Journal of proteome research

pubs.acs.org/jpr

Article

HRaDeX: R Package and Web Server for Computing High-Resolution Deuterium Uptake Rates for HDX−MS Data

Published as a part of the Journal of Proteome Research special issue "Software Tools and Resources 2025".

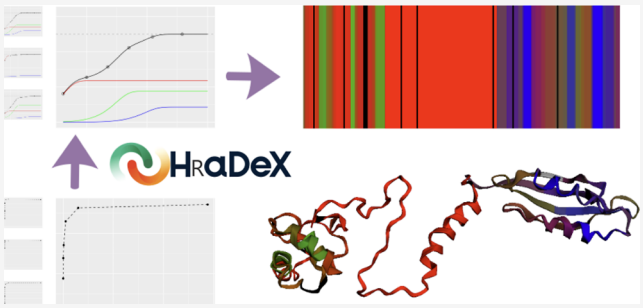
Weronika Puchała, Michał Kistowski, Liliya Zhukova, Michał Burdukiewicz, and Michał Dadlez*

Cite This: *J. Proteome Res.* 2025, 24, 1688−1700

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Hydrogen−deuterium exchange monitored by mass spectrometry (HDX−MS) is a well-established and powerful technique used to study protein dynamics and stability by capturing local and global unfolding events in protein structures. However, in this technique, obtaining region-specific information requires proteolytic digestion that breaks the protein into peptide fragments, causing the HDX data to reflect averages over these fragments rather than individual amino acids. We propose a new computational method that provides deuterium uptake kinetic parameters with high resolution, considering deuterium uptake trajectories of superimposed peptides. Our algorithm, HRaDeX, is available as a web server and an R package capable of processing data from single-state and comparative HDX−MS studies. Utilizing eight benchmark data sets, we demonstrate that HRaDeX reaches an average root-mean-square error of 7.15% in the reconstitution of experimental normalized deuterium uptake curves.



KEYWORDS: *hydrogen−deuterium exchange, HDX−MS, structural proteomics, web server*

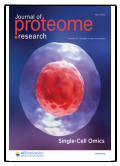
INTRODUCTION

Hydrogen−deuterium exchange monitored by mass spectrometry (HDX−MS)[1] is a technique for studying protein structural dynamics that provides a unique insight into the dynamic properties of protein chains entangled in more or less stable spatial structures. In a commonly used experimental setup for HDX, the protein is incubated in D_2O for selected periods. During this process, hydrogens and protons in the protein are exchanged for deuterons with different kinetics, depending on their position within the protein scaffold, entanglement in hydrogen-bonding networks, and solvent accessibility. For analysis, at selected times, the exchange is practically stopped by rapidly decreasing the pH and temperature of the sample. Subsequently, the protein is digested into peptides by a protease active at low pH, and their deuterium uptake can be approximated by measuring the increase in their masses by using a mass spectrometer. By measuring peptide masses at various time points (usually between seconds and 24 h), the kinetics of the exchange of main chain amide protons to deuteria can be investigated in a region-specific way. The MS-based analysis assesses the exchange by analyzing the details of the isotopic envelope or its centroid (weighted mass average).[2]

The level of deuterium incorporation, measured by comparison to nondeuterated control samples (deuterium incorporation profile), provides insight into the level of protection against the exchange of amide protons in a given region. Slower exchange is mainly caused by two significant factors: the involvement of an amide in H-bonding networks and impaired solvent accessibility. Therefore, HDX probes the level of entanglement of these protons in more or less stable elements of secondary, tertiary, or quaternary protein structures. In other words, HDX allows sampling of frequency and degree of local or global unfolding episodes in a wide dynamic range of several orders of magnitude of different timeframes of structural events. Thus, not all hydrogens are equally available for exchange. Some, especially those buried in the protein core, may become available only during global unfolding events when the protein fully opens its structure, while others require only local fluctuations. Accordingly, a two-state kinetic model is commonly used to describe the HDX of an amide proton[3]

$$NH_{closed} \underset{k_{close}}{\overset{k_{open}}{\leftrightarrow}} NH_{open} \overset{k_{int}}{\leftrightarrow} NH_{open} \underset{k_{open}}{\overset{k_{close}}{\leftrightarrow}} ND_{closed}$$

Received: August 16, 2024
Revised: February 4, 2025
Accepted: March 4, 2025
Published: March 19, 2025



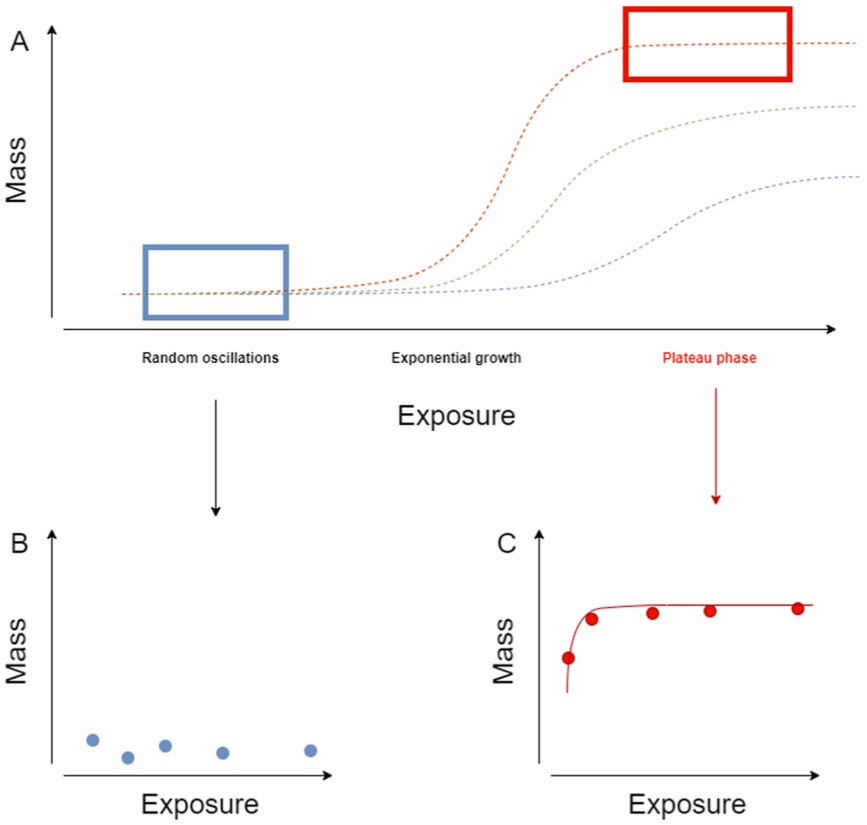


Figure 1. (A) Exponential nature of HDX exemplified by a typical time course of mass increase after deuterium uptake (vertical axis) according to a ZS model, which assumes that each amide can be assigned to one of the three exchange regimes, fast, intermediate, or slow. The exposure time axis (horizontal) is in the logarithmic scale and presents the incubation time. During the experiment, the mass of a peptide increases as three groups of amides (marked with red, green, and blue colors) exchange their hydrogens for deuterium with, respectively, high, intermediate, and low rates. After all available amides are deuterated, the mass of a peptide stabilizes during the plateau phase. (B) In the case of an extremely slow exchange, its time course is best represented by small linear growth in the time frame of experimental observation. The random mass oscillations may prevent identifying a visible trend. (C) During the plateau phase, the mass is stable because the rates of both exchanges (HD and DH) are in equilibrium. However, exponential growth is still appropriate to describe this process, provided a proper observation window is used.

The overall exchange rate (k_{ex}) is limited by the opening and closing rate constants (k_{open} and k_{close}, respectively) of the protein structure. Moreover, k_{ex} depends on the exchange rate of the unstructured peptide (k_{int}). In this model, a single proton exchanges exponentially with a unique rate, k_{ex}, while the exchange rate measured for a peptide is a linear combination of single amide rates.

MS-based HDX experimental settings result in three aspects of data complexity: spectrum analysis, peptide-level resolution, and time trajectory. The most fundamental one is the analysis of the full spectrum of the signals within the shifted isotope envelope of the peptide, which provides complete information on the exchange and allows the study of the kinetic limits of the exchange.[4] The second level, most commonly observed in proteomics studies, relies on converting spectra to their centroids, the intensity-weighted average of an isotopic envelope. While it loses much information about the distribution of exchanging proton subpopulations, limiting the ability to interpret HDX−MS measurements in depth, the researchers often value the simplicity offered by this approach.[5]

Analyzing deuterium uptake, even when centroids are considered only, remains challenging because the resolution of its measurements is limited to the peptide level. As a result, any variations in deuterium uptake along the peptide length are undetectable, making it difficult to pinpoint specific subregions or residues involved in dynamic processes or interactions. This limitation hinders linking changes in peptide uptake to residue-level structural events occurring locally within a protein.

Classic HDX−MS suffers from resolution limited to the peptide level, as exchange is analyzed by measuring whole peptide masses by the mass spectrometer. Hybrid mass spectrometers enable fragmentation of the peptide, but soon it became clear that the most commonly used collisional fragmentation technique (CID—or other slow-heating techniques) leads to redistribution of protons/deuterons. This phenomenon, leading to complete randomization of in-solution deuterium label distribution along the sequence, has been named "H/D scrambling" (see for instance[6,7]). Efforts to overcome H/D scrambling by using alternative, prompt fragmentation methods (ETD/ECD/MALDI ISD/UVPD), were carried out,[8] but they have their own limitations, like the availability of specialized instrumentation or the inferior efficiency of fragmentation; therefore, they did not go into routine use up to now.

The third aspect of HDX−MS complexity arises from the fact that peptide deuterium uptake is measured at specific time points. Since analysis of a single time point could be misleading, it is vital to consider the trajectory of the deuterium uptake, representing it as a time-dependent variable. However, converting a series of measurements into a single value (e.g., deuterium uptake rate or rate superposition) adds another layer of complexity.[9]

https://doi.org/10.1021/acs.jproteome.4c00700
J. Proteome Res. 2025, 24, 1688−1700

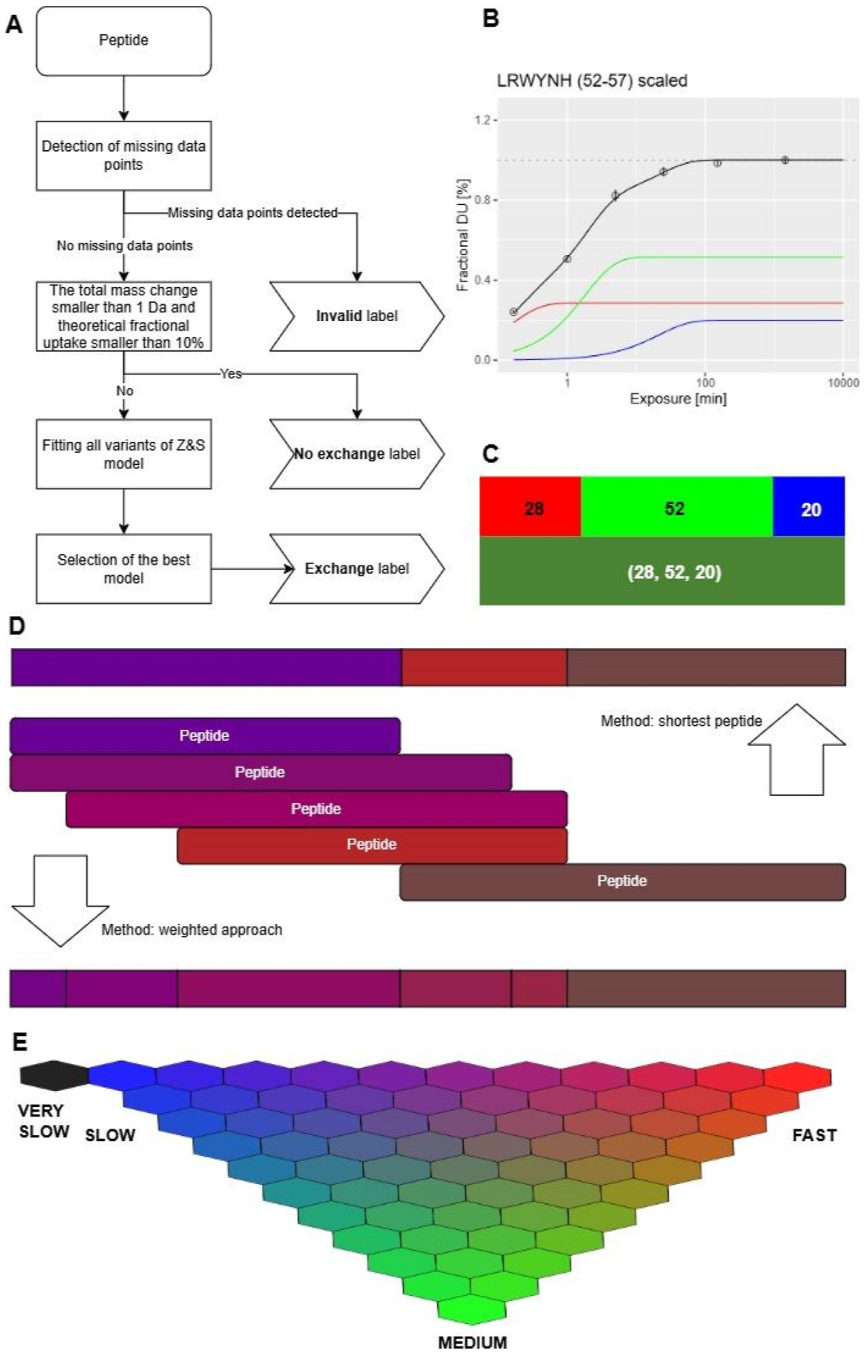


Figure 2. (A) HRaDeX algorithm. First, the data set is checked for completeness of time data points. Peptides with missing measurements or lacking uptake control are marked as "invalid" and excluded from further analysis. Next, we detect uptake trajectories with no or a very slow exchange and assign them with a "no exchange" label. We fit the extended ZS model to all curves not assigned to these categories. Then, we select the best model, considering the BIC, and annotate the peptide with the label "exchange". (B) Uptake curve for exemplary peptide LRWYNH with the fitted model (black line). The normalized experimental uptake is marked with circles. The colored lines present the components of the fitted model (red—fast, green—medium, and blue—slow). (C) Three colored blocks represent the percentage of the total abundance of each component of peptide LRWYNH, and the final color is constructed using the RGB palette, as described in eq 5. (D) Simplified scheme of the high-resolution HDX−MS data analysis. The high-resolution exchange rates are computed using two methods: "shortest peptide" (each residue is annotated with the exchange groups of the shortest peptide that contains it) and "weighted average" (each exchange group is computed as the weighted average of exchange groups of all covering peptides, with weights inverse to their theoretical maximum uptake). (E) Color RGB palette of possible classification results.

HDX−MS can also be used for comparative studies of unfolding events in two different biological states of a protein. The differences between the two HDX profiles reflect the impact of the biological conditions on protein dynamics. Examples cover scenarios as diverse as the impact of the point mutation on protein stability,[10] determination of binding sites,[11] or confirming allosteric regulation of enzymes.[12] However, the comparative studies suffer from the limitations of classic HDX−MS.

https://doi.org/10.1021/acs.jproteome.4c00700
J. Proteome Res. 2025, 24, 1688−1700

Thus, we propose a novel method, HRaDeX, for high-resolution HDX−MS data analysis for both single-state and comparative studies that addresses the multidimensional aspect of HDX−MS data. HRaDex is a pipeline consisting of several building blocks, merging elements already known in the community with new approaches combined innovatively. It presents an easy-to-use workflow that reflects the full complexity of uptake curves in a set of parameters of a simple mathematical model and visualizes them on a single panel. It keeps the interpretability of the analysis based on the weighted mass center, unearths information hidden in overlapping peptides, and summarizes the deuterium uptake trajectory into the deuterium uptake rate. To facilitate high-throughput analysis, we share our tool as an R package, but to accommodate less programmatically fluent users, we have also developed a web server.

To test HRaDex, we carried out analyses using several data sets available in public repositories. In total, the ability of HRaDex to reconstruct experimental uptake curves was tested for more than 4000 peptides.

METHODS

Extension of the ZS Model

Several mathematical models describe HDX, treating the mass of a peptide as a function of time. One of them, a Zhang and Smith model (named here as ZS), groups the exchanging amides into three exchange groups based on their exchange rates: slow, intermediate, and fast exchange.[13,14] Employing this three-component model facilitates using intuitively understood and easily interpretable deuterium uptake rates formalism directly translatable to deuterium uptake levels.[15]

In the ZS model, mass increase during HDX is defined as

$$D = n_1(1 - \exp^{(-k_1 t)}) + n_2(1 - \exp^{(-k_2 t)}) + n_3(1 - \exp^{(-k_3 t)}) \quad (1)$$

where D is the mass increase in a chosen unit (preferably fractional), and t is the exposure (time of the exchange—usually in minutes). This model explicitly assumes that the uptake kinetics can be approximated by three subpopulations of amides (n_1, n_2, and n_3) with respectively fast, intermediate, and slow exchange rates (k_1, k_2, and k_3) (Figure 1A). The original presentation of ZS assumes that the third (slow) component is approximated by using a linear function

$$D = n_1(1 - \exp^{(-k_1 t)}) + n_2(1 - \exp^{(-k_2 t)}) + n_3 k_3 t \quad (2)$$

However, we discovered that the assumption of three exchange rate groups is not necessary in cases where a majority of amides undergo HDX at a comparable rate. Instead of addressing this issue by introducing a linear component (as provided in eq 2), we consider scenarios where one or two exchange groups (n_1, n_2, or n_3) could be nonexistent, leading to the models with only two or one exponential components

$$D = n_A(1 - \exp^{(-k_A t)}) + n_B(1 - \exp^{(-k_B t)}) \quad (3)$$

$$D = n(1 - \exp^{(-kt)}) \quad (4)$$

This modification resulted in a flexible set of models that could be applied in any situation where the curve exhibits even the smallest growth. Despite that, the ZS model cannot handle the situation with a lack of observable exchange. We define the "no exchange" variant as a scenario where, over the whole course of the HDX experiment, the mass of a peptide increased by less than 1 Da and did not exceed 10% of theoretical deuterium uptake to not reject shorter peptides. In such a situation, we observe that the measurement points oscillate randomly around the mass of a peptide, but there is no growth tendency, or it is negligible (Figure 1B). Thus, the ZS could not be fitted, as its exponential nature requires at least minimal growth. This problem does not exist in the opposite case (immediate exchange), as such curves exhibit growth (Figure 1C).

HRaDeX Algorithm

The idea behind the HRaDeX algorithm is to describe the H/D exchange using the extended ZS model and compute the exchange rates for all available overlapping peptides. First, we detected the peptides with missing measurements. Since the community recommends using exchange controls,[16] we label all peptides without them as "invalid" (Figure 2A). Similarly, we annotate as "invalid" peptides with less than two measurements (excluding controls) as they often lead to erroneous fits.

Next, we proceed to detect the lack of or extremely slow exchange ("no exchange"), which we define as the situation where the mass uptake in the last time point is lower than 1 Da and lower than 10% of theoretical exchange (using maximal possible uptake based on the peptide sequence). Additionally, we also mark subfragments containing only proline as "no exchange".

For all peptides that are not annotated as "invalid" or "no exchange", we fit the extended ZS model using the Levenberg−Marquardt nonlinear least-squares algorithm, imposing the lower limit for the n value equal to the inverse of the number of exchangeable amides in a peptide (Figure 2B). The deuterium uptake is normalized to consider back-exchange according to the equation: $(m - m_0)/(m_{100} - m_0)$, where m is the mass of a peptide, m_0 is the peptide mass before the exchange, and m_{100} is the mass of a fully deuterated control. Despite the recent controversies surrounding this method, we decided to employ it, as there are still no other satisfactory solutions.[17]

The suggested workflow (3/2/1) tries to fit models with three, two, and one exponential components. However, a workflow that considers only a subset of these models (3/1 and 2/1) is possible and especially applicable for data sets with a small number of time points.

The values of exchange rates k delimiting the low, intermediate, and high exchange have default values resulting from our experience with analyzing multiple HDX−MS data sets. However, we allow users to alter these default exchange rates, k depending on the nature of the exchange in the analyzed protein system. In the spirit of reproducible analysis, the limits of k should be reported in the extended HDX−MS data summary.[16] To find which model reflects the HDX of a specific peptide and avoid overfitting, we use the Bayesian Information Criterion (BIC), which allows for a comparison of models with different numbers of parameters.

To enable visualization of the distribution of exchange groups (n_1, n_2, and n_3) on a single panel along the entire protein sequence, we decided to use an RGB (red, green, and blue) color code. The rates are converted to colors using the following equation

https://doi.org/10.1021/acs.jproteome.4c00700
J. Proteome Res. 2025, 24, 1688−1700

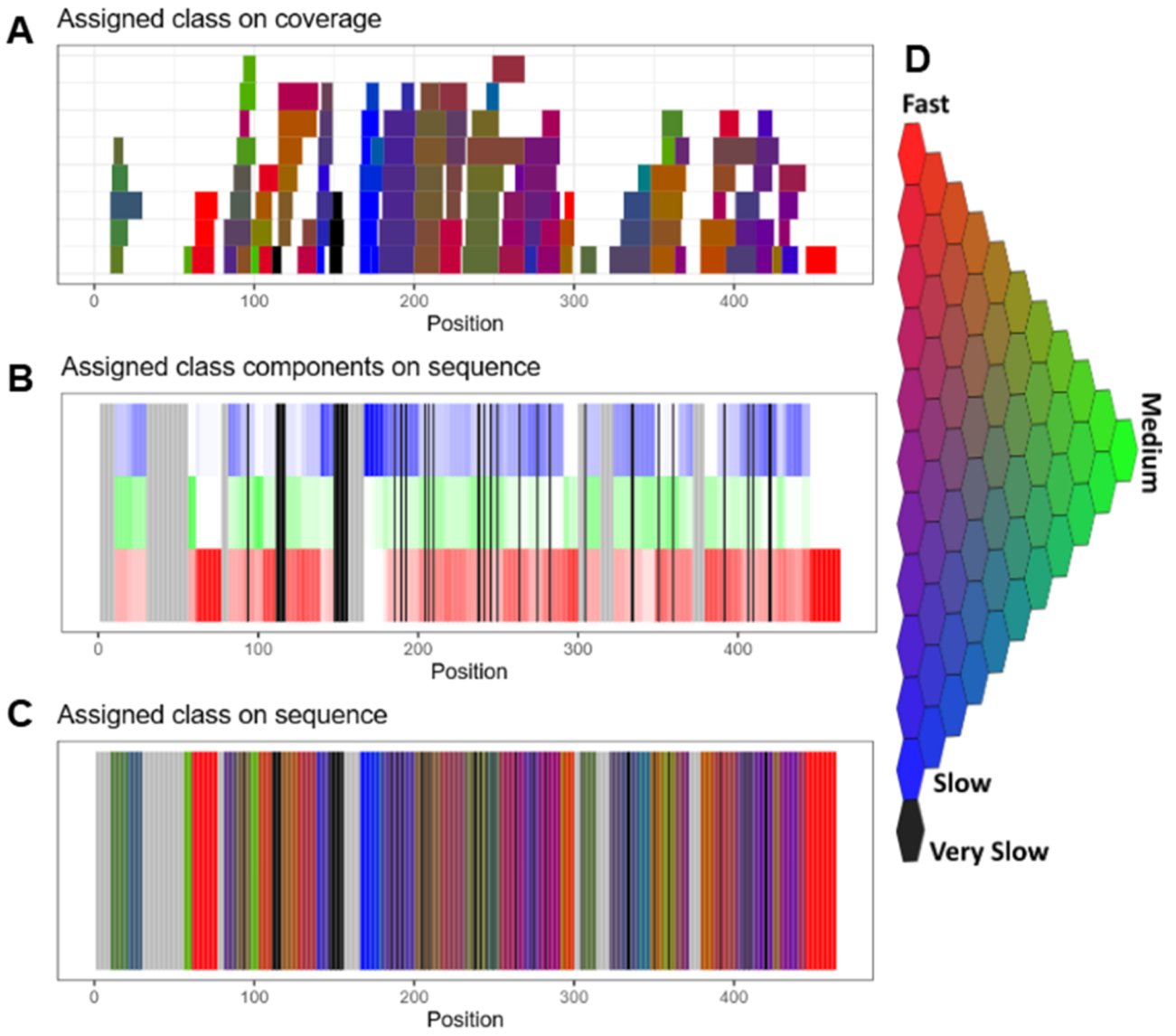


Figure 3. Example of the overlay of color-coded HDX results on protein sequence. Data for the eEF1A1 protein (PRIDE PXD056814) are shown. Panel D shows a color palette enabling the description of fast (red), intermediate (green), or slow (blue) components as well as mixed states whenever the uptake curve indicates in a given peptide the presence of exchange rates assigned to different exchange classes. Color palette can be used directly on peptides (panel A) or after the aggregation step (see text) into a single position barcode, either with three color channels separated (panel B) or aggregated using the panel D palette (panel C) and one of the two aggregation methods described in the text. Results using the "weighted average" approach are shown. The bar code shown in panel C can be subsequently used to overlay on the structural protein model.

$$(r, g, b) = \frac{(n_1, n_2, n_3)}{n_1 + n_2 + n_3} \quad (5)$$

Thus, the main color indicates the dominant population, and the changes in its hue signal the presence of other populations (Figure 2C). Because of the constraint stemming from eq 2, the peptides could be assigned only particular colors instead of the whole color spectrum described by the RGB scale in Figure 3 panel A. Thanks to that, we utilize otherwise unused colors (gray and black, respectively) to represent the "Invalid" or "No exchange" labels.

The color encoding offers a quick visual summary of the exchange rate distribution along the protein sequence (Figure 3A). When there are fewer than three exponential components in the final model, HRaDeX checks which exchange groups (low, intermediate, and high exchange) the resulting rates belong to and treats the other exchange group abundances as zeros. The computation of the color encoding is described in detail in Supporting Information S1 (Figures S1−S4).

The aggregation step collates exchange rates of overlapping peptide fragments to increase the resolution of the final group assignments (Figure 2D). HRaDeX provides two alternative ways to solve this problem. The first approach, the so-called "shortest peptide", assigns the HDX parameters of the shortest peptide that covers a given residue. If multiple peptides of equal length overlap a single residue, we prioritize the peptide closest to the N-terminus of a protein.

The second approach, "weighted average", uses residue averaging to assess the HDX of a single residue.[18] In the first step, this method computes the weighted average of all of the exchange groups assigned to the peptides. The weights are inversely proportional to the theoretical maximum uptake. Following the original implementation, which omits the two residues in the peptide sequence to correct for the back

https://doi.org/10.1021/acs.jproteome.4c00700
J. Proteome Res. 2025, 24, 1688−1700

exchange (deuterium/hydrogen exchange), we provide this functionality as an option.

The final output of the HRaDeX algorithm takes the form of a set of exchange rates (k) and a fraction of amides belonging to each exchange group (n) assigned to each of the residues of the sequence. In principle, a single amide should be assigned to one of the three groups, fast, intermediate, or slow, but the data do not allow for such an assignment to be obtained. Peptide-level data can be explained by different distributions of exchange rates for all amides along the sequence. Therefore, during the aggregation process, the most probable exchange group is indicated at each residue, as judged by the highest fractional population of an exchange group assigned. We leverage this information on a so-called barcode plot, where we assign a resulting color to each position in the protein sequence covered by at least one peptide (Figure 4C).

The HRaDeX algorithm can be effortlessly extended to comparative analysis, the results of which are traditionally visualized on a Woods plot. Since HRaDeX provides high-resolution results only for a single biological state, we introduce a companion algorithm, compaHRaDeX, to facilitate comparative analysis. To do so, we define d as the distance between the two vectors of exchange groups abundance in states A and B for each residue

$$d = \sqrt{(n_{1B} - n_{1A})^2 + (n_{2B} - n_{2A})^2 + (n_{3B} - n_{3A})^2} \quad (6)$$

The value of d correlates with the difference in the exchange group assignments, indicating the difference in the exchange pattern of a residue in two biological states. Although this approach could enable the comparison of biological states with different digestion patterns, we strongly recommend using identical peptides for comparative studies.

Graphical User Interface

To streamline the usage of the HRaDeX algorithm, we developed a user-friendly graphical interface (see Figure S5). The interface allows users to perform the HRaDeX analysis on their protein using the data in cluster file format (as offered by the DynamX software from Waters Corporation). The graphical user interface (GUI) facilitates in-depth analysis of the results, starting with the high-level overview, including barcode plots. To increase the accessibility of our software, each element is enhanced by tooltips that provide the exact information on the fit and resulting exchange rates. Moreover, the results are available in a tabular format. HRaDeX also contains advanced functions for investigating exact fits and their BICs if there is a need for more low-level data analysis.

The analytical workflow is fully customizable, as users can influence the complete procedure of fitting the HDX models, e.g., by redefining the boundaries of the exchange groups. If the user provided the protein structure in a PDB file format, the computed exchange rates could also be presented on each residue belonging to the three-dimensional (3D) protein model. All of the tables and figures are fully downloadable to facilitate reporting the results.

For comparative analysis of HRaDeX outputs, we developed a dedicated compaHRaDeX tool (Supporting Information Note S1, Figure S6). It allows comparisons of either the same protein in different biological states or different HRaDeX fits in the same biological state. Currently, compaHRaDeX supports a comparison of two HRaDeX results.

Since some users might prefer programmatic access to HRaDeX, we also share it as an open-source R package (https://github.com/hadexversum/HRaDeX). The package contains all of the functionalities available in the web server and might also be used to deploy a local instance of the HRaDeX and compaHRaDeX GUIs. In addition, the open-source distribution of HRaDeX provides users with a transparent codebase and allows for external code validation.

HDX Experiment with Cytochrome C

Prior to HDX experiments, nondeuterated fractions of protein served as a source of peptide lists. For this purpose, the LC−MS analysis was carried out with all steps the same as described below for HDX runs, but in this case, D_2O, used for exchange, was substituted by H_2O. Peptides were identified using ProteinLynx Global Server Software (Waters).

For the HDX reaction, the starting stock protein concentration of cytochrome *C* (from equine heart) was 50 μM. A stock solution of cytochrome *C* (Fluka Analytical, purity ≥90%, #30396) was prepared with Iron(III) chloride. HDX exchange incubations were performed at six time points (10 s, 1 min, 5 min, 30 min, 2 h, 5 h, and 24 h) in 4 replicates each. 5 μL aliquots of protein stocks were added to 45 μL of deuterated buffer (50 mM sodium phosphate buffer pH 7.0, uncorrected meter reading) at room temperature (20 °C). Final concentrations of protein in the deuteration reactions were 5 μM.

The H/D exchange reactions were quenched by moving the exchange aliquots to precooled tubes (on ice) containing 10 μL of quenching buffer (2 M glycine, in 99.95% D_2O, pD 2.4). After quenching, samples were frozen immediately in liquid nitrogen and kept at −80 °C until MS measurement. Samples were thawed directly before measurement and injected manually onto the nano ACQUITY UPLC BEH C18 column (130 Å, 1.7 μm, 2.1 mm × 50 mm, 60 °C; Waters, #186002350) system equipped with HDX−MS Manager (Waters). Proteins were digested on 2.1 × 20 mm columns with immobilized Nepenthesin-2 (AffiPro), for 1.5 min at 20 °C and eluted with 0.07% formic acid in water at a flow rate of 200 μL/min. Digested peptides were passed directly to the ACQUITY BEH C18 VanGuard precolumn, from which they were eluted onto the reversed-phase ACQUITY UPLC BEH C18 column (Waters) using a 10−35% gradient of acetonitrile in 0.01% of formic acid at a flow rate of 90 μL/min at 0.5 °C. Samples were measured on the SYNAPTG2 HDX−MS instrument (Waters). The instrument parameters for MS detection were as follows: ESI - positive mode; capillary voltage − 3 kV; sampling cone voltage − 35 V; extraction cone voltage − 3 V; source temperature − 80 °C; desolvation temperature − 175 °C; and desolvation gas flow − 800 L/h.

Two control experiments were conducted to assess the minimum and maximum full deuteration (FD) H−D exchange level. For obtaining the minimal exchange of each peptide (Mmin), 10 μL of a quench buffer was mixed with 45 μL of D_2O reaction buffer (50 mM sodium phosphate pH 7.0) prior to the addition of 5 μL of protein stocks and analyzed by LC−MS. To obtain the FD H/D exchange control, the deuteration reaction was conducted in a neutral pH buffer on lyophilized cytochrome *C* peptides collected from the immobilized Nepenthesin column. The FD sample was then processed in the LC−MS system as all other samples. The control experiments were also performed in quadruplicate.

The peptide lists obtained using nondeuterated protein samples were used to analyze the exchange data using DynamX 3.0 (Waters) software. The PLGS peptide list was filtered by

https://doi.org/10.1021/acs.jproteome.4c00700
J. Proteome Res. 2025, 24, 1688−1700

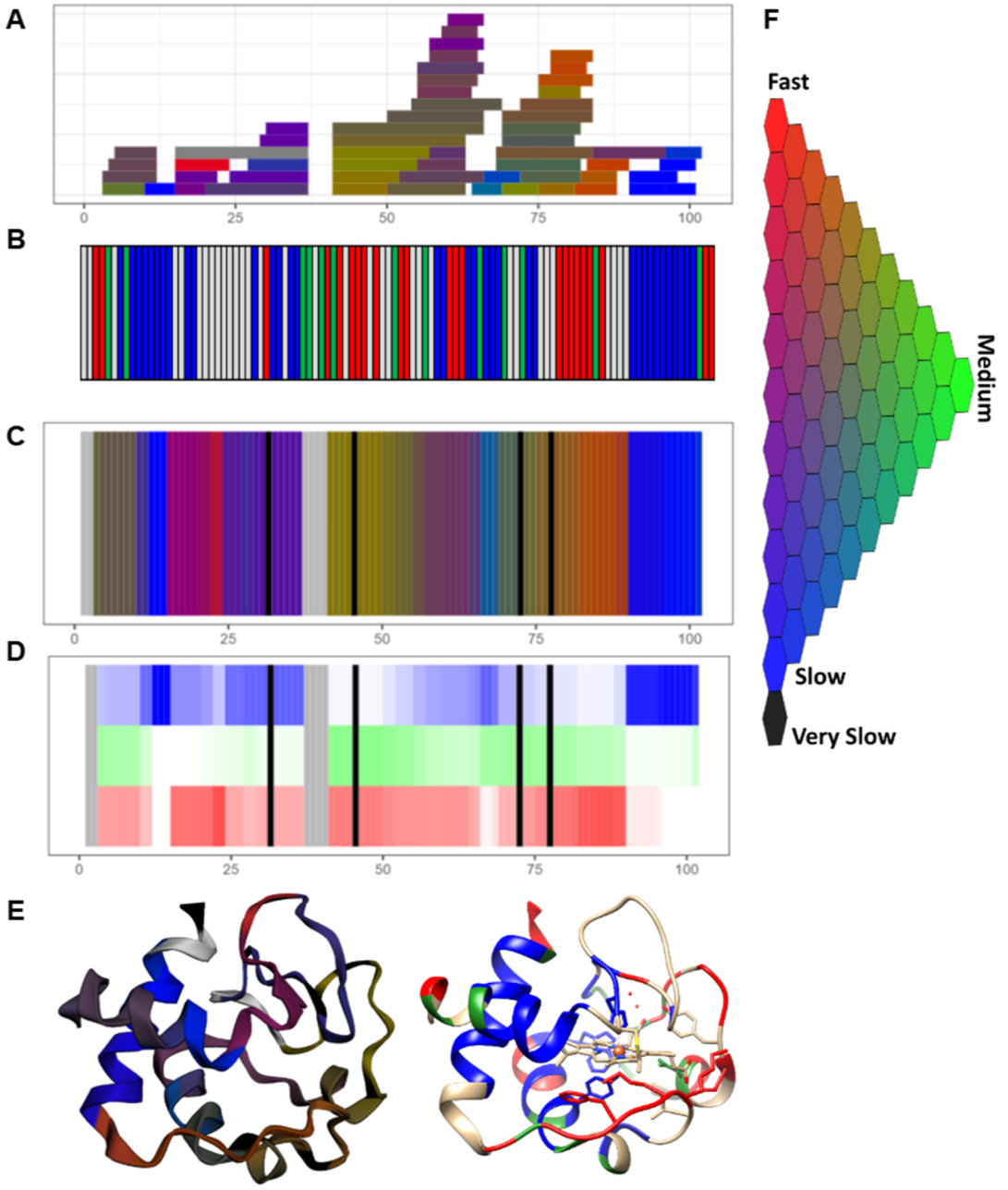


Figure 4. Comparison of color-coded assignment of exchange class to regions of cytochrome *C*. The assignment was carried out either by using HRaDex or a set of single-site exchange rates measured by NMR or ETD−MS−HDX previously (see text). Single residue rates measured by NMR or ETD−MS−HDX were assigned to one of the three classes used in this work: (1) fast (k_{ex} > 0.2 min^{-1}) in red, (2) intermediate (0.01 < k_{ex} < 0.2 min^{-1}) in green, and (3) slow (k_{ex} <0.01 min^{-1}) in blue. (A) HRaDex results obtained on the peptide level for all detected peptides along the cytochrome *C* sequence. (B) NMR/ETD results for all positions along the sequence. Regions of slow single-site rates indicated by NMR/ETD (blue), in panel A, correspond to peptides for which the dominating color is blue, therefore assigned as "slow". (C) Results of HRaDex analysis on subregion level (aggregated using "weighted approach") in the form of a merged RGB color coding strip. Black sticks mark prolines. (D) Results of HRaDex analysis on subregion level with separation of each color component intensity into the three color channels—red, green, and blue. Also, in this type of analysis, the correspondence between stable and dynamic regions assigned by HRaDex and previous NMR/ETD data is clear. (E) Comparison between HRaDex subregion analysis (left) and previous NMR/ETD data (right), by overlay on the cytochrome *C* structure (PDB code 1 hc), illustrating the localization of the protein core within a spatially proximal three helix system (left side in diagrams), while the rest of the protein is characterized by increased dynamics. Irrelevant to the type of HRaDex analysis, the agreement between HRaDex and previous data is satisfactory. The movie demonstrating these results is available as a Supporting Information Movie S1. (F) RGB color palette.

minimum intensity criteria—3000 and minimal product per amino acid—0.3. All raw files were processed and analyzed with DynamX 3.0 software. All MS assignments in Dynamix were inspected manually.

RESULTS

To illustrate the capabilities of the HRaDeX framework, we present results obtained for two proteins in two case studies sections.

https://doi.org/10.1021/acs.jproteome.4c00700
J. Proteome Res. 2025, 24, 1688−1700

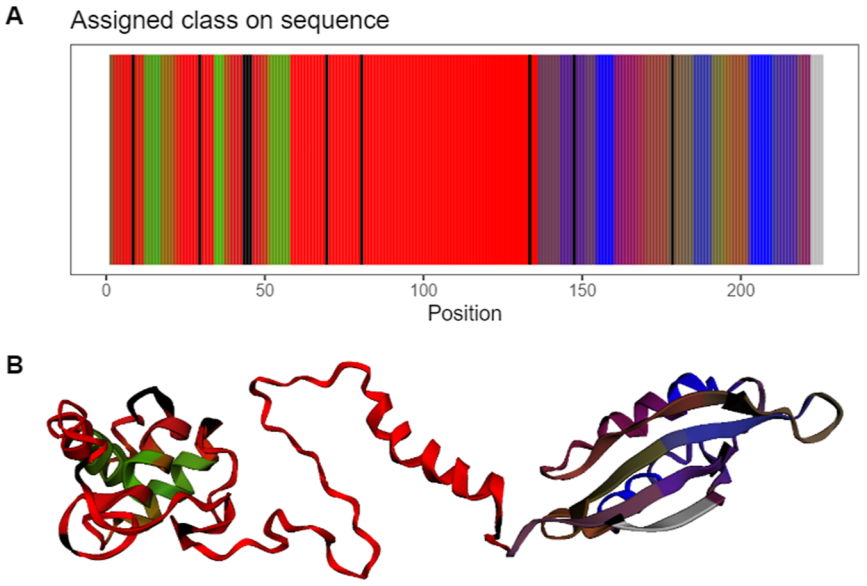


Figure 5. Results of the HRaDeX analysis of the eEF1Bα HDX−MS experiment. (A) Barcode plot of eEF1Bα. (B) HRaDeX output is visualized on the predicted structure of eEF1Bα.[21]

Case Study 1

We have tested HRaDex using HDX data obtained by us for cytochrome *C*, one of the typical, easily available proteins, for which also single-site HDX k_{ex} values have previously been measured using other methods.[19,20] HRaDex results were then compared with previous data by inspecting their overlay on the cytochrome *C* sequence (Figure 4 panels A,B) and on the cytochrome *C* structure (PDB code 1 hc—Figure 4 panel E). The comparison shows an agreement in the placement of most stable regions in a three-helical core of the protein detected in the same regions by HRaDex analysis and localized by previous results. Also in agreement are regions of increased dynamics.

Case Study 2

To showcase the capabilities of the HRaDeX framework on a protein of higher dynamics, we have reanalyzed the HDX−MS measurements of the eEF1Bα subunit of the human guanine-nucleotide exchange factor (GEF) complex (eEF1B).[21] The raw data was obtained from the ProteomeXchange Consortium (data set id: PXD031783) and processed using the DynamX 3.0.0 program (Waters) with the following acceptance criteria: minimum intensity threshold of 3000, minimum products per amino acid of 0.0, minimum score of 7.5, and the deviation from the theoretical value for parent ion mass below 10 ppm and without any limits for the peptide length. The preliminary data analysis was conducted with the HaDeX software[22] and presented in Supporting Information Note S1 (Figures S7−S8).

Next, we conducted HRaDeX analysis of the eEF1B subunit α (eEF1Bα), using the default settings of HRaDeX, and the results of this analysis are presented in Figure 5A. The high-resolution analysis confirms stability in distal domains and high exchange levels in the linker region (residues 58−136). However, the HRaDeX result also highlights the presence of several prone-to-exchange regions of the N-terminal and C-terminal domains, for both of which the X-ray structure is known since the model of full-length eEF1Bα was constructed using the following X-ray domain structures: isolated GEF domain (PDB ID: 1B64) and its N-terminal domain in complex with the GST-like domain of eEF1Bγ (PDB ID: 5DQS) and a highly homologous C-terminal region of eEF1Bβ (PDB ID: 2N51) as templates. In agreement with HRaDex, the analysis of protein structure indicates the presence of an exposed linker in the interdomain region, as presented in the structure of eEF1Bα predicted using X-ray structures as templates (Figure 5B). The detailed analysis, including the intermediate steps and the rationale behind the barcode plot, is available in Supporting Information 1 (Figures S9−S11).

The benefit from the high-resolution analysis is also evident in the comparative analysis. eEF1Bα undergoes conformational changes upon binding to the eEF1Bγ subunit in the eEF1B complex. CompaHRaDeX reveals a decrease in the protection of the C-terminal domain and an increase in the protection of the N-terminal domain (Figure 6). The authors of the original publication reporting these data indicated that this opposite effect might be caused by the involvement of the C-terminal domain in the binding interface. Supporting Information Note S1 (Figures S12−S14) discusses the comparative analysis and describes in detail the computation of the distance metric (also consult eq 6).

Validation of the High-Resolution Analysis

To verify the accuracy of the high-resolution analysis results, we conducted a retrospective analysis of several HDX−MS experiments (PXD057294, PXD057066, PXD057057, PXD056814, PXD053893, PXD039682, and PXD057902) enhanced with the pyHDX data set.[23]

We analyzed (a) the overall quality of ZS model fits (peptide-level validation) and (b) the ability to reconstruct original trajectories of HDX from the high-resolution data (subpeptide-level validation).

Peptide Level Validation

For peptide-level validation, the residual sum of squares (RSS) was calculated between experimental deuterium uptake values at time points of the uptake curve and those from ZS model fit parameters using the following formula

$$\mathrm{RSS} = \sum_{i=1}^{N} (x_i - \widehat{x_i})^2 \tag{7}$$

https://doi.org/10.1021/acs.jproteome.4c00700
J. Proteome Res. 2025, 24, 1688−1700

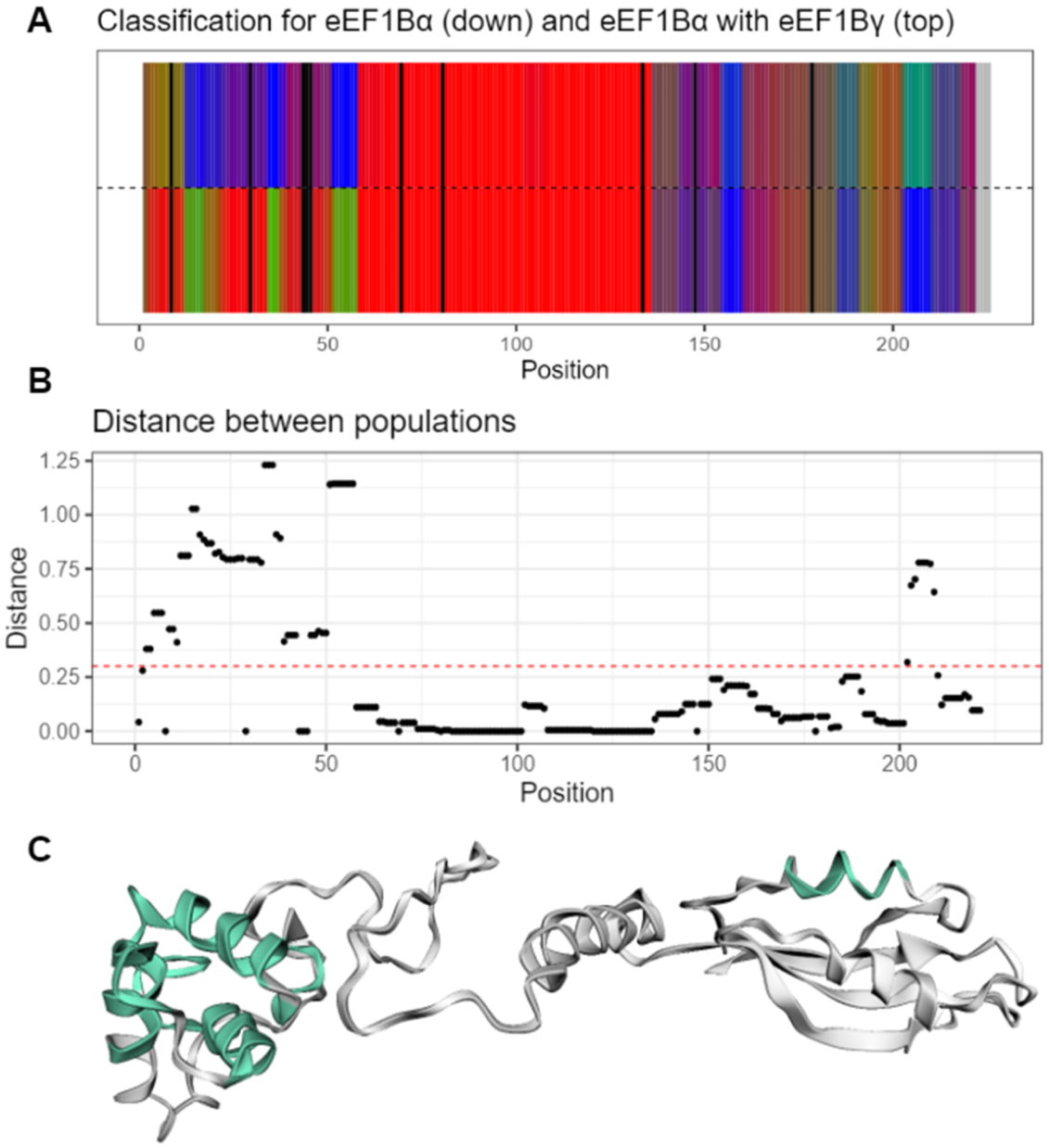


Figure 6. (A) Barcode plot of the compaHRaDeX analysis of eEF1Bα exchange in isolation and in complex with eEF1Bγ subunit, computed with the weighted average approach. (B) Distance (eq 6) between assigned exchange classes of eEF1Bα exchange in isolation and in complex with eEF1Bγ subunit. The red dashed lines indicate the exemplary threshold of 0.3. (C) Regions with an exchange distance above the threshold of 0.3 are marked with the cyan color on the predicted structure of eEF1Bα.[21]

where x_i is the normalized experimental deuterium uptake at a given time point and $\hat{}_{x_i}$ is the recovered deuterium uptake at a given time point.

Results for all data sets are presented in Figure 7 divided into panels representing RSS of fit to three-rates variant (eq 1), two-rates fit (eq 3), and one-rate fit (eq 4). Average RSS is the lowest for the first variant, but in general, the results show that presented here the implementation of the ZS model is suitable for the representation of experimental uptake curves.

RSS, commonly used to measure the overall variance that the model cannot explain, showcases that our extended ZS mode fits the HDX data very precisely. Although we expected this result considering the widespread acceptance of the ZD model, it is vital to highlight that the average RSS for 1-, 2-, and 3-component models are 0.0469, 0.0064, and 0.0001. Moreover, these fits result from running HRaDeX with default parameters, suggesting that manual supervision can lead to even better performance.

Subpeptide Level Validation

We grouped the HRaDeX deuterium uptake rates of amino acids belonging to each peptide, reconstituted the resultant uptake trajectory, and calculated the root-mean-square error between the recovered peptide deuterium uptake and the experimental peptide deuterium uptake.

We computed RMSE using the following formula

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{N} (x_i - \hat{x}_i)^2}{N}} \tag{8}$$

https://doi.org/10.1021/acs.jproteome.4c00700
J. Proteome Res. 2025, 24, 1688−1700

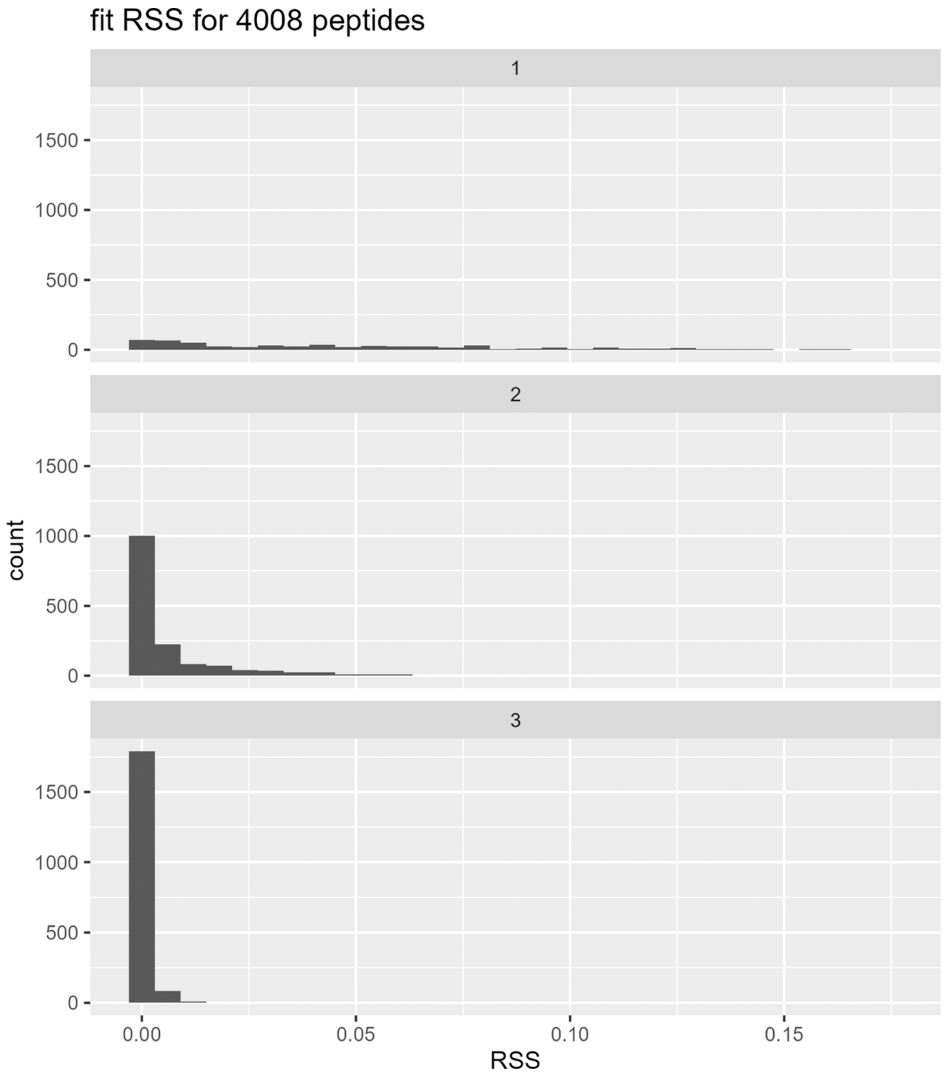


Figure 7. RSS of fits for peptides considered in HRaDeX validation. Each facet represents a different ZS model variant (eqs 1, 3, and 4, respectively).

where N is the number of time points, x_i is the experimental deuterium uptake at a given time point, and $\hat{x}_i$ is the recovered deuterium uptake at a given time point. The average RMSE for the "weighted average" aggregation method is 4.16%, and for "shortest peptide", it is 4.49%, while the normalized data range from 0 to 100%.

To expand the performance discourse, we present the results of retrospective high-resolution validation of a set of experimental data that consists of 8 different data sources: 7 from the PRIDE database (PXD057294, PXD057066, PXD057057, PXD056814, PXD053893, PXD039682, and PXD057902) and an exemplary data set from pyHDX. The collected data contain 19 source files, 41 biological states, and 4224 peptides. The number of times of incubation varies between the data sets. Figure 8 presents the average RMSE between curves reconstituted from high-resolution analysis parameters and experimental uptake curves for all data sets tested for both aggregation methods (complete analysis available in Supporting Information, Figures S15−S66, Tables S1−S2).

DISCUSSION

The high-resolution analysis of HDX−MS data has been a long-term goal that has been pursued for many years. Although the first solutions were available only as MATLAB scripts (Kan et al., 2013), several tools are now available as open web servers that readily provide their output in real-time.[23−27] However, these approaches do not provide straightforward web servers, and we would like to complement these approaches by providing a simpler and easier-to-interpret high-resolution analysis of HDX−MS data. The straightforwardness of our tool is defined by three factors: setup, input, and data processing (Table 1).

The setup of our tool is as minimalistic as reaching out to the HRaDeX web server, which is supported by most modern browsers. However, we also share our tool as an open R package. Besides keeping the code readable and accessible to all researchers who want to scrutinize it, integration with the R ecosystem streamlines the local deployment of HRaDeX for high-throughput analyses.

Although the existing tools for high-resolution HDX−MS use inputs as varied as isotopic envelopes or centroids,[25] we opted to choose weighted mass center data as it is a more common means of exchanging HDX−MS information. More-

https://doi.org/10.1021/acs.jproteome.4c00700
J. Proteome Res. 2025, 24, 1688−1700

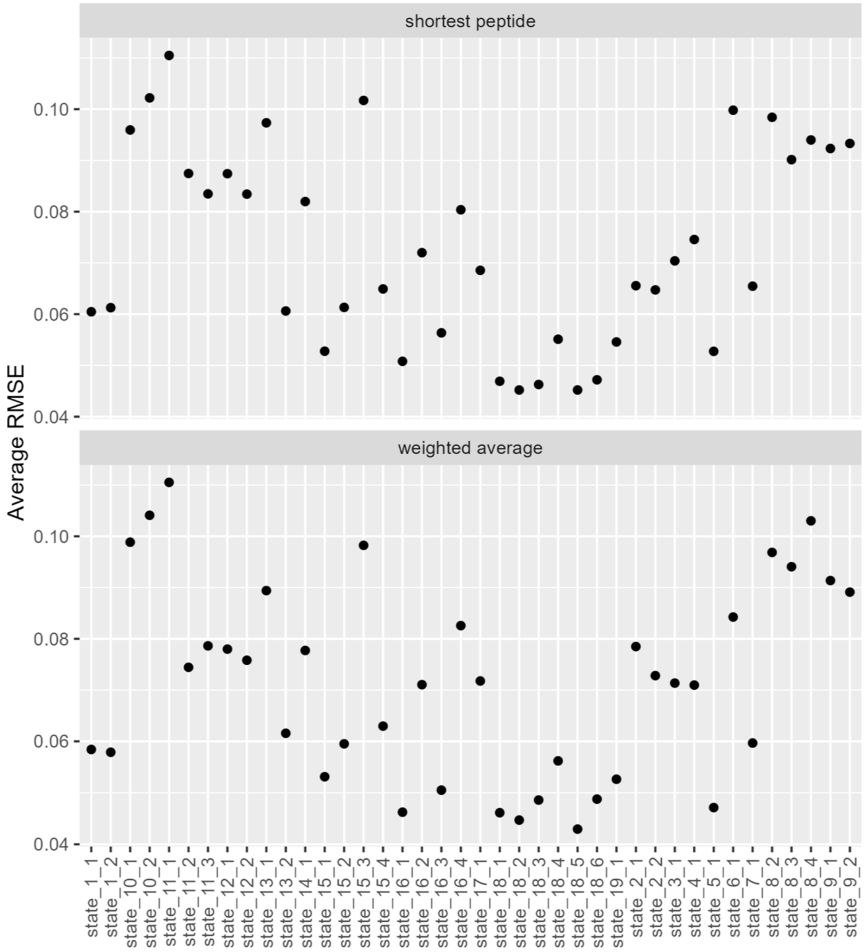


Figure 8. Summary of averaged RMSE for all gathered data sets, obtained in the HraDeX analysis using default parameters, calculated between experimental uptake curves and curves reconstituted from high-resolution analysis parameters obtained either by the "shortest peptide" (upper panel) or "weighted average" (lower panel) approach of Keppel and Weis.[18] Details of the process are available in the Supporting Information Note S1 (Figures S67–S68).

Table 1. Summary of HRaDeX Functionalities

property	description
Input	single-state experiment: experimental data file in "cluster" format from DynamX (Waters)
	comparative experiment: two output files from HRaDeX, processed separately for the experimental options
Functionality	high-resolution analysis of HDX–MS data using the extended ZS model
Output	single-state experiment: high-resolution results are presented in tabular form. HRaDeX offers both 2D and 3D visualization of its output
	comparative experiment: differences between high-resolution results, in tabular, numerical form, and visualizations in linear or 3D form

over, analysis on this level considers the changes introduced during manual curation, which is one of the most crucial stages of HDX–MS data analysis.[28]

One of the main advantages of our tool lies in its reliance on the expanded ZS model. By using this well-studied and widely approved approximation of HDX, we provide the output as directly interpretable deuterium uptake rates instead of a 2D matrix of exchange levels at incubation times. Moreover, our approach is easy to validate as the user can inspect every algorithm step and, for example, pinpoint an incorrect fit of the exchange curve to the experimental data. The capabilities of this manual validation extend also to the computation of the high-resolution deuterium uptake trajectory. A user can also recover the original deuterium uptake data and decide if the chosen aggregation method introduces unexpected artifacts. These validation tools were used in this work to benchmark HRaDex efficiency both in the case of peptide-level analysis and on the level of subpeptide, high-resolution, results. For peptide level analysis, we calculated RSS between the experimental uptake curve and the one reconstituted from ZS model fit parameters for the whole test data set containing uptake curves of >4000 peptides. The overwhelming majority of RSS values fall below 0.015, which indicates that obtained fits are very good. In conclusion, these tests, carried out on numerous data sets from various sources, verified the applicability of the ZS approach to approximate HDX results, which, in the majority, can be well reflected by a simple set of one to three rates. For subpeptide analysis tests, we calculated RMSE values between the experimental deuterium uptake curve and the curve recovered from a set of single residue parameters obtained either by the "shortest peptide" or "weighted average" approach. With an average RMSE of 7.15%, subpeptide analysis results also reflect experimental data sets very well.

Despite these advantages, the ZS model shares limitations typical of mathematical modeling. The fit quality is also

https://doi.org/10.1021/acs.jproteome.4c00700
J. Proteome Res. 2025, 24, 1688–1700

strongly influenced by the number of time points and data points (technical replicates or charge states) available for each time point. To address this, we identify cases where the user provides data with fewer than six time points per peptide and recommend using ZS with a reduced number of exponential components, effectively resulting in a scenario where one of the exchange groups is nonexistent. Unfortunately, other data processing practices may also impede the fitting of the exchange curve, such as selecting only a single charge state per peptide while ignoring other available charge states.

Moreover, similar to other high-resolution methods, HRaDeX is severely limited by the sequence coverage and peptide redundancy. If a given protein region is covered by only one peptide, then HRaDeX cannot extrapolate any additional information from neighboring peptides. This approach would have to consider the structural information and could be a valuable direction to extend our algorithm.[29] Even if the lack of redundant peptides hampers the high-resolution aspect of HRaDeX analysis, it remains valid as an analytic tool. In such scenarios, HRaDeX provides a simple numerical output that accurately reproduces even complex trajectories of all HDX uptake curves, aggregating information collected at all time points. Such information could be useful for tools presenting HDX−MS data as HDX-Viewer[30] or producing HDX−MS-aware structural ensembles as HDXer.[31]

The performance of our algorithm is highly dependent on various experiment- and protein-specific factors, such as digestion patterns, exchange speed, and number of time points. Thus, to further optimize our algorithm, we require highly diverse data sets to establish a reliable benchmark for other tools in the field. Additionally, these data sets would advance high-resolution HDX−MS analysis by providing a standardized framework for comparison of HRaDeX with other algorithms, ensuring that new developments are rigorously validated across a wide range of experiments.

CONCLUSIONS

In conclusion, HRaDeX provides an intuitive approach to the high-resolution analysis of HDX−MS data as deuterium uptake rates rather than in terms of exchange levels at different incubation times. Upon conversion of this information by color-coding the distribution of exchange groups in RGB format, it allows a single-panel representation of the whole experiment either on protein sequence or on structural model, illustrating exchange kinetic regimes dominating in different protein regions. Alternatively, in the compaHRaDex version, the differences in rate distribution between two states are easily marked on the sequence or structure. We tailored this tool to streamline the analysis of complicated, high-dimensional data and support researchers in generating publication-ready, reproducible reports from their studies.

CODE AVAILABILITY

HRaDeX and compaHRaDeX are freely available as the R packages:

- HRaDeX
  - Command-line interface: https://github.com/hadexversum/HRaDeX.
  - Documentation of command-line interface: https://hadexversum.github.io/HRaDeX/index.html.
  - GUI: https://github.com/hadexversum/HRaDeXGUI.
  - Actively maintained web server: https://hradex.mslab-ibb.pl/.
- compaHRaDeX
  - GUI: https://github.com/hadexversum/compahradex.
  - Actively maintained web server: https://compahradex.mslab-ibb.pl/.

The data used for case studies in this manuscript is preloaded as the exemplary data in a web server, available in the HRaDeX repository, and published as the PXD031783 PRIDE data set. The Cytochrome C data set is published as the PXD057902 PRIDE data set.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jproteome.4c00700.

Color-coded assignment of exchange class to regions of cytochrome *C* (MP4)

Detailed results of the benchmark and tutorial document with detailed step-by-step instructions for performing HDX−MS analysis with HRaDeX; graphical examples of computation of color encoding; HaDeX data processing; rationale behind the barcode plot; comparative analysis of eEF1Bα; average RMSE between curves reconstituted from high-resolution analysis parameters and experimental uptake curves for all data sets; averaged RMSE for all gathered data sets; HRaDeX GUI; compaHRaDeX GUI; and HRaDeX parameters and the benchmark results (PDF)

AUTHOR INFORMATION

Corresponding Author

Michał Dadlez − *Institute of Biochemistry and Biophysics, Polish Academy of Sciences, Warsaw 02-106, Poland;* orcid.org/0000-0002-8749-9224; Email: michald@ibb.waw.pl

Authors

Weronika Puchała − *Institute of Biochemistry and Biophysics, Polish Academy of Sciences, Warsaw 02-106, Poland;* orcid.org/0000-0003-2163-1429

Michał Kistowski − *Institute of Biochemistry and Biophysics, Polish Academy of Sciences, Warsaw 02-106, Poland*

Liliya Zhukova − *Institute of Biochemistry and Biophysics, Polish Academy of Sciences, Warsaw 02-106, Poland*

Michał Burdukiewicz − *Clinical Research Centre, Medical University of Białystok, Białystok 15-369, Poland; Institute of Biotechnology, Vilnius University, Vilnius 10257, Lithuania;* orcid.org/0000-0001-8926-582X

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jproteome.4c00700

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The authors would like to thank members of the Laboratory of Mass Spectrometry (Institute of Biochemistry and Biophysics

Polish Academy of Sciences) and Krystyna Grzesiak (University of Wrocław) for helpful discussions. W.P. and M.D. acknowledge funding from the National Science Centre Poland, PRELUDIUM BIS 1 [2019/35/O/NZ2/03745]. W.P. acknowledges funding from the Polish National Agency for Academic Exchange [PPN/STA/2021/1/00025]. M.B. acknowledges funding from the National Science Centre Poland, SONATA 19 [2023/51/D/NZ7/02847].

REFERENCES

(1) Englander, S. W. Hydrogen Exchange and Mass Spectrometry: A Historical Perspective. *J. Am. Soc. Mass Spectrom.* **2006**, *17*, 1481−1489.
(2) Vinciauskaite, V.; Masson, G. R. Fundamentals of. *Essays Biochem.* **2023**, *67*, 301.
(3) Hvidt, A.; Nielsen, S. O. Hydrogen exchange in proteins. *Adv. Protein Chem.* **1966**, *21*, 287−386.
(4) Weis, D. D.; Wales, T. E.; Engen, J. R.; Hotchko, M.; Ten Eyck, L. F. Identification and Characterization of EX1 Kinetics in H/D Exchange Mass Spectrometry by Peak Width Analysis. *J. Am. Soc. Mass Spectrom.* **2006**, *17*, 1498−1509.
(5) Zhang, J.; Ramachandran, P.; Kumar, R.; Gross, M. L. H/D Exchange Centroid Monitoring is Insufficient to Show Differences in the Behavior of Protein States. *J. Am. Soc. Mass Spectrom.* **2013**, *24*, 450−453.
(6) Johnson, R. S.; Krylov, D.; Walsh, K. A. Proton mobility within electrosprayed peptide ions. *J. Mass Spectrom.* **1995**, *30*, 386−387.
(7) Jørgensen, T. J. D.; Gårdsvoll, H.; Ploug, M.; Roepstorff, P. Intramolecular Migration of Amide Hydrogens in Protonated Peptides upon Collisional Activation. *J. Am. Chem. Soc.* **2005**, *127*, 2785−2793.
(8) Rand, K. D.; Zehl, M.; Jørgensen, T. J. D. Measuring the hydrogen/deuterium exchange of proteins at high spatial resolution by mass spectrometry: overcoming gas-phase hydrogen/deuterium scrambling. *Acc. Chem. Res.* **2014**, *47*, 3018−3027.
(9) Woods, V. L.; Hamuro, Y. 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.* **2001**, *84*, 89−98.
(10) Ayaz, P.; Lyczek, A.; Paung, Y.; Mingione, V. R.; Iacob, R. E.; de Waal, P. W.; Engen, J. R.; Seeliger, M. A.; Shan, Y.; Shaw, D. E. Structural mechanism of a drug-binding process involving a large conformational change of the protein target. *Nat. Commun.* **2023**, *14*, 1885.
(11) Santos, J.; et al. A Targetable N-Terminal Motif Orchestrates α-Synuclein Oligomer-to-Fibril Conversion. *J. Am. Chem. Soc.* **2024**, *146*, 12702−12711.
(12) Gong, G. Q.; et al. A small-molecule PI3Kα activator for cardioprotection and neuroregeneration. *Nature* **2023**, *618*, 159−168.
(13) Zhang, Z.; Post, C. B.; Smith, D. L. Amide Hydrogen Exchange Determined by Mass Spectrometry: Application to Rabbit Muscle Aldolase. *Biochemistry* **1996**, *35*, 779−791.
(14) Zhang, Z.; Smith, D. L. Determination of amide hydrogen exchange by mass spectrometry: A new tool for protein structure elucidation. *Protein Sci.* **1993**, *2*, 522−531.
(15) Hoofnagle, A. N.; Resing, K. A.; Ahn, N. G. Protein analysis by hydrogen exchange mass spectrometry. *Annu. Rev. Biophys. Biomol. Struct.* **2003**, *32*, 1−25.
(16) Masson, G. R.; et al. Recommendations for performing, interpreting and reporting hydrogen deuterium exchange mass spectrometry (HDX-MS) experiments. *Nat. Methods* **2019**, *16*, 595−602.
(17) Konermann, L.; Scrosati, P. M. Hydrogen/Deuterium Exchange Mass Spectrometry: Fundamentals, Limitations, and Opportunities. *Mol. Cell. Proteomics* **2024**, *23*, 100853.
(18) Keppel, T. R.; Weis, D. D. 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*, 547−554.
(19) Hamuro, Y. Determination of Backbone Amide Hydrogen Exchange Rates of Cytochrome c Using Partially Scrambled Electron Transfer Dissociation Data. *J. Am. Soc. Mass Spectrom.* **2018**, *29*, 989−1001.
(20) Milne, J. s.; Mayne, L.; Roder, H.; Wand, A. j.; Englander, S. W. Determinants of protein hydrogen exchange studied in equine cytochrome c. *Protein Sci.* **1998**, *7*, 739−745.
(21) Bondarchuk, T. V.; et al. Quaternary organization of the human eEF1B complex reveals unique multi-GEF domain assembly. *Nucleic Acids Res.* **2022**, *50*, 9490−9504.
(22) Puchała, W.; et al. HaDeX: an R package and web-server for analysis of data from hydrogen−deuterium exchange mass spectrometry experiments. *Bioinformatics* **2020**, *36*, 4516−4518.
(23) Smit, J. H.; et al. Probing Universal Protein Dynamics Using Hydrogen−Deuterium Exchange Mass Spectrometry-Derived Residue-Level Gibbs Free Energy. *Anal. Chem.* **2021**, *93*, 12840−12847.
(24) Crook, O. M.; Chung, C.; Deane, C. M. Empirical Bayes functional models for hydrogen deuterium exchange mass spectrometry. *Commun. Biol.* **2022**, *5*, 588−610.
(25) Gessner, C.; Steinchen, W.; Bédard, S.; Skinner, J.; Woods, V. L.; Walsh, T. J.; Bange, G.; Pantazatos, D. P. Computational method allowing Hydrogen-Deuterium Exchange Mass Spectrometry at single amide Resolution. *Sci. Rep.* **2017**, *7*, 3789.
(26) Salmas, R. E.; Borysik, A. J. HDXmodeller: an online webserver for high-resolution HDX-MS with auto-validation. *Commun. Biol.* **2021**, *4*, 199.
(27) Skinner, S. P.; Radou, G.; Tuma, R.; Houwing-Duistermaat, J. J.; Paci, E. Estimating Constraints for Protection Factors from HDX-MS Data. *Biophys. J.* **2019**, *116*, 1194−1203.
(28) Filandr, F.; Sarpe, V.; Raval, S.; Crowder, D. A.; Khan, M. F.; Douglas, P.; Coales, S.; Viner, R.; Syed, A.; Tainer, J. A.; et al. Automating data analysis for hydrogen/deuterium exchange mass spectrometry using data-independent acquisition methodology. *Nat. Commun.* **2024**, *15*, 2200.
(29) Woods, V. A., Abzalimov, R. R.; Keedy, D. A. Native dynamics and allosteric responses in PTP1B probed by high-resolution HDX-MS. *Protein Sci.*. 2023. 33. e5024. .
(30) Bouyssié, D.; et al. HDX-Viewer: interactive 3D visualization of hydrogen-deuterium exchange data. *Bioinformatics* **2019**, *35*, 5331.
(31) Bradshaw, R. T.; Marinelli, F.; Faraldo-Gómez, J. D.; Forrest, L. R. Interpretation of HDX Data by Maximum-Entropy Reweighting of Simulated Structural Ensembles. *Biophys. J.* **2020**, *118*, 1649−1664.

https://doi.org/10.1021/acs.jproteome.4c00700
J. Proteome Res. 2025, 24, 1688−1700