analytical chemistry

Article

Cite This: Anal. Chem. XXXX, XXX, XXX−XXX

pubs.acs.org/ac

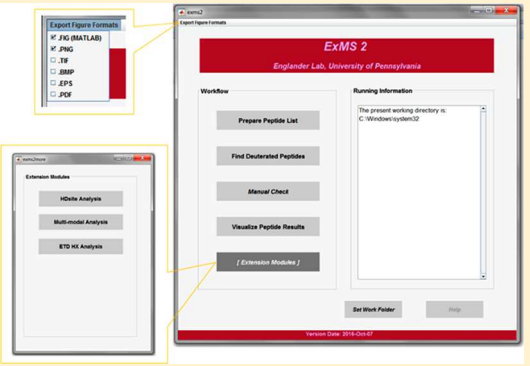
ExMS2: An Integrated Solution for Hydrogen−Deuterium Exchange Mass Spectrometry Data Analysis

Zhong-yuan Kan,*,† Xiang Ye, John J. Skinner, Leland Mayne, and S. Walter Englander*

Johnson Research Foundation, Department of Biochemistry and Biophysics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania 19104, United States

Supporting Information

ABSTRACT: Hydrogen−deuterium exchange mass spectrometry (HDX MS) has become an important technique for the analysis of protein structure and dynamics. Data analysis remains a bottleneck in the workflow. Sophisticated computer analysis is required to scan through the voluminous MS output in order to find, identify, and validate many partially deuterated peptides, elicit the HDX information, and extend the results to higher structural resolution. We previously made available two software suites, ExMS for identification and analysis of peptide isotopic envelopes in the HDX MS raw data and HDsite for residue-level resolution. Further experience has led to advances in the usability and performance of both programs. Also, newly added modules deal with ETD/ECD analysis, multimodal mass spectra analysis, and presentation options. These advances have been integrated into a stand-alone software solution named ExMS2. The package has been successfully tested by many workers in fine scale epitope mapping, in protein folding studies, and in dissecting structure and structure change of large protein complexes. A description and tutorial for this major upgrade are given here.



Hydrogen−deuterium exchange experiments (HDX) take advantage of the naturally occurring exchange between water hydrogens and protein main chain amide hydrogens, one on every amino acid (except proline) on every protein molecule.[1−6] Hydrogen bonded structure protects amide NHs and can slow their exchange by many orders of magnitude.[7] The exchange reaction depends on dynamic H-bond-breaking reactions which connect to resolved structural equilibria and structural stabilization free energy.[8] Therefore, the degree and description of HX slowing can inform on structure, structure change, interactions, and protein dynamics, and quantify real thermodynamic energies, all of this, in principle, at amino acid resolution.[9−21]

We focus here on the bottom-up fragment separation approach.[22] HDX can be performed in a continuous measurement mode over a time scale of milliseconds to days which allows the many orders of magnitude of HDX rate distribution of native proteins to be measured.[23] Alternatively, a pulse mode allows the detailed study of transiently present species such as intermediates during protein folding.[24] Timed protein samples are analyzed for carried deuterium by proteolysis into small peptides (∼6 to 25 residues) and their separation using LC and mass spectrometry. Current online proteolysis/HPLC/MS systems routinely measure hundreds of sequentially overlapping peptide fragments with variable D content. Each peptide monitors the structure and dynamics of the protein segment that it represents. The method can handle large complex protein systems using only tiny amounts of protein, 50 pmols at 1 μM concentration per HX time point.[25−27] ETD methods can be seen as another proteolysis option.[28−34]

Much effort has gone into developing HDX MS analysis software in both academia and industry (as most recently reviewed in ref 35). For this purpose, we previously released the ExMS program[36] which culls through the ∼1000 MS scans at each HX time point to find and verify many peptides and accurately determine their D-content. To move toward residue resolution we developed the HDsite program.[37] HDsite uses isotopic envelope shape from the ExMS output and comparative information from peptide overlaps to extract residue-level D occupancies and assign them to individual residues or small groups of "switchable" residues. It has been successfully applied in studying fine scale epitope mapping,[38] protein folding,[39] and structure change of large protein complexes.[25,26] Shortcomings in the earlier versions of ExMS and HDsite include dependence on commercial software (MATLAB), user interface convenience, speed, limited display options, and a few remaining false identifications. A great deal of experience with these programs has now been collected into a major upgrade that integrates the data analysis into a stand-alone software package named ExMS2.

Received: April 5, 2019
Accepted: May 13, 2019
Published: May 13, 2019

DOI: 10.1021/acs.analchem.9b01682
Anal. Chem. XXXX, XXX, XXX−XXX

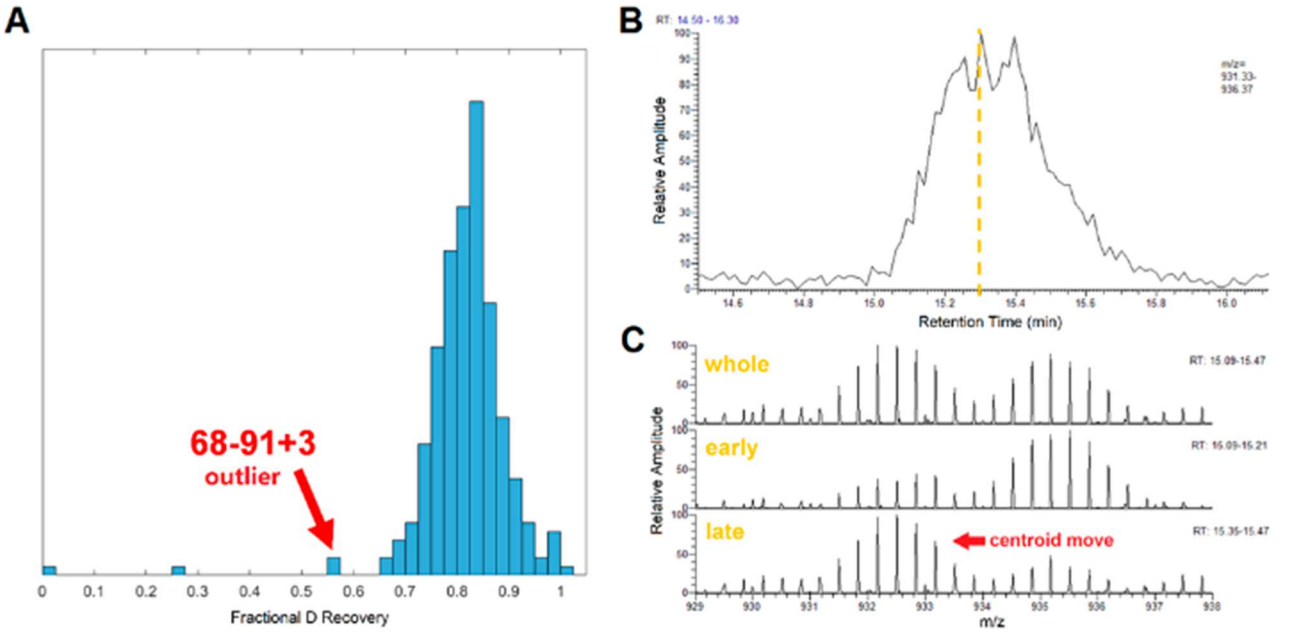


Figure 1. Misidentification signals for SNase peptide 68−91 + 3. A. Back exchange check. The peptide is an outlier in the back exchange histogram of fully deuterated peptides. B. Elution RT width check. The m/z range is normal. C. The all-D mass spectrum (upper panel) fails the unimodality check. It fails the centroid shift check, as the centroids of early and late RT scan subranges are different. It fails (not shown) the deltaD comparison check with other peptide 68−91 charge states. It conflicts with 11 other peptides in the common-end centroid subtraction check.

RESULTS

The ExMS analysis starts with a list of peptides produced in the proteolysis step, prepared by standard MS/MS methods for all-H (nondeuterated) protein samples. In the present illustrations we use results obtained for Staphylococcal nuclease (SNase, 149 residues, 363 peptides) and hexameric Hsp104 (6 × 908 residues, 841 peptides), obtained with LTQ XL and Exactive Plus EMR Orbitrap mass spectrometers with calibrated accuracy of 1 to 0.1 ppm, respectively. Both peptide sets provide 100% protein coverage with high redundancy and have sequence overlaps that intrinsically specify near residue resolution. Here we use experimental results to demonstrate setting up and running ExMS2, summarize the performance, and show example figures.

Methods for producing many peptides were described previously.[27] For online HDX analysis, ExMS takes advantage of high-resolution MS data to search for each peptide in turn through the many scans of an LC MS run, applies a number of tests to ensure correct peptide identification, and records the degree and shape of each isotopic envelope, all automatically once optional parameters have been set. The delta mass centroid and the isotopic envelope shape of each peptide as a function of exchange time measures the number and rate of D atoms bound by that protein segment and its segmental dynamics and can resolve results to near amino acid resolution.

Finding Deuterated Peptides (ExMS). Peptides produced by enzymatic digestion are initially identified using all-H protein and standard data dependent CID MS/MS, analyzed with SEQUEST, and collected into an initial peptide list. Options are provided to handle search results from other analysis programs (Figure S1). Excel spreadsheet output from SEQUEST/Bioworks software is directly imported into ExMS2, and peptide hits are merged and filtered by a user-settable peptide score cutoff (P_{pep} < 0.1) and new built-in prefilters.

An HDX data set contains multiple exchange time points and includes also a nondeuterated (all-H) control sample to specify chromatographic RT for that day. Each sample is passed through the online LC MS analysis. ExMS2 is then used to search MS scans within the known RT window for the individual isotopic peaks of each MS envelope of each peptide in the peptide list, find the partially deuterated peptides, and record the m/z values and relative amplitudes for each peak to define envelope shapes. Peptides in the list are searched for in turn in experimental HDX MS samples by the ExMS2 program.

Figure S2 shows the dedicated window for setting up data files and program parameters. The important user-adjustable program parameters, listed in the right panel, are RT window size, m/z tolerance, and peak intensity threshold. Here we lower the intensity threshold (from 500 down to 300 and 1500 down to 1000) to maximally find lower abundance peptides. We find that 2 ppm m/z tolerance is about the optimal value for both data sets from spectrometers with different levels of accuracy (Tables S1 and S2), suggesting that the major source of m/z uncertainty for peak searching is the D vs ^{13}C mass defect difference (in the heavier peaks).

ExMS2 calculates the D occupancy (delta mass centroid) for each peptide and optionally generates a figure (Figure S3) showing the envelope profiles found. After processing the whole peptide list, a result table (the "final table") is generated with information for each peptide (Figure S4) in both pure text and Excel spreadsheet format. These tables contain all the information needed for displaying D incorporation level or for further analysis (e.g., back exchange correction, multimodal analysis, and HDsite analysis).

Checking the Peptides Found. To guarantee correct peptide identification, ExMS2 applies the same six tests described previously[36] and six new ones. The original version of ExMS used six tests (combined as the Autocheck) to confirm the accuracy of peptide assignments of each

DOI: 10.1021/acs.analchem.9b01682
Anal. Chem. XXXX, XXX, XXX−XXX

deuterated peptide one at a time. The previously developed tests are as follows: (1) Summed peak intensity threshold test. (2) Intensity out of possible m/z range (−1 or max+1) test. (3) Wrong charge state test (intervening peaks). (4) Smooth intensity distribution test. (5) RT consistency test (for deuterated samples). (6) Isobaric/overlapping test.

The first three tests are applied in analyzing each MS scan. For each peptide envelope, the scans that pass all three are summed. Then all six tests are applied to the summed mass spectrum. Performance statistics are in Tables S1 and S2.

New Peptide Checks. The original tests were designed on a per peptide basis. They have proven effective but some false positives are revealed by their inconsistency with other peptides. We have added new checks mostly based on multipeptide comparisons. These checks alert the user to a possible problem with that peptide, which can then be individually checked (optional Manual Check function) or simply rejected if other peptides are sufficient. (7) Unimodality check: The isotopic envelopes of all peptides should have no more than their theoretical maximum width. An overly wide or bimodal isotopic envelope suggests that an interfering peptide or peptide carries over from a previous run, or real multimodal exchange (e.g., due to EX1). (8) Back exchange check: All peptides should show back exchange that changes gradually only with retention time (RT). Outliers suggest misidentified peptides. (9) RT width check: All peptides should have similar chromatographic widths. Outliers suggest coelution of interfering peptide or misidentification. (10) Charge states comparison check. Many peptides are seen in multiple charge states. They should have close RT and mass centroid. (11) Common-end comparison check: Many proteolytic peptides share a common N- or C-terminus. A longer common-end peptide should not carry fewer deuterons or more additional D's than the increased number of residues.(12) Centroid shift check: A significant shift of the mass distribution through a peptide elution profile suggests peptides that both coelute and overlap in m/z.

Peptides that fail one or more tests are flagged, and can be eliminated or kept, or can be examined in manual check function and either repaired or eliminated. The results of each of these new checks as well as the old checks are written into the final table (Figure S4, panel A). Panel B gives some further annotations of the table columns including formula/criteria/value explanation of the checks. An example is shown in Figure 1. This peptide passed all the old tests but failed multiple new population-level checks, illustrating their sensitivity. Total pass/fail statistics of the new checks are in Table 1.

In addition to the automatic process for finding deuterated peptides, ExMS2 also provides an independent module for manually checking the results for any individual peptide that appears questionable (Figure S5) and for testing and resetting its retention time range and updating the result table. However, Manual Check is time-consuming. Given the large number of peptides normally found, Manual Check is advisible only for particularly important peptides.

Other Improvements to ExMS. *Speed.* ExMS2 adds a preprocessing step to convert scan profile data to centroid data using the ExMS peak picking routine. Previously this was done as needed for each peptide on the search list. Because of the considerable overlap among peptides in both RT and m/z this resulted in peak picking being run more than once for many MS scans.

Table 1. Pass/Fail Statistics for Initially Fully Deuterated Sample of SNase and HSP104[a]

	SNase (total 363 peptides in list)		HSP104 (total 841 peptides in list)	
peptides found by ExMS2	280	77.1%	692	82%
peptides that passed all six Autocheck tests	244	67%	472	56%
new check failure rates				
check #1 (unimodality)	6	2%	9	1%
check #2 (back exchange consistency)	8	2%	31	4%
check #3 (elution RT width)	4	1%	2	0.2%
check #4 (RT/deltaD consistency)	0	0	2	0.2%
check #5 (common-end subtraction)	4	1%	15	2%
check #6 (centroid shift within RT)	0	0	5	0.6%
peptides survived (entering HDsite analysis)	224	62%	424	50%

[a]The m/z tolerance is set at 2 ppm. Full statistics with different m/z tolerance comparisons are in Tables S1 and S2.

Filters. Two prefilters were added to check the initial MS/MS peptide identification for potential errors. They check whether the chromatographic RT range for multiple charge states for the same peptide agree and whether the peptide list contains isobaric peptides that have close RT values. In these cases we keep the peptide with the best peptide score.

Isotopic Peak Acceptance. ExMS searches for the individual isotopic peaks of each peptide in the initial peptide list. The m/z value of the monoisotopic peak can be calculated precisely so the allowed tolerance need only reflect instrumental accuracy. For higher mass isotopic peaks m/z, uncertainty is introduced due to the difference in additional mass between ^{13}C and ^{2}H isotopes, with mass 1.003355 vs 1.006277, respectively (0.002922 difference). Previously ExMS used a wider m/z tolerance in order to accept a heavier peak as possibly belonging to the candidate peptide. ExMS2 calculates and uses the narrower isotopic peak m/z tolerance for each mass peak (Figure S9).

Residue-Level Analysis (HDsite). HDsite analysis starts with the peptide level HDX results just described (Figure S8) and uses iterative fitting to find the best match between the measured data and the individual amide HX rates that might have produced that data. We have shown that best results are obtained by fitting isotopic envelope shapes plus peptide overlap information rather than using overlaps alone.[37]

The user selects the final table text files, inputs the corresponding HDX times that were experimentally measured, and also includes the fully deuterated control if the back exchange correction is used. Back exchange less than ∼20% makes little difference in the HDsite result. The HDsite fitting options to be specified include the fitting level (centroid level or envelope level), the fitting algorithm, termination criteria to be used in the iterative fitting, and the possible range of deuteration. Here we use the default options: fitting level 2 − Envelope; fitting algorithm 1 − *LsqNonLin*. The range of deuteration is between 0 and 0.9 because our H to D exchange samples contain 90% D_2O.

Two different modes of HDX rate fitting can be employed. One calculates D-occupancy per amide at each exchange time point and then fits the results for each amide with a single exponential time course. Alternatively, one can attempt a direct rate fitting to calculate individual amide HX rates based on data for a whole set of peptides and exchange times. In either

DOI: 10.1021/acs.analchem.9b01682
Anal. Chem. XXXX, XXX, XXX−XXX

case it is useful to fit separate subsets of overlapping peptides rather than the whole protein all at once, which would take much longer.

Improvements to HDsite. Three major improvements have been made. (1) The convolution calculation step in the deuterated mass spectrum calculation now uses a Fast Fourier Transform (FFT) method, which reduces time-consumption from about half of the overall run time to a negligible value. Table 2 compares two example data sets. (2) A parallel computing option is provided for multicore processors. Due to the overhead communication time among the processors this option increases processing speed only for data sets with more than ~500 peptides (Table 2). (3) A new function has been added (see Figure S8) to autodefine appropriate fitting subregions based on overlaps in the input peptide set. Fitting smaller independent subsequences greatly reduces runtime. For the example in Table 2, the fitting of two different subregions cost 9.5 and 7345 s, respectively (with the older algorithm and nonparallel computing). If the two subregions are fit together, runtime increases to 12589 s (~2× slower).

Table 2. HDsite Fitting Runtime Comparison[a]

	smaller data set (88 peptides)	larger data set (923 peptides)
old algorithm and nonparallel computing	9.5 s	7345 s
new algorithm and nonparallel computing	4.4 s	5372 s
old algorithm and parallel computing	34.8 s	6744 s
new algorithm and parallel computing	26.7 s	4471 s

[a]The test dataset is from a native SNase experiment with 10 HDX time-points. The smaller dataset fit the subregion residue number 88 to 103, containing 88 peptides combined from all time points. The larger dataset fit the subregion 61 to 92, containing 923 peptides combined from all time points. The used CPU is Intel Core i7-4790 (3.60 GHz × 2) with 16.0 GB RAM.

An optional prefilter based on the results of ExMS checks has been added to exclude all problematic peptides.

The combination of all currently implemented ExMS checks may occasionally not catch a peptide misidentification. The error will often become apparent as an HDsite fitting outlier, especially when fitting multiple HDX time points together. Although these outlier peptides usually have only minor impact on the fitting, accuracy can be improved by filtering them out. For this, an option for two-rounds of fitting has been added: The first-round fitting is done with the full initial peptide set and nonstrict fitting termination criteria for a quick run. Outlier peptides are automatically identified by comparing the centroids of the experimental and fitted mass spectra. With the outlier peptides removed, the second-round fitting is done with more strict fitting termination criteria to get a more accurate result. The fitting termination criteria include *TolX*, *TolFun*, *MaxIter*, *MaxFunEvals*, etc., as provided in MATLAB.

In the original HDsite program, fitting was performed for each single HDX time point to get the D occupancy at that time. The time course of D occupancy could be later fitted with a single exponential function to obtain the HDX rate of individual residues. We have added an option to simultaneously fit a time course to directly get the HDX rates of individual residues. In principle, rate fitting (RF) has the advantage of more data for each adjustable parameter. It also avoids the problem of sorting switchable residues between time points and potentially increases site resolution. In practice, testing found that D-fitting is usually faster than rate fitting, and it is more tolerant of data noise. In real cases, one or the other mode may perform better depending on the data set when using the default *LsqNonLin* algorithm (Figure S10). Rate fitting is more likely to be trapped in local minima in the complex fitting landscape. Different start points may end with different answers (Figure S11). A more advanced global optimization algorithm may improve these results. While this manuscript was in preparation, Y. Hamuro presented an analysis of cytochrome c data using a similar rate fitting approach.[40]

The choice of the fitting algorithm is another important setting in HDsite. For both D-occupancy fitting and direct rate fitting, the currently available algorithms include the following: (1) the default nonlinear least-squares fitting with "trust-region-reflective" algorithm; (2) pattern search; (3) multistart; (4) global search; (5) simulated annealing; (6) genetic algorithm. In our experience, the multistart algorithm seems to work best for direct rate fitting (Figure S12).

To evaluate the accuracy of HDsite fitting results, optional routines for error estimation have been added. In the advanced settings of HDsite, the user may input the number of simulated replicate data sets to be sequentially fit following the normal data set. Each simulated data set is generated by (1) adding Monte Carlo noise with user specific signal-to-noise level and/or (2) bootstrapping on peptides with multiple charge states. At the end, the confidence intervals for HDsite can be calculated and the error bars can be plotted.

Along with the above algorithmic revisions, a thorough code cleanup has been done to additionally speed up the analysis. The compilation of MATLAB scripts into a standalone application, as well as the performance enhancements made by newer versions of MATLAB, also help speed up our software in general.

Data Presentation. ExMS2 provides several useful modes for visualization of HDX results (Figures 2, S6, and S7), with the option to compare up to four different HDX experiments or conditions. For the presentation of HDsite results, the site-resolved rate or protection factor output can be illustrated in various ways. Rate results at a site-resolved or switchable-set level can be printed out numerically and compared with structural features and experimental changes therein. More qualitatively, Figure 3 shows a compact rate color map using the 3D PyMOL structure of SNase.

Usability Enhancements. ExMS2 eliminates the requirement for a MATLAB license by compiling program scripts into a standalone application based on the free MATLAB runtime library. Also, we have implemented a graphical user interface (GUI) that guides the user through the data analysis workflow (Figure 4).

To make the use of our software as universal as possible, we standardized the format of the input and output files. We use the instrument vendor-neutral mzXML format for mass spec data input. The peptide list containing to-be-searched peptide information is a simple text table (Figure S1) which can be compiled from any MS2 search result. The ExMS peptide search results are saved in a text table containing peptide HDX MS related information (Figure S4), which can be directly used for presentation or further analysis. HDsite output is saved in a text table containing residue level HDX information with standardized column format, which can be used for

DOI: 10.1021/acs.analchem.9b01682
Anal. Chem. XXXX, XXX, XXX−XXX

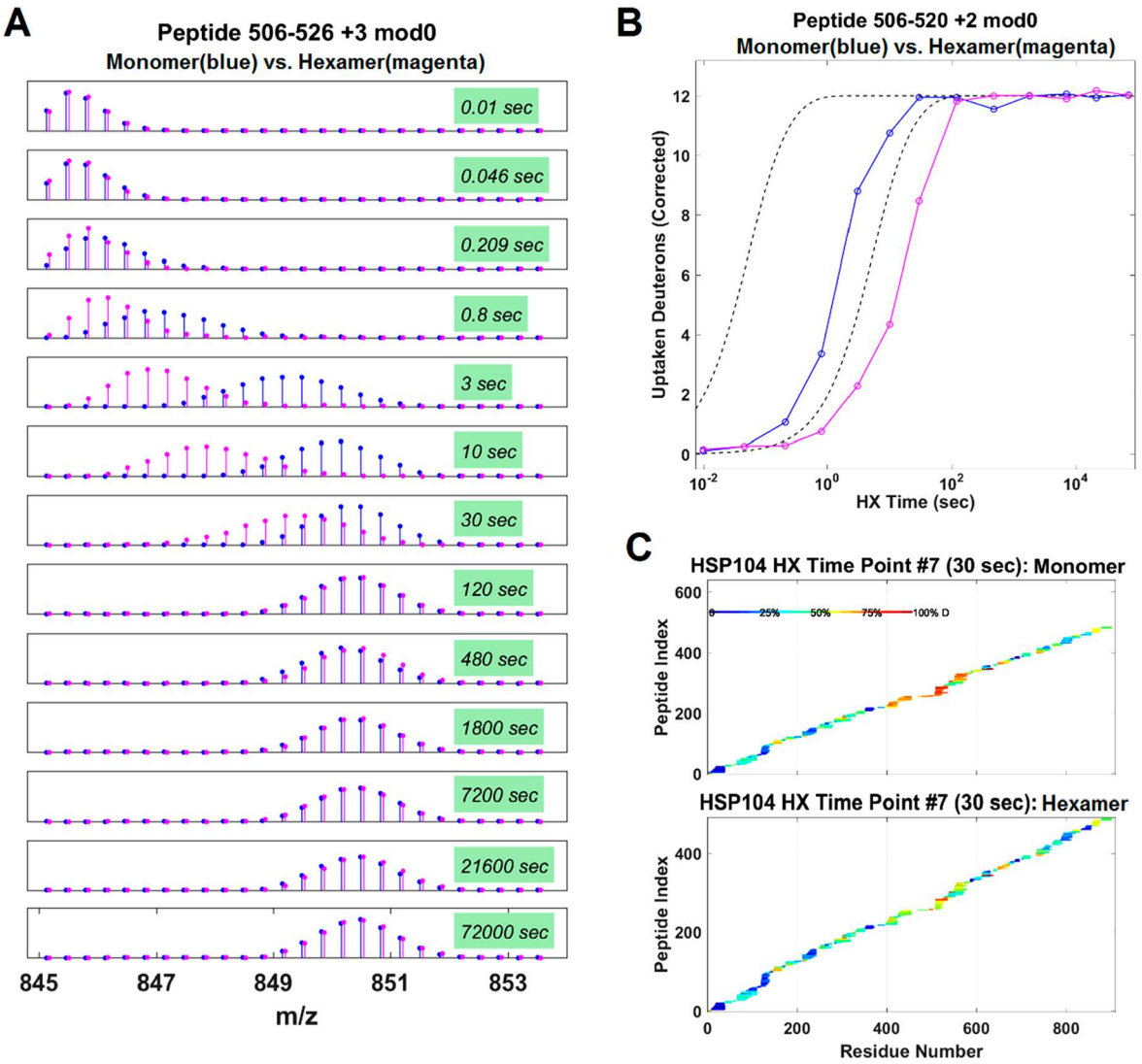


Figure 2. Visualization of HDX results. A. D-labeling time points for peptide 506−526 in Hsp104. B. The corresponding mass centroid plots. Reference curves (dashed) are calculated for HDX of the peptide when unprotected[2] and with a protection factor of 100. C. Whole protein peptide coverage map with color-coding indicating amount of exchange at the 30 s time point. Other display options include heat map and butterfly plots (Figure S7).

plotting or can be written in a PDB file for structural HDX visualization with PyMOL (Figure 3).

Other Modules. *ETD HDX Analysis.* ETD (electron transfer dissociation) mass spectrometry[28] is a setting to perform HDX experiments with the potential to extract site-resolved hydrogen exchange information by simply subtracting the incorporated deuterons of highly overlapping daughter ions. The ETD process essentially provides an alternative to proteolysis for protein fragmentation, generating *c*- and *z*-type fragment ions instead of proteolytic peptides. In the ETD HDX analysis module (Figure S13), the same ExMS and HDsite algorithms have been adapted to find and fit *c*/*z* ion data, with very similar principles as for regular peptide analysis.

Multimodal Analysis. HDX samples may show multimodal peptide mass spectra, almost always bimodal in our experience that arise from EX1 exchange or from a heterogeneous protein region that can occupy different conformations with different HX properties. ExMS was already able to find multimodal peptides, and the newly added Unimodality check can flag and keep them. ExMS2 implements a new module to perform multimodal analysis intended to quantify the relative population fractions, given by the MS amplitudes of the two differently labeled populations and to determine the D uptake of each.[25] As in Figure S14, the multimodal analysis workflow starts by importing peptide level HDX results from the ExMS output table. Each peptide's mass spectrum is fitted with the user-chosen number of populations and one of the available fitting functions:

(1) Varied binomials. This mode uses the same computation as in HDsite (the convolution of binomial terms for the percentage of D at each amide site) to fit the subspectrum of each HDX population. This is the most accurate approach. (2) Uniform binomials + Gaussian(s). This mode has fewer fitting parameters than for varied binomials. The uniform binomials function is used to fit the envelope of the less exchanged population which is far from Gaussian at low D uptake. (3) Gaussians: The multimodal spectra are fit by Gaussians. (4) Reference shapes. When appropriate control samples display the pure subspectrum of one HDX population, it can be used

DOI: 10.1021/acs.analchem.9b01682
Anal. Chem. XXXX, XXX, XXX−XXX

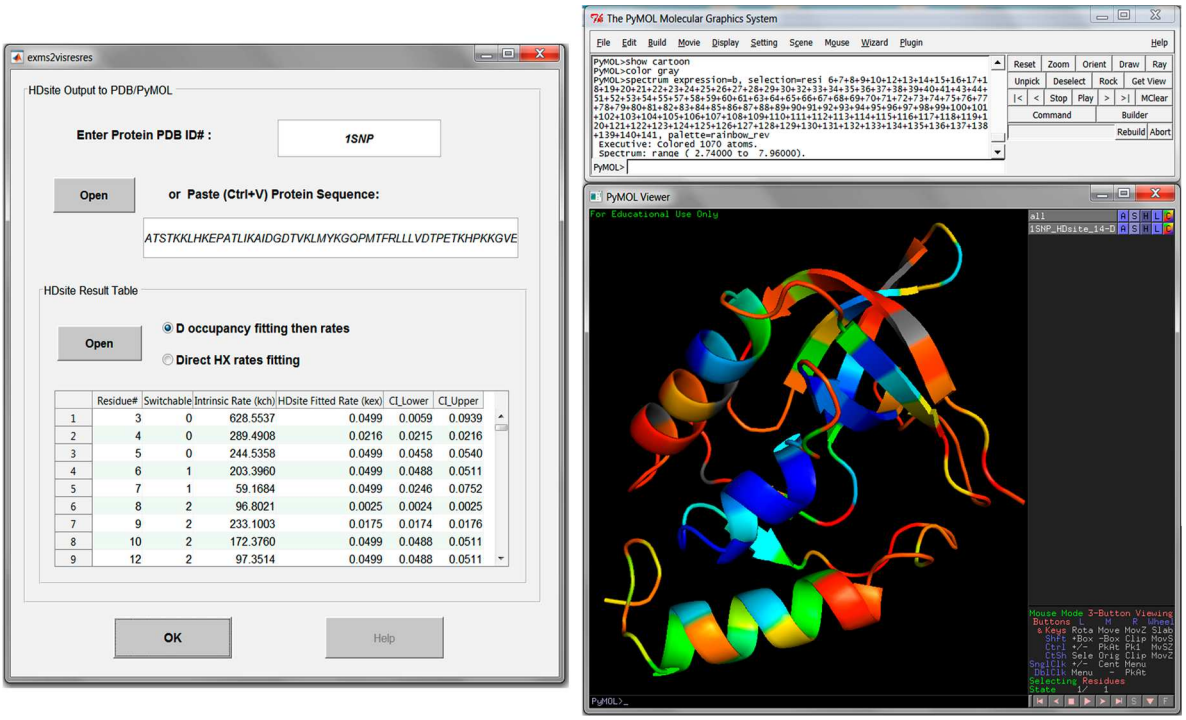


Figure 3. Visualization of residue-level HDX results for SNase (HDsite output). Left panel is the ExMS2 window where user may import and list HDsite results. One can then overwrite the PDB file (1SNP for SNase), generate a PyMOL script ("_HDsite4PyMOL.txt") containing the PyMOL commands for automatically opening the saved modified PDB file, and set up the desired molecular visualization including position, cartoon style, and the color spectrum-encoded HDX result. In PyMOL (right panel), user may simply run this single script to get the final visualization.

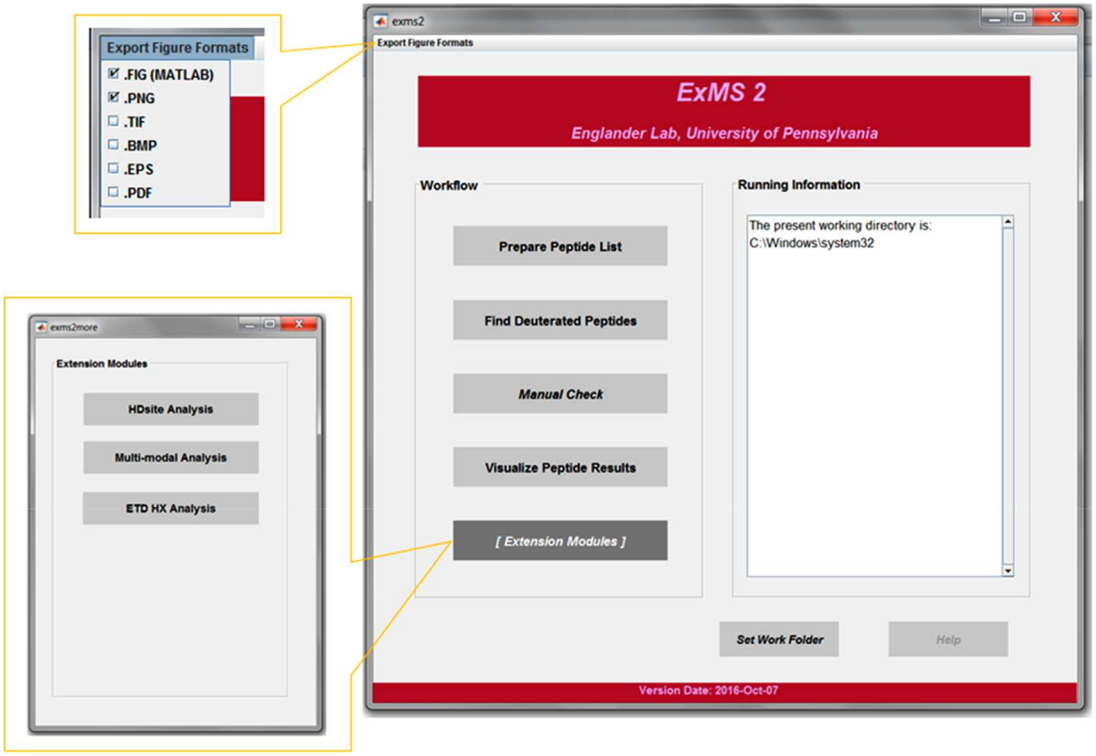


Figure 4. ExMS2 main window. The left panel lists the essential and extended functional modules through the HDX MS data analysis workflow. Each module has a pop-up window for the respective functionality. The right panel displays real-time run information, and there is a separate auxiliary console window to display more detail. In the main window, users may set work folder for fast access to input data, for saving results, and to get general software help information. From the drop-down menu, users may choose output figure formats before analysis.

as a reference to fit the multimodal spectra by only adjusting their relative amplitudes.

The fitting options can be set once for the whole data set, so all the peptides will be sequentially fitted and the results will be

DOI: 10.1021/acs.analchem.9b01682
Anal. Chem. XXXX, XXX, XXX−XXX

plotted together (right panel of Figure S14). In addition, the user can easily go through the fitted peptides and change any fitting option to update the result. Finally, all the multimodal analysis results are saved appended to the ExMS final table (Figure S4B, from column no. 21 + maxND + maxD onward). This appended table can be used later to easily extract the contained HX subpopulation information for any downstream analysis.

DISCUSSION

Hydrogen exchange mass spectrometry is a powerful analytical tool that is being increasingly employed in academia and the biopharmaceutical industry. Efficient deep data analysis is a critical part of this information-rich technique. Many groups have been developing HDX MS analysis software. We believe ExMS2 provides significant advantages that will benefit the HDX MS community. Much experience shows that ExMS is able to find, validate, and analyze very many peptides even in the presence of difficult conditions including broadened isotopic envelopes due to partial deuteration, major m/z overlap, unpredictable m/z shifts, and multimodal envelopes due to static or dynamic structural heterogeneity or EX1 behavior.

Potentially useful directions for future development can be considered. The setting of program parameters can have a large effect on performance. An algorithm might be designed to examine the input data set and automatically find optimal values for both peptide-level and residue-level analysis. The set of checks currently developed and tested for validating deuterated peptide envelopes can be highly successful. However, users have to understand each of these checks in order to use them to judge whether a peptide is really correct, which may not always be clear from the checks' raw results. An additional subroutine might helpfully judge which findings are really correct and which are not, based on the summed results of all of the checks as reported in the peptide final table. A machine learning method would be appropriate to adopt here. One might split the peptides into training and prediction sets and utilize the Manual Check function for model training.

So far our analysis depends on a precompiled target peptide list from the MS2 experiment. Alternatively one might also consider adding a de novo feature detection in the HDX MS data. The found peptide features (usually many more than the MS2-confirmed peptides) can be assigned to theoretical and known modified and unmodified peptides, which would greatly help to fill gaps in some difficult proteins.

The analysis workflow might be expanded to search also for complex modifications (disulfides, glycosylation), for integrating covalent-labeling or ion mobility spectrometry (IMS) experiments[41] (which could provide additional peptide checks among other advantages), for top-down and targeted structural analysis and for quantification. Last, about the distribution of our software, we expect to make the integrated ExMS2 package available on request.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.analchem.9b01682.

Figures S1−S14 and Tables S1 and S2 (PDF)

AUTHOR INFORMATION

Corresponding Authors

*E-mail: kanpenn@gmail.com.

*E-mail: engl@pennmedicine.upenn.edu.

ORCID

Zhong-yuan Kan: 0000-0001-6689-0593

Present Address

†Sorrento Therapeutics Inc., San Diego, CA 92121.

Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We thank all test users of ExMS2 during its development and upgrade. This work was supported by research grants from the National Institutes of Health (GM031847), the National Science Foundation (MCB1020649), and the G. Harold and Leila Y. Mathers Charitable Foundation.

REFERENCES

(1) Molday, R. S.; Englander, S. W.; Kallen, R. G. *Biochemistry* **1972**, *11*, 150−158.

(2) Bai, Y.; Milne, J. S.; Mayne, L.; Englander, S. W. *Proteins: Struct., Funct., Genet.* **1993**, *17*, 75−86.

(3) Guttman, M.; Lee, K. K. *Methods Enzymol.* **2016**, *566*, 405−426.

(4) Oganesyan, I.; Lento, C.; Wilson, D. J. *Methods* **2018**, *144*, 27−42.

(5) Masson, G. R.; Jenkins, M. L.; Burke, J. E. *Expert Opin. Drug Discovery* **2017**, *12*, 981−994.

(6) Engen, J. R.; Wales, T. E. *Annu. Rev. Anal. Chem.* **2015**, *8*, 127−148.

(7) Skinner, J. J.; Lim, W. K.; Bedard, S.; Black, B. E.; Englander, S. W. *Protein Sci.* **2012**, *21*, 996−1005.

(8) Englander, S. W. *J. Am. Soc. Mass Spectrom.* **2006**, *17*, 1481−1489.

(9) Wales, T. E.; Engen, J. R. *Mass Spectrom. Rev.* **2006**, *25*, 158−170.

(10) Kaltashov, I. A.; Eyles, S. J. *Mass Spectrom. Rev.* **2002**, *21*, 37−71.

(11) Tang, L.; Hopper, E. D.; Tong, Y.; Sadowsky, J. D.; Peterson, K. J.; Gellman, S. H.; Fitzgerald, M. C. *Anal. Chem.* **2007**, *79*, 5869−5877.

(12) Garcia, R. A.; Pantazatos, D.; Villarreal, F. J. *Assay Drug Dev. Technol.* **2004**, *2*, 81−91.

(13) Tsutsui, Y.; Wintrode, P. L. *Curr. Med. Chem.* **2007**, *14*, 2344−2358.

(14) Gallagher, E. S.; Hudgens, J. W. *Methods Enzymol.* **2016**, *566*, 357−404.

(15) Huang, R. Y.; Chen, G. *Anal. Bioanal. Chem.* **2014**, *406*, 6541−6558.

(16) Rand, K. D.; Zehl, M.; Jorgensen, T. J. *Acc. Chem. Res.* **2014**, *47*, 3018−3027.

(17) Konermann, L.; Rodriguez, A. D.; Sowole, M. A. *Analyst* **2014**, *139*, 6078−6087.

(18) Majumdar, R.; Middaugh, C. R.; Weis, D. D.; Volkin, D. B. *J. Pharm. Sci.* **2015**, *104*, 327−345.

(19) Wang, L.; Chance, M. R. *Mol. Cell. Proteomics* **2017**, *16*, 706−716.

(20) Kostyukevich, Y.; Acter, T.; Zherebker, A.; Ahmed, A.; Kim, S.; Nikolaev, E. *Mass Spectrom. Rev.* **2018**, *37*, 811−853.

(21) Benhaim, M.; Lee, K. K.; Guttman, M. *Protein Pept. Lett.* **2019**, *26*, 16−26.

DOI: 10.1021/acs.analchem.9b01682
Anal. Chem. XXXX, XXX, XXX−XXX

(22) Englander, J. J.; Rogero, J. R.; Englander, S. W. *Anal. Biochem.* **1985**, *147*, 234−244.
(23) Konermann, L.; Pan, J.; Liu, Y. H. *Chem. Soc. Rev.* **2011**, *40*, 1224−1234.
(24) Englander, S. W.; Mayne, L.; Kan, Z. Y.; Hu, W. *Annu. Rev. Biophys.* **2016**, *45*, 135−152.
(25) Ye, X.; Lin, J.; Mayne, L.; Shorter, J.; Englander, S. W. *Proc. Natl. Acad. Sci. U. S. A.* **2019**, *116*, 7333−7342.
(26) Ye, X.; Mayne, L.; Kan, Z. Y.; Englander, S. W. *Proc. Natl. Acad. Sci. U. S. A.* **2018**, *115*, 519−524.
(27) Mayne, L.; Kan, Z. Y.; Chetty, P. S.; Ricciuti, A.; Walters, B. T.; Englander, S. W. *J. Am. Soc. Mass Spectrom.* **2011**, *22*, 1898−1905.
(28) Rand, K. D.; Zehl, M.; Jensen, O. N.; Jorgensen, T. J. *Anal. Chem.* **2009**, *81*, 5577−5584.
(29) Huang, R. Y.; Garai, K.; Frieden, C.; Gross, M. L. *Biochemistry* **2011**, *50*, 9273−9282.
(30) Landgraf, R. R.; Chalmers, M. J.; Griffin, P. R. *J. Am. Soc. Mass Spectrom.* **2012**, *23*, 301−309.
(31) Song, H.; Olsen, O. H.; Persson, E.; Rand, K. D. *J. Biol. Chem.* **2014**, *289*, 35388−35396.
(32) Pan, J.; Zhang, S.; Chou, A.; Hardie, D. B.; Borchers, C. H. *Anal. Chem.* **2015**, *87*, 5884−5890.
(33) Seger, S. T.; Breinholt, J.; Faber, J. H.; Andersen, M. D.; Wiberg, C.; Schjodt, C. B.; Rand, K. D. *Anal. Chem.* **2015**, *87*, 5973−5980.
(34) Masson, G. R.; Maslen, S. L.; Williams, R. L. *Biochem. J.* **2017**, *474*, 1867−1877.
(35) Claesen, J.; Burzykowski, T. *Mass Spectrom. Rev.* **2017**, *36*, 649−667.
(36) Kan, Z. Y.; Mayne, L.; Chetty, P. S.; Englander, S. W. *J. Am. Soc. Mass Spectrom.* **2011**, *22*, 1906−1915.
(37) Kan, Z. Y.; Walters, B. T.; Mayne, L.; Englander, S. W. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, 16438−16443.
(38) Casina, V. C.; Hu, W.; Mao, J. H.; Lu, R. N.; Hanby, H. A.; Pickens, B.; Kan, Z. Y.; Lim, W. K.; Mayne, L.; Ostertag, E. M.; Kacir, S.; Siegel, D. L.; Englander, S. W.; Zheng, X. L. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112*, 9620−9625.
(39) Hu, W.; Walters, B. T.; Kan, Z. Y.; Mayne, L.; Rosen, L. E.; Marqusee, S.; Englander, S. W. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, 7684−7689.
(40) Hamuro, Y.; E, S. Y. *J. Am. Soc. Mass Spectrom.* **2018**, *29*, 989−1001.
(41) Majuta, S. N.; Maleki, H.; Kiani Karanji, A.; Attanyake, K.; Loch, E.; Valentine, S. J. *Curr. Opin. Chem. Biol.* **2018**, *42*, 101−110.

DOI: 10.1021/acs.analchem.9b01682
Anal. Chem. XXXX, XXX, XXX−XXX