D, Forest E. 2003. Use of different proteases working in acidic conditions to improve sequence coverage and resolution in hydrogen/deuterium exchange of large proteins. Rapid Commun Mass Spectrom 17:2387–2393.

Deng Y, Pan H, Smith DL. 1999. Selective isotope labeling demonstrates that hydrogen exchange at individual peptide amide linkages can be determined by collision induced dissociation mass spectrometry. J Am Chem Soc 121:1966–1967.

Dudoit S, van der Laan MJ. 2008. Multiple testing procedures with applications to genomics. New York: Springer Series in Statistics.

Eng JK, McCormack AL, Yates JR. 1994. An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. J Am Soc Mass Spectrom 5:976–989.

Engen JR. 2009. Analysis of protein conformation and dynamics by hydrogen/deuterium exchange MS. Anal Chem 81:7870–7875.

Englander JJ, Rogero JR, Englander SW. 1985. Protein hydrogen exchange studied by the fragment separation method. Anal Biochem 147: 234–244.

Englander SW. 2006. Hydrogen exchange and mass spectrometry: A historical perspective. J Am Soc Mass Spectrom 17:1481–1489.

Englander SW, Mayne L, Bai Y, Sosnick TR. 1997. Hydrogen exchange: The modern legacy of Linderstrfim-Lang. Protein Sci 6:1101–1109.

Englander SW, Sosnick TR, Englander JJ, Mayne L. 1996. Mechanisms and uses of hydrogen exchange. Curr Opin Struct Biol 6:18–23.

Ernst RR, Bodenhausen G, Wokaun A. 1988. Principles of nuclear magnetic resonance in one and two dimensions. Oxford: Clarendon Press.

Eyles SJ, Dresch T, Gierasch LM, Kaltashov IA. 1999. Unfolding dynamics of a beta-sheet protein studied by mass spectrometry. J Mass Spectrom 34:1289–1295.

Fajer PG, Bou-Assaf GM, Marshall AG. 2012. Improved sequence resolution by global analysis of overlapped peptides in hydrogen/deuterium exchange mass spectrometry. J Am Soc Mass Spectrom 23:1202–1208.

Freitas MA, Hendrickson CL, Emmett MR, Marshall AG. 1999. Gas-phase bovine ubiquitin cation conformations resolved by gas-phase hydrogen/deuterium exchange rate and extent. Int J Mass Spectrom 185:565–575.

Geller O, Lifshitz C. 2004. Applying a new algorithm to H/D exchange of multiply protonated cytochrome c. Int J Mass Spectrom 223:125–129.

Geller O, Lifshitz C. 2005. A fast ow tube study of gas phase H/D exchange of multiply protonated ubiquitin. J Phys Chem A 109:2217–2222.

Guttman M, Weis DD, Engen JR, Lee KK. 2013. Analysis of overlapped and noisy hydrogen/deuterium exchange mass spectra. J Am Soc Mass Spectrom 2013:1906–1912.

Hamuro Y, Coales SJ, Molnar KS, Tuske SJ, Morrow JA. 2008. Specificity of immobilized porcine pepsin in H/D exchange compatible conditions. Rapid Commun Mass Spectrom 22:1041–1046.

Hamuro Y, Coales SJ, Morrow JA, Molnar KS, Tuske SJ, Southern MR, Griffin PR. 2006. Hydrogen/Deuterium-exchange (H/D-Ex) of PPAR LBD in the presence of various modulators. Protein Sci 15:1883–1892.

Hamuro Y, Coales SJ, Southern MR, Nemeth-Cawley JE, Stranz DD, Griffin PR. 2003. Rapid analysis of protein structure and dynamics by hydrogen/deuterium exchange mass spectrometry. J Biomol Tech 14:171–182.

Hoofnagle AN, Resing KA, Ahn NG. 2003. Protein analysis by hydrogen exchange mass spectrometry. Annu Rev Biophys Biomol Struct 32:1–25.

Horn DM, Zubarev RA, McLafferty FW. 2000. Automated reduction and interpretation of high resolution electrospray mass spectra of large molecules. J Am Soc Mass Spectrom 11:320–332.

Horn JR, Kraybill B, Petro EJ, Coales SJ, Morrow JA, Hamuro Y, Kossiakoff AA. 2006. The role of protein dynamics in increasing binding affinity for an engineered protein-protein interaction established by H/D exchange mass spectrometry. Biochemistry 45:8488–8498.

Hotchko M, Anand GS, Komives EA, Ten Eyck LF. 2006. Automated extraction of backbone deuteration levels from amide 1H/2H mass spectrometry experiments. Protein Sci 15:583–601.

Houde D, Berkowitz SA, Engen JR. 2011. The utility of hydrogen/deuterium exchange mass spectrometry in biopharmaceutical comparability. J Pharm Sci 100:2071–2086.

Hsu YH, Johnson DA, Traugh JA. 2008. Analysis of conformational changes during activation of protein kinase Pak2 by amide hydrogen/deuterium exchange. J Biol Chem 52:36397–36405.

Hvidt A. 1964. A discussion of the pH dependence of the hydrogen-deuterium exchange of proteins. C R Trav Lab Carlsberg 34:299–317.

Hvidt A, Nielsen SO. 1966. Hydrogen exchange in proteins. Adv Protein Chem 21:287–386.

Iacob RE, Engen JR. 2012. Hydrogen exchange mass spectrometry: Are we out of the quicksand? J Am Soc Mass Spectrom 23:1003–1010.

Jaswal SS. 2013. Biological insights from hydrogen exchange mass spectrometry. Biochim Biophys Acta 1834:1188–1201.

Kaltashov IA, Bobst CE, Abzalimov RR. 2009. H/D exchange and mass spectrometry in the studies of protein conformation and dynamics: Is there a need for a top-down approach? Anal Chem 81:7892–7899.

Kaltashov IA, Eyles SJ. 2005. Mass spectrometry in biophysics: Conformation and dynamics of biomolecules. Hoboken, NJ: John Wiley.

10982787, 2017, 5, Downloaded from https://analyticalsciencejournals.onlinelibrary.wiley.com/doi/10.1002/mas.21519 by Universiteit Hasselt, Wiley Online Library on [01/06/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Kan ZY, Mayne L, Chetty PS, Englander SW. 2011. ExMS: Data analysis for HX-MS experiments. J Am Soc Mass Spectrom 22:1906–1915.

Kan ZY, Walters BT, Mayne L, Englander SW. 2013. Protein hydrogen exchange at residue resolution by proteolytic fragmentation mass spectrometry analysis. Proc Natl Acad Sci USA 110:16438–16443.

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

Kavan D, Man P. 2011. MSTools—Web based application for visualization and presentation of HXMS data. Int J Mass Spectrom 302: 53–58.

Kazazic S, Zhang HM, Schaub TM, Emmett MR, Hendrickson CL, Blakney GT, Marshall AG. 2010. 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.

Keppel TR, Weis DD. 2015. Mapping residual structure in intrinsically disordered proteins a residue resolution using millisecond Hydrogen/Deuterium exchange and residue averaging. J Am Soc Mass Spectrom 26:547–554.

Khakinejad M, Kondalaji SG, Maleki H, Arndt JR, Donohoe GC, Valentine SJ. 2014. Combining ion mobility spectrometry with hydrogen-deuterium exchange and top-down MS for peptide ion structure analysis. J Am Soc Mass Spectrom 25:2103–2115.

Kim MY, Maier CS, Reed DJ, Deinzer ML. 2001. Site-specific amide hydrogen/deuterium exchange in *E. coli* thioredoxins measured by electrospray ionization mass spectrometry. J Am Chem Soc 123:9860–9866.

Konermann L, Pan J, Liu YH. 2011. Hydrogen exchange mass spectrometry for studying protein structure and dynamics. Chem Soc Rev 40: 1224–1234.

Konermann L, Simmons DA. 2003. Protein-folding kinetics and mechanisms studied by pulse-labeling and mass spectrometry. Mass Spectrom Rev 22:1–26.

Konermann L, Tong X, Pan Y. 2008. Protein structure and dynamics studied by mass spectrometry: H/D exchange, hydroxyl radical labeling, and related approaches. J Mass Spectrom 43:1021–1036.

Kossiakoff AA. 1982. Protein dynamics investigated by the neutron diffraction-hydrogen exchange technique. Nature 296:713–721.

Kreshuk A, Stankiewicz M, Lou X, Kirchner M, Hamprecht FA, Mayer MP. 2011. Automated detection and analysis of bimodal isotope peak distributions in H/D exchange mass spectrometry using HeXicon. Int J Mass Spectrom 302:125–131.

Linderstrøm-Lang KU. 1958. Deuterium exchange and protein structure. In symposium on protein structures. 23–24.

Lindner R, Lou X, Reinstein J, Shoeman RL, Hamprecht FA, Winkler A. 2014. Hexicon 2: Automated processing of hydrogen-deuterium exchange mass spectrometry data with improved deuteration distribution estimation. J Am Soc Mass Spectrom 25:1018–1028.

Liu Q, Easterling ML, Agar JN. 2014. Resolving isotopic fine structure to detect and quantify natural abundance- and hydrogen/deuterium exchange-derived isotopomers. Anal Chem 86:820–825.

Liu S, Liu L, Uzuner L, Zhou X, Gu M, Shi W, Zhang Y, Dai SY, Yuan JS. 2011. HDX-analyzer: A novel package for statistical analysis of protein structure dynamics. BMC Bioinformatics 12:S43.

Liu T, Pantazatos D, Li S, Hamuro Y, Hilser VJ, Woods VL. 2012. Quantitative assessment of protein structural models by comparison of H/D exchange MS data with exchange behavior accurately predicted by DCOREX. J Am Soc Mass Spectrom 23:43–56.

Lou X, Kirchner M, Renard BY, Köthe U, Boppel S, Graf C, Lee CT, Steen JAJ, Steen H, Mayer MP, Hamprecht FA. 2010. Deuteration distribution estimation with improved sequence coverage for HX/MS experiments. Bioinformatics 26:1535–1541.

Marcsisin SR, Engen JR. 2010. Hydrogen exchange mass spectrometry: What is it and what can it tell us? Anal Bioanal Chem 397:967–972.

Mason SA, Bentley GA, McIntyre GJ. 1983. Neutrons in Biology, vol. 27. Springer US, 1984, ch. Deuterium exchange in lysozyme at 1.4A resolution. pp. 323–334.

Miller DE, Prasannan CB, Villar MT, Fenton AW, Artigues A. 2012. HDXFinder: Automated analysis and data reporting of deuterium/hydrogen exchange mass spectrometry. J Am Soc Mass Spectrom 23:425–429.

Morgan CR, Engen JR. 2009. Investigating solution-phase protein structure and dynamics by hydrogen exchange mass spectrometry. Curr Protoc Protein Sci 58:17.6.117.6.17.

Morháč M. 2006. Deconvolution methods and their applications in the analysis of gamma-ray spectra. nuclear instruments and methods in physics research section a accelerators spectrometers detectors and associated equipment. Nucl Instrum Methods Phys Res Sect A 559:119–123.

Palmblad M, Buijs J, Håkansson P. 2001. Automatic analysis of Hydrogen/Deuterium exchange mass spectra of peptides and proteins using calculations of isotopic distributions. J Am Soc Mass Spectrom 12:1153–1162.

Pascal BD, Chalmers MJ, Busby SA, Griffin PR. 2009. HD Desktop: An integrated platform for the analysis and visualization of H/D exchange data. J Am Soc Mass Spectrom 20:601–610.

Pascal BD, Chalmers MJ, Busby SA, Mader CC, Southern MR, Tsinoremas NF, Griffin PR. 2007. The Deuterator: Software for the determination of backbone amide deuterium levels from H/D exchange MS data. BMC Bioinformatics 8:156.

Pascal BD, Willis S, Lauer JL, Landgraf RR, West GM, Marciano D, Novick S, Goswami D, Chalmers MJ, Griffin PR. 2012. HDX Workbench: Software for the analysis of H/D Exchange MS data. J Am Soc Mass Spectrom 23:1512–1521.

Percy AJ, Rey M, Burns KM, Schriemer DC. 2012. Probing protein interactions with hydrogen/deuterium exchange and mass spectrometry—A review. Anal Chim Acta 721:7–21.

Perkins DN, Pappin DJ, Creasy DM, Cottrell JS. 1999. Probability-based protein identification by searching sequence databases using mass spectrometry data. Electrophoresis 20:3551–3567.

Pirrone GF, Iacob RE, Engen JR. 2015. Applications of hydrogen/deuterium exchange MS from 2012 to 2014. Anal Chem 87:99–118.

Rand KD, Lund FW, Amon S, Jørgensen TJD. 2011a. Investigation of amide hydrogen back-exchange in Asp and His repeats measured by hydrogen (1H-2H) exchange mass spectrometry. Int J Mass Spectrom 302: 110–115.

Rand KD, Pringle SD, Morris M, Brown JM. 2012. Site-specific analysis of gas-phase Hydrogen/Deuterium exchange peptides and proteins by electron transfer dissociation. Anal Chem 84:1931–1940.

Rand KD, Pringle SD, Morris M, Engen JR, Brown JM. 2011b. ETD in a traveling wave ion guide at tuned z-spray ion source conditions allows for site-specific hydrogen/deuterium exchange measurements. J Am Soc Mass Spectrom 22:1784–1793.

Rand KD, Zehl M, Jensen ON, Jørgensen TJD. 2009. Protein hydrogen exchange measured at single-residue resolution by electron transfer dissociation mass spectrometry. Anal Chem 81:5577–5584.

Renard BY, Kirchner M, Steen H, Steen JAJ, Hamprecht FA. 2008. NITPICK: Peak identification for mass spectrometry data. BMC Bioinformatics 9:355.

Reuben BG, Ritov Y, Geller O, McFarland MA, Marshall AG, Lifshitz C. 2003. Applying a new algorithm for obtaining site specific rate constants for H/D exchange of the gas phase proton-bound arginine dimer. Chem Phys Lett 380:88–94.

Rey M, Sarpe V, Burns KM, Buse J, Baker CAH, van Dijk M, Wordeman L, Bonvin AMJJ, Schriemer DC. 2014. Mass spec studio for integrative structural biology. Structure 22:1538–1548.

Rockwood AL, Haimi P. 2006. Efficient calculation of accurate masses of isotopic peaks. J Am Soc Mass Spectrom 17:415–419.

Rosa JJ, Richards FM. 1979. An experimental procedure for increasing the structural resolution of chemical hydrogen-exchange measurements on

10982787, 2017, 5, Downloaded from https://analyticalsciencejournals.onlinelibrary.wiley.com/doi/10.1002/mas.21519 by Universiteit Hasselt, Wiley Online Library on [01/06/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

proteins: Application to ribonuclease S peptide. J Mol Biol 133:399–416.

Salisbury JP, Liu Q, Agar JN. 2014. QUDeX-MS: Hydrogen/deuterium exchange calculation for mass spectra with resolved isotopic fine structure. BMC Bioinformatics 15:403.

Schaaff TG, Stephenson JL, McLuckey SA. 2000. Gas phase H/D exchange kinetics: DI versus D2O. J Am Soc Mass Spectrom 11:167–171.

Senko MW, Beu SC, McLafferty FW. 1995. Determination of monoisotopic masses and ion populations for large biomolecules from resolved isotopic distribution. J Am Soc Mass Spectrom 6:229–233.

Sheff JG, Rey M, Schriemer DC. 2013. Peptide-column interactions and their influence on back exchange rates in hydrogen/deuterium exchange-MS. J Am Soc Mass Spectrom 24:1006–1015.

Sheff JG, Schriemer DC. 2014. Toward standardizing deuterium content reporting in hydrogen exchange-MS. Anal Chem 86:11962–11965.

Slysz GW, Baker CAH, Bozsa BM, Dang A, Percy AJ, Bennett M, Schriemer DC. 2009. Hydra: Software for tailored processing of H/D exchange data from MS or tandem MS analyses. BMC Bioinformatics 10:162.

Smith DL, Deng Y, Zhang Z. 1997. Probing the non-covalent structure of proteins by amide hydrogen exchange and mass spectrometry. J Mass Spectrom 32:135–146.

Street TO, Lavery LA, Verba K, Lee CT, Mayer MP, Agard DA. 2012. Cross monomer substrate contacts reposition the Hsp90 N-terminal domain and prime the chaperone activity. J Mol Biol 415:3–15.

Suckau D, Shi Y, Beu SC, Senko MW, Quinn JP, Wampler FM, McLafferty FW. 1993. Coexisting stable conformations of gaseous protein ions. Proc Natl Acad Sci USA 90:790–793.

Tiyanont K, Wales TE, Aste-Amezaga M, Aster JC, Engen JR, Blacklow SC. 2011. Evidence for increased exposure of the Notch1 metalloprotease cleavage site upon conversion to an activated conformation. Structure 19:546–554.

Tsutsumi S, Mollapour M, Prodromou C, Lee CT, Panaretou B, Yoshida S, Mayer MP, Neckers LM. 2012. Charged linker sequence modulates eukaryotic heat shock protein 90 Hsp90 chaperone activity. Proc Natl Acad Sci USA 109:2937–2942.

Venable JD, Scuba W, Brock A. 2013. Feature based retention time alignment for improved HDXMS analysis. J Am Soc Mass Spectrom 24:642–645.

Wagner G, Wuthrich K. 1982. Amide proton exchange and surface conformation of BPTI in solution: Studies with 2D NMR. J Mol Biol 160:343–361.

Wales TE, Eggertson MJ, Engen JR. 2013. Considerations in the analysis of hydrogen exchange mass spectrometry data. Methods Mol Biol 1007:263–288.

Wales TE, Engen JR. 2006. Hydrogen exchange mass spectrometry for the analysis of protein dynamics. Mass Spectrom Rev 25:158–170.

Wang L, Pan H, Smith DL. 2002. Hydrogen exchange-mass spectrometry: Optimization of digestion conditions. Mol Cell Proteomics 1:132–138.

Wei H, Ahn J, Yu YQ, Tymiak A, Engen JR, Chen G. 2012. Using hydrogen/deuterium exchange mass spectrometry to study conformational changes in granulocyte colony stimulating factor upon PEGylation. J Am Soc Mass Spectrom 23:498–504.

Weis DD, Engen JR, Kass IJ. 2006. Semi-automated data processing of hydrogen exchange mass spectra using HX-Express. J Am Soc Mass Spectrom 17:1700–1703.

Weis DD, Wales TE, Engen JR. 2006. Identification and characterization of EX1 kinetics in H/D exchange mass spectrometry by peak width analysis. J Am Soc Mass Spectrom 17:1498–1509.

Winger BE, Light-Wahl KJ, Rockwood AL, Smith RD. 1992. Probing qualitative conformation differences of multiply protonated gas-phase proteins *via* H/D isotopic exchange with D2O. J Am Chem Soc 114:5897–5898.

Winkler HDF, Dzyuba EV, Schalley CA. 2011. Gas-phase H/D-exchange experiments in supramolecular chemistry. New J Chem 35:529–541.

Wood TD, Chorush RA, Wampler FM, Little DP, O'Connor PB, McLafferty FW. 1995. Gas-phase folding and unfolding of cytochrome c cations. Proc Natl Acad Sci USA 92:2451–2454.

Woods VL, Hamuro Y. 2001. High resolution, high-throughput amide deuterium exchange mass spectrometry (DXMS) determination of protein binding site structure and dynamics: Utility in pharmaceutical design. J Cell Biochem S37:89–98.

Wyttenbach T, Bowers MT. 1999. Gas phase conformations of biological molecules: The hydrogen/deuterium exchange mechanism. J Am Soc Mass Spectrom 10:9–14.

Zhang J, Ramachadran P, Kumar R, Gross ML. 2013. H/D exchange centroid monitoring is insufficient to show differences in the behavior of protein states. J Am Soc Mass Spectrom 24:450453.

Zhang Z. 2004. Prediction of low-energy collision-induced dissociation spectra of peptides. Anal Chem 76:3908–3922.

Zhang Z. 2005. Prediction of low-energy collision-induced dissociation spectra of peptides with three or more charges. Anal Chem 77:6364–6373.

Zhang Z. 2009. Large-scale identification and quantification of covalent modifications in therapeutic proteins. Anal Chem 81:8354–8364.

Zhang Z. 2010. Prediction of electron-transfer/capture dissociation spectra of peptides. Anal Chem 82:1990–2005.

Zhang Z. 2011. Prediction of collision-induced-dissociation spectra of peptides with post-translational or process-induced modifications. Anal Chem 83:8642–8651.

Zhang Z. 2012a. Improved protein Hydrogen/Deuterium exchange mass spectrometry platform with fully automated data processing. Anal Chem 84:4942–4949.

Zhang Z. 2012b. Retention time alignment of LC/MS data by a divide-and-conquer algorithm. J Am Soc Mass Spectrom 23:764–772.

Zhang Z, Guan S, Marshall AG. 1997. Enhancement of the effective resolution of mass spectra of high-mass biomolecules by maximum-entropy based deconvolution to eliminate the isotopic natural abundance distribution. J Am Soc Mass Spectrom 8:659–670.

Zhang Z, McElvain JS. 1999. Optimizing spectroscopic signal-to-noise ratio in analysis of data collected by a chromatographic/spectroscopic system. Anal Chem 71:39–45.

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

10982787, 2017, 5, Downloaded from https://analyticalsciencejournals.onlinelibrary.wiley.com/doi/10.1002/mas.21519 by Universiteit Hasselt, Wiley Online Library on [01/06/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License