INSTITUTE OF BIOCHEMISTRY AND BIOPHYSICS
POLISH ACADEMY OF SCIENCES

Doctoral Dissertation

Weronika Puchała

High-resolution analysis of hydrogen/deuterium exchange monitored by mass spectrometry experiment data

Wysokorozdzielcza analiza danych z eksperymentów wymiany proton-deuter monitorowanych spektrometrią mas

Field of science: Natural sciences

Discipline: biological sciences

Thesis Supervisor

prof. dr hab. Michał Dadlez

Institute of Biochemistry and Biophysics PAS

Auxiliary Supervisor

dr Michał Burdukiewicz

Medical University of Białystok

Vilnius University

Warsaw, 2025

The research presented in this doctoral thesis was financed by the following project:

1) Fundacja na rzecz Nauki Polskiej, program TEAM TECH CORE FACILITY/2016-2/2

2) Narodowe Centrum Nauki, projekt PRELUDIUM BIS 2019/35/O/NZ2/03745 + NAWA Preludium BIS 1 PPN/STA/2021/1/00025

Abstract

Hydrogen deuterium exchange, monitored by mass spectrometry (HDX-MS), is a well-established method for exploring protein structural properties. This method provides unique insight into the dynamic properties of protein structures. The principle is relatively simple: measure mass increase during the exchange between hydrogens from the main chain and deuterium from the aqueous solvent. However, the simplicity ends with the experiment, as the results and interpretations may be complex. In the routine HDX-MS, the analysed protein is digested into peptides, often overlapping with each other. Although this method enables pinpointing observed changes in specific regions, there is an obvious need to integrate information from peptides to high resolution, effectively representing information at the single amino acid residue level.

The first step in HDX-MS data analysis, following manual curation of spectral data, is a routine assessment of exchange levels at different incubation times, which can be conducted using the data analysis tool I created, HaDeX. As the input data needs preprocessing to present informative values, HaDeX does all the necessary calculations alongside the measurement uncertainty to ensure reproducible results. For the general overview of the experiment, many commonly used forms of publication-quality visualisations are available, along with precise statistical evaluation. HaDeX follows community recommendations and continues to expand to stay up-to-date with the field.

I developed and implemented an entirely novel approach in HRaDeX. Using a literature-derived description of the exchange process, HRaDeX matches the deuterium uptake curve with an appropriately parametrised base function. This method allows for flattening the dimensions and effectively describes the exchange throughout the entire time course as a simple set of numbers, then colour-coded using an RGB scale. This representation enables the overlay of the whole, concentrated HDX information onto the 3D protein structure. Combining the experimental exchange time course and spatial information gives a unique insight into the protein structural dynamics. In addition, as HRaDeX utilises standard input data, re-examining previously conducted experiments provides no-cost supplementary insight. Apart from the single-state analysis condoned by HRaDeX, a comparative study between biological states is possible due to the dedicated tool compaHRaDeX.

HaDeX and HRaDeX combined create a unique and robust workflow, enabling the complex examination of HDX-MS data, from preprocessing experimental data to the peptide and

high-resolution levels, not only for single-state but also for comparative analysis. They are designed as transparent, open-source, and user-friendly solutions.

Streszczenie

Wymiana wodór-deuter monitorowana spektrometrią mas (HDX-MS) jest uznaną metodą badań strukturalnych. Metoda umożliwia unikalny wgląd w dynamikę strukturalną w opozycji do klasycznej statycznej struktury. Koncepcja eksperymentu jest stosunkowo prosta: polega na porównaniu przyrostu masy podczas wymiany, która zachodzi między między wodorami z głównego łańcucha a deuterem znajdującym się w roztworze zawierającym ciężką wodę. Jednak przejrzystość kończy się na etapie eksperymentu, ponieważ jego wyniki i konsekwentnie ich interpretacja, mogą być złożone. W typowym eksperymencie HDX-MS, badane białko jest poddane trawieniu w wyniku którego otrzymujemy nakładające się na siebie peptydy. Co prawda ten sposób umożliwia przyporządkowanie zmian do konkretnych regionów, jednak konieczna jest metoda integracji informacji z poziomu peptydowego na poziom wysokorozdzielczy, faktycznie pokazujący dane osobno dla reszt aminokwasowych.

Pierwszym krokiem analizy danych HDX-MS, po ręcznej kuracji danych, jest przyporządkowanie poziomów wymiany dla różnych czasów pomiarowych, co może zostać zrealizowane za pomocą opracowanego przeze mnie narzędzia do analizy danych HaDeX. Jako że dane wejściowe wymagają przeprocesowania celem uzyskania interpretowalnych rezultatów, HaDeX wykonuje wszystkie niezbędne obliczenia jak również rachunek niepewności, otrzymując powtarzalne wyniki. Dla całościowego obrazu eksperymentu, dostępne są metody wizualizacji o jakości publikacyjnej, jak i rachunek istotności statystycznej. HaDeX stosuje zalecenia praktyków oraz jest rozwijany aby nadążyć za potrzebami środowiska.

Opracowałam i zaimplementowałam całkowicie nowatorskie podejście w narzędziu HRaDeX. Korzystając z osadzonego w literaturze opisu procesu wymiany, HRaDeX dopasowywuje do krzywych wymiany peptydów odpowiednio sparametryzowaną funkcją bazową. Ta metoda umożliwia spłaszczenie wymiarowości zbioru danych i efektywne przedstawienie wyników z różnych punktów czasowych za pomocą zestawu kilku liczb, którym przyporządkowana jest wartość ze skali RGB. Ten sposób reprezentacji pokazuje całą, skondensowaną informację pochodzącą z eksperymentu HDX na strukturze 3D białka. Połączenie zarówno krzywych wymiany jak i informacji przestrzennych umożliwia unikalny wgląd w charakterystykę białka. Ponadto, ponieważ HRaDeX wykorzystuje standardowe dane wejściowe, można przeanalizować uprzednio zgromadzone dane otrzymując potencjalnie nowe informacje. Oprócz analizy jednostanowej przeprowadzanej przy użyciu

narzędzia HRaDeX, dwustanowa analiza porównawcza jest możliwa za pomocą dedykowanego narzędzia compaHRaDeX.

HaDeX oraz HRaDeX tworzą razem unikalny i perspektywiczny zestaw narzędzi umożliwiających kompleksową analizę danych HDX-MS, począwszy od wstępnego procesowania danych eksperymentalnych na poziomie peptydowym, aż do wysokiej rozdzielczości - nie tylko dla jednego stanu ale w ujęciu porównawczym. Oba narzędzia zostały zaprojektowane jako łatwe do użycia rozwiązania open-source.

Introduction

Proteins and their structure

The crucial role of proteins in building and orchestrating the whole biological world is undeniable. Though present everywhere in tremendous amounts, proteins have still not been fully characterised due to their complexity. The structure-function relationship is an accepted dogma (Osadchy & Kolodny, 2011; Shoichet et al., 1995; Whisstock & Lesk, 2003), increasing the importance of structural aspects of protein studies. A wide range of factors, including post-translational modifications (PTMs), alternative splicing, proteolytic processing, ligand binding, and environmental conditions, significantly increase the number of proteoforms, resulting in a complex structural landscape. Additionally, varying biological conditions can alter the conformation of proteins and their functions, further adding to this complexity. Proteomics, as a field of protein study, faces numerous challenges and continues to necessitate the development of new methods.

Proteins, functional components of living organisms, are themselves composed of smaller and well-defined building blocks: amino acids. There are 20 amino acids (technically, one is an imino acid) with the same structural core: a central carbon atom connected with an amino functional group (-NH2), a carboxylic acid functional group (-COOH), a hydrogen atom, and an amino acid-specific residue (side chain). The essential structural exception is the single imino acid, proline. In this case, the secondary amino group joins the side chain with the carbon atom, and the amino functional group effectively forms a pyrrolidine loop. This makes proline more rigid than other amino acids and eliminates the hydrogen from the main chain.

A protein is a polymer of linked amino acids. When the amino group of one amino acid monomer reacts with the carboxylate carbon of another amino acid, it forms an amide linkage called a peptide bond. Depending on the chain length, we call the resulting linear polymer either a peptide (shorter) or a protein (longer). Although the nomenclature is not very strict, proteins are usually defined as chains of at least 50 amino acids and are the primary study focus. What is particularly important is that the protein sequence has a polypeptide chain with a repeating part called the main chain, rich in hydrogen bonding potential (Berg et al., 2002). The sequence is conventionally denoted in order from N- to C-terminus, as is presented in Figure 1.

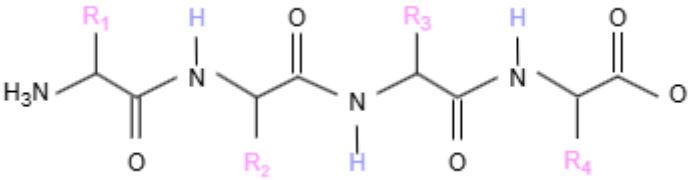


Figure 1. Primary structure of the exemplary peptide with four amino acid residues coloured in pink, and hydrogens from the main chain coloured in violet.

The linear protein sequence is merely an initial step in describing the protein structure. The chain folds into a spatial shape that can be defined on different levels, leading to secondary, tertiary and quaternary structure.

Said chains can fold into several types of secondary structures, of which the most common are alpha helix and beta sheet. In an alpha helix, the main chain amide protons are engaged in intramolecular hydrogen bonds of different amino acids spaced by four positions in the sequence. In contrast, the beta sheet is created by the hydrogen bonding between backbone atoms from different segments, forming a rigid structure. Nevertheless, hydrogen bonds are key to stabilising the basic secondary structure elements.

The next step of complexity is how those regular secondary structures are spatially aligned with respect to each other, creating a tertiary structure. Experimental characterisation of 3D structure is challenging. In addition, the fourth-level structure is a protein complex of two or more proteins in spatial entanglements and dependencies, further complicating the image. Not only do the individual proteins need to be studied separately, but it is also vital to learn how they interact spatially with each other or with other molecules. A good example is a protein interacting with a ligand, where, in addition to examining the structures of both the protein and the ligand, it is also essential to study their binding site.

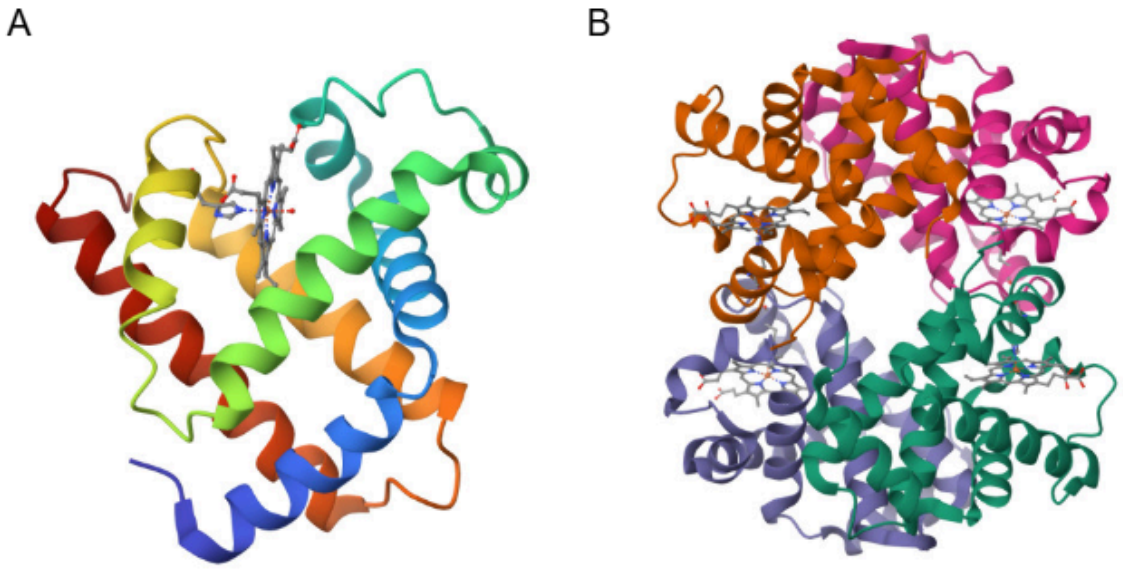


Figure 2. Examples of tertiary and quaternary structure. A. Myoglobin (figure from (Watson & Kendrew, 1976)) has a single polypeptide chain of 153 amino acids folded compactly into eight alpha helices forming a globular structure, connected by flexible loops and stabilised by hydrophobic residues. The prosthetic heme group, centrally buried within the hydrophobic core, is responsible for oxygen binding. As myoglobin is a single-chain protein, it has no quaternary structure. B. Haemoglobin (figure from (Fan et al., 2021)) is a protein complex with four subunits, two alpha and two beta. The subunits are arranged in a symmetrical, globular structure, each containing a heme group, resembling the myoglobin structure. Thanks to this structural arrangement, it may carry up to four oxygen molecules. It exhibits cooperative binding - when one heme group binds to oxygen, the protein undergoes conformational change, exhibiting the subunits without oxygen, burying one with oxygen inside the structure.

Furthermore, the protein structure is not static; atoms constantly oscillate around the centre of mass, even when embedded in the crystal lattice and more so in solution. The amplitudes and frequencies of these motions span several orders of magnitude, from extremely stable regions with conformations of half-lifetimes of tens and hundreds of hours to completely flexible loops changing their position on a microsecond timescale, all combined within a single protein or even domain. Proteins constantly undergo various-scale structural changes, such as side chain rotations, domain opening, or transitions from active to inactive states. Each protein has its own specific kinetics, with a unique profile of dynamic changes. The extreme case of highly dynamic proteins is the class of intrinsically disordered regions or proteins (IDRs), which lack detectable structures. These proteins lack a fixed structure and exist in the form of a conformational ensemble, constantly fluctuating. When interacting with other proteins, they can bind with them, folding and becoming structured only in a specific

complex. However, the lack of a stable structure does not result in a lack of function. The discovery of functional IDRs is a matter of the last 25 years so the studies of the dynamic axis of protein structurome are still in development, also because classic methods are either not fit for revealing the structural dynamics (like X-ray) or extremely time consuming and not fit for larger proteins (like nuclear magnetic resonance - NMR) as described in dedicated latter section.

Myoglobin and haemoglobin (examples of tertiary and quaternary structure in Figure 2) are classic examples of proteins that have been extensively studied, primarily because of their essential function. However, a vast number of proteins have yet to attract significant scientific attention. The human genome consists of 30,000 - 40,000 protein-coding genes (Lander et al., 2001), in addition expressed in different proteoforms, but only approximately 4,000 protein structures are deposited in the PDB database, with *Homo sapiens* origin. For many of these lesser-known proteins, our current understanding of their structural characteristics and functional roles is still limited. Although enormous progress has been made in investigating the structure and its relationship with protein function thanks to classic structural approaches, protein systems of increased dynamics escape characterisation through these methods. Moreover, many protein sequences still do not code stable products (Peng et al., 2014), effectively being excluded from novel modelling tools like AlphaFold. As the only way to investigate them is through experimental methods, the need to expand the spectrum of classic structural characterisation methods becomes more and more pressing.

Structural methods

There is an increasing need for precise structural studies of various proteins to address these challenges. The scientific community widely employs a set of established structural methods, such as X-ray crystallography, nuclear magnetic resonance (NMR) spectroscopy or cryo-electron microscopy (cryo-EM).

X-ray crystallography is a well-established structural method that provides high-resolution results. This method utilises the diffraction of X-rays by the electron density of crystallised molecules, allowing the creation of structural models. However, it can only be used for proteins that form a static crystalline state structure. This excludes regions of high dynamics, which are often truncated to enable successful crystallisation, and thus vanish from our view, although their functional and structural role should not be neglected. The X-ray technique is

demanding, subject to a high failure rate, and disqualifies a broad group of proteins with intrinsically disordered regions.

Nuclear magnetic resonance (NMR) spectroscopy (Hu et al., 2021) is a time-tested method capable of analysing in-solution 3D and 4D structures of proteins at the atomic level. It allows the analysis of the protein under native conditions, and the dynamic regions do not have to be excluded from the study. The protein is placed in a strong magnetic field, and radiofrequency radiation is applied in order to characterise the resonant frequency of the atomic nuclei of the studied sample. The latest technological improvements have made NMR more accessible and informative. However, this method still has limitations, such as high sample concentration and size not exceeding 50 kDa. The analysis of the data is complex and time-consuming.

Cryo-electron microscopy (cryo-EM) is a high-resolution spatial method that often provides results at the atomic level, with electrons serving as the imaging source. Samples are frozen in the native state, allowing the analysed material to retain its structural heterogeneity. However, electron radiation used in microscopy may damage the sample. The sample preparation is complex and costly, as is the experiment itself.

Although their successes are indisputable, classical methods cannot be applied for various reasons to protein systems that are highly dynamic, structurally heterogeneous, or oligomerising. Therefore, new methods are still necessary, capable of mapping protein dynamics, as the dynamic aspects of protein structure deserve more attention than they get in the structurome dominated by stable protein structures. This leaves a massive opportunity for alternative, faster, and cheaper methods, such as HDX-MS, to step up.

HDX-MS experiment

HDX as a structural method was first introduced by Kai U. Linderstrøm-Lang in the 1950s (Englander et al., 1997). Using simple laboratory equipment, he was able to demonstrate the slowdown of HDX exchange in native versus denatured proteins, thereby verifying Linus Pauling's hypothesis that hydrogen bonds stabilise protein structures. Indeed, when H-bonded, protons cannot exchange with solvent, while those not H-bonded, accessible to solvent protons exchange rapidly. Therefore, measuring the efficiency of HD exchange seems to be a perfect probe for the presence of H-bonds, for instance, in protein structures, secondary, tertiary or quaternary. Apart from H-bonding, other factors, like solvent

accessibility, are also taken into account. Depending on the hydrogen position in the primary and higher-level structure, exchange may occur with rates differing by several orders of magnitude. The exchange conditions may change in time with unfolding and folding processes.

The exchange model assumes that a hydrogen atom oscillates between an “open state”, with hydrogens susceptible to exchange and a “closed state”, with hydrogens blocked from exchange. In this model, the exchange mechanism between hydrogen and deuterium from the solvent is described as follows:

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

When hydrogen dwells in a closed state, it does not exchange. Still, protein structure opens occasionally (locally or globally), H-bonds are temporarily broken, and hydrogen may exchange with deuterium if solvent-accessible. The closed state is a folded and exchange-incompetent form, while the open state is solvated and exchange-competent.

Let’s start the process description while the hydrogen is in the closed state (NH_{closed}). Then, the structure opens with the opening rate k_{open}, changing into an open state (NH_{open}). Then, it undergoes exchange with deuterium from the solvent with its intrinsic exchange rate k_{int}. Next, the deuterium in the previous position of hydrogen in the open state (ND_{open}) changes the state to the closed one (ND_{closed}) with the closing rate k_{close}. Depending on the conditions, this process can occur in both directions, with an exchange from deuterium back to hydrogen called back-exchange.

Each amino acid has its intrinsic exchange rate (k_{int}), which describes the exchange in the absence of secondary, tertiary or quaternary structures, but accounts for the nature of neighbouring residues in the sequence (Bai et al., 1993; Nguyen et al., 2018). It is also affected by temperature, pH, pressure, and ionic strength of the buffer (Vinciauskaite & Masson, 2023). However, the structural context greatly influences the exchange rate, usually diminishing it, serving as a protection against the exchange. As a result, the experimental exchange rate (k_{exp}) may differ from the intrinsic rate.

The experimental exchange rate can be measured by several analytical methods, including the most precise - NMR. However, I focus on the most frequently used nowadays - mass spectrometry (MS).

Both H-bonding breakage and solvent exposure depend on structural stability in a given region; therefore, tracking deuterium incorporation over time offers an indirect means to infer structural features and kinetic processes.

The labile hydrogens (bound to heteroatoms, in this case, nitrogen) continuously exchange with the solvent (Jensen & Rand, 2016), a phenomenon utilised in this experimental approach. In the most common “bottom-up” approach (Fig. 3), the protein sample is labelled in the aqueous solvent containing heavy water (D_2O), for a specified period of time, after which exchange is quenched by lowering the pH and temperature and subjected to proteolysis by a protease active at low pH. Then, the mass measurement is done using mass spectrometry (MS). Therefore, HDX-MS resolution is in principle limited to the level of peptides (protein fragments obtained in the digestion process); the deuterium uptake that is measured is the superposition of monoexponential growth functions of the number of amide protons in a peptide (minus one, as the N-terminal one undergoes back-exchange very fast). The mass difference between hydrogen and deuterium is roughly 1 dalton, so the experimentally observed mass difference may be translated directly to the number of exchanged hydrogens. The mass measurement on the peptide level allows for pinpointing the exchange within protein regions. The stages of “bottom-up” HDX-MS are presented schematically in Figure 3.

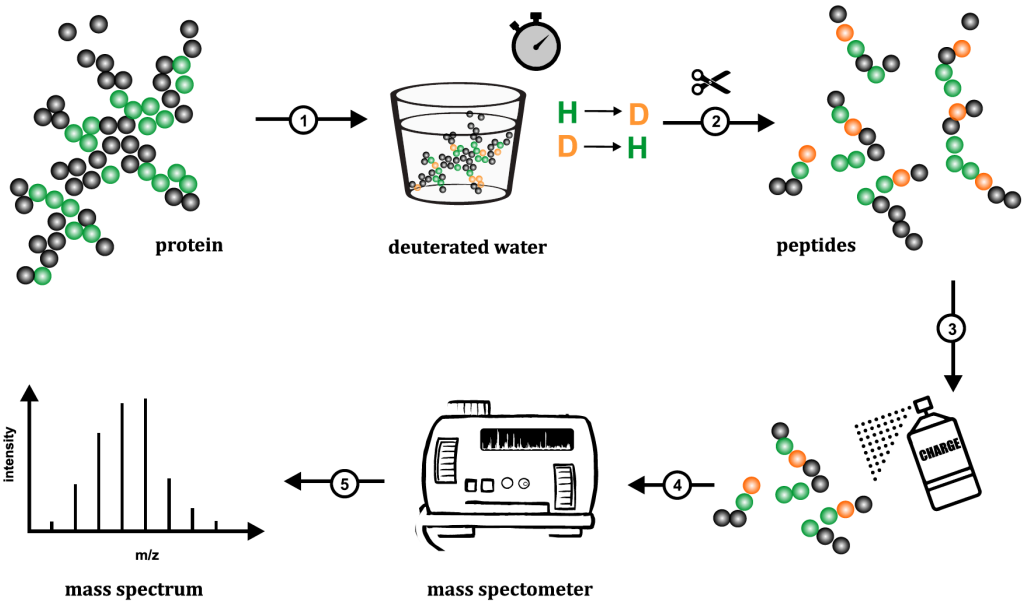


Figure 3. The stages of the HDX-MS experiment, figure from (Grzesiak, 2021). 1) The prepared protein sample is placed in the solvent in native buffer conditions. 2) The exchange and back-exchange occur for the desired incubation time, but due to an excess of deuterium ions in the solvent, deuteration prevails. After the exchange is quenched, the digestion process is conducted to obtain fragments of protein. 3) Samples are ionised as a preparation step for the mass spectrometry. 5) Mass measurement conducted by mass spectrometry. 6) The result of mass measurement - isotopic envelope, here for a single mass measurement.

To observe the kinetics of exchange, the measurements are planned for multiple incubation times with D_2O, their choice depending on the aim of the experiment, e.g., with time points cumulated either in the first phase for fast-exchanging proteins or a later phase for slowly exchanging ones, or distributed evenly between seconds and hours. Then, after the exchange process is quenched, the protein is digested using acidic proteases to obtain a set of peptides. There is a set of proteases active at low pH, such as pepsin or nepenthesin, which affects the digestion pattern and, as a result, coverage and redundancy.

In a routine HDX-MS, only the exchange of labile hydrogens from the main chain (i.e. amide protons) can be tracked. Other labile protons, as in side chains, exchange more rapidly; therefore, they undergo fast back-exchange during LC, which precedes MS and their deuteration cannot be observed. Due to its imino acidic nature, proline does not participate in the exchange due to the missing hydrogen in the main chain.

HDX-MS is often used due to its relatively small time requirements for experiments and data analysis, where obtaining results may take a single day, and its superior sensitivity. The direct result of mass measurement done by mass spectrometry is in the form of an isotopic envelope, a spectrum containing signals from all isotopic variants of the molecule. Still, the mass centre is shifted to higher values due to the presence of deuterium. NMR can also measure HDX rates, but this method has its limitations in comparison with MS, such as lower sensitivity caused by higher protein concentration requirements, a smaller protein size, and a much longer measurement and data analysis process.

When we measure the changes in mass over time for specific peptides, we may notice differences in regions; in some, the exchange may be fast, in some slow, while in others, there may be a superposition of different regimes. As the ratio between opening and closing rates determines the exchange process, we use it to distinguish between two extreme exchange regimes. Exchange regime EX1 fulfils the dependency of $k_{int} \gg k_{close}$, and thus $k_{exp} \approx k_{open}$. In this case, one opening event is sufficient to complete the exchange. This results in the partition of the population undergoing the exchange into molecules that exchanged in full, and the rest of the molecules that did not exchange at all. The opposite dependency means that the opening event is insufficient for the successful exchange and is described as EX2. In this regime, exchange goes on gradually with partial, incremental exchange after each opening event. The exchange regimes EX1 and EX2 are easily distinguishable when comparing the obtained isotopic envelopes. In EX1, the exchange occurs during a single opening event, resulting in a bimodal distribution. This introduces heterogeneity in the mass spectrum, leading to problematic assessment of the mass centre. In contrast, EX2 is more straightforward to interpret as it results in a uniform distribution of m/z.

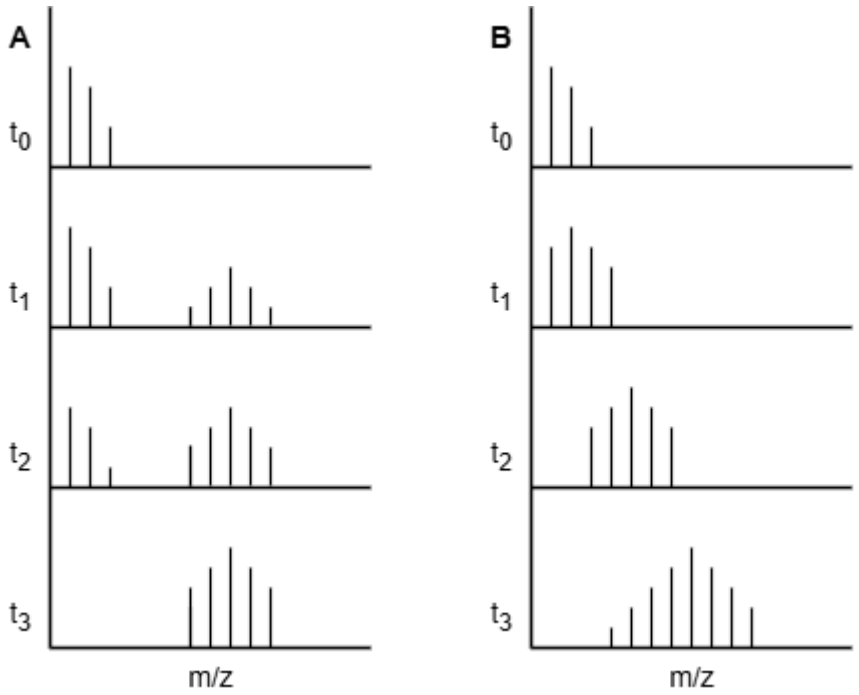


Figure 4. Change schemas in isotopic envelopes in two exchange regimes, EX1 (A) and EX2 (B), over time. The top trace shows the undeuterated sample, with lower traces showing longer incubation times.

The “bottom-up” is not the only way to conduct an HDX-MS experiment, although it is predominantly used as an easy method to obtain information on the peptide level. In the “top-down” approach, the protein undergoes the exchange and mass measurement follows directly afterwards. Only the global change is determined, without information on where it occurred, mainly to check if the exchange is possible.

A vital aspect of conducting an HDX-MS experiment is preparing undeuterated and fully deuterated controls. As we operate on mass uptake, we need an undeuterated reference point. The undeuterated control provides a peptide average mass before exchange and allows us to calculate deuterium uptake in daltons (where approximately one Dalton of uptake signifies one exchanged hydrogen). Unfortunately, the value in daltons may be challenging to interpret, as it provides a different context depending on the peptide length. The community recommends (Masson et al., 2019) preparing a fully deuterated control (FD) to obtain fractional deuterium uptake. A fully deuterated control is necessary, as the back-exchange during low-pH LC separation is sequence-dependent and thus differs for different peptides. Whenever back-exchange needs to be taken into account, fully deuterated control is a must.

HDX-MS, unlike other structural methods, effectively avoids limitations of protein size and high requirements for sample preparation. Therefore, the barriers to conducting a pilot experiment are relatively low, and it is possible to assess the chances before planning a proper set of experiments. The method is widely used, with over 300 articles (searched by keywords “HDX-MS” or “hydrogen deuterium exchange”) published in the SCOPUS database annually in the last decade.

HDX-MS enables studies of large protein complexes (Lesne et al., 2020) (e.g., proteasome (Kochert et al., 2018), elongator factor eEF1B (Bondarchuk et al., 2022)), membrane proteins (Javed et al., 2023), and proteins with post-translational modifications or intrinsically disordered regions (Parson et al., 2022). In addition, combining or cross-referencing HDX-MS information with results obtained using different methods may be highly effective, e.g., there are attempts to use HDX-MS data in molecular simulations (Devaurs et al., 2022). There are multiple community efforts to ensure access to gathered data, such as dedicated databases (Pancsa et al., 2016).

Data processing

The measurement result from mass spectrometry is a set of isotopic envelopes for each peptide from the peptide pool, along with their corresponding charge values, for a given time point of incubation under specified biological conditions. Every detected isotopic envelope has to be assigned to the peptide. This is not trivial, as the peptide mass is no longer equal to the undeuterated mass. Therefore, other parameters (like retention time or ion mobility) are used to ensure peptide identity. Following this, the weighted mass centre of the isotopic envelope (centroid) is usually calculated along with mass shift. This number serves as the basis for further analysis in the routine HDX approach, providing the number of exchanged amides at a given incubation time. Although these data analysis steps are automated by existing tools (DynamX™, HDeXaminer™), the automated data processing is far from perfect and manual curation is needed to double-check that the spectra are correctly labelled to peptides from the pool, and their centroid is properly calculated. This task is especially meticulous, time-consuming, and requires extreme caution.

Though isotopic envelopes are precise images of exchange status, they are rarely used in routine HDX due to their complexity. Centroid values are most often sufficient to fulfil the purposes set by HDX-MS; however, once the centroid-level analysis is exploited, a more detailed spectra analysis can be carried out, including the analysis of isotopic envelope

shape, width and homogeneity, bringing additional interpretative potential. For instance, the EX1 regime is characterised by bimodal isotopic envelopes, which necessitate dedicated data analysis tools, typically performed as an extra step to standard data analysis.

Once the assignment of isotopic envelopes is verified, each centroid is calculated in the routine HDX data analysis workflow. Finally, the data are exported as m/z centroid values or m/z values are transformed into mass values, and then those for different charge values are aggregated into one value per replicate. As the experiment is recommended to be conducted at least in triplicate, the values from replicates are averaged, and their measurement uncertainty should be calculated. The steps of obtaining data are summarised in Figure 5.

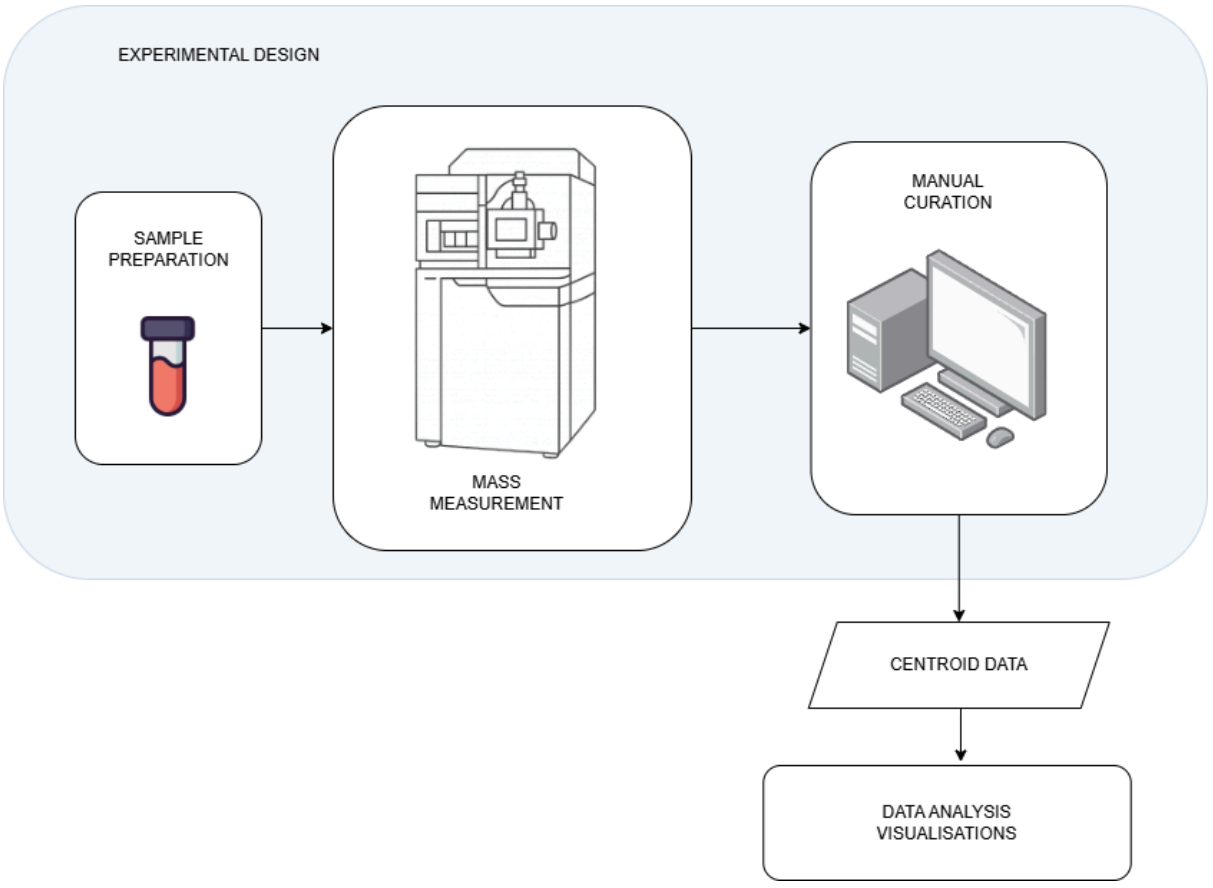


Figure 5. Data workflow and necessary analysis steps in classic HDX

Even if the results are simplified into centroids, presenting such a complex, multidimensional dataset becomes a significant challenge. Deuterium uptake is calculated for numerous, sometimes partially overlapping peptides for multiple incubation time points, multiple technical replicates, and for varying biological conditions. Figure 6 presents common forms of visualisation and discusses choices leading to the loss of data dimensions. Each plot type

has its advantages and disadvantages, but none of them allows for the full complexity of the dataset to be represented. All of them are based on presenting exchange levels at different times of incubation, none operates on exchange rates.

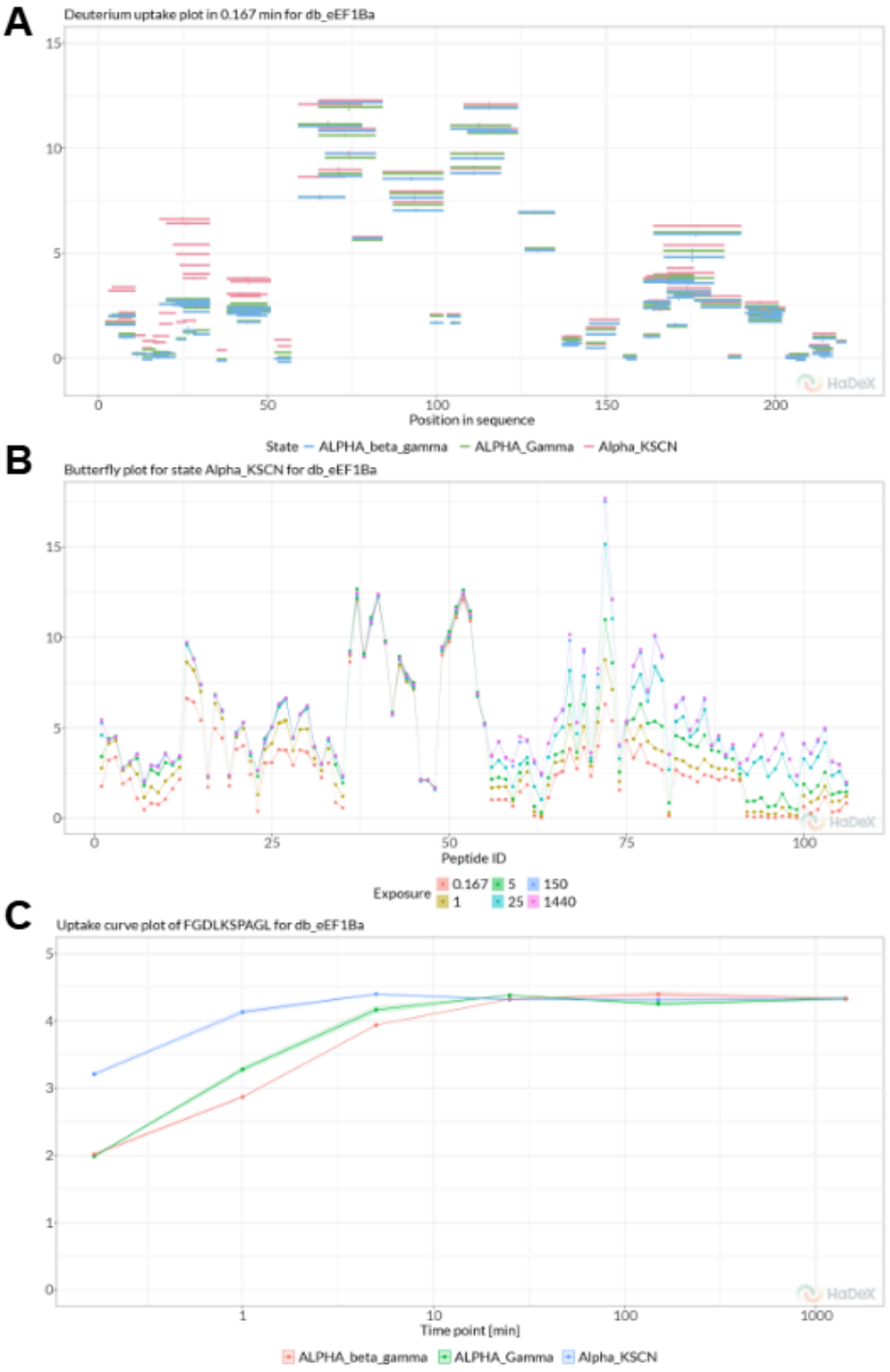


Figure 6. Most popular methods of data visualisation. A. Comparison plot presents deuterium uptake for peptides, but only for a single time point of measurement. It shows the exchange for each peptide along with its position and length. B. The Chiclet plot presents deuterium uptake for multiple time points of measurement. It loses information on the position and length of the peptides; instead, it labels them with their ID values. The disadvantage is that regions of high redundancy are overrepresented. C. The uptake curve presents the exchange process in time, but only for a single peptide (here for three biological contexts), omitting the overlapping and neighbouring peptides, with no information on the position of this peptide in the whole protein sequence.

In terms of software for mass spectrometry data acquisition, Thermo™ equipment's reach has grown, with its team actively expanding the native software, exceeding the limited functionalities offered by DynamX (Waters™). However, the DynamX apparatus remains widely used in many laboratories worldwide, and there is still a significant need for dedicated deuterium uptake analysis tools.

Conversion to centroids is just the first step of the analysis. Separate tools are necessary for subsequent steps of integrating results from the entire experiment, visualising them on different plot types, and extracting biologically relevant information, such as identifying regions with stable structure, intermediate flexibility, and fully flexible regions. Especially, the differences between biological states of interest need to be calculated with proper statistical significance calculus. For this purpose, a wide range of tools has been proposed for peptide-level analysis over the past few years, due to the growing demand for HDX-MS data analysis tools. Each of them has different functionalities and therefore different advantages. Dueteros (A. M. Lau et al., 2020; A. M. C. Lau et al., n.d.) allows data visualisation using an efficient web server, but lacks programmatic access to its core functions for personalised analysis. HD-eXplosion (N. Zhang et al., 2021) allows only two forms of visualisation: the chiclet and volcano plots. HDX-Viewer (Bouyssié et al., 2019) is an invaluable tool for interactively presenting data on the 3D structure, but the HDX dataset has to be processed elsewhere beforehand. The extensive functionalities of HDXBoxer (Janowska et al., 2024) are available only in the form of an R package, which requires potential users to have a programming background. One complex, easy-to-use tool covering all the functionalities with options to be widened by adjusting its workflow on the code level is still lacking.

Another challenge is converting a set of measured exchange levels at different times into a set of exchange rates that fully characterise the whole uptake curve and decrease the dimensionality of the dataset. This allows for the compression of complex result visualisation into a single panel. For peptide-level analysis, exchange should be best characterised by a

superposition of several (up to the number of exchangeable amides in a peptide) exponential growth kinetics. Such models are accepted in literature, but no software has been made available for the wider public to fulfil this purpose.

The ultimate goal of HDX analysis is to obtain single-residue (or sub-peptide) information on the exchange of the smallest possible set of amides or a single amide, rather than peptide-level analysis. Hence, a natural step from peptide-level analysis is high-resolution level - assigning the exchange information directly to the amide position in the protein sequence, for instance, based on redundant information from overlapping peptides. To achieve this, several approaches have been discussed over time, some of which have served as a proof of concept (Gessner et al., 2017; Skinner et al., 2019). Other solutions have ready-to-use implementations, such as HDX Modeller (Salmas & Borysik, 2021), and pyHDX (Smit et al., 2020). However, due to a lack of documentation and low user-friendliness, these solutions are cumbersome to use, requiring a lot of attention and trial and error before being utilised. Another aspect is the lack of transparency and complications with understanding and interpreting the data transformation steps, especially when using an opaque decision-making machine learning approach like in HDX Modeller. pyHDX utilises a complex set of equations that can only be solved using a sophisticated computing system, such as pyTorch. It assumes that a single amide rate can be reliably estimated based on peptide-level data of variable redundancy, and such an assumption may not always be accurate or lead to significant estimation errors. Still, in pyHDX specification of the calculated intrinsic exchange rate as an additional restraint is required, leading to possible overparameterisation of the fitting procedure. Moreover, all proposed solutions try to assess the experimental exchange rate for each amino acid, but there is insufficient experimental data volume to ensure reliable results at a single amino acid level. Also, they lack a way to validate the high-resolution results by comparison with the initial experimental uptake data. In conclusion, there is no consensus within the HDX-MS practitioners community regarding the preferred high-resolution method, and experimentalists rarely conduct such analyses.

Aim of the work

The primary aim is to facilitate HDX-MS data analysis and visualisation for centroided mass shift data after exchange. It is realised in three ways:

1. peptide level deuterium uptake calculations with their visualisations (HaDeX),
2. high-resolution exchange rate classification based on deuterium uptake curves for single state (HRaDeX),
3. comparative high-resolution analysis of HRaDeX results (compaHRaDeX).

HaDeX - peptide level analysis

The great need for high-quality HDX-MS analytical tools was met by developing the tools mentioned in the previous section; however, almost all of them were developed after the launch of HaDeX (Puchała et al., 2020), which preceded many of the tools described in the section *Data processing*. Moreover, I have introduced several good analytical practices, including the propagation of measurement uncertainty, and my groundwork on this topic has been acknowledged by others (Weis, 2021). In addition, recent review articles (Engen et al., 2020; James et al., 2022; Stofella et al., 2024) mention HaDeX as an exemplary HDX analytical tool due to its comprehensiveness and wide applicability.

As described in the introductory section, the output data from the HDX-MS experiment is typically in a tabularised, difficult-to-analyse format, as a .csv table exported from DynamX (Waters™), a popular tool for obtaining centroids from HDX-MS experimental data. The DynamX export file contains unaggregated mass measurements from technical replicates for different charge values within each replicate. HaDeX package integrates information on both levels, charge values and replicates. The GUI presents only the final result, aggregated across technical replicates with calculated uncertainty. Automating the data aggregation and uncertainty calculation process is crucial to ensure the assessment of workflow reproducibility. The components responsible for this data transformation are described in the documentation, tested, and open-source to prevent doubt about their correctness. HaDeX ensures a transparent process to fully understand data transformation and produce reproducible results, with their visualisation and uncertainty calculation.

The measured peptide mass shift in Daltons at a given time lacks a reference point. The same absolute shift in Da may indicate full exchange in a short peptide, while a minor exchange in a long peptide. On the other hand, a slight change in mass may originate from fast back-exchange, rather than from high protection. Therefore, it is advisable to normalise data for the peptide length at least. HaDeX provides different options to calculate normalised deuterium uptake, with reference either to experimental (a full deuteration control including back-exchange) or theoretical (the maximum number of exchangeable protons in a peptide) maximum uptake controls to obtain fractional values instead of absolute deuterium uptake values in Daltons.

Once the deuterium uptake is calculated with the chosen parameters, the deuterium uptake values accompanied by their respective uncertainty may be presented in a comparison plot

or Woods plot for differential study, as shown in Figure 7. Those plots present information for a specific time point, presenting deuterium uptake values and details about peptide length and position in the protein sequence. The error bars present uncertainty of measurement, or propagated uncertainty for comparative analysis (calculated using the law of propagation (Weis, 2021)). Moreover, the difference in uptake between state A and state B is accompanied by statistical analysis; the Houde test is conducted to determine whether the values are statistically significant. Those two methods of visualisation present results for the whole protein sequence. For broader information on specific peptides, an uptake curve showing the deuterium uptake changes in time for one selected peptide is available. The numeric data table used to create plots is accompanying each plot. Moreover, the plots can be downloaded in SVG format for further personalisation.

Another critical piece of information provided by HaDeX is the coverage and redundancy plots, which present the peptide pool and its overlaps, possibly serving as a quality control. A summary of the experiment is also recommended as part of experiment reporting (Masson et al., 2019). All the plots and data can be easily compiled into an HTML report. These features make HaDeX an accessible tool for studying deuterium uptake through single-state or comparative analysis. The documentation describes all the data transformations, and the code is open-source, facilitating transparency. The user-friendly interface is accessible directly from a web browser, ensuring comfortable use even for users without programming skills.

HaDeX was developed alongside the community of HDX-MS practitioners. They used the software and provided necessary feedback to ensure the quality and versatility of the tool's functionalities. Since the time of publication, HaDeX was used, among others, in a study of CHN-1 protein, focused on finding the binding site of UFD-2 (Das et al., 2021), in a study of full-length RAGE in complex with S100B (Moysa et al., 2021), in a study of Ssq1, alone and in complex (Jelen et al., 2023), in a study of Nucleobindin-2 (Lenda et al., 2024), and in a study of phosphatase PP1 (Fatalska et al., 2024).

Availability

There are multiple ways to use HaDeX. The R package is uploaded on CRAN (https://cran.r-project.org/web/packages/HaDeX/index.html) or available in the repository (https://github.com/hadexversum/HaDeX). The dedicated GUI is hosted as a web server (https://hadex.mslab-ibb.pl/) or may be run locally using the R package. The standalone software is also available (https://sourceforge.net/projects/hadex/).

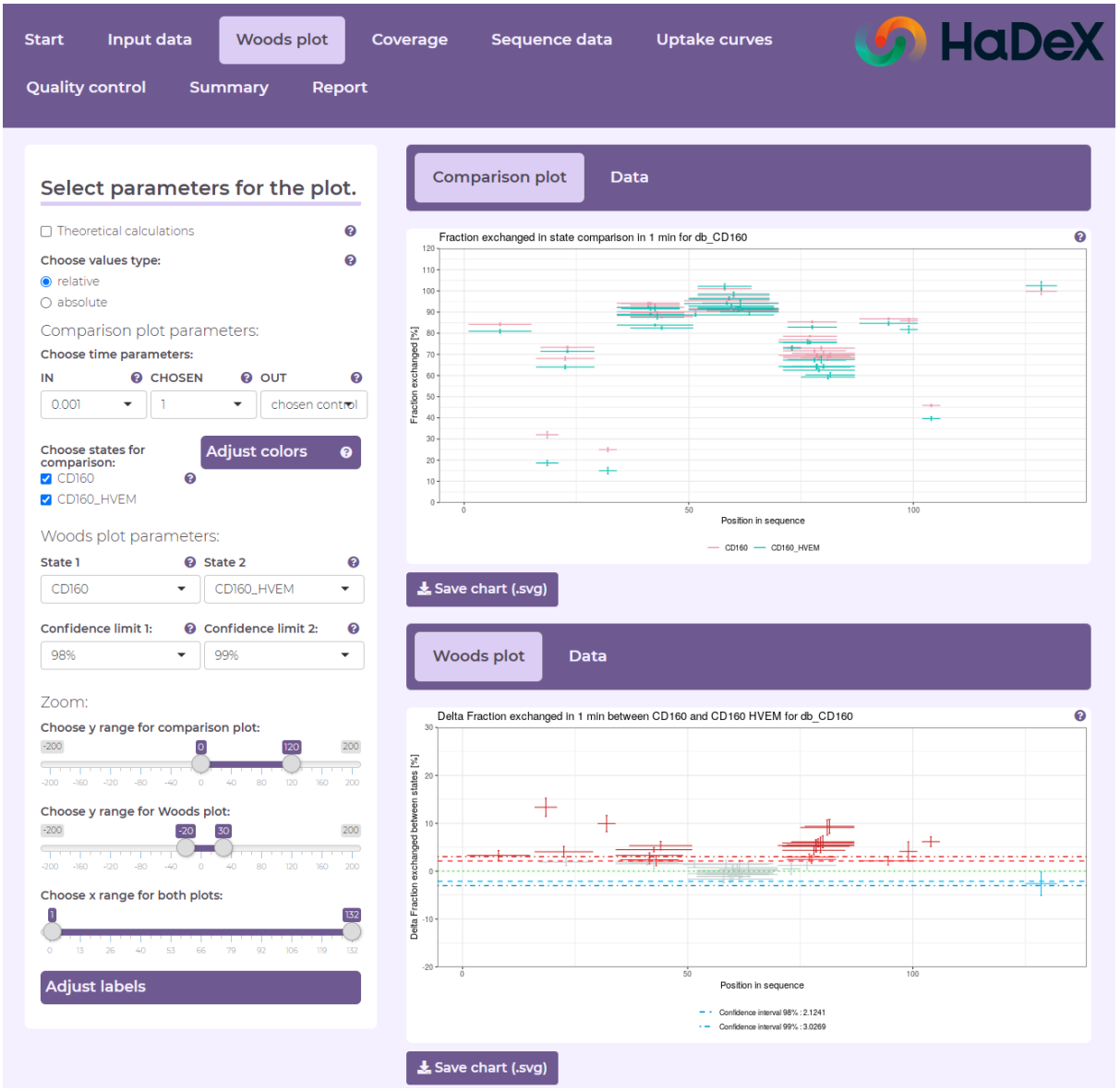


Figure 7. The HaDeX application panel, showing the comparison plot and the Woods differential plot. The upper panel presents a comparison plot, showing the deuterium uptake values for two biological states. The lower panel presents Wood's plot, showing the difference between the two chosen biological states with statistical significance analysis. Both plots present values for one specific time point of measurement, selected in the left panel.

HRaDeX - high-resolution level analysis

The scientific community has long been interested in obtaining high-resolution data from HDX-MS. Experimental approaches, providing high-resolution single amide exchange data, such as using ETD-type fragmentation, have a high entry threshold in terms of demanding fragmentation efficiency or highly specialised instrumentation. Therefore, many computational attempts have been made to extract single-amide exchange parameters from peptide-level exchange data of classic HDX-MS. None of them fulfils the list of desired features: transparency of the method and its implementation, user-friendliness, complete workflow, straightforward interpretation of the results, and methods of validation. None of these tools has been accepted for routine use by the community. For instance, based on peptide-level data, previous solutions have attempted to provide exchange rates for each single amide proton, resulting in high uncertainty in the results. Therefore, I proposed a different approach, combining all data dimensions into dense information that is available to practitioners or bioinformaticians, thereby addressing all the items on the aforementioned list of requirements by creating HRaDeX (Puchała et al., 2025).

The idea is based on the classification of the uptake curves. The uptake curve visualises changes in deuterium uptake over time for a single peptide. How can this whole curve be flattened to just a small set of numbers? By assessing carefully designed metrics.

The starting point is the literature-based equation, implementing the Zhang-Smith model of approximation of the uptake curves, dividing the exchanged hydrogens into three groups of exchange - fast, medium, and slow (Z. Zhang et al., 1996; Z. Zhang & Smith, 1993):

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

Where D is the deuterium uptake level described as a function of three exchange rates (k_1, k_2, k_3), coupled to the fraction of protons exchanging by a given rate (n_1, n_2, n_3). Therefore, the superposition of n exchange rates in a peptide of n amides is approximated by the superposition of three uptake processes with three rates, regardless of peptide length. This approach reduces the possibility of overparameterisation, as datasets in practice rarely exceed 7 time points.

In this method, we do not asses an exchange rate for each amino acid, as we do not have sufficient data, but we assume that we can assign an exchange regime for each amide into one of three groups of exchange: fast, medium, and slow, by least-squares fitting of the above equation to the uptake curve (shown in Figure 8E, both as separate fitted groups and as their superposition in black). Then, we treat the obtained n_i values as a proportion of amides following each exchange group. On top of that, we take these three n values and, using an RGB colour palette (Figure 8F), we describe a vector (n_1,n_2,n_3) with a single value - colour. In such a way, the distribution of exchange groups in a given peptide can be described by three numbers and easily visualised in the sequence or structure. One or two of the three groups may be nonexistent depending on the experimental data.

A three-element vector resulting in colour is assigned to each peptide from the peptide pool (Figure 8A). Further on this information can be aggregated into high-resolution level (Figure 8B) using one of two possible aggregation methods: shortest (using the shortest peptide covering said amino acid) or weighted approach (Keppel & Weis, 2015) (as a weighted mean of all peptides covering said amino with weights inverse proportional to their lengths). As an intermediate step, the classification colour parameter set is plotted to present the intensity of each group in regions (Figure 8C). Once we have the high-resolution values, they can be presented on the 3D protein structural model, if available (Figure 8D), thereby merging experimental insights with available spatial information. This visualisation method is exceptionally dense, providing a possible broad overview of all the information gained at a glance. The application was tested on 16 datasets (totalling over 4,000 peptides), both obtained in our laboratory and from PRIDE repositories, with the general conclusion that the domination of the slow regime is rather an exception than the rule, and the substantial dynamics of protein chains is widespread.

Moreover, the HRaDeX algorithm does not require any assumptions about the exchange process or setting hard-to-estimate parameters, apart from setting a flexible definition of the exchange groups that can be modified in repeated runs.

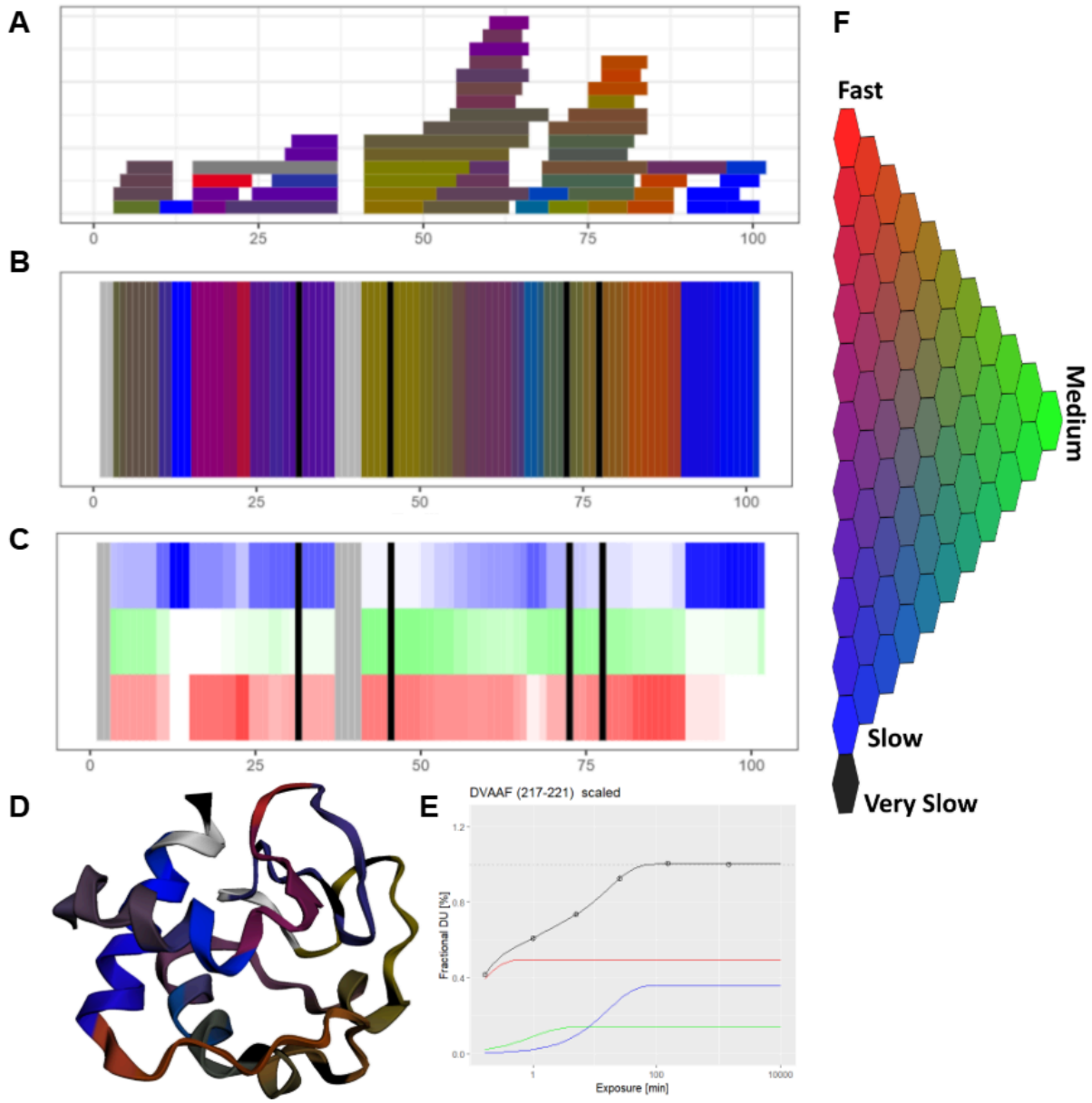


Figure 8. The intermediate and final steps in the HRaDeX workflow are shown in the case study of Cytochrome C. A. The colour-coded classification results for the peptides, with redundancy. B. The aggregated classification results for each amide are obtained using a weighted approach. C. Plot with intensities of each group to show which group is dominating in regions. D. Aggregated classification results presented on the 3D structure. E. Example of fitting result for one peptide. F. RGB colour palette. The Figure is adapted from (Puchała et al., 2025).

In the HRaDeX package, an essential tool for verifying the quality of reproducing experimental uptake curves using a set of calculated *k* and *n* values has been implemented. Dedicated functionality allows going back from the high-resolution values to recreate experimental uptake curves of peptides. This serves as a control method to assess the goodness of the fitting process by comparing it with experimental data and testing the algorithm's performance. It helps the user to interpret the results and check whether they are

satisfactory. It is a crucial feature to ensure that the algorithm behaves correctly and that the results reliably reflect experimental data.

We avoid calling this algorithm single-amino resolution, as the resolution depends strictly on the peptide coverage obtained in the experiment. To achieve single-residue resolution, very high peptide redundancy would be required, with peptides differing by one residue in length in all cases, which is not feasible experimentally.

HRaDeX only analyses a single dataset representing a single biological condition. However, as a comparative study is essential to HDX-MS analysis, I created compaHRaDeX, a dedicated tool for comparing HRaDeX results for two states. It enables the comparison of different aspects of HRaDeX results between these datasets, for instance, representing different biological conditions. As a result, the regions of difference can be identified and presented on the 3D structure to provide a broader picture.

HRaDeX was published recently and has not been cited yet. However, it has been invaluable in several ongoing HDX-MS analyses in the Mass Spectrometry Laboratory, the results of which will be published soon. HRaDeX is used daily in ongoing projects conducted in the Mass Spectrometry Laboratory, IBB including at present studies on: metalloproteins, performed by the Kevin Waldron Lab (IBB), the complex of S100B-Abeta(1-40) peptide from Igor Zhukov Lab, or a study of AlkB dioxygenase and AlkA glycosylase following the interest of Jarosław Poznański Lab (IBB). The case study of Cytochrome C, presented in Figure 8, was conducted and designed to demonstrate HRaDeX usage (Puchała et al., 2025).

HRaDeX offers a method of data aggregation to present all possible aspects of the experiment in a condensed form, allowing for visualisation of data on the 3D structure for a maximally broad overview of the exchange process. Additionally, HRaDeX is an entirely transparent tool with comprehensive documentation. Each intermediate step is described and may be used to double-check the classification workflow. Utilising a well-known equation, all the data transformation is clear to follow and straightforward. Availability in multiple forms, focusing on a ready-to-use, user-friendly web server, ensures a comfortable and personalised approach to addressing users' needs, distinguishing HRaDeX from other tools.

Availability

HRaDeX is available as an R package (https://github.com/hadexversum/HRaDeX) for the programmatic level. HaDeXGUI is available both as an R package (https://github.com/hadexversum/HRaDeXGUI) for local use or as a web server (https://hradex.mslab-ibb.pl/). Similarly, compaHRaDeX is available as an R package (https://github.com/hadexversum/compaHRaDeX) and as a web server (https://compahradex.mslab-ibb.pl/).

List of articles

1. Puchała, W. et al. HaDeX: an R package and web-server for analysis of data from hydrogen–deuterium exchange mass spectrometry experiments. Bioinformatics 36, 4516–4518 (2020).

Contribution: In this work, W.P. researched up-to-date methods of calculation and visualisation, designed and created HaDeX application and its scope, tested the code and wrote the unit tests, and wrote the code and workflow documentation, co-wrote the manuscript draft, and supplementary information.

2. Puchała, W., Kistowski, M., Zhukova, L., Burdukiewicz, M. & Dadlez, M. HRaDeX: R Package and Web Server for Computing High-Resolution Deuterium Uptake Rates for HDX–MS Data. J. Proteome Res. 24, 1688–1700 (2025).

Contribution: In this work, W.P. tested and created the original classification workflow, improved the algorithm upon testing, designed and created the HRaDex and compaHRaDeX applications, wrote the code and algorithm documentation, and co-wrote the article manuscript. W.P. prepared an extensive comparison between experimental data and algorithm results, available as supplementary information to the article.

Conclusions

HDX-MS is a promising and fast-developing field of structural studies. With the method's growing recognition and new ways of utilising it, there is a constant need for novel and creative approaches to data visualisation. New tools should not only provide an easy-to-use interface but also perform analysis in a transparent and reproducible way. To address this popular demand, I developed a comprehensive workflow for data processing at both peptide and high-resolution levels. To my knowledge, it is currently the only actively maintained solution, adapting to growing needs and changing opportunities. HaDeX and HRaDeX are compatible, with HaDeX providing a more general overview and HRaDeX offering a deep dive into details. Both tools were thoroughly discussed with method practitioners, both meritocratically and for the sake of user comfort.

The first step is HaDeX, which verifies that the experiment was conducted correctly and that the obtained data does not raise any concerns. Then, we can assess the trends in uptake, both for single peptides and for the whole peptide pool. For comparative analysis, we can examine the uptake differences between states and identify significant changes, as determined by statistical significance. Sometimes, this information may be sufficient for the experimental aim. HaDeX provides an easy way to pinpoint the regions of interest.

HRaDeX offers a novel approach to detailed analysis, but it requires more involvement and a deep understanding of the exchange process. Choosing the correct definition of exchange classes may take some time, but it enables a deep dive into the uptake phenomenon. HRaDeX presents dense insight, paired with structural information if possible. However, these condensed results can be analysed at different stages to ensure a profound understanding of the exchange specific to the studied protein, or subtle changes that may have been overlooked in the general approach. Additionally, HRaDeX can be used to re-analyse historical data, potentially broadening the understanding of conducted experiments without requiring any additional laboratory work.

Both tools are actively maintained to ensure they are always ready for use. The documentation is kept up to date to assure users that the created “hadexversum”, a family of tools and resources dedicated to the community, is a long-term, pampered project.

Figure 9 presents the data workflow from Figure 5, extended by the tools I built, to help contextualise their use.

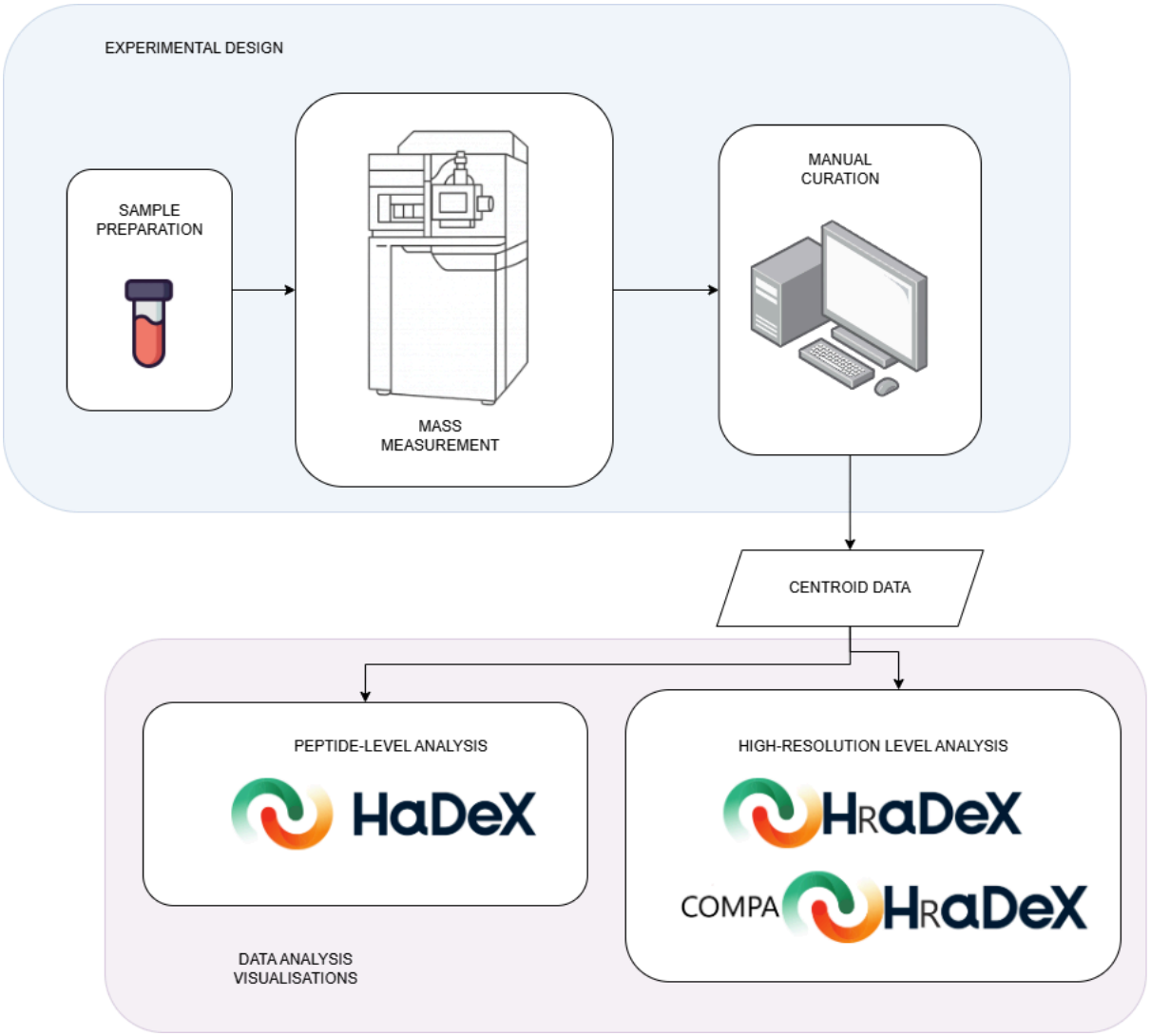


Figure 9. Workflow from experimental data to visualised results.

HaDeX2

HaDeX proved to be extremely useful for the daily work of HDX-MS practitioners. Over time, however, some functionalities offered by HaDeX became unsatisfactory, and new features emerged as expertise grew. This way, the natural need to expand HaDeX has grown, aiming to eliminate the resource gap. HaDeX2 is a comfortable umbrella tool for commonly used visualisation methods and provides novel and unique data presentation methods, conceived alongside field development. The summary of visualisation methods is presented in Figure 10.

Apart from the functionalities implemented in HaDeX, HaDeX2 provides additional options for deuterium uptake visualisation, such as a chiclet plot (presenting colour-coded uptake values for peptides identified by their ID over time) and a butterfly plot (presenting deuterium uptake for peptides identified by their ID at different time points). Moreover, the peptide-level deuterium uptake known from the comparison plot is aggregated into high-resolution and presented on the heatmap and - if possible - on the 3D structure. Every deuterium uptake plot is available in its differential aspect, with recommended hybrid testing (Hageman & Weis, 2019). Quality control is assured by checking, e.g., an uncertainty plot - presenting the uncertainty for each measurement to see if it is acceptable. The coverage heatmap presents the back exchange if a complete deuterated control is provided. If the value exceeds this, it is a sign to check the experimental conditions.

Moreover, HaDeX2 is accompanied by comprehensive documentation that describes the stages of data processing and the elements of the GUI. The function code is well-documented and transparent, with prepared examples of use to ensure correct usage. Additional articles discussing data aggregation and processing can be found in the tool documentation.

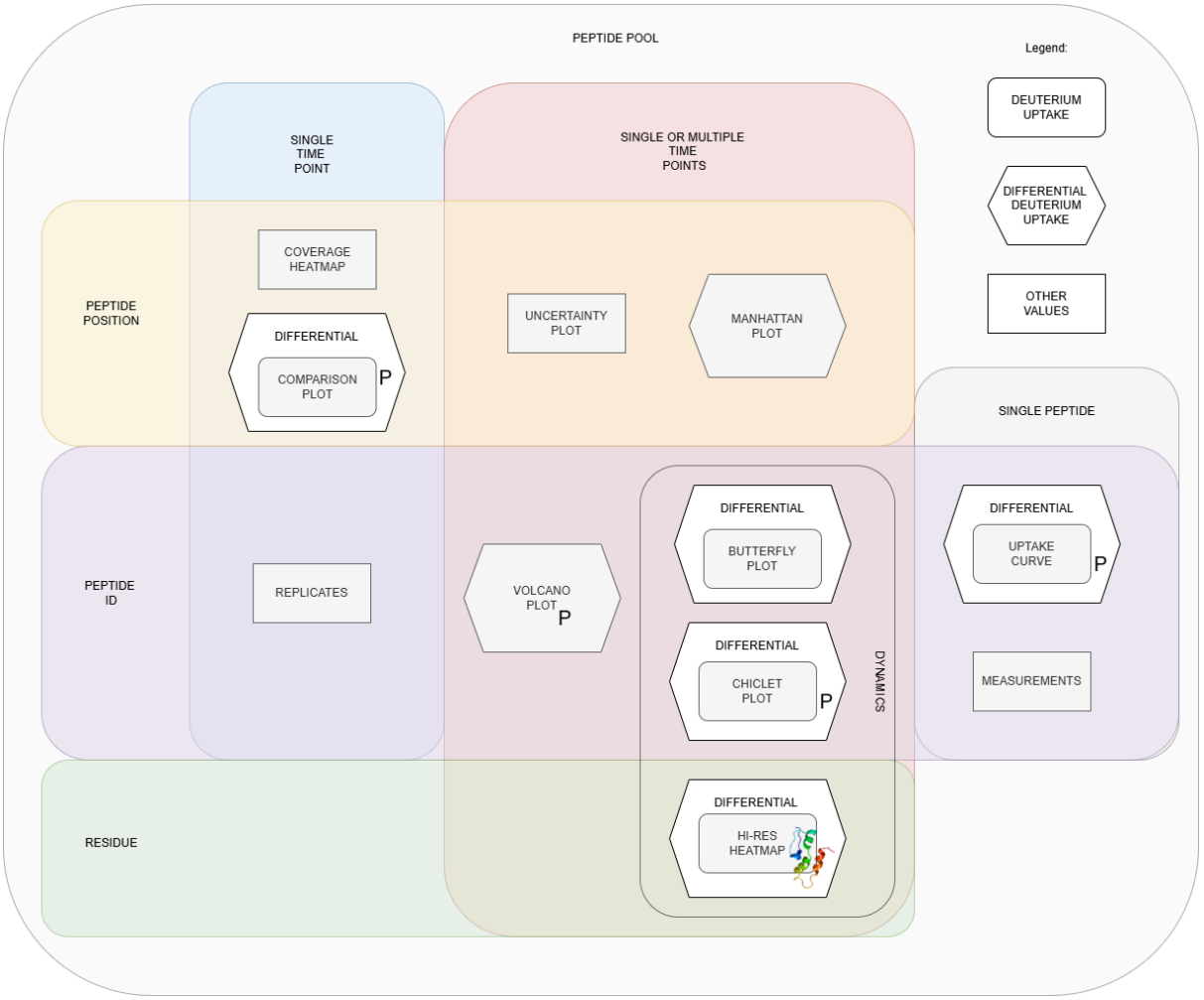


Figure 10. Scheme illustrating the complex spectrum of HDX data visualisation choices.

The field of HDX-MS data analysis remains in need of new methods that provide additional and unique insights into experimental results. Although HaDeX and HRaDeX are filling the gap, there are still ways to improve them and broaden their functionalities with new features that have not yet been described. Future directions include, for instance, isotopic envelope shape, width, and homogeneity analysis, which allows for in-depth analysis of more complex processes, such as EX1 exchange.

Bibliography

Bai, Y., Milne, J. S., Mayne, L., & Englander, S. W. (1993). Primary Structure Effects on Peptide Group Hydrogen Exchange. *Proteins*, *17*(1), 75–86. https://doi.org/10.1002/prot.340170110

Berg, J. M., Tymoczko, J. L., & Stryer, L. (2002). *Biochemistry, Fifth Edition*. W.H. Freeman. https://books.google.pl/books?id=uDFqAAAAMAAJ

Bondarchuk, T. V., Shalak, V. F., Lozhko, D. M., Fatalska, A., Szczepanowski, R. H., Liudkovska, V., Tsuvariev, O. Y., Dadlez, M., El'skaya, A. V., & Negrutskii, B. S. (2022). Quaternary organization of the human eEF1B complex reveals unique multi-GEF domain assembly. *Nucleic Acids Research*, *50*(16), 9490–9504. https://doi.org/10.1093/nar/gkac685

Bouyssié, D., Lesne, J., Locard-Paulet, M., Albigot, R., Burlet-Schiltz, O., & Marcoux, J. (2019). HDX-Viewer: Interactive 3D visualization of hydrogen–deuterium exchange data. *Bioinformatics*, *35*(24), 5331–5333. https://doi.org/10.1093/bioinformatics/btz550

Das, A., Thapa, P., Santiago, U., Shanmugam, N., Banasiak, K., Dabrowska, K., Nolte, H., Szulc, N. A., Gathungu, R. M., Cysewski, D., Krüger, M., Dadlez, M., Nowotny, M., Camacho, C. J., Hoppe, T., & Pokrzywa, W. (2021). *Heterotypic Assembly Mechanism Regulates CHIP E3 Ligase Activity* (p. 2021.08.20.457118). bioRxiv. https://doi.org/10.1101/2021.08.20.457118

Devaurs, D., Antunes, D. A., & Borysik, A. J. (2022). Computational Modeling of Molecular Structures Guided by Hydrogen-Exchange Data. *Journal of the American Society for Mass Spectrometry*, *33*(2), 215–237. https://doi.org/10.1021/jasms.1c00328

Engen, J. R., Botzanowski, T., Peterle, D., Georgescauld, F., & Wales, T. E. (2020). Developments in Hydrogen/Deuterium Exchange Mass Spectrometry. *Analytical Chemistry*. https://doi.org/10.1021/acs.analchem.0c04281

Englander, S. W., Mayne, L., Bai, Y., & Sosnick, T. R. (1997). Hydrogen exchange: The

modern legacy of Linderstrøm-Lang. *Protein Science: A Publication of the Protein Society*, *6*(5), 1101–1109. https://doi.org/10.1002/pro.5560060517

Fan, H. C., Zhang, Y., & Sun, F. (2021). *3.6 A structure of the human hemoglobin: 7vde* [Dataset]. https://doi.org/10.2210/pdb7vde/pdb

Fatalska, A., Hodgson, G., Freund, S. M. V., Maslen, S. L., Morgan, T., Thorkelsson, S. R., Slegtenhorst, M. van, Lorenz, S., Andreeva, A., Kaat, L. D., & Bertolotti, A. (2024). Recruitment of trimeric eIF2 by phosphatase non-catalytic subunit PPP1R15B. *Molecular Cell*, *84*(3), 506-521.e11. https://doi.org/10.1016/j.molcel.2023.12.011

Gessner, C., Steinchen, W., Bédard, S., Skinner, J. J., Woods, V. L., Walsh, T. J., Bange, G., & Pantazatos, D. P. (2017). Computational method allowing Hydrogen-Deuterium Exchange Mass Spectrometry at single amide Resolution. *Scientific Reports*, *7*(1), 3789. https://doi.org/10.1038/s41598-017-03922-3

Grzesiak, K. (2021). *A novel semiparametric model for hydrogen-deuterium exchange monitored by mass spectrometry data*.

Hageman, T. S., & Weis, D. D. (2019). Reliable Identification of Significant Differences in Differential Hydrogen Exchange-Mass Spectrometry Measurements Using a Hybrid Significance Testing Approach. *Analytical Chemistry*, *91*(13), 8008–8016. https://doi.org/10.1021/acs.analchem.9b01325

Hu, Y., Cheng, K., He, L., Zhang, X., Jiang, B., Jiang, L., Li, C., Wang, G., Yang, Y., & Liu, M. (2021). NMR-Based Methods for Protein Analysis. *Analytical Chemistry*, *93*(4), 1866–1879. https://doi.org/10.1021/acs.analchem.0c03830

James, E. I., Murphree, T. A., Vorauer, C., Engen, J. R., & Guttman, M. (2022). Advances in Hydrogen/Deuterium Exchange Mass Spectrometry and the Pursuit of Challenging Biological Systems. *Chemical Reviews*, *122*(8), 7562–7623. https://doi.org/10.1021/acs.chemrev.1c00279

Janowska, M. K., Reiter, K., Magala, P., Guttman, M., & Klevit, R. E. (2024). HDXBoxeR: An R package for statistical analysis and visualization of multiple Hydrogen–Deuterium Exchange Mass-Spectrometry datasets of different protein states. *Bioinformatics*,

40(8), btae479. https://doi.org/10.1093/bioinformatics/btae479

Javed, W., Griffiths, D., & Politis, A. (2023). Hydrogen/deuterium exchange-mass spectrometry of integral membrane proteins in native-like environments: Current scenario and the way forward. *Essays in Biochemistry*, *67*(2), 187–200. https://doi.org/10.1042/EBC20220173

Jelen, M., Grochowina, I., Grabinska-Rogala, A., Ciesielski, S. J., Dabrowska, K., Tomiczek, B., Nierzwicki, L., Delewski, W., Schilke, B., Czub, J., Dadlez, M., Dutkiewicz, R., Craig, E. A., & Marszalek, J. (2023). Analysis of Reconstituted Tripartite Complex Supports Avidity-based Recruitment of Hsp70 by Substrate Bound J-domain Protein. *Journal of Molecular Biology*, *435*(21), 168283. https://doi.org/10.1016/j.jmb.2023.168283

Jensen, P. F., & Rand, K. D. (2016). Hydrogen Exchange. In *Hydrogen Exchange Mass Spectrometry of Proteins* (pp. 1–17). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781118703748.ch1

Keppel, T. R., & Weis, D. D. (2015). Mapping residual structure in intrinsically disordered proteins at residue resolution using millisecond hydrogen/deuterium exchange and residue averaging. *Journal of the American Society for Mass Spectrometry*, *26*(4), 547–554. https://doi.org/10.1007/s13361-014-1033-6

Kochert, B. A., Iacob, R. E., Wales, T. E., Makriyannis, A., & Engen, J. R. (2018). Hydrogen-Deuterium Exchange Mass Spectrometry to Study Protein Complexes. *Methods in Molecular Biology (Clifton, N.J.)*, *1764*, 153–171. https://doi.org/10.1007/978-1-4939-7759-8_10

Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., Funke, R., Gage, D., Harris, K., Heaford, A., Howland, J., Kann, L., Lehoczky, J., LeVine, R., McEwan, P., … The Wellcome Trust: (2001). Initial sequencing and analysis of the human genome. *Nature*, *409*(6822), 860–921. https://doi.org/10.1038/35057062

Lau, A. M. C., Ahdash, Z., Martens, C., & Politis, A. (n.d.). *Deuteros: Software for rapid*

analysis and visualiza- tion of data from differential hydrogen deuterium exchange-mass spectrometry. 4.

Lau, A. M., Claesen, J., Hansen, K., & Politis, A. (2020). Deuteros 2.0: Peptide-level significance testing of data from hydrogen deuterium exchange mass spectrometry. *Bioinformatics*, *btaa677*. https://doi.org/10.1093/bioinformatics/btaa677

Lenda, R., Zhukova, L., Ożyhar, A., & Bystranowska, D. (2024). Deciphering the dual nature of nesfatin-1: A tale of zinc ion's Janus-faced influence. *Cell Communication and Signaling*, *22*(1), 298. https://doi.org/10.1186/s12964-024-01675-x

Lesne, J., Locard-Paulet, M., Parra, J., Zivković, D., Menneteau, T., Bousquet, M.-P., Burlet-Schiltz, O., & Marcoux, J. (2020). Conformational maps of human 20S proteasomes reveal PA28- and immuno-dependent inter-ring crosstalks. *Nature Communications*, *11*(1), Article 1. https://doi.org/10.1038/s41467-020-19934-z

Masson, G. R., Burke, J. E., Ahn, N. G., Anand, G. S., Borchers, C., Brier, S., Bou-Assaf, G. M., Engen, J. R., Englander, S. W., Faber, J., Garlish, R., Griffin, P. R., Gross, M. L., Guttman, M., Hamuro, Y., Heck, A. J. R., Houde, D., Iacob, R. E., Jørgensen, T. J. D., … Rand, K. D. (2019). Recommendations for performing, interpreting and reporting hydrogen deuterium exchange mass spectrometry (HDX-MS) experiments. *Nature Methods*, *16*(7), 595–602. https://doi.org/10.1038/s41592-019-0459-y

Moysa, A., Steczkiewicz, K., Niedzialek, D., Hammerschmid, D., Zhukova, L., Sobott, F., & Dadlez, M. (2021). A model of full-length RAGE in complex with S100B. *Structure*. https://doi.org/10.1016/j.str.2021.04.002

Nguyen, D., Mayne, L., Phillips, M. C., & Walter Englander, S. (2018). Reference Parameters for Protein Hydrogen Exchange Rates. *Journal of the American Society for Mass Spectrometry*, *29*(9), 1936–1939. https://doi.org/10.1007/s13361-018-2021-z

Osadchy, M., & Kolodny, R. (2011). Maps of protein structure space reveal a fundamental relationship between protein structure and function. *Proceedings of the National Academy of Sciences*, *108*(30), 12301–12306. https://doi.org/10.1073/pnas.1102727108

Pancsa, R., Varadi, M., Tompa, P., & Vranken, W. F. (2016). Start2Fold: A database of hydrogen/deuterium exchange data on protein folding and stability. *Nucleic Acids Research*, *44*(D1), D429–D434. https://doi.org/10.1093/nar/gkv1185

Parson, M. A. H., Jenkins, M. L., & Burke, J. E. (2022). Investigating how intrinsically disordered regions contribute to protein function using HDX-MS. *Biochemical Society Transactions*, *50*(6), 1607–1617. https://doi.org/10.1042/BST20220206

Peng, Z., Yan, J., Fan, X., Mizianty, M. J., Xue, B., Wang, K., Hu, G., Uversky, V. N., & Kurgan, L. (2014). Exceptionally abundant exceptions: Comprehensive characterization of intrinsic disorder in all domains of life. *Cellular and Molecular Life Sciences: CMLS*, *72*(1), 137–151. https://doi.org/10.1007/s00018-014-1661-9

Puchała, W., Burdukiewicz, M., Kistowski, M., Dąbrowska, K. A., Badaczewska-Dawid, A. E., Cysewski, D., & Dadlez, M. (2020). HaDeX: An R package and web-server for analysis of data from hydrogen-deuterium exchange mass spectrometry experiments. *Bioinformatics (Oxford, England)*, *36*(16), 4516–4518. https://doi.org/10.1093/bioinformatics/btaa587

Puchała, W., Kistowski, M., Zhukova, L., Burdukiewicz, M., & Dadlez, M. (2025). HRaDeX: R Package and Web Server for Computing High-Resolution Deuterium Uptake Rates for HDX–MS Data. *Journal of Proteome Research*, *24*(4), 1688–1700. https://doi.org/10.1021/acs.jproteome.4c00700

Salmas, R. E., & Borysik, A. J. (2021). HDXmodeller: An online webserver for high-resolution HDX-MS with auto-validation. *Communications Biology*, *4*(1), Article 1. https://doi.org/10.1038/s42003-021-01709-x

Shoichet, B. K., Baase, W. A., Kuroki, R., & Matthews, B. W. (1995). A relationship between protein stability and protein function. *Proceedings of the National Academy of Sciences*, *92*(2), 452–456. https://doi.org/10.1073/pnas.92.2.452

Skinner, S. P., Radou, G., Tuma, R., Houwing-Duistermaat, J. J., & Paci, E. (2019). Estimating Constraints for Protection Factors from HDX-MS Data. *Biophysical Journal*, *116*(7), 1194–1203. https://doi.org/10.1016/j.bpj.2019.02.024

Smit, J. H., Krishnamurthy, S., Srinivasu, B. Y., Karamanou, S., & Economou, A. (2020). PyHDX: Derivation and visualization of protection factors from Hydrogen-Deuterium Exchange Mass Spectrometry at near residue resolution. *bioRxiv*, 2020.09.30.320887. https://doi.org/10.1101/2020.09.30.320887

Stofella, M., Grimaldi, A., Smit, J. H., Claesen, J., Paci, E., & Sobott, F. (2024). Computational Tools for Hydrogen–Deuterium Exchange Mass Spectrometry Data Analysis. *Chemical Reviews*, *124*(21), 12242–12263. https://doi.org/10.1021/acs.chemrev.4c00438

Vinciauskaite, V., & Masson, G. R. (2023). Fundamentals of HDX-MS. *Essays in Biochemistry*, *67*(2), 301–314. https://doi.org/10.1042/EBC20220111

Watson, H. C., & Kendrew, J. C. (1976). *The stereochemistry of the protein myoglobin: 1mbn* [Dataset]. https://doi.org/10.2210/pdb1mbn/pdb

Weis, D. D. (2021). Recommendations for the Propagation of Uncertainty in Hydrogen Exchange-Mass Spectrometric Measurements. *Journal of the American Society for Mass Spectrometry*, *32*(7), 1610–1617. https://doi.org/10.1021/jasms.0c00475

Whisstock, J. C., & Lesk, A. M. (2003). Prediction of protein function from protein sequence and structure. *Quarterly Reviews of Biophysics*, *36*(3), 307–340. https://doi.org/10.1017/S0033583503003901

Zhang, N., Yu, X., Zhang, X., & D'Arcy, S. (2021). HD-eXplosion: Visualization of hydrogen–deuterium exchange data as chiclet and volcano plots with statistical filtering. *Bioinformatics*, *37*(13), 1926–1927. https://doi.org/10.1093/bioinformatics/btaa892

Zhang, Z., Post, C. B., & Smith, David. L. (1996). Amide Hydrogen Exchange Determined by Mass Spectrometry: Application to Rabbit Muscle Aldolase. *Biochemistry*, *35*(3), 779–791. https://doi.org/10.1021/bi952227q

Zhang, Z., & Smith, D. L. (1993). Determination of amide hydrogen exchange by mass spectrometry: A new tool for protein structure elucidation. *Protein Science*, *2*(4), 522–531. https://doi.org/10.1002/pro.5560020404

Bioinformatics, 36(16), 2020, 4516–4518
doi: 10.1093/bioinformatics/btaa587
Advance Access Publication Date: 24 June 2020
Applications Note

Structural bioinformatics

HaDeX: an R package and web-server for analysis of data from hydrogen–deuterium exchange mass spectrometry experiments

Weronika Puchała[1,†], Michał Burdukiewicz [2,†], Michał Kistowski[1], Katarzyna A. Dąbrowska[1], Aleksandra E. Badaczewska-Dawid [3], Dominik Cysewski [1,*] and Michał Dadlez[1,*]

[1]Institute of Biochemistry and Biophysics, Polish Academy of Sciences, Warsaw 02-106, Poland, [2]Faculty of Mathematics and Information Science, Warsaw University of Technology, Warsaw 00-662, Poland and [3]Faculty of Chemistry, Biological and Chemical Research Center, University of Warsaw, Warsaw 02-093, Poland

*To whom correspondence should be addressed.

Associate Editor: Yann Ponty

[†]The authors wish it to be known that, in their opinion, the first two authors should be regarded as Joint First Authors.

Received on August 29, 2019; revised on May 8, 2020; editorial decision on June 12, 2020; accepted on June 16, 2020

Abstract

Motivation: Hydrogen–deuterium mass spectrometry (HDX-MS) is a rapidly developing technique for monitoring dynamics and interactions of proteins. The development of new devices has to be followed with new software suites addressing emerging standards in data analysis.

Results: We propose HaDeX, a novel tool for processing, analysis and visualization of HDX-MS experiments. HaDeX supports a reproducible analytical process, including data exploration, quality control and generation of publication-quality figures.

Availability and implementation: HaDeX is available primarily as a web-server (http://mslab-ibb.pl/shiny/HaDeX/), but its all functionalities are also accessible as the R package (https://CRAN.R-project.org/package=HaDeX) and standalone software (https://sourceforge.net/projects/HaDeX/).

Contact: michald@ibb.waw.pl

Supplementary information: Supplementary data are available at *Bioinformatics* online.

1 Introduction

The understanding of interactions between proteins and other molecules is crucial for studying complex biological systems. Among the methods for characterization of conformational dynamics of proteins and their complexes, hydrogen–deuterium mass spectrometry (HDX-MS) has proven to be both rapid and sensitive (Konermann et al., 2011). This technique is especially important in the case of proteins that are difficult to study with other methods such as membrane proteins, oligomerizing proteins or intrinsically disordered proteins (Goswami et al., 2013).

Hydrogen–deuterium exchange monitors an exchange of amide hydrogens in peptide bonds. Protein incubation in D_2O leads to the exchange of hydrogen to deuterium atoms in amides. The speed of such exchange depends mostly on the stability of hydrogen bonds formed by the hydrogen amides and also by the accessibility to the solvent. The importance of the influence of these factors is discussed, but it is considered that the dominant component is the stability of hydrogen bonds and not the availability of solvent. Thus, this process may be slowed due to structural factors: stability, flexibility and accessibility. Therefore the method, by measuring the level of protection of amides at different times of incubation in D_2O, recognizes the stability of hydrogen bonding networks and regions of protein with limited solvent accessibility. HDX probes the dynamic nature of proteins systems, as opposed to static structures, offered by X-ray crystallography. It also allows mapping the regions affected by the interaction between the proteins concerned.

The main scheme of local (continuous-labeling, bottom-up) HDX-MS experiments consist of: the incubation of a protein in a D_2O solution, buffered to native or native-like conditions, followed by exchange quench, proteolytic digestion and mass spectrometry measurement of resulting peptide's masses.

Results generated by HDX-MS are complex and demanding in terms of analysis, interpretation and visual presentation. While there

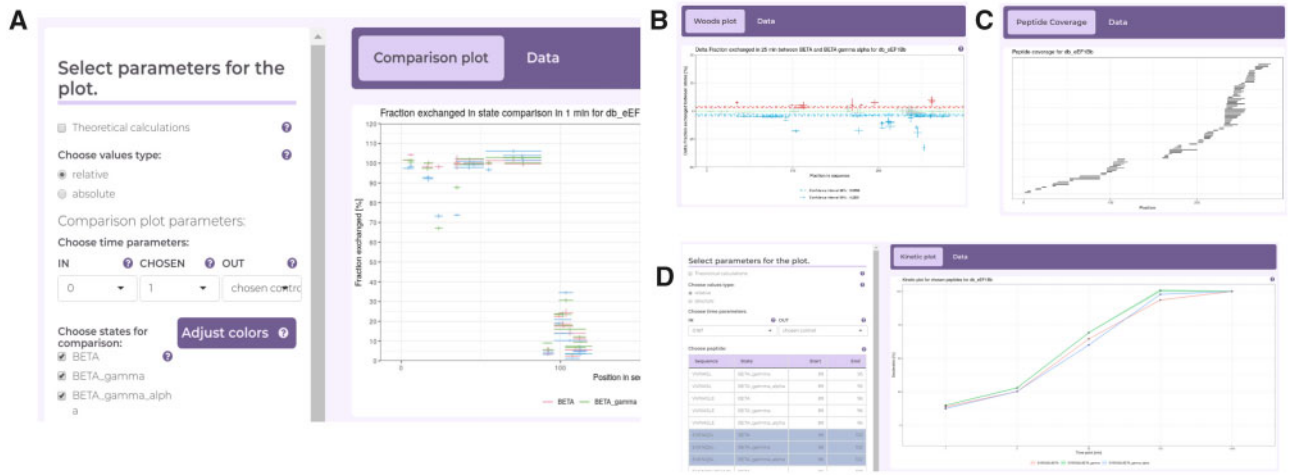


Fig. 1. The core functionalities of the HaDeX GUI. (**A**) Multistate discriminative analysis of peptides; (**B**) Woods plots; (**C**) coverage of the protein sequence by peptides measured with mass spectrometry and (**D**) kinetics of hydrogen–deuterium exchange

are many open-source and free to use software packages addressing the challenges of HDX-MS data analysis (Hourdel et al., 2016; Kavan and Man, 2011; Lau et al., 2019; Lumpkin and Komives, 2019), HaDeX aims to cover post-processing workflow, where results of the experiment are analyzed and presented in a publication-friendly format (for comparison of mentioned tools see Supplementary Information SI1). It forces HDX-MS users to rely on several pieces of software, thus making the already laborious process even more time-consuming. Yet, another challenge of the HDX-MS data analysis is a proper visualization of results on 3D protein structure which is supported by HDX-Viewer (Bouyssié et al., 2019). As the field HDX-MS is still growing, researchers introduce new standards, including data analysis and reporting (Masson et al., 2019). The available software does not allow automatic generation of reports according to the new guidelines, which adds work for experimentalists.

To alleviate these issues, we propose HaDeX, a comprehensive software suite for analysis of HDX-MS data. The aim of HaDeX is not only to provide a comprehensive way to study results of HDX-MS experiments but also to report their results in a reproducible way by including all parameters relevant to data analysis as the size of confidence intervals.

2 Materials and methods

HaDeX dissects work into three steps: (i) general properties of a sequence reconstructed from measured peptides, (ii) uncertainty of measurements and their significance and (iii) visualization of results (Fig. 1).

The only input necessary to start work with HaDeX is an exported data as a.csv datafile (in the Cluster format) produced by the DynamX^{TM} 2.0 or 3.0 (Waters Corp.) Our software does not require any external preprocessing, which not only streamlines the whole workflow but also increases its reproducibility.

HaDeX uses a well-established method to compute confidence intervals for measured peptides (Houde et al., 2011) (see Supplementary Information SI2.5). Additionally, we enhanced this functionality by providing uncertainties derived by error propagation (Joint Committee for Guides in Metrology, 2008) (see Supplementary Information SI2.2).

Known in the literature as Woods charts (Woods and Hamuro, 2001), these types of plots are used to visually inspect results of HDX-MS studies (Kupniewska-Kozak et al., 2010). A user can further enhance these charts by indicating specified confidence levels (Fig. 1B). All figures are exportable in vector formats.

HaDeX provides a highly customizable report generation module, which increases the reproducibility of its analytic workflow. The report not only contains partial results of the analysis, but also the additional input provided by the user (e.g. altered significance levels).

3 Conclusion and availability

HaDeX supports a large part of the analytic workflow of HDX-MS data (Claesen and Burzykowski, 2017), from the quality control of the data to publication-ready visualizations. However, our tool does not provide any high-resolution output (as 3D visualizations or deuteration heatmaps based on single residues) thus we refer the user to methods addressing this problem (Gessner et al., 2017). Thanks to the input of HDX-MS users, our software is not only a convenient re-implementation of already existing methods but also provides unique functionalities unavailable elsewhere: novel, ISO-based uncertainty computations, multistate analysis and downloadable reports produced according to the novel guidelines (Masson et al., 2019). As our tool is targeted at experimentalists, it is available as a web server and a standalone GUI (Windows only). However, bioinformaticians can access HaDeX functions programatically, as it is also available as an R package. We hope that thanks to its comprehensiveness and reproducibility-oriented features HaDeX can satisfy requirements of users from both academia and industry.

Acknowledgements

The authors would like to thank Julien Marcoux (IPBS-Toulouse) and members of the Mass Spectrometry Core Facility (Institute of Biochemistry and Biophysics Polish Academy of Sciences) for suggestions with software features.

Funding

Foundation of Polish Science TEAM TECH CORE FACILITY/2016-2/2 Mass Spectrometry of Biopharmaceuticals—improved methodologies for qualitative, quantitative and structural characterization of drugs, proteinaceous drug targets and diagnostic molecules.

Conflict of Interest: none declared.

References

Bouyssié,D. *et al.* (2019) HDX-Viewer: interactive 3D visualization of hydrogen–deuterium exchange data. *Bioinformatics*, **35**, 5331–5333.

Claesen,J. and Burzykowski,T. (2017) Computational methods and challenges in hydrogen/deuterium exchange mass spectrometry. *Mass Spectr. Rev.*, **36**, 649–667.

Gessner,C. *et al.* (2017) Computational method allowing hydrogen–deuterium exchange mass spectrometry at single amide resolution. *Sci. Rep*., **7**, 3789.
Goswami,D. *et al.* (2013) Time window expansion for HDX analysis of an intrinsically disordered protein. *J. Am. Soc. Mass Spectr*., **24**, 1584–1592.
Houde,D. *et al.* (2011) The utility of hydrogen/deuterium exchange mass spectrometry in biopharmaceutical comparability studies. *J. Pharm. Sci*., **100**, 2071–2086.
Hourdel,V. *et al.* (2016) MEMHDX: an interactive tool to expedite the statistical validation and visualization of large HDX-MS datasets. *Bioinformatics*, **32**, 3413–3419.
Joint Committee for Guides in Metrology. (2008) JCGM 100: evaluation of measurement data – guide to the expression of uncertainty in measurement. *Technical report*, JCGM.
Kavan,D. and Man,P. (2011) MSTools—web based application for visualization and presentation of HXMS data. *Int. J. Mass Spectr*., **302**, 53–58.
Konermann,L. *et al.* (2011) Hydrogen exchange mass spectrometry for studying protein structure and dynamics. *Chem. Soc. Rev*., **40**, 1224–1234.
Kupniewska-Kozak,A. *et al.* (2010) Intertwined structured and unstructured regions of exrage identified by monitoring hydrogen–deuterium exchange. *J. Mol. Biol*., **403**, 52–65.
Lau,A.M.C. *et al.* (2019) Deuteros: software for rapid analysis and visualization of data from differential hydrogen deuterium exchange-mass spectrometry. *Bioinformatics*, **35**, 3171–3173.
Lumpkin,R. and Komives,E.A. (2019) DECA, a comprehensive, automatic post-processing program for HDX-MS data. *Mol. Cell. Proteomics*., **18**, 2516–2523.
Masson,G.R. *et al.* (2019) Recommendations for performing, interpreting and reporting hydrogen deuterium exchange mass spectrometry (HDX-MS) experiments. *Nat. Methods*, **16**, 595–602.
Woods,V.L. and Hamuro,Y. (2001) High resolution, high-throughput amide deuterium exchange-mass spectrometry (DXMS) determination of protein binding site structure and dynamics: utility in pharmaceutical design. *J. Cell. Biochem*., **37**, 89–98.

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".

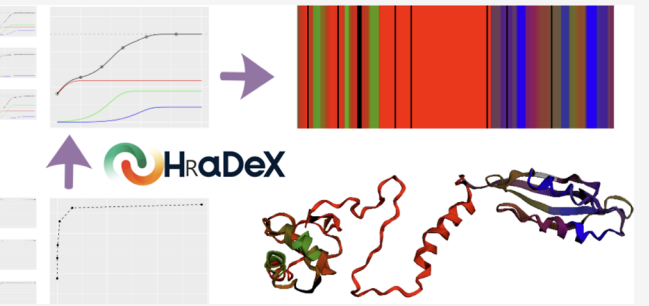
Weronika Puchała, Michał Kistowski, Liliya Zhukova, Michał Burdukiewicz, and Michał Dadlez*

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*

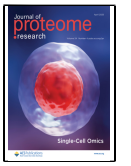
INTRODUCTION

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



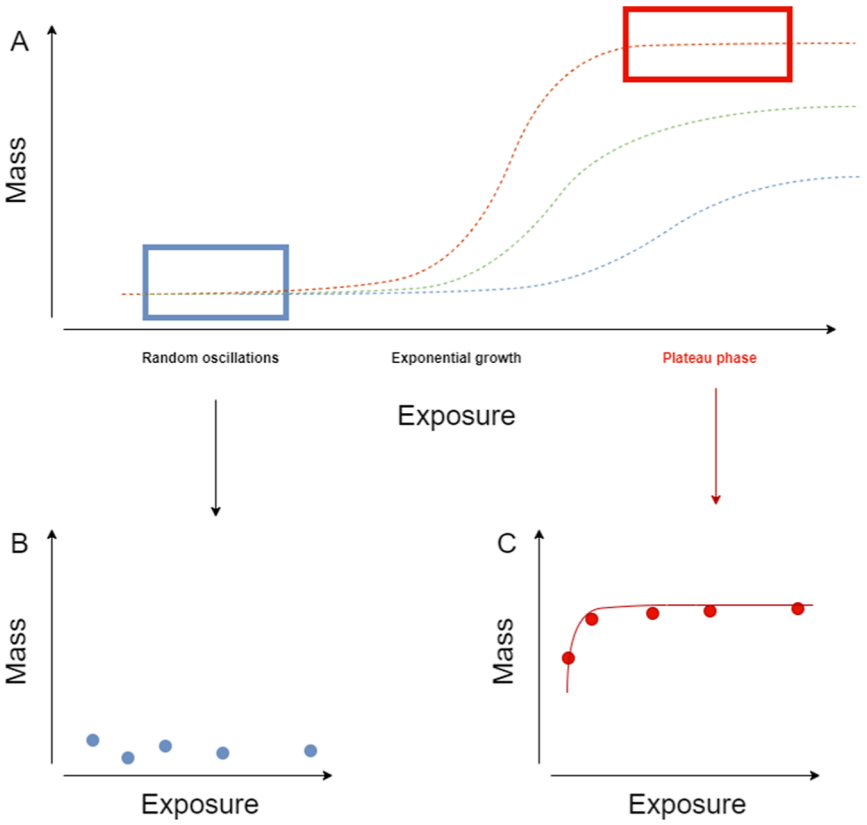


Figure 1. (A) Exponential nature of HDX exemplified by a typical time course of mass increase after deuterium uptake (vertical axis) according to a ZS model, which assumes that each amide can be assigned to one of the three exchange regimes, fast, intermediate, or slow. The exposure time axis (horizontal) 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

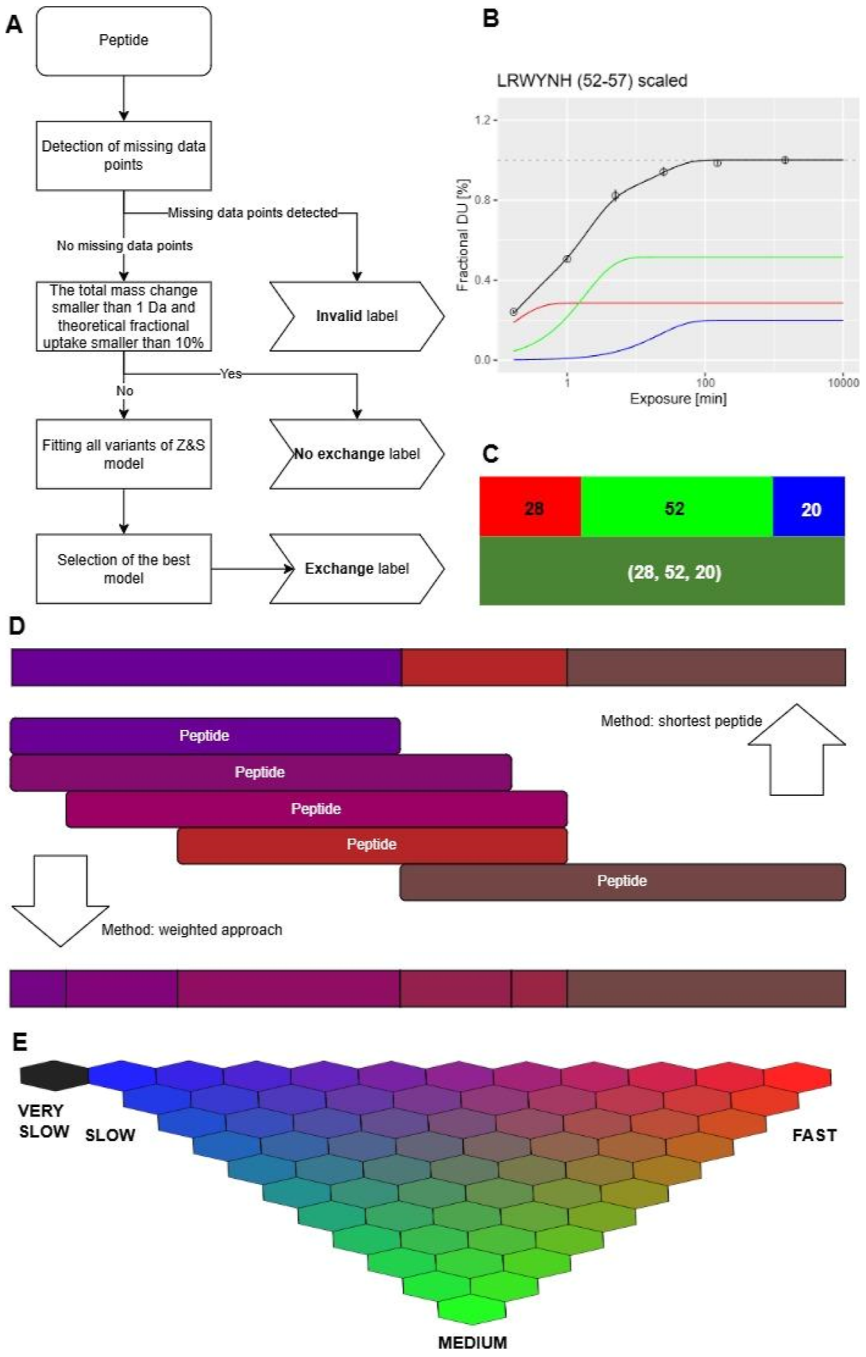


Figure 2. (A) HRaDeX algorithm. First, the data set is checked for completeness of time data points. Peptides with missing measurements or lacking uptake control are marked as "invalid" and excluded from further analysis. Next, we detect uptake trajectories with no or a very slow exchange and assign 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

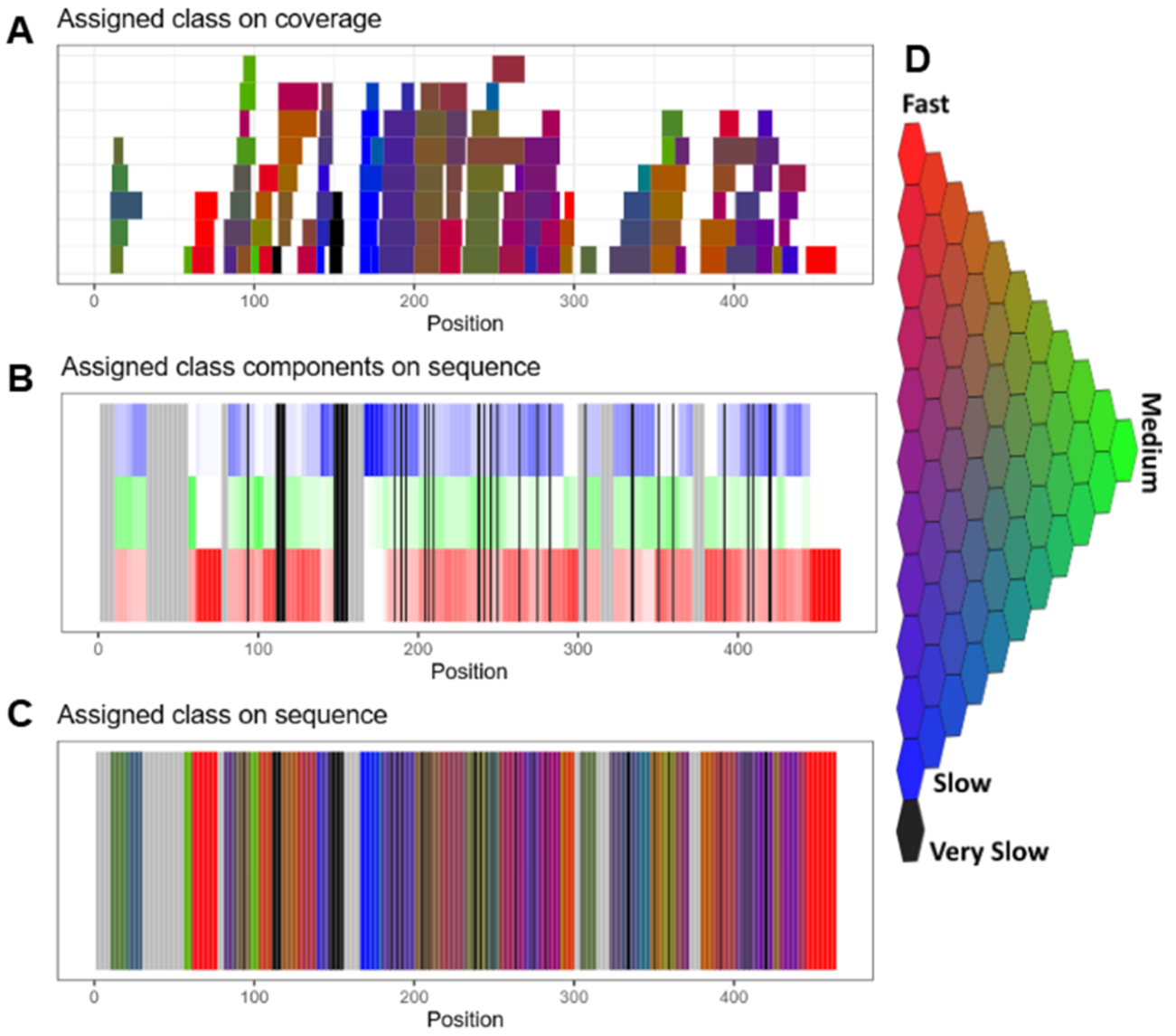


Figure 3. Example of the overlay of color-coded HDX results on protein sequence. Data for the eEF1A1 protein (PRIDE PXD056814) are shown. Panel D shows a color palette enabling the description of fast (red), intermediate (green), or slow (blue) components as well as mixed states whenever the uptake curve 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

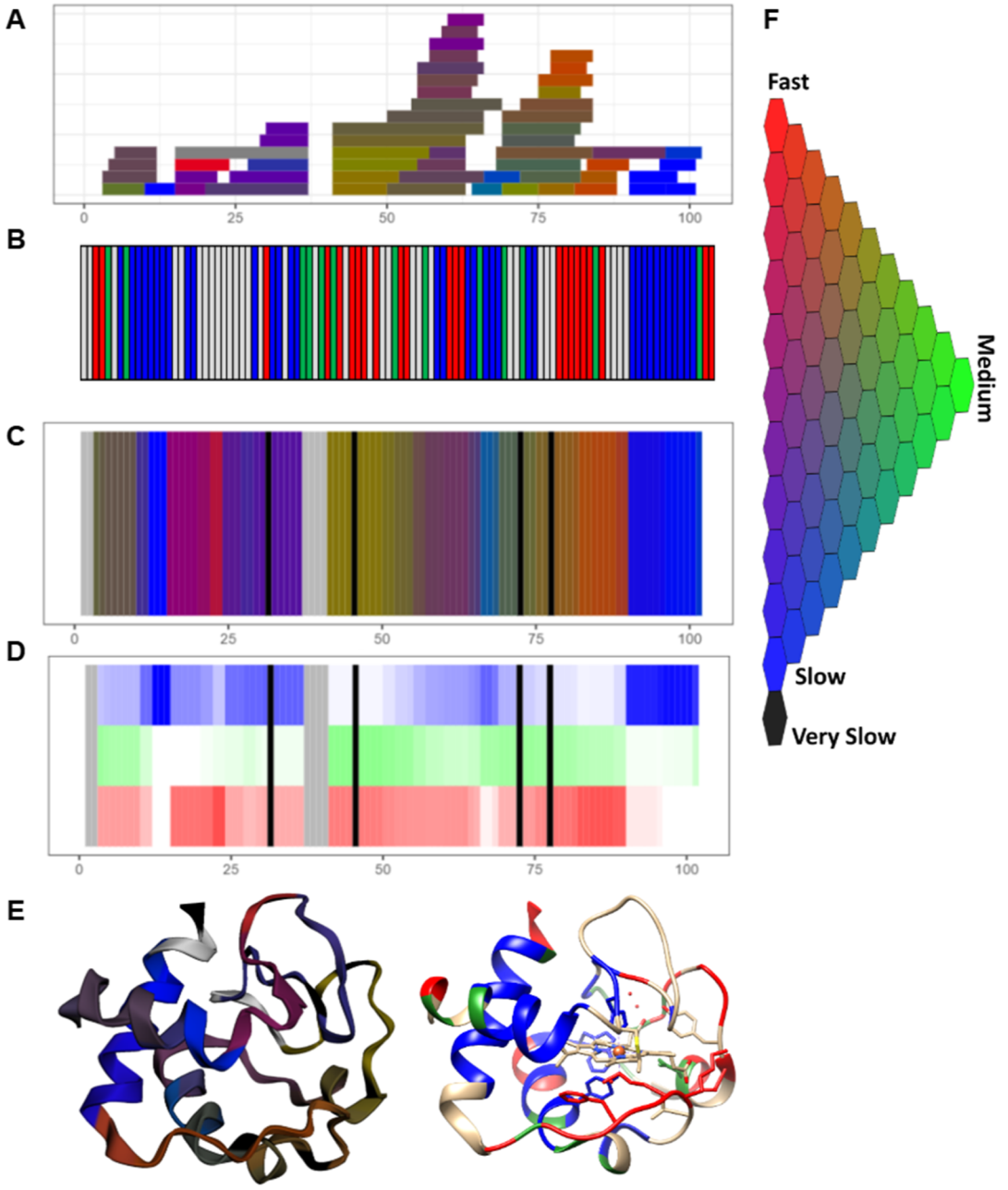


Figure 4. Comparison of color-coded assignment of exchange class to regions of cytochrome *C*. The assignment was carried out either by using HRaDex or a set of single-site exchange rates measured by NMR or ETD−MS−HDX previously (see text). Single residue rates measured by NMR or ETD−MS−HDX were assigned 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

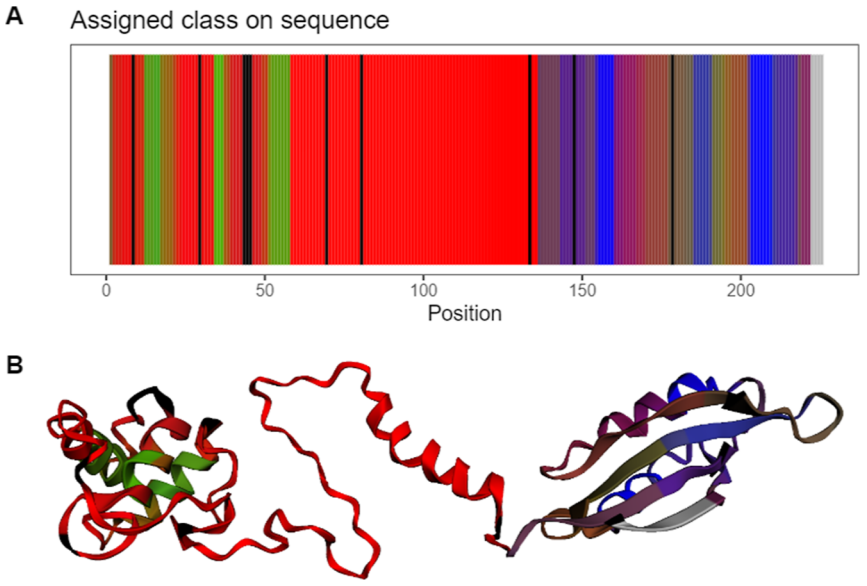


Figure 5. Results of the HRaDeX analysis of the eEF1Bα HDX−MS experiment. (A) Barcode plot of eEF1Bα. (B) HRaDeX output is visualized on the predicted structure of eEF1Bα.[21]

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

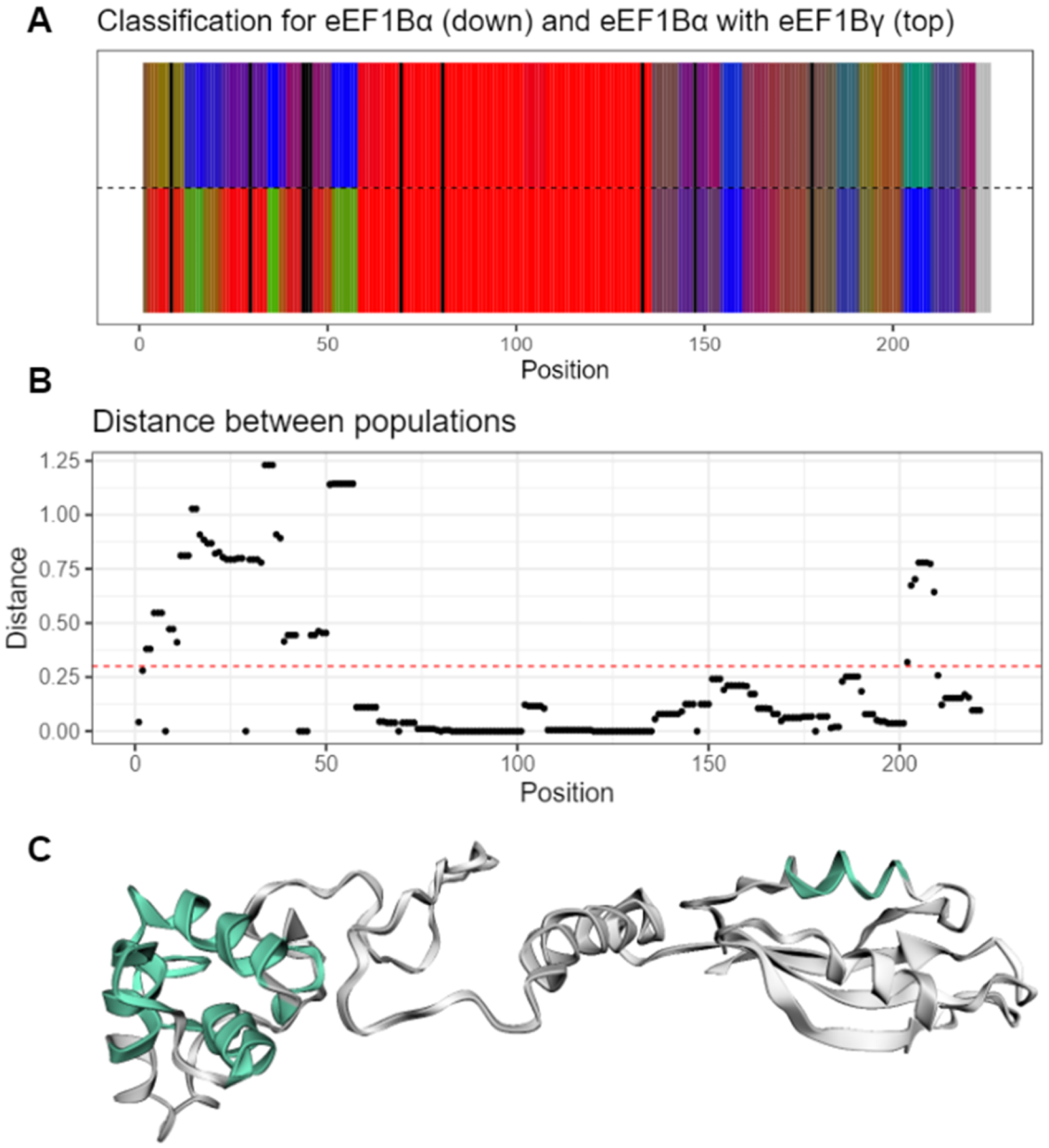


Figure 6. (A) Barcode plot of the compaHRaDeX analysis of eEF1Bα exchange in isolation and in complex with eEF1Bγ subunit, computed with the weighted average approach. (B) Distance (eq 6) between assigned exchange classes of eEF1Bα exchange in isolation and in complex with eEF1Bγ subunit. The red dashed 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

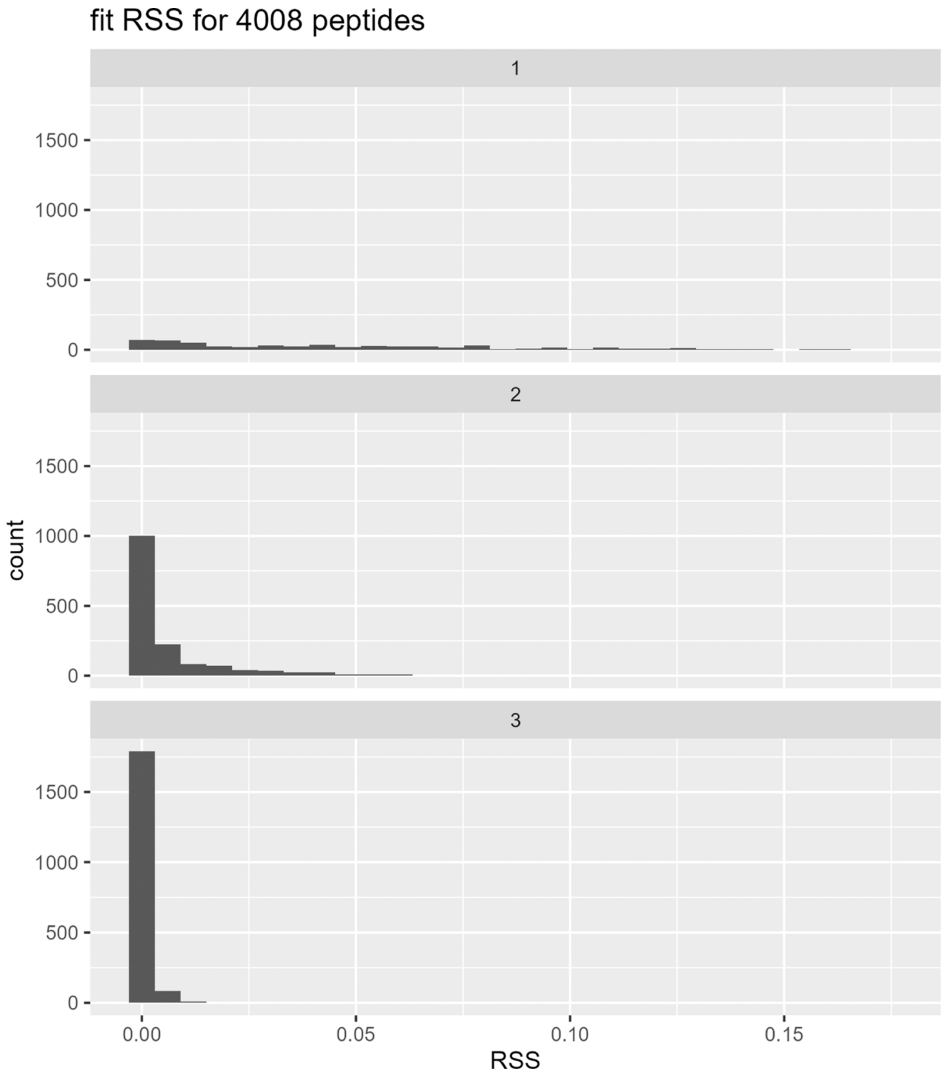


Figure 7. RSS of fits for peptides considered in HRaDeX validation. Each facet represents a different ZS model variant (eqs 1, 3, and 4, respectively).

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

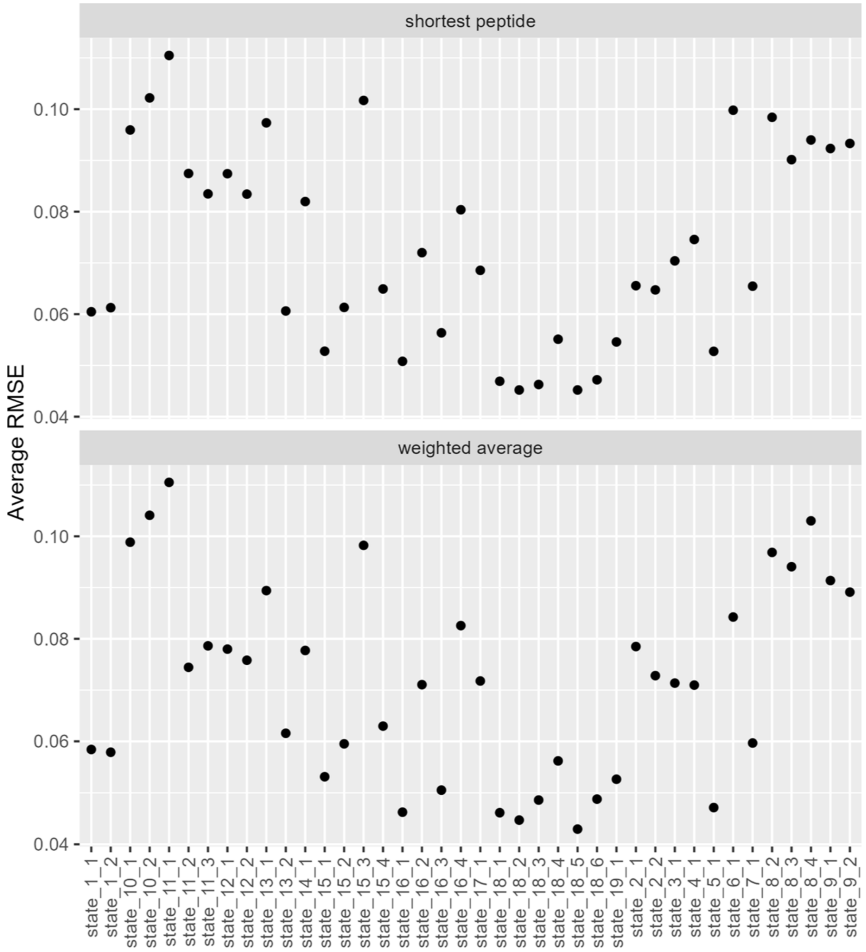


Figure 8. Summary of averaged RMSE for all gathered data sets, obtained in the HraDeX analysis using default parameters, calculated between experimental uptake curves and curves reconstituted from high-resolution analysis parameters obtained either by the "shortest peptide" (upper panel) or "weighted 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