HDf1eX: Software for flexible high structural resolution of hydrogen/deuterium-exchange mass spectrometry data

Neeleema Seetaloo¹, Monika Kish¹, Jonathan J. Phillips^{1,2*}

¹Living Systems Institute, Department of Biosciences, University of Exeter, Stocker Road, Exeter, UK

²Alan Turing Institute, British Library, London, UK

*Correspondence to: jj.phillips@exeter.ac.uk

ABSTRACT Hydrogen/deuterium-exchange mass spectrometry (HDX-MS) experiments on protein structures can be performed at three levels: (1) by enzymatically digesting labelled proteins and analyzing the peptides (bottom-up), (2) by further fragmenting peptides following digestion (middle-down), and (3) by fragmenting the intact labelled protein (top-down), using soft gas-phase fragmentation methods, such as electron transfer dissociation (ETD). However, to the best of our knowledge, the software packages currently available for the analysis of HDX-MS data do not enable the peptide- and ETD-levels to be combined – they can only be analyzed separately. Thus, we developed HDf1eX – a standalone application for the analysis of flexible high structural resolution of HDX-MS data, which allows data at any level of structural resolution (intact protein, peptide, fragment) to be merged. HDf1eX features rapid experimental data fitting, robust statistical significance analyses and optional methods for theoretical intrinsic calculations and a novel empirical correction for comparison between solution conditions.

Hydrogen/deuterium-exchange mass spectrometry (HDX-MS) is a widely-used analytical technique that enables the probing of protein conformational dynamics and interactions on a wide timescale, from milliseconds to hours, by monitoring the isotopic exchange at the backbone amide hydrogens with the surrounding solvent^{1,2}. HDX-MS is growing in popularity as an orthogonal structural biology tool since it can analyze proteins that are not amenable to other techniques, such as antibodies^{3,4} and intrinsically disordered proteins^{5,6}. In order to localize the exchange kinetics to specific regions of a protein, it is crucial to obtain sub-molecular hydrogen/deuterium-exchange data. This is currently a major limitation in the mass spectrometry detection of hydrogen/deuterium-exchange and, until recently, the structural resolution of any bottom-up HDX-MS experiment relied on the number and position of proteolytic peptides identified^{5,7}. By spectral assignment of the digested peptides onto the protein sequence, the changes in deuterium uptake of each peptide can be mapped back to the primary sequence, thus giving insights into the behavior of localized regions of the protein.

Much recent interest has been directed to achieve higher structural resolution by fragmenting the digested peptides further using gas-phase soft fragmentation techniques – termed middle-down HDX-MS^{8,9}. However, it is important that the fragmentation does not induce any hydrogen/deuterium scrambling (H/D scrambling)^{10,11}. Methods shown to be amenable to this include electron-based dissociation methods (ExD) with HDX-MS, such as electron-transfer dissociation (ETD)^{3,12,13}, electron-capture dissociation (ECD)^{11,14}, and ultra-violet photodissociation (UVPD)^{15,16}. These forms of fragmentation can be tuned to have analyte energy profiles that do not promote proton

mobility and so result in low H/D scrambling under the right conditions¹⁷. This has made HDX-MS possible in top-down (fragmentation of entire proteins into smaller fragments)¹³ and middle-down (fragmentation of digested peptides into even smaller fragments)⁸ fashions, which each provide varying levels of structural resolution.

Nevertheless, presently, each of these experiments and the resulting global and local HDX-MS data can only be analyzed and interpreted separately since there is currently no widely available process that allows the combination of data at all levels^{18–21}. This is especially confounding, as soft-fragmentation reactions are typically inefficient and yield sparse data which might usefully be complementary to conventional bottom-up data. In an effort to remedy this, we hereby present HDf1eX: Software for flexible structural resolution of HDX-MS data, which allows combination of data from bottom-up, middle-down, and top-down experiments. HDf1eX can (1) Perform non-linear regression fitting of deuterium uptake data against labelling time for each peptide or fragment, (2) Correct for back-exchange using internal or external measures of maximal deuteration, (3) Normalize between different pH and salt conditions, (4) Calculate intrinsic rate and protection factor (P_f), (5) Combine peptide and ExD fragments, and visualize the absolute deuterium uptake as histograms for individual protein states, (6) Perform hybrid significance testing using various methods of global significance testing and doing a Welch's t-test, and (7) Visualize significant differences on difference and volcano plots.

Figure 1: Flowchart summarizing the processes performed by HDf1eX.

METHODS

HDf1eX was written in MATLAB release 2021b (MathWorks, USA) and made into a graphical user interface with MATLAB App Designer. The software performs a series of steps summarized in Figure 1 and detailed below. An in-depth flowchart of the process can also be found in SI Figures 1-3.

Data import and preparation. Prior to HDf1eX analysis, an input file containing all the information about the HDX-MS experiment needs to be created. Following review of isotopic assignments, a comma-separated variable (.csv) file is generated, which is the first user-input to HDf1eX. The file must be comma-separated and have the nine columns in the order: protein name, sequence start number, sequence end number, sequence, modification, fragment, maximum possible uptake, mass of monoisotopic species, state name, exposure time, file name, charge, retention time, intensity, centroid. After HDf1eX parses the input file, the user selects the protein and states to plot, type of back-exchange correction to perform (if at all), number of replicates collected, number of phases into which to fit the uptake curves, pH of reference state, pH of each experimental state, time format between milliseconds (ms), seconds (s) and minutes (min), and the time window over which the fitted uptake curve will be evaluated. In cases where the replicates are not named trivially (e.g., file_01, file_02, etc.), the user can manually assign files to the appropriate states and replicates. Charge deconvolution and deuterium uptake per peptide and ETD fragment are determined as described in Wales et al²².

Uptake curve fitting and area under the uptake curve. Fitting the HDX-MS data from equilibrium experiments can provide crucial information about the exchange kinetics. We fit the data points for each peptide and ETD fragment either by interpolation or fitting to a stretched exponential. The deuterium uptake at labelling time t , is fitted to a stretched exponential given by Equation 1²³, where $nExp$ is a user-defined number of exponential

phases, N is the maximum number of labile hydrogens, k_{obs} is the observed exchange rate constant and β is a stretching factor.

$$D_t = \sum_{i=1}^{nExp} N \left[1 - e^{-(k_{obs}t)^\beta} \right] \quad \text{Equation 1}$$

The non-linear fitting is performed using the fit function from the Curve Fitting Toolbox in MATLAB. If $nExp = 1$, N is fixed at the maximum number of labile hydrogens for that particular peptide/fragment if back-exchange correction has been applied. Otherwise, only an upper bound of N is applied. k_{obs} and β are both limited between 0 and 1. The resulting fitted curve equation is subsequently evaluated and integrated over the user-defined time window to give the area under the uptake curve, HDX_{Area} . In the case of data that reach a plateau (i.e., maximal deuteration achieved at later timepoints), a further variable called X_{plat} (SI Figure 4) is calculated, which is used as the upper limit of the time window of integration. This prevents significant differences from being minimized when overly wide time windows have been used for any individual peptide.

There is also an option to fit the data points using an interpolation method. Here, we used the Piecewise cubic Hermite interpolation (PCHIP from MATLAB's Curve Fitting Toolbox). However, this method does not provide a k_{obs} and β , and thus cannot be used for empirical adjustments, observed exchange rate constant and protection factor analyses.

The entire HDf1eX workflow following uptake curve fitting, such as data flattening and statistical significance analysis, can be performed with uptake at a time t , uptake area up to a time t , the observed exchange rate constant or protection factors.

Back-exchange correction. Despite efforts to minimize back-exchange in HDX-MS experiments²⁴⁻²⁶, there is always some percentage of the on-exchanged deuterium that exchanges back to hydrogen. For some types of

Figure 2: Calibrating the chemical exchange rate using bradykinin. (A) Bradykinin uptake curves in four states, back-exchange corrected and fitted in one exponential phase according to equation 1; (B) pH adjustment applied to the timepoints of bradykinin in each state prior to fitting; (C) Stretching factor β against the observed rate constant k_{obs} for each buffer condition. (D) Empirical adjustment factor applied to bradykinin conditions, showing all four curves now perfectly overlaid (Δ uptake = 0 Da).

analysis, such as the calculation of protection factors²⁷, this needs to be corrected for. Prior to curve fitting, the user can choose whether to perform back-exchange correction²⁸ or not. HDf1eX allows for two methods of back-exchange correction, from: (1) an experimentally determined maximum deuteration (maxD) value, where the experimental maxD is included in the cluster data input file, and (2) a maximum deuteration value derived from the plateau value of the fitted uptake curve. The plateau maxD option can only be used on proteins and peptides that can undergo full exchange within the experimental time window.

pH adjustment factor. The rate of HDX is affected by several factors, including pH and ionic strength^{27,29,30}. When normalizing between different pH, HDf1eX alters the experimental labelling timepoints according a pH adjustment factor, pH_{adjust_factor} , according to Equation 2³¹, where pH_{ref} is the reference state pH and pH_{exp} is the experimental pH of the state being compared. pH_{adjust_factor} is subsequently multiplied to all the experimental labelling timepoints for that particular state.

$$pH_{adjust_factor} = 10^{pH_{exp} - pH_{ref}} \quad \text{Equation 2}$$

To test the pH adjustment, we used bradykinin in four different buffers (Table S1: Buffers 1-4). We chose bradykinin peptide (Arg1-Pro2-Pro3-Gly4-Phe5-Ser6-Pro7-Phe8-Arg9) as it has been found to adopt random-coil conformational states in aqueous media³²⁻³⁴. Due to its unstructured nature in aqueous medium, any changes in HDX in bradykinin between the conditions being compared can be attributed to the solution effects. HDX time courses on bradykinin with three replicates and seven timepoints were collected for each buffer condition being compared (see SI Methods for experimental details). The data was fitted and corrected for back-exchange as discussed above (Figure 2A). We first applied the pH adjustment to the four different conditions to normalize the solution effects among them (Figure 2B). If the solution effects differed only by pH, we would expect to see the uptake curves to perfectly overlay each other. However, our data showed that despite applying the pH correction, there were still differences up to 0.43 Da between the uptake curves, which is above the calculated global threshold of 0.09 Da for this peptide. This confirms that the salts used also influenced the intrinsic exchange kinetics³⁵.

Empirical adjustment factor. As we saw in bradykinin, both pH and ionic composition can affect the intrinsic

exchange rate³⁶. In this case, it becomes difficult to discern the structural changes from the intrinsic exchange rate changes. To allow accurate comparison between conditions of different pH and ionic composition, we created a new empirical method which uses the unstructured peptide bradykinin to calibrate the intrinsic exchange rate effect.

Following back-exchange correction and curve-fitting, each state has two fitted parameters, k_{obs} and β , associated with it. In Figure 2C, a plot of the stretching factor β against the observed rate constant k_{obs} shows the k_{obs} and β for each buffer condition. The first step is to choose a reference state (Buffer 1). A 2D scaling factor is determined for each coordinate in Figure 2C to transform to the reference coordinate. This scaling factor, consisting of a value for k_{obs} and one for β is multiplied to the fitted parameters for all the buffer conditions. For bradykinin, this results in the uptake curves to perfectly overlay each other (Figure 2D), with differences of 0 Da between the curves, i.e., Δ uptake = 0 Da. The empirical factor correction for each condition can then be applied to the uptake curves of the protein of interest, thereby normalizing the intrinsic exchange rate effects, allowing the structural effects to be clearly distinguished.

Statistical significance testing. It is important to carry out statistical significance analyses in differential HDX experiments to determine significant differences across the protein between different experimental conditions. HDf1eX takes the hybrid significance testing approach brought forward by Hageman and Weis for statistical significance testing³⁷.

Obtaining data distribution by separating the replicates. In the first stages of HDf1eX, all the replicates were pooled together to perform the calculations. To calculate the global significance threshold for the HDX parameters (uptake area, observed rate constant and protection factors) except for the uptake, we need to know their distribution. Thus, it is necessary to perform curve fitting of different combinations of replicates. Here, we describe three methods by which HDf1eX creates the distributions from the raw data for each peptide/fragment. SI Figure 5 shows a graphical explanation of each method. (1) Simple per replicate number combinations: At each timepoint, one replicate is chosen in numerical order until all the replicates have been processed. In cases of missing replicates at any particular timepoints, the latter will not include that replicate number in the curve fitting. The mean and SD of each of these distributions is calculated, (2) Random

replicate combinations: At each timepoint, a random replicate is chosen from the set of replicates available. This is repeated n times, where n is the original number of replicates in the experiment. The mean and SD of each of these distributions is calculated, (3) Bootstrapping method of combinations: All the possible combinations of timepoints and replicates are initially calculated. A number s of random combinations is sampled uniformly from all the possible combinations, with replacement, using the MATLAB function `datasample`. Thereafter, 10,000 bootstrap data samples are drawn from s using the `bootstrp` function, and the mean and SD of each of these distributions is calculated.

Global significance threshold. A global threshold is necessary to eliminate negligible differences that appear significant from the Welch's t-test described in the next section, which would otherwise lead to false positives. Three methods of calculating the global significance threshold are available in HDfLeX. The first one is based on the calculation of the confidence interval from experimental standard deviations (SD) as described previously by Hageman and Weis³⁷. We have extended their method to include ETD fragments, to remove outliers and to be also used on the area under the curve, in addition to the uptake. The second method of calculating the global threshold, applying only to (1) and (2) data distribution methods above, initially involves calculating the pairwise differences between replicates within each state for the experimental data (SI Figure 6A). The pairwise difference is calculated rather than the difference between each replicate and the mean because we did not want any outliers in the dataset to skew the mean, which should be zero as the differences being calculated are within states. The outliers in the pairwise differences are optionally removed using the MATLAB function `isoutlier`³⁸, with a method chosen by the user (mean, median, quartiles, grubbs, gesd), prior to fitting the pairwise replicate differences to a normal distribution. The upper confidence interval limit of this normal distribution is the global significance threshold. For the uptake area, the replicate differences and standard deviations are calculated from the integral of the fitted curves, whereas for the uptake, they are calculated from the experimental data points. When the bootstrapping method of combinations is selected (SI Figure 6B), the global threshold is calculated as simply being the upper limit of confidence interval of the all the means across the peptides and ETD fragments.

Welch's t-test. A Welch's t-test is performed between the peptides and fragments of the two states being compared to give a p-value. The MATLAB function `ttest2` is used with the Satterthwaite's approximation for the effective degrees of freedom³⁹. To be significant, any differences between the peptides and fragments of the two states being compared must be above the global significance threshold and have a p-value lower than the alpha confidence level (e.g. $p < 0.05$)³⁷.

Data flattening over amino acid sequence. Instead of analyzing the HDX-MS on a per peptide/fragment basis, HDfLeX flattens the data across all the available peptides or fragments to give high structural resolution HDX-MS data. The uptake area or uptake at time t at any peptide/fragment is divided by the number of exchangeable hydrogens (q). This number is subsequently applied across the amino acid residues covering that peptide as shown in

SI Figure 7, except for the N-terminal residue and any prolines, which are fixed at 0. Note, for the ETD fragments, we attempted an intuitive subtraction of the consecutive fragments in the c/z series to obtain the deuterium uptake/area of highly resolved overlapping fragments, however this introduced large errors and was ultimately rejected (SI Figure 8).

Visualizations of differences generated by HDfLeX. HDfLeX provides several ways to visualize the difference between states: difference and volcano plots at peptide resolution (un-flattened data) and amino acid residue resolution (flattened data). The differences between the means of the uptake area or uptake at any time t , are calculated by subtraction of one state from the other at each amino acid/peptide. Those differences are then plotted against the peptide/amino acid number to give a difference plot. Volcano plots for both flattened and un-flattened data are also generated for the p-value vs difference in uptake/uptake area³⁷.

3D structural visualization of HDX differences. The differences calculated by HDfLeX for uptake, uptake area, protection factor and observed rate constant are commonly visualized on the protein structure. HDfLeX generates .pml script files at user-selected timepoints that can be run in PYMOL⁴⁰.

Calculation of intrinsic rates and protection factors. The intrinsic rates and protection factors were fitted and calculated as described previously⁴¹. Briefly, intrinsic uptake curves are generated, fitted into one- or two stretched exponentials. The intrinsic rate constant is extracted from the fitted parameters and the protection factor obtained using Equation 3, where P_f is the protection factor, k_{int} is the intrinsic amide hydrogen exchange rate constant from published values and k_{exp} is the experimental rate constant from the experimental fits.

$$P_f = \frac{k_{int}}{k_{exp}} \quad \text{Equation 3}$$

RESULTS

To test HDfLeX, we used datasets containing HDX-MS peptide and ETD information on the purified wild-type intrinsically disordered protein alpha-synuclein (aSyn) in various solution conditions (Table S1: Buffer conditions 1-4). Briefly, the millisecond timescale HDX was performed using a fully automated, millisecond HDX labelling and online quench-flow instrument, ms2min (Applied Photophysics, UK), connected to an HDX manager (Waters, USA). Mass spectrometry (MS) data were acquired on a Waters Synapt G2-Si Q-IM-TOF instrument with ETD capabilities. Both bottom-up and middle-down data were collected.

Improved structural resolution when ETD fragments included with peptides. The results of HDfLeX processing of HDX-MS data for aSyn shows an increase in structural resolution of MS data when ETD data are included with peptide data (Figure 3). Bottom-up data were combined with middle-down ETD data, resulting in a 50% increase (from 20 to 30) of single amino acids with fully resolved HDX data and a 150% and 230% increase in HDX data resolved at the dipeptides and tripeptides level, respectively (Figure 4). Overall, the merging of the ETD fragments and peptides have greatly improved the structural resolution. It was not possible to test any top-down or ECD data for this

Figure 3: Inclusion of ETD fragments to peptide data improves structural resolution for sample data of aSyn HDX. Top sequence corresponds to peptide-only data (pink) and bottom sequence corresponds to peptide + ETD fragments data (green). The blue ticks represent the merged resolution obtained by combining peptides and c/z fragments.

study, but these data should in principle be compatible with the HDflex analysis to provide even higher structural resolution.

Validation of global significance threshold methods. The importance of a global threshold for significance has recently been stressed by various groups for the correct statistical analysis of HDX-MS data^{30,37,42}. HDflex calculates the global threshold for the uptake (in Da) in the same way devised by Hageman and Weis³⁷. For the uptake area global threshold, three methods are available: calculation of a confidence interval from (1) the standard deviations^{37,43}, (2) the pairwise differences between replicates within states and (3) bootstrapped replicates.

Figure 4: Comparison of the merged resolution between peptide only (pink) and peptide + ETD (green) data.

To validate the three methods, we carried out a null experiment on the aSyn protein in Buffer 2 for four replicates and created six groups of two replicates from the same dataset (SI Figure 9). The difference in uptake area between the groups were calculated for all peptides, shown as scatter density plots in Figure 5. By definition, for a null experiment, the mean of the differences should be zero and this is what we observe. The spread of the differences is affected by outliers, as can be seen from the Null 3 dataset, for example. The presence of outliers alters the calculated global threshold for significant differences. HDflex allows for the optional removal of outliers from the calculation of the global threshold using the MATLAB function `isoutlier`³⁸. Without removing any outliers, we can see that for the 95% confidence intervals (Figure 5A), the bootstrap replicates and standard deviation-based confidence interval calculations result in no false positives. The other methods show significant differences due to false

positives. Increasing the confidence levels to 99% (Figure 5B) removes false positives from the pairwise difference-based calculation using the simple per replicate data distribution method. However, the increase in confidence interval disproportionately alters the standard deviation and bootstrapping-based calculations and could result in false negatives, causing significant differences to be missed. Therefore, we consider that the bootstrap replicates and standard deviation-based calculations are highly robust at removing false positives whilst minimizing the risk of false negatives.

Empirical adjustment as a novel normalization method for HDX-MS experiments. In order to normalize data between different buffer conditions (i.e. salts) that would

Figure 5: Comparison of the three methods of global threshold calculation on three null datasets derived from the protein aSyn in one condition only. The circles represent the difference in uptake between the two groups in the null dataset. Each null comparison was unique. The dashed lines represent the global thresholds calculated from each of the methods available in HDflex. CI: confidence interval.

Figure 6: Glycogen phosphorylase HDX-MS data analyzed with HDf1eX. (A) Example deuterium uptake plot showing the experimental and theoretical fitting of the glycogen phosphorylase State_A and State_B; (B) Volcano plot representation of the hybrid significance testing comparison of State_A vs State_B; (C) Uptake difference plot of State_A vs State_B at peptide resolution; (D) $\ln(P_f)$ difference plot of State_A vs State_B at peptide resolution; (E) Uptake difference plot of State_A vs State_B at amino acid residue resolution; (F) $\ln(P_f)$ difference plot of State_A vs State_B at amino acid residue resolution. Significant differences are denoted by a red asterisk.

alter intrinsic rates in a manner with no established solution, we used the unstructured peptide bradykinin^{32,33} as an internal standard. In this way, we aim to remove significant differences that are a result of buffer-derived changes to intrinsic rates and better differentiate those derived from changes to protein structural dynamics. The protein aSyn was on-labelled in the same buffers as bradykinin in separate experiments and the data kinetically fitted in HDf1eX to obtain a rate constant (k) and stretch term (β), shown in Table S2. The empirical adjustment factors for k and β to convert to a reference state (here Buffer 1) are globally set values derived from bradykinin that are multiplied to the fitted parameters of each peptide and fragment of aSyn (SI Figure 10). It is important to note that this method is entirely empirical and has only been tested on the unstructured peptide bradykinin and intrinsically disordered protein aSyn in single-exponential phases, and therefore, we cannot assure it will work with more structured proteins.

Intrinsic rate calculation and protection factors. HDf1eX also calculates the theoretical intrinsic rates for peptides⁴¹ at the experimental pH and temperature. aSyn is incompatible with the Bai and Englander model for the calculation of intrinsic rates^{27,44} as all the peptides exchanged faster than the predicted rates. To demonstrate the ability of HDf1eX to calculate intrinsic rates and protection factors, we analyzed previously described

datasets on glycogen phosphorylase a and b (State_A and State_B, respectively)⁴¹. The uptake plot in Figure 6A shows the expected theoretical (black) and experimentally measured (blue and green) deuterium uptake for a selected peptide for both protein states, from which the P_f is calculated. There is an upper limit of detection for P_f , as determined by the measurable labelling times (for this dataset $\ln(P_f) < 20$). Therefore, a P_f threshold can be set to discard unquantifiable, highly protected, slow-exchanging peptides. Following this filtering, 224 peptides were selected for the differential HDX analysis. To determine the significant differences between the two states, we performed hybrid significance testing as described earlier, the results of which are represented as a volcano plot (Figure 6B). Significant peptide uptake differences were found in 53% of the peptides and fall in the red shaded region. Furthermore, differences between State_A and State_B can be visualized as difference plots per peptide (Figure 6C and Figure 6D) and per amino acid residue (Figure 6E and 6F), with significant differences denoted with a red asterisk. We compared the uptake area difference plot per peptide generated by HDf1eX to the butterfly plot generated by DynamX (SI Figure 11). Multiple features were found to be (in)significant by HDf1eX, which would have the opposite evaluation without fitting and statistical analysis. Here, we show an important distinction when using protection factors is that ‘hidden’ features with low

uptake but with high protection can be detected (shown as blue shaded region in Figure 6C and 6D).

CONCLUSION

In this paper, we presented HDfLeX, a new standalone application for the analysis of HDX-MS data post-processing, allowing the combination of data at different levels: peptides and ETD fragments. Other currently available programs do not allow the combination of data at different fragmentation levels, limiting the analysis of HDX-MS data and the resulting structural resolution. With HDfLeX, we saw an improvement in structural resolution of the HDX-MS data for areas where ETD data was collected, up to a 50% increase in single residues.

We have implemented a robust significance testing method, adapting the hybrid significance testing approach by Hageman and Weis³⁷. The adjustments we included, with respect to outlier removal make the calculation of the global significance testing more robust to the presence of outliers. Throughout, we use the ‘uptake area’, as opposed to the more commonly used ‘uptake’, as our main parameter. Other groups have proposed the use of uptake area in differential HDX studies^{37,45}, but with no global threshold calculation method available due to the added complexity of the different combinations of replicate data points – a problem now solved by HDfLeX. As such, not only did we expand the hybrid significance testing method to work on uptake areas, but we created two new methods of calculating the global threshold for the uptake area (based on replicate variability and bootstrapping). This enables users the option to use uptake area easily and effortlessly in their analyses, whilst maintaining robust elimination of false positives and false negatives.

HDfLeX generates the theoretical intrinsic uptake plots and protection factor, based on the intrinsic rates devised by Bai et al²⁷. We also introduced a novel empirical approach to normalize between various conditions (e.g., salts) and for proteins (e.g., intrinsically disordered proteins) in a way not accounted for by established correction²⁷. We tested this approach on the unstructured peptide standard bradykinin and the intrinsically disordered aSyn. We hope that this new method, when used with scrutiny, will expand the types of differential HDX-MS experiments that can be done, as at present, it is rare to see comparisons between factors that affect the intrinsic rate (such as pH, temperature, ionic strength)^{31,46}.

HDfLeX is available to download from <http://hdl.handle.net/10871/127982> and does not require a MATLAB license to use.

ASSOCIATED CONTENT

The Supporting Information is available free of charge on the ACS Publications website.

Text, Additional method details (Bradykinin HDX-MS experimental details); Figures, flowcharts describing import and fitting stage, data flattening and statistical significance analysis; plateau back-exchange determination; methods of data distribution by separating the replicates; error distribution of the entire dataset; schematic of data flattening process; testing weights for data flattening; deuterium uptake plots for null experiment on aSyn; application of empirical adjustment to an aSyn peptide; comparison of difference plots generated by DynamX vs HDfLeX; Tables, Bradykinin buffer chemical

compositions; determination of empirical adjustment factor from bradykinin fitted parameters.

ACKNOWLEDGEMENTS

NS is funded by the University Council Diamond Jubilee Scholarship. JJP and MK are supported by a UKRI Future Leaders Fellowship [Grant number: MR/T02223X/1]. We are grateful for fruitful discussions with Dr S. Subramanian.

REFERENCES

1. Katta V, Chait BT. Conformational changes in proteins probed by hydrogen-exchange electrospray-ionization mass spectrometry. *Rapid Commun Mass Spectrom.* 1991;5(4):214-217. doi:10.1002/rcm.1290050415
2. Englander JJ, Del Mar C, Li W, et al. Protein structure change studied by hydrogen-deuterium exchange, functional labeling, and mass spectrometry. *Proc Natl Acad Sci U S A.* 2003;100(12):7057-7062. doi:10.1073/pnas.1232301100
3. Phillips JJ, Buchanan A, Andrews J, et al. Rate of Asparagine Deamidation in a Monoclonal Antibody Correlating with Hydrogen Exchange Rate at Adjacent Downstream Residues. *Anal Chem.* 2017;89(4):2361-2368. doi:10.1021/acs.analchem.6b04158 PMID - 28194941
4. Kim M, Sun Z-YJ, Rand KD, et al. Antibody mechanics on a membrane-bound HIV segment essential for GP41-targeted viral neutralization. *Nat Struct Mol Biol.* 2011;18(11):1235-1243. doi:10.1038/nsmb.2154
5. Keppel TR, Weis DD. Mapping Residual Structure in Intrinsically Disordered Proteins at Residue Resolution Using Millisecond Hydrogen/Deuterium Exchange and Residue Averaging. *J Am Soc Mass Spectrom.* 2015;26(4):547-554. doi:10.1007/s13361-014-1033-6 PMID - 25481641
6. Stephens AD, Zacharopoulou M, Moons R, et al. Extent of N-terminus exposure of monomeric alpha-synuclein determines its aggregation propensity. *Nat Commun.* (2020). doi:10.1038/s41467-020-16564-3
7. Masson GR, Maslen SL, Williams RL. Analysis of phosphoinositide 3-kinase inhibitors by bottom-up electron-transfer dissociation hydrogen/deuterium exchange mass spectrometry. *Biochem J.* 2017;474(11):1867-1877. doi:10.1042/bcj20170127 PMID - 28381646
8. Karch KR, Coradin M, Zandarashvili L, et al. Hydrogen-Deuterium Exchange Coupled to Top- and Middle-Down Mass Spectrometry Reveals Histone Tail Dynamics before and after Nucleosome Assembly. *Structure.* 2018;26(12):1651-1663.e3. doi:10.1016/j.str.2018.08.006 PMID - 30293810
9. Pan J, Zhang S, Chou A, Borchers CH. Higher-order structural interrogation of antibodies using middle-down hydrogen/deuterium exchange mass spectrometry. *Chem Sci.* 2016;7(2):1480-1486. doi:10.1039/C5SC03420E
10. Rand KD, Jørgensen TJD. Development of a Peptide Probe for the Occurrence of Hydrogen (1 H/ 2 H) Scrambling upon Gas-Phase Fragmentation. *Anal Chem.* 2007;79(22):8686-8693. doi:10.1021/ac0710782 PMID - 17935303
11. Rand KD, Adams CM, Zubarev RA, Jørgensen TJD. Electron Capture Dissociation Proceeds with a Low Degree of Intramolecular Migration of Peptide Amide Hydrogens. *J Am Chem Soc.* 2008;130(4):1341-1349. doi:10.1021/ja076448i PMID - 18171065
12. Zehl M, Rand KD, Jensen ON, Jørgensen TJD. Electron transfer dissociation facilitates the measurement of deuterium incorporation into selectively labeled peptides with single residue resolution. *Scopus.* 130(51):17453-17459. doi:10.1021/ja805573h PMID - 19035774
13. Abzalimov RR, Kaplan DA, Easterling ML, Kaltashov IA. Protein conformations can be probed in top-down HDX MS experiments utilizing electron transfer dissociation of protein ions without hydrogen scrambling. *J Am Soc Mass Spectrom.* 2009;20(8):1514-1517. doi:10.1016/j.jasms.2009.04.006 PMID - 19467606
14. Zubarev RA, Horn DM, Fridriksson EK, et al. Electron Capture Dissociation for Structural Characterization of Multiply Charged Protein Cations. *Anal Chem.*

- 2000;72(3):563-573. doi:10.1021/ac990811p
15. Mistarz UH, Bellina B, Jensen PF, Brown JM, Barran PE, Rand KD. UV Photodissociation Mass Spectrometry Accurately Localize Sites of Backbone Deuteration in Peptides. *Anal Chem.* 2018;90(2):1077-1080. doi:10.1021/acs.analchem.7b04683
16. Brodie NI, Huguet R, Zhang T, et al. Top-Down Hydrogen-Deuterium Exchange Analysis of Protein Structures Using Ultraviolet Photodissociation. *Anal Chem.* 2018;90(5):3079-3082. doi:10.1021/acs.analchem.7b03655
17. Rand KD, Pringle SD, Morris M, Engen JR, Brown JM. 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.* 2011;22(10):1784. doi:10.1007/s13361-011-0196-7 PMID - 21952892
18. Kan Z, Ye X, Skinner JJ, Mayne L, Englander SW. ExMS2: An Integrated Solution for Hydrogen-Deuterium Exchange Mass Spectrometry Data Analysis. *Anal Chem.* 2019;91(11):7474-7481. doi:10.1021/acs.analchem.9b01682
19. Lau AMC, Ahdash Z, Martens C, Politis A. Deuteros: software for rapid analysis and visualization of data from differential hydrogen deuterium exchange-mass spectrometry. *Bioinformatics.* 2019;35(17):3171-3173. doi:10.1093/bioinformatics/btz022
20. Hamuro Y, Coales SJ, Southern MR, Nemeth-Cawley JF, Stranz DD, Griffin PR. Rapid analysis of protein structure and dynamics by hydrogen/deuterium exchange mass spectrometry. *J Biomol Tech.* 2003;14(3):171-182.
21. Pascal BD, Willis S, Lauer JL, et al. HDX Workbench: Software for the Analysis of H/D Exchange MS Data. *J Am Soc Mass Spectrom.* 2012;23(9):1512-1521. doi:10.1007/s13361-012-0419-6 PMID - 22692830
22. Wales TE, Eggertson MJ, Engen JR. Considerations in the analysis of hydrogen exchange mass spectrometry data. *Methods Mol Biol.* 2013;1007:263-288. doi:10.1007/978-1-62703-392-3_11
23. Chetty PS, Mayne L, Lund-Katz S, Stranz D, Englander SW, Phillips MC. Helical structure and stability in human apolipoprotein A-I by hydrogen exchange and mass spectrometry. *Proc Natl Acad Sci U S A.* 2009;106(45):19005-19010. doi:10.1073/pnas.0909708106 PMID - 19850866
24. Hamuro Y, Coales SJ. Optimization of Feasibility Stage for Hydrogen/Deuterium Exchange Mass Spectrometry. *J Am Soc Mass Spectrom.* 2018;29(3):623-629. doi:10.1007/s13361-017-1860-3 PMID - 29299838
25. Walters BT, Ricciuti A, Mayne L, Englander SW. Minimizing Back Exchange in the Hydrogen Exchange-Mass Spectrometry Experiment. *J Am Soc Mass Spectrom.* 2012;23(12):2132-2139. doi:10.1007/s13361-012-0476-x PMID - 22965280
26. Venable JD, Okach L, Agarwalla S, Brock A. Subzero Temperature Chromatography for Reduced Back-Exchange and Improved Dynamic Range in Amide Hydrogen/Deuterium Exchange Mass Spectrometry. *Anal Chem.* 2012;84(21):9601-9608. doi:10.1021/ac302488h
27. Bai Y, Milne JS, Mayne L, Englander SW. Primary structure effects on peptide group hydrogen exchange. *Proteins.* 1993;17(1):75-86. doi:10.1002/prot.340170110
28. Zhang Z, Smith DL. Determination of amide hydrogen exchange by mass spectrometry: a new tool for protein structure elucidation. *Protein Sci.* 1993;2(4):522-531. doi:10.1002/pro.5560020404
29. Marcisin SR, Engen JR. Hydrogen exchange mass spectrometry: what is it and what can it tell us? *Anal Bioanal Chem.* 2010;397(3):967-972. doi:10.1007/s00216-010-3556-4 PMID - 20195578
30. Masson GR, Burke JE, Ahn NG, et al. Recommendations for performing, interpreting and reporting hydrogen deuterium exchange mass spectrometry (HDX-MS) experiments. *Nat Methods.* 2019;16(7):595-602. doi:10.1038/s41592-019-0459-y PMID - 31249422
31. Li J, Rodnin M V, Ladokhin AS, Gross ML. Hydrogen-deuterium exchange and mass spectrometry reveal the pH-dependent conformational changes of diphtheria toxin T domain. *Biochemistry.* 2014;53(43):6849-6856. doi:10.1021/bi500893y
32. Bonechi C, Ristori S, Martini G, Martini S, Rossi C. Study of bradykinin conformation in the presence of model membrane by Nuclear Magnetic Resonance and molecular modelling. *Biochim Biophys Acta - Biomembr.* 2009;1788(3):708-716. doi:10.1016/j.bbamem.2008.12.019
33. Kyle DJ, Martin JA, Farmer SG, Burch RM. Design and Conformational Analysis of Several Highly Potent Bradykinin Receptor Antagonists. *J Med Chem.* 1991;34(3):1230-1233. doi:10.1021/jm00107a052
34. Cai X, Dass C. Conformational analysis of proteins and peptides. *Curr Org Chem.* 2003;7(18):1841-1854. doi:10.2174/1385272033486161
35. Christoffersen M, Bolvig S, Tüchsen E. Salt Effects on the Amide Hydrogen Exchange of Bovine Pancreatic Trypsin Inhibitor. *Biochemistry.* 1996;35(7):2309-2315. doi:10.1021/bi951711q
36. Murphree TA, Vorauer C, Brzoska M, Guttman M. Imidazolium Compounds as Internal Exchange Reporters for Hydrogen/Deuterium Exchange by Mass Spectrometry. *Anal Chem.* 2020;92(14):9830-9837. doi:10.1021/acs.analchem.0c01328
37. Hageman TS, Weis DD. Reliable Identification of Significant Differences in Differential Hydrogen Exchange-Mass Spectrometry Measurements Using a Hybrid Significance Testing Approach. *Anal Chem.* 2019;91(13):8008-8016. doi:10.1021/acs.analchem.9b01325 PMID - 31099554
38. MathWorks. isoutlier. Accessed May 11, 2021. <https://uk.mathworks.com/help/matlab/ref/isoutlier.html>
39. MathWorks. ttest2. Accessed May 11, 2021. https://uk.mathworks.com/help/stats/ttest2.html?s_tid=doc_ta
40. L. S. The PyMOL Molecular Graphics System. Published online 2015.
41. Kish M, Smith V, Subramanian S, et al. Allosteric regulation of glycogen phosphorylase solution phase structural dynamics at high spatial resolution. *bioRxiv.* Published online 2019. doi:10.1101/654665
42. Houde D, Berkowitz SA, Engen JR. The utility of hydrogen/deuterium exchange mass spectrometry in biopharmaceutical comparability studies. *J Pharm Sci.* 2011;100(6):2071-2086. doi:10.1002/jps.22432 PMID - 21491437
43. Hageman TS, Weis DD. A Structural Variant Approach for Establishing a Detection Limit in Differential Hydrogen Exchange-Mass Spectrometry Measurements. *Anal Chem.* 2019;91(13):8017-8024. doi:10.1021/acs.analchem.9b01326 PMID - 31099552
44. Del Mar C, Greenbaum EA, Mayne L, Englander SW, Woods VLJ. Structure and properties of alpha-synuclein and other amyloids determined at the amino acid level. *Proc Natl Acad Sci U S A.* 2005;102(43):15477-15482. doi:10.1073/pnas.0507405102
45. Mazur SJ, Weber DP. The Area Between Exchange Curves as a Measure of Conformational Differences in Hydrogen-Deuterium Exchange Mass Spectrometry Studies. *J Am Soc Mass Spectrom.* 2017;28(5):978-981. doi:10.1007/s13361-017-1615-1 PMID - 28236290
46. Tajoddin NN, Konermann L. Analysis of Temperature-Dependent H/D Exchange Mass Spectrometry Experiments. *Anal Chem.* 2020;92(14):10058-10067. doi:10.1021/acs.analchem.0c01828

A**RPPGFSPFR** **B****(1-9) RPPGFSPFR** **C** **D****(1-9) RPPGFSPFR**

A **B**

- Bootstrap replicates
- Random replicates + CI standard deviation
- Simple per replicate + CI standard deviation
- Simple per replicate + CI Pairwise difference
- Random replicates + CI Pairwise difference