

Protein structure prediction using residue-resolved protection factors from hydrogen-deuterium exchange NMR

Tung T. Nguyen,² Daniel R. Marzolf,¹ Justin T. Seffernick,¹ Sten Heinze,¹ and Steffen Lindert^{1,*}

¹Department of Chemistry and Biochemistry, Ohio State University, Columbus, OH, 43210

²Department of Chemistry and Biochemistry, Denison University, Granville, OH, 43023

Summary

Hydrogen-deuterium exchange (HDX) measured by NMR provides structural information for proteins relating to solvent-accessibility and flexibility. While this structural information is beneficial, the data cannot be used exclusively to elucidate structures. However, the structural information provided by the HDX-NMR data can be supplemented by computational methods. In previous work, we developed an algorithm in Rosetta to predict structures using qualitative HDX-NMR data (categories of exchange rate). Here we expand on the effort, and utilize quantitative protection factors (PFs) from HDX-NMR for structure prediction. From observed correlations between PFs and solvent accessibility/flexibility measures, we present a scoring function to quantify the agreement with HDX data. Using a benchmark set of 10 proteins, an average improvement of 5.13 Å in root-mean-square deviation (RMSD) is observed for cases of inaccurate Rosetta predictions. Ultimately, 7/10 predictions are accurate without including HDX data, and 9/10 are accurate when using our PF-based HDX Score.

* Lead contact: Steffen Lindert

Department of Chemistry and Biochemistry, Ohio State University

2114 Newman & Wolfrom Laboratory, 100 W. 18th Avenue, Columbus, OH 43210

614-292-8284 (office), 614-292-1685 (fax)

lindert.1@osu.edu

Introduction

Protein function is encoded in the structure of the protein, allowing for understanding and manipulation of both mechanisms and function if the structure can be determined experimentally. However, experimental structure determination is challenging, particularly for more complex and disordered proteins, with some proteins eluding high resolution structure determination due to their intrinsic properties. As a result, the gap between known protein sequences, which can be quickly determined, and known protein structures increases by the day. This gap is fueled by the difficulty of structural biology experiments, such as X-ray crystallography, cryogenic electron microscopy (cryo-EM), and ¹³C-, ¹⁵N-edited nuclear magnetic resonance (NMR) experiments, all of which have limitations associated with them (Ardenkjaer-Larsen et al., 2015; Lyumkis, 2019; Srivastava et al., 2018).

While tertiary structure has traditionally been determined using the aforementioned methodologies, in recent decades the advent of computational methods has led to potentially far higher throughput structure prediction. With these predictions, however, come limitations such as increased probability of inaccurate *ab initio* structure predictions as protein size increases (Kim et al., 2009). The accuracy of computational methods can be significantly improved by the inclusion of experimental data (Seffernick and Lindert, 2020). Inclusion efforts began in the 1980s with NMR and X-ray crystallography, with more recent studies aiming towards including data from lower-resolution methods, such as electron paramagnetic resonance (EPR), mass spectrometry (MS), and cryo-EM (Alexander et al., 2008; Aprahamian et al., 2018; Aprahamian and Lindert, 2019; Biehn and Lindert, 2021; Bowers et al., 2000; DiMaio et al., 2015; Harvey et al., 2019; Lindert et al., 2012; Pilla et al., 2017; Roberts et al., 2017; Seffernick et al., 2019a; Srivastava et al., 2018; van Zundert et al., 2015). Ideally, these types of experimental data would be high-throughput and require a smaller sample size.

Biomolecular hydrogen-deuterium exchange (HDX) experiments, originating in the 1970s, have typically been used to map exchange rates onto atomic-resolution structures, elucidating dynamic properties using static models (Choe et al., 1998; Hooke et al., 1994; Palmer, 1997; Rosa and Richards, 1979). These studies have probed native-state dynamics on a diverse set of proteins, as well as mapped protein folding pathways by performing the experiments as the molecule is (un)folding (Di Paolo et al., 2010; Rogov et al., 2004; Schulman et al., 1995). Determination of dynamic properties stem from a consensus that HDX rates are generally governed by exposure and flexibility of the amide protons, thus regions of faster exchange can be correlated to regions of higher flexibility and/or solvent exposure. Additionally, a wide variety of studies have demonstrated the importance of accounting for dynamic properties by using sophisticated sampling methods (such as molecular dynamics simulations) to match structures to experimental data, as well as using these methods to better understand the factors influencing HDX (Best and Vendruscolo, 2006; Devaurs et al., 2017; Devaurs et al., 2018; Hilser and Freire, 1996; Makarov et al., 2020; Martens et al., 2019; McAllister and Konermann, 2015; Mohammadiarani et al., 2018; Petruk et al., 2013; Vendruscolo et al., 2003; Wan et al., 2020). Studies to simulate HDX data (Liu et al., 2012), reweigh model ensembles (Bradshaw et al., 2020; Craig et al., 2011; Wan et al., 2020), and evaluate protein-protein interactions (Borysik, 2017) require structural information and long, computationally expensive trajectories obtained from molecular dynamics simulations.

In absence of existing structural models, the data can also be used as restraints in computational protein structure prediction. Multiple studies have demonstrated that modeling based on agreement to HDX data measured from MS can be beneficial to enhance structural understanding of specific systems. For structure prediction of monomeric proteins from sequence,

approaches using homology modeling along with HDX-MS data can be successful, even for systems that are difficult to crystallize (Ramsey *et al.*, 2018; Zhang *et al.*, 2014). Protein-protein docking has also been performed in combination with HDX-MS data, allowing for the prediction of complex structures in agreement with the data (Borysik, 2017; Roberts *et al.*, 2017; Zhang *et al.*, 2019).

In previous work, we have demonstrated that structural features relating to exposure and flexibility of amide protons are correlated to experimental HDX rates measured using NMR (Marzolf *et al.*, 2021). Using a benchmark set of 38 proteins, a correlation between HDX rate and four calculated parameters related to exposure and flexibility was established. In brief, exposure was quantified using a calculated amide proton neighbor count (NC) and relative solvent accessible surface area (RelSASA) (Aprahamian *et al.*, 2018). Flexibility was estimated using the order score (OS) (Kim *et al.*, 2018; Seffernick *et al.*, 2019b), a window-averaged residual Rosetta score, and hydrogen-bonding energies involving the amide proton. After developing a score term that evaluated agreement of the four parameters to experimental HDX rate categories in a large set of decoy structures, we were able to improve prediction accuracy, reducing root-mean-square-deviation (RMSD) substantially. However, this scoring term used strength categories which corresponded to HDX rates rather than the quantitative exchange rates themselves, introducing a potential source of uncertainty.

In this report, we extend the methodologies from the previous work to directly employ quantitative HDX rates during the protein modeling using a benchmark set of 10 proteins. The quantitative rates, reported as protection factors, were used to predict parameters which related to exposure and flexibility. The predicted values were then compared to calculated values in models generated by Rosetta, with the difference in the values incorporated in a Rosetta scoring term. This method differs from our previous work (using quantitative protection factors, rather than categories of exchange) and other methods in the literature, which are generally used to match agreement for a single system (monomer or complex) using HDX-MS. In this report, we directly incorporated HDX rates measured from NMR experiments into a general method, available in Rosetta, to predict monomer structure from sequence alone. Using this new scoring term, protein structure prediction accuracy increased, moving from 7/10 to 9/10 of the benchmark proteins predicting an accurate model (< 5.5 Å RMSD to the native model) when the HDX data was included.

Results and Discussion

Previous work showed HDX-NMR data provide useful information to structure prediction

Experimental data from hydrogen-deuterium exchange (HDX) measured by NMR provide structural information on proteins. While this information alone cannot be used to fully elucidate the structures, it can be utilized as a supplement to computational prediction methods. Recently, we developed a method to use HDX-NMR data to predict tertiary structure from sequence using Rosetta *ab initio* structure prediction (Marzolf *et al.*, 2021). In this previous work, the experimental data were categorized as *strong*, *medium*, or *weak*, corresponding to slow, medium, and fast exchange, respectively. The HDX score term was dependent on the following four residue-resolved metrics that were expected to influence the HDX rate: neighbor count (NC), relative solvent accessible surface area (RelSASA), hydrogen bond score (HB), and Rosetta order score [OS, using ResidueDisorder (Kim *et al.*, 2018; Seffernick *et al.*, 2019b)]. In the previous study, structure prediction results improved when qualitative HDX data were included (Marzolf *et al.*, 2021).

Protection factors provided more information for modeling and trends followed expected structural hypotheses

While results from previous work were promising and showed that a very small amount of experimental data can meaningfully improve prediction results, there was a potential to include more information from HDX-NMR experiments into structure prediction. In this study, rather than using qualitative data (where measured exchange rates were categorized into *strong*, *medium*, and *weak*), we used quantitative protection factors (PFs) measured from HDX-NMR experiments. Protection factors are defined as the ratio of the sequence-dependent intrinsic HDX rate constant to observed exchange rate constant. Residues (specifically backbone amide hydrogens) with higher protection factor (corresponding to lower relative exchange rate) are expected to be less flexible (i.e., participate strongly in hydrogen bonding) and/or have less exposure to the solvent. Similar to previous work, here we used HB and OS to quantify flexibility, and NC and RelSASA to quantify solvent exposure, but have done so using protection factors.

Quantitative protection factors were not as readily available as qualitative exchange rate categories for a large number of proteins. Our benchmark dataset contained 10 proteins (ranging from 76-223 residues; summarized in Table S1) for which experimental protection factors were provided in the literature. For this dataset, we first sought to test our hypotheses (high exposure and high flexibility leading to increased exchange) by examining the correlations between the calculated residue-resolved parameters and PF. Figure 1 shows the NC, RelSASA, HB, and OS as a function of $\log(\text{PF})$ for residues in all 10 proteins (a total of 431 residues). The values of slopes and intercepts for each correlation are provided in Table S3. While the individual correlations did not appear to be exceptionally strong (R^2 of 0.12, 0.05, 0.05, and 0.04 respectively), the hypothesized trends were observed for all four parameters. The relatively weak correlations were unsurprising because no individual parameter can be expected to govern exchange in a complex process such as HDX. Nonetheless, based on these native distributions, we developed a scoring function to be used to quantify the quality of a predicted protein structure based on its agreement with the NMR-HDX data.

Inclusion of HDX-NMR PF was beneficial for structure prediction

Rosetta *ab initio* structure prediction was used to predict structures from sequence for the majority of proteins in the dataset (8/10). However, due to poor observed sampling, homology models were instead generated for 1SNO and 2ETL using RosettaCM. Regardless of method, 10,000 models were generated for each protein in the dataset and scored using the Rosetta REF2015 scoring function. For each protein, within the respective pool of models, structures with RMSD (root-mean-square deviation with respect to the crystal structures) of less than 4 Å were sampled. The predicted structure without using HDX-NMR data was the lowest scoring model by Rosetta score. The accuracy of the predicted structures varied, with the RMSDs ranging from 0.67 Å to 14.75 Å, with 7/10 less than 5.5 Å (see Figure 2).

To test the ability of HDX-NMR protection factors to facilitate protein structure prediction, structures were also scored based on agreement with PF from HDX-NMR in addition to scoring with Rosetta (HDX Score). A detailed summary of the score is provided in the Methods section, but in short, for each residue with a measured experimental PF, observed regression lines from Figure 1 (Table S3) were used to predict a value for all exposure and flexibility parameters respectively (NC, RelSASA, HB, and OS). Next, each predicted parameter was compared to the calculated parameter using the structure of the model. Based on the difference between observed and predicted parameters, each residue was scored in the range of -1 (good agreement, fully

rewarded) and 0 (bad agreement, not rewarded). Using this function on all one hundred thousand generated models, ~20% of residues were fully rewarded (score of -1), ~35% were intermediately rewarded (score between -1 and 0), and ~45% of residues were not rewarded (score of 0). This distribution was similar to the distributions of the proteins individually, showing no strong bias for any one protein in the benchmark set. The HDX component of the scores for full protein structures was then derived by aggregating the agreement scores (referring to Equation 4) for each residue (with PF) and each parameter (NC, RelSASA, OS, and HB). While the scoring results using all four computational parameters are described in detail in the following paragraphs, we note that scoring with individual scoring terms was not as successful. This was not surprising, since single features were not expected to govern the complicated process of HDX. The scoring with individual terms was similar or slightly worse than scoring with Rosetta only (average RMSD of Rosetta only: 5.4 Å, average RMSD range using individual terms: 5.1-7.7 Å).

Ultimately, structures were scored using a combined score including both the Rosetta score and HDX component scores (based on agreement with HDX-NMR data) to get a consensus between the Rosetta energy function and experimental data (HDX Score, referring to Equation 5). However, the HDX-dependent component of the HDX Score performed remarkably well by itself as well (referring to the sum of HDX terms in Equation 5). For the majority of proteins in the dataset (7/10), the top scoring model using only the HDX-dependent component of the HDX score was less than 5.5 Å, the same number as when scoring with Rosetta only (although a different set of 7). When using the HDX Score, the RMSDs of all (10/10) predicted structures were improved or were within 0.75 Å of the result obtained by scoring without experimental data. We also note that the results based on the relative weights between the four terms (NC, SASA, OS, and HB) in Equation 5 were relatively stable with respect to small changes in weights. Figure 2A shows the RMSD of the predicted structures with and without including HDX data. Overall, the number of predicted structures with less than 5.5 Å RMSD improved to 9/10 when HDX protection factors were included (7/10 without PF). The top scoring models were rewarded more by the individual HDX components of the score compared to the complete set of decoys: 27% fully rewarded, 40% intermediately rewarded, and 33% not rewarded.

Furthermore, two important features were observed: (1) for the seven proteins where Rosetta alone predicted an accurate structure (< 5.5 Å RMSD), inclusion of HDX data did not significantly change prediction results (selected model RMSD changed by an average of 0.07 Å) and (2) for the three proteins where Rosetta alone predicted an inaccurate structure (1AYI, 1FS3, and 3BLG with RMSDs of 12.72 Å, 10.63 Å, and 14.75 Å, respectively), the inclusion of HDX data markedly improved the prediction results. The RMSD of the predicted model for each of these proteins improved by more than 3 Å with an average improvement of 5.13 Å when HDX data were included. This improvement compared favorably to our previous work (Marzolf *et al.*, 2021) which used qualitative HDX-NMR data to predict structures (3.63 Å). We acknowledge that the dataset is too small (only 3 proteins met this criterion) to draw any universal conclusions. However, we hypothesize that more structural information is encoded in the quantitative HDX PF used in this work than the qualitative categories of exchange rates used in previous work. The additional information provided in the data may have contributed to the improved results. For two of these improved predictions (1AYI and 1FS3), accurate structures were predicted when using HDX data, i.e., the RMSD of the final predicted structure was less than 5.5 Å. Figures 2B and 2D show the scores as a function of RMSD without (red) and with (blue) HDX data used for scoring 1AYI and 1FS3, respectively. The predicted structures (i.e., models with the lowest score, corresponding to Figure 2A) are indicated with a star and those structures are shown aligned to the native crystal

structures in Figures 2C and 2E. Similar score vs. RMSD plots for the remaining 8 proteins in the dataset are shown in Figure S1. From the *ab initio* sampling, the best RMSD structures were typically in the range of 2-5 Å (depending on the system). To demonstrate that the HDX Score also ranked native-like structures well, we calculated the scores for 100 relaxed native structures (shown in the plots as gray circles). These data show that the HDX Score did indeed identify these ~1 Å RMSD structures as native-like.

Confidence metric separated most native-like predictions from others

We sought to define a metric to assess the confidence of a prediction in the absence of a crystal structure. The confidence metric was defined as the average RMSD of the top 10 scoring models to the top scoring model. We hypothesized that this value would be lower when accurate predictions were made because a larger number of structures similar to the predicted structure would also score well. Using this metric, we were able to categorize our predictions as high confidence (< 6 Å), or low confidence (> 6 Å), as shown in Figure 3. This metric identified the top six predictions (all with RMSD less than 3 Å) as high confidence, while identifying the only prediction with RMSD > 5.5 Å as low confidence. For high confidence models, the average RMSD of the predicted structure was 2.06 Å, compared to 6.21 Å for low confidence models.

Conclusion

Hydrogen-deuterium exchange rates of backbone amide hydrogens measured by NMR provide structural information such as flexibility and solvent accessibility of backbone amide protons. In previous work, we developed a scoring function to utilize qualitative HDX-NMR data for protein structure prediction. In this study, we extended this methodology to HDX-NMR data containing quantitative protection factors. We demonstrated that these protection factors correlated with structural features of proteins exhibiting the expected trends when modeled onto native crystal structures.

We then used these observed correlations of each parameter to develop a scoring function to quantify agreement between modeled structures and PF. When predicting tertiary structures of proteins using Rosetta, accurate structures (< 5.5 Å RMSD of the predicted model) were predicted in 7/10 cases. However, when quantitative protection factors from HDX were included, 9/10 of these benchmark cases were predicted accurately. For two cases, the RMSD of the predicted model improved by more than 6.5 Å when experimental data were included and none of the accurate predictions were adversely affected by the HDX data (only changed by an average of 0.07 Å RMSD). Finally, we developed a confidence metric to identify the cases where our predicted structure was most likely to be native-like in the absence of knowledge of the crystal structure. This confidence metric correctly flagged the top six predictions as high confidence.

This work is an important extension of our previously developed HDX-NMR algorithm for protein structure prediction. These two algorithms now enable researchers to extract data from HDX-NMR as either qualitative categories or quantitative protection factors to elucidate structures of proteins. Based on the relative ease of data collection for HDX-NMR, this computational method is beneficial. In this benchmark set, for all predictions where an inaccurate structure was predicted using Rosetta, the scoring based on HDX-NMR data improved the accuracy of the predicted structure. Based on our data, we speculate that the protection factors from HDX-NMR provide information that can discriminate between correct and incorrect topology, but do not contain enough information to refine near-atomic resolution models. A tutorial on how to perform qualitative and quantitative HDX scoring with our algorithm in Rosetta has been included in the

SI (Method S1). In future work, we will extend our HDX methodology to study protein complexes. Additionally, we aim to utilize HDX data measured from mass spectrometry (which has become more popular in recent years) and incorporate multiple types of experimental data to make more accurate predictions.

Acknowledgements

We thank the members of the Lindert lab for many useful discussions. We would like to thank the Ohio Supercomputer Center for valuable computational resources (Ohio Supercomputer Center, 1987). Research reported in this publication, including the virtual summer Lindert lab internship by T.T.N., was supported by NSF (CHE 1750666).

Author Contributions

T.T.N. collected data and analyzed results. D.M.H., J.T.S., and S.L. designed the project and assisted with data interpretation. S.H. and D.M.H. wrote and implemented Rosetta applications. All authors contributed to the manuscript preparation by way of writing (J.T.S., T.T.N., and D.M.H), editing (J.T.S., T.T.N., D.M.H, S.H., and S.L), and generating figures (J.T.S., T.T.N., and D.M.H).

Declaration of Interests

The authors declare no competing interests.

Figure 1: *Experimental data correlations to structure.* Correlations between the calculated structural parameters (neighbor count [NC], relative solvent-accessible surface area [RelSASA], hydrogen bond energy [HB], and order score [OS]) with the log base 10 of HDX protection factor. Linear regression lines are shown in red along with the R^2 values. All trends matched structural hypotheses.

Figure 2: *Comparison of prediction accuracy with and without HDX data.* (A) RMSD of top scoring models with (y-axis) and without (x-axis) HDX data. Circles indicate structures predicted with *ab initio* and triangles indicate homology models. (B) and (D) show score vs. RMSD plots for 1AYI and 1FS3 for Rosetta (red) and HDX score (blue). The predicted structure (lowest scoring model) is represented as a black star and 100 relaxed native structures are represented by gray circles. (C) and (E) Predicted structures (Rosetta score: red; HDX Score: blue) aligned to native (shown in gray).

Figure 3: *Confidence metric for analyzing prediction results.* RMSD of predicted structure using HDX data for each protein shown along with the confidence metric (avg RMSD of top 10 scoring models to top scoring model). Circles indicate structures predicted with *ab initio* and triangles indicate homology models. Vertical line shows separation between confidence regions: high confidence (less than 6 Å) and low confidence (greater than 6 Å). The confidence metric separated predictions well without knowledge of the native structure.

STAR METHODS

RESOURSE AVAILABILITY

Lead Contact

Further information and requests for resources should be directed to and will be fulfilled by the lead contact, Steffen Lindert (lindert.1@osu.edu).

Materials Availability

This study did not generate new unique reagents.

Data and Software Availability

The HDXEnergy application has been contributed to Rosetta software, and a tutorial for running the application is available in the Supplemental Information (Method S1). All model generation and residual property calculations were performed within the Rosetta software framework, available free for educational purposes at <https://www.rosettacommons.org/software>. All protein structure and sequence files are available in the Protein Data Bank (PDB) at <https://www.rcsb.org/>. Residual protection factors can be found in the Excel file provided in the Supplemental Information (Data S1). All software and algorithms used in this paper are available and listed in the key resources table. Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

All data were generated from the datasets provided in the KRT.

METHOD DETAILS

Benchmark Dataset

In this work, we used HDX-NMR protection factors (PFs), which are dependent on solvent accessibility and structural flexibility of specific residues (Bai *et al.*, 1993). For a residue, protection factor is defined as a ratio between the intrinsic HDX rate constant (k_{int}) and the observed HDX rate constant (k_{ex}) as shown in Equation 1 (Bai *et al.*, 1993; Perrett *et al.*, 1995):

$$PF = \frac{k_{int}}{k_{ex}} \quad (\text{Eq. 1})$$

The benchmark dataset for this study contained ten proteins with reported residue-resolved HDX-NMR protection factors which were all consistent with the definition in Equation 1. While a detailed summary of the benchmark set is shown in Table S1, in short, the proteins used in this study were barnase (PDB: 1A2P), colicin E immunity protein 7 (PDB: 1AYI), barstar (PDB: 1BTA), che Y (PDB: 1CYE), bovine pancreatic ribonuclease A (PDB: 1FS3), staphylococcal nuclease (PDB: 1SNO), chymotrypsin inhibitor 2 (PDB: 1TM1), ubiquitin (PDB: 1UBQ), ubiquitin carboxyl-terminal hydrolase isozyme L1 (PDB: 2ETL), and beta-lactoglobulin (PDB: 3BLG). All ten proteins were monomers and experimentally determined structures were available in the Protein Data Bank (PDB). The total number of residues of these proteins ranged from 64 to 223 (Table S1). For 1A2P, 1FS3, and 2ETL, the HDX experiments were conducted at various pHs. The protection factors for the most acidic pH were selected for each protein [pH = 6.5 for 1A2P

(Perrett *et al.*, 1995), pH = 6.5 for 1FS3 (Wang *et al.*, 1995), and pH = 7.6 for 2ETL (Lou *et al.*, 2016)]. In our benchmark dataset, five proteins were reported with the raw protection factors [1A2P (Perrett *et al.*, 1995), 1BTA (Bhuyan and Udgaonkar, 1998), 1FS3 (Wang *et al.*, 1995), 1UBQ (Pan and Briggs, 1992), and 3BLG (Forge *et al.*, 2000)], three proteins were reported with the logarithmic base 10 of protection factors [1CYE (Lacroix *et al.*, 1997), 1SNO (Devaurs *et al.*, 2017; Skinner *et al.*, 2012), and 2ETL (Lou *et al.*, 2016)], and two proteins were reported with the natural logarithm of protection factors [1AYI (Devaurs *et al.*, 2017) and 1TM1 (Devaurs *et al.*, 2017)]. Thus, to standardize protection factors and calculations, all the values were converted into logarithm with base 10 of raw protection factors (log[PF]).

Model Generation

To assess the competence of the developed HDX scoring function to distinguishing near-native models from other decoy models, the decoy set must have a wide distribution of RMSDs (root-mean square deviation to crystal structures), ranging from near-native models (<4 Å RMSD) to models with incorrect topology (>8 Å RMSD). For the majority of proteins in the benchmark dataset, to achieve this purpose, we generated 10,000 decoy models using the standard Rosetta AbinitioRelax protocol (Leaver-Fay *et al.*, 2011; Simons *et al.*, 1997). The Robetta Web server was used to generate the 3-mer and 9-mer fragment files (Kim *et al.*, 2004). The REF2015 scoring function (Alford *et al.*, 2017) was used to score each model for its conformational stability, and the RMSD to the relaxed native structure was calculated for each of the generated models. For 1BTA, 1CYE, 1TM1, 1UBQ, and 3BLG, we used the fragment files generated by excluding homologs. However, for 1A2P, 1AYI, and 1FS3, due to poor model sampling, (none of the models generated were considered near-native) fragment files were generated by including homologs.

Likewise, because of poor observed sampling of the Rosetta AbinitioRelax protocol (both with and without homologous fragments), for 1SNO and 2ETL, the RosettaCM protocol was applied to generate 10,000 decoy homology models of the target proteins (Song *et al.*, 2013). The Standard Protein BLAST server (using the Position-Specific Iterated BLAST) was used to search for different protein sequences with various query coverage and percent identity to the target protein (Johnson *et al.*, 2008). For 1SNO, 3SK6 and 2W8U were selected as templates, while for 2ETL, the templates were 1UCH and 3IHR. Percent identity, coverage, and similarity for each template are shown in Table S2. The templates were aligned with the targets (1SNO and 2ETL) using the Clustal Omega server (Sievers and Higgins, 2018). From the alignment, the sequence of the target was threaded onto each template. Finally, 10,000 decoy models were generated by hybridizing the threaded target sequence on templates with different weights (Table S2).

Calculations of HDX-dependent metrics

The HDX rate for each residue was expected to correlate with flexibility and solvent exposure (Vadas and Burke, 2015). Neighbor count (NC) and relative solvent accessible surface area (RelSASA) were used to quantify solvent accessibility, and order score (OS) and hydrogen bond energy (HB) were used to quantify flexibility. Calculations of each of these four metrics were thoroughly described in our previous work (Marzolf *et al.*, 2021). In short, NC was based on the number of oxygen atoms in the proximity of the amide protons (based on distance and angular cutoffs). RelSASA quantified the solvent accessibility of the side-chain atoms by calculating the ratio between the measured SASA and the theoretical maximum SASA. OS was a window averaged Rosetta residue score, representing the disorder of a residue (Kim *et al.*, 2018; Seffernick

et al., 2019b). Finally, HB was the hydrogen bonding energy of interactions between the amide proton and other residues. All four metrics were calculated using Rosetta applications.

HDX agreement scores

Similar to the previous work, we developed the HDX agreement score based on the observed correlation between the flexibility and solvent exposure metrics and the experimental HDX-NMR protection factors (Marzolf *et al.*, 2021). However, here, the quantitative protection factors were incorporated directly into the scoring function. From the crystal structures of the ten proteins, for each residue with experimental PF (total of 431), four different parameters (NC, RelSASA, OS, and HB) were calculated. For each metric, the correlation between the parameters and PF of ten proteins was examined with linear regression lines as shown in Equation 2.

$$\text{Predicted parameter} = \log(\text{PF}) * \alpha + \beta \quad (\text{Eq. 2})$$

The trends between each parameter and $\log(\text{PF})$ matched our hypotheses (sign of slope, α), thus they were used to score predicted structures based on HDX agreement. To score sampled structures from Rosetta *ab initio* or RosettaCM, we first used Rosetta to calculate the four parameters (NC, RelSASA, OS, and HB) for each of the residues with HDX PF. Next, the values of the parameters were predicted based on the PF (using Equation 2). The slopes and y-intercepts of the linear regressions for the four metrics are reported in Table S3. The HDX agreement score was developed to score residues by quantifying the absolute differences (Δ_i) between the Rosetta calculated values and the predicted values of a residue i using the corresponding PF. In short, residues with a poor agreement were scored as 0, residues with strong agreement were scored as -1, and a linear function was used for those between. The specific form of the score function is shown in Equation 3.

$$\text{Score per residue}_i = \begin{cases} -1, \Delta_i \leq c_1 \\ \frac{\Delta_i}{2} - \left(\frac{c_1 + c_2}{4}\right) - 0.5, c_1 < \Delta_i \leq c_2 \\ 0, c_2 < \Delta_i \end{cases} \quad (\text{Eq. 3})$$

For each metric, there were two cutoffs used to determine whether the residue followed the HDX agreement. Cutoff 1 (c_1) was the rewarding cutoff. Specifically, the residue was fully rewarded with -1, if Δ_i was less than c_1 , which showed that the observed parameters were consistent with the predicted parameters. Values of c_1 for each different metric were determined based on the range of numerical values of each metric ($c_1 = 1$ for NC, $c_1 = 0.1$ for RelSASA, $c_1 = 0.2$ for OS, and $c_1 = 0.2$ for HB). Cutoff 2 (c_2) was the non-rewarding cutoff. If Δ_i was greater than c_2 , the calculated parameters were inconsistent with the predicted values. Hence, the residue would not be rewarded (0). From the standard deviations of native models of 10 proteins for NC, RelSASA, OS, and HB (2.31 for NC, 0.278 for RelSASA, 0.623 for OS, and 0.836 for HB), values of c_2 were determined ($c_2 = 2.0$ for NC, $c_2 = 0.3$ for RelSASA, $c_2 = 0.6$ for OS, and $c_2 = 0.8$ for HB). If Δ_i fell between c_1 and c_2 , the residue would be linearly rewarded based on Δ_i as shown in Equation 3.

Using this scoring algorithm, the score per residue for each parameter would range between -1 for a fully rewarded residue to 0 for a non-rewarded residue. For a model, the HDX component

score (S) for each individual parameter was defined as the sum of the residue scores for all residues with experimental PF, as shown in Equation 4.

$$S = \sum_i^{\# \text{ residues with PF}} \text{Score per residue}_i \quad (\text{Eq. 4})$$

By including the HDX agreement score, the HDX score was defined as the linear combination of the REF2015 Rosetta score (RS) and model score (S) of each metric, as showed in Equation 5. The HDX terms in Equation 5 are defined as the HDX component of the score.

$$\text{HDX Score} = RS + 2 * S(\text{NC}) + 3 * S(\text{RelSASA}) + 4 * S(\text{OS}) + 3 * S(\text{HB}) \quad (\text{Eq. 5})$$

Confidence metric

A confidence metric was developed to assess the extent of confidence in near-native model generation without knowledge of a crystal structure. The confidence metric was defined as the average RMSD of the top 10 scoring models to the top scoring model. If the confidence metric was less than 6 Å, the prediction was considered to be the high confidence. For these proteins, the predicted model had a high probability to be a near-native model. On the other hand, if the confidence metric was greater than 6 Å, the prediction was considered low confidence, such that a near-native model was not expected to be predicted.

QUANTIFICATION AND STATISTICAL ANALYSIS

The correlation between the flexibility and solvent exposure metrics and the experimental HDX-NMR protection factors of 431 residues, for which we had protection factors, were examined using linear regression (Figure 1). The R^2 values were reported in Figure 1, while the slopes and y-intercepts were reported in Table S3. The standard deviations of native models of 10 proteins for NC, RelSASA, OS, and HB of 431 residues were calculated and reported in the Method Details section. All calculations and statistical analysis were performed using Python v3.8.

SUPPLEMENTAL INFORMATION

SI_PF_HDX_NMR_Structure_resubmission_final_final_v3.pdf

Document S1. Method S1, Table S1-S3, and Figure S1. Related to STAR Methods.

PF_HDX_Data.xlsx

Data S1. Residue-resolved protection factor values of the benchmark dataset. Related to STAR Methods.

1A2P_tutorial.zip

Data S2. Inputs for the scoring tutorial in the Supplemental Information. Related to Method S1 and STAR Methods.

References

Alexander, N., Bortolus, M., Al-Mestarihi, A., McHaourab, H., and Meiler, J. (2008). De novo high-resolution protein structure determination from sparse spin-labeling EPR data. *Structure* *16*, 181-195. 10.1016/j.str.2007.11.015.

Alford, R.F., Leaver-Fay, A., Jeliazkov, J.R., O'Meara, M.J., DiMaio, F.P., Park, H., Shapovalov, M.V., Renfrew, P.D., Mulligan, V.K., Kappel, K., et al. (2017). The Rosetta All-Atom Energy Function for Macromolecular Modeling and Design. *J Chem Theory Comput* *13*, 3031-3048. 10.1021/acs.jctc.7b00125.

Aprahamian, M.L., Chea, E.E., Jones, L.M., and Lindert, S. (2018). Rosetta Protein Structure Prediction from Hydroxyl Radical Protein Footprinting Mass Spectrometry Data. *Anal Chem* *90*, 7721-7729. 10.1021/acs.analchem.8b01624.

Aprahamian, M.L., and Lindert, S. (2019). Utility of Covalent Labeling Mass Spectrometry Data in Protein Structure Prediction with Rosetta. *J Chem Theory Comput* *15*, 3410-3424. 10.1021/acs.jctc.9b00101.

Ardenkjaer-Larsen, J.H., Boebinger, G.S., Comment, A., Duckett, S., Edison, A.S., Engelke, F., Griesinger, C., Griffin, R.G., Hilty, C., Maeda, H., et al. (2015). Facing and Overcoming Sensitivity Challenges in Biomolecular NMR Spectroscopy. *Angew Chem Int Ed Engl* *54*, 9162-9185. 10.1002/anie.201410653.

Bai, Y., Milne, J.S., Mayne, L., and Englander, S.W. (1993). Primary structure effects on peptide group hydrogen exchange. *Proteins* *17*, 75-86. 10.1002/prot.340170110.

Best, R.B., and Vendruscolo, M. (2006). Structural interpretation of hydrogen exchange protection factors in proteins: characterization of the native state fluctuations of CI2. *Structure* *14*, 97-106. 10.1016/j.str.2005.09.012.

Bhuyan, A.K., and Udgaonkar, J.B. (1998). Two structural subdomains of barstar detected by rapid mixing NMR measurement of amide hydrogen exchange. *Proteins* *30*, 295-308.

Biehn, S.E., and Lindert, S. (2021). Accurate protein structure prediction with hydroxyl radical protein footprinting data. *Nat Commun* *12*, 341. 10.1038/s41467-020-20549-7.

Borysik, A.J. (2017). Simulated Isotope Exchange Patterns Enable Protein Structure Determination. *Angew Chem Int Ed Engl* *56*, 9396-9399. 10.1002/anie.201704604.

Bowers, P.M., Strauss, C.E.M., and Baker, D. (2000). De novo protein structure determination using sparse NMR data. *Journal of Biomolecular NMR* *18*, 311-318. 10.1023/A:1026744431105.

Bradshaw, R.T., Marinelli, F., Faraldo-Gómez, J.D., and Forrest, L.R. (2020). Interpretation of HDX Data by Maximum-Entropy Reweighting of Simulated Structural Ensembles. *Biophys J* *118*, 1649-1664. 10.1016/j.bpj.2020.02.005.

Choe, S.E., Matsudaira, P.T., Osterhout, J., Wagner, G., and Shakhnovich, E.I. (1998). Folding kinetics of villin 14T, a protein domain with a central beta-sheet and two hydrophobic cores. *Biochemistry* *37*, 14508-14518. 10.1021/bi980889k.

Craig, P.O., Lätzer, J., Weinkam, P., Hoffman, R.M., Ferreiro, D.U., Komives, E.A., and Wolynes, P.G. (2011). Prediction of native-state hydrogen exchange from perfectly funneled energy landscapes. *J Am Chem Soc* *133*, 17463-17472. 10.1021/ja207506z.

Devaurs, D., Antunes, D.A., Papanastasiou, M., Moll, M., Ricklin, D., Lambris, J.D., and Kavraki, L.E. (2017). Coarse-Grained Conformational Sampling of Protein Structure Improves the Fit to Experimental Hydrogen-Exchange Data. *Front Mol Biosci* *4*, 13. 10.3389/fmoleb.2017.00013.

Devaurs, D., Papanastasiou, M., Antunes, D.A., Abella, J.R., Moll, M., Ricklin, D., Lambris, J.D., and Kavraki, L.E. (2018). Native State of Complement Protein C3d Analysed via Hydrogen

Exchange and Conformational Sampling. *Int J Comput Biol Drug Des* 11, 90-113. 10.1504/ijcbdd.2018.090834.

Di Paolo, A., Balbeur, D., De Pauw, E., Redfield, C., and Matagne, A. (2010). Rapid collapse into a molten globule is followed by simple two-state kinetics in the folding of lysozyme from bacteriophage λ . *Biochemistry* 49, 8646-8657. 10.1021/bi101126f.

DiMaio, F., Song, Y., Li, X., Brunner, M.J., Xu, C., Conticello, V., Egelman, E., Marlovits, T., Cheng, Y., and Baker, D. (2015). Atomic-accuracy models from 4.5- \AA cryo-electron microscopy data with density-guided iterative local refinement. *Nat Methods* 12, 361-365. 10.1038/nmeth.3286.

Forge, V., Hoshino, M., Kuwata, K., Arai, M., Kuwajima, K., Batt, C.A., and Goto, Y. (2000). Is folding of beta-lactoglobulin non-hierarchical? Intermediate with native-like beta-sheet and non-native alpha-helix. *J Mol Biol* 296, 1039-1051. 10.1006/jmbi.1999.3515.

Harvey, S.R., Seffernick, J.T., Quintyn, R.S., Song, Y., Ju, Y., Yan, J., Sahasrabuddhe, A.N., Norris, A., Zhou, M., Behrman, E.J., et al. (2019). Relative interfacial cleavage energetics of protein complexes revealed by surface collisions. *Proc Natl Acad Sci U S A* 116, 8143-8148. 10.1073/pnas.1817632116.

Hilser, V.J., and Freire, E. (1996). Structure-based calculation of the equilibrium folding pathway of proteins. Correlation with hydrogen exchange protection factors. *J Mol Biol* 262, 756-772. 10.1006/jmbi.1996.0550.

Hooke, S.D., Radford, S.E., and Dobson, C.M. (1994). The refolding of human lysozyme: a comparison with the structurally homologous hen lysozyme. *Biochemistry* 33, 5867-5876. 10.1021/bi00185a026.

Johnson, M., Zaretskaya, I., Raytselis, Y., Merezhuk, Y., McGinnis, S., and Madden, T.L. (2008). NCBI BLAST: a better web interface. *Nucleic Acids Res* 36, W5-9. 10.1093/nar/gkn201.

Kim, D.E., Blum, B., Bradley, P., and Baker, D. (2009). Sampling bottlenecks in de novo protein structure prediction. *J Mol Biol* 393, 249-260. 10.1016/j.jmb.2009.07.063.

Kim, D.E., Chivian, D., and Baker, D. (2004). Protein structure prediction and analysis using the Robetta server. *Nucleic Acids Res* 32, W526-531. 10.1093/nar/gkh468.

Kim, S.S., Seffernick, J.T., and Lindert, S. (2018). Accurately Predicting Disordered Regions of Proteins Using Rosetta ResidueDisorder Application. *J Phys Chem B* 122, 3920-3930. 10.1021/acs.jpcb.8b01763.

Lacroix, E., Bruix, M., López-Hernández, E., Serrano, L., and Rico, M. (1997). Amide hydrogen exchange and internal dynamics in the chemotactic protein CheY from *Escherichia coli*. *J Mol Biol* 271, 472-487. 10.1006/jmbi.1997.1178.

Leaver-Fay, A., Tyka, M., Lewis, S.M., Lange, O.F., Thompson, J., Jacak, R., Kaufman, K., Renfrew, P.D., Smith, C.A., Sheffler, W., et al. (2011). ROSETTA3: an object-oriented software suite for the simulation and design of macromolecules. *Methods Enzymol* 487, 545-574. 10.1016/b978-0-12-381270-4.00019-6.

Lindert, S., Hofmann, T., Wötzl, N., Karakaş, M., Stewart, P.L., and Meiler, J. (2012). Ab initio protein modeling into CryoEM density maps using EM-Fold. *Biopolymers* 97, 669-677. 10.1002/bip.22027.

Liu, T., Pantazatos, D., Li, S., Hamuro, Y., Hilser, V.J., and Woods, V.L., Jr. (2012). Quantitative assessment of protein structural models by comparison of H/D exchange MS data with exchange behavior accurately predicted by DXCOREX. *J Am Soc Mass Spectrom* 23, 43-56. 10.1007/s13361-011-0267-9.

Lou, S.C., Wetzel, S., Zhang, H., Crone, E.W., Lee, Y.T., Jackson, S.E., and Hsu, S.D. (2016). The Knotted Protein UCH-L1 Exhibits Partially Unfolded Forms under Native Conditions that Share Common Structural Features with Its Kinetic Folding Intermediates. *J Mol Biol* 428, 2507-2520. 10.1016/j.jmb.2016.04.002.

Lyumkis, D. (2019). Challenges and opportunities in cryo-EM single-particle analysis. *J Biol Chem* 294, 5181-5197. 10.1074/jbc.REV118.005602.

Makarov, A.A., Iacob, R.E., Pirrone, G.F., Rodriguez-Granillo, A., Joyce, L., Mangion, I., Moore, J.C., Sherer, E.C., and Engen, J.R. (2020). Combination of HDX-MS and in silico modeling to study enzymatic reactivity and stereo-selectivity at different solvent conditions. *J Pharm Biomed Anal* 182, 113141. 10.1016/j.jpba.2020.113141.

Martens, C., Shekhar, M., Lau, A.M., Tajkhorshid, E., and Politis, A. (2019). Integrating hydrogen-deuterium exchange mass spectrometry with molecular dynamics simulations to probe lipid-modulated conformational changes in membrane proteins. *Nat Protoc* 14, 3183-3204. 10.1038/s41596-019-0219-6.

Marzolf, D.R., Seffernick, J.T., and Lindert, S. (2021). Protein Structure Prediction from NMR Hydrogen-Deuterium Exchange Data. *J Chem Theory Comput*. 10.1021/acs.jctc.1c00077.

McAllister, R.G., and Konermann, L. (2015). Challenges in the interpretation of protein h/d exchange data: a molecular dynamics simulation perspective. *Biochemistry* 54, 2683-2692. 10.1021/acs.biochem.5b00215.

Mohammadiarani, H., Shaw, V.S., Neubig, R.R., and Vashisth, H. (2018). Interpreting Hydrogen-Deuterium Exchange Events in Proteins Using Atomistic Simulations: Case Studies on Regulators of G-Protein Signaling Proteins. *J Phys Chem B* 122, 9314-9323. 10.1021/acs.jpcb.8b07494.

Ohio Supercomputer Center. (1987). Ohio Supercomputer Center.

Palmer, A.G., 3rd (1997). Probing molecular motion by NMR. *Curr Opin Struct Biol* 7, 732-737. 10.1016/s0959-440x(97)80085-1.

Pan, Y., and Briggs, M.S. (1992). Hydrogen exchange in native and alcohol forms of ubiquitin. *Biochemistry* 31, 11405-11412. 10.1021/bi00161a019.

Perrett, S., Clarke, J., Hounslow, A.M., and Fersht, A.R. (1995). Relationship between equilibrium amide proton exchange behavior and the folding pathway of barnase. *Biochemistry* 34, 9288-9298. 10.1021/bi00029a003.

Petruk, A.A., Defelipe, L.A., Rodríguez Limardo, R.G., Bucci, H., Marti, M.A., and Turjanski, A.G. (2013). Molecular Dynamics Simulations Provide Atomistic Insight into Hydrogen Exchange Mass Spectrometry Experiments. *J Chem Theory Comput* 9, 658-669. 10.1021/ct300519v.

Pilla, K.B., Gaalswyk, K., and MacCallum, J.L. (2017). Molecular modeling of biomolecules by paramagnetic NMR and computational hybrid methods. *Biochim Biophys Acta Proteins Proteom* 1865, 1654-1663. 10.1016/j.bbapap.2017.06.016.

Ramsey, K.M., Narang, D., and Komives, E.A. (2018). Prediction of the presence of a seventh ankyrin repeat in IkBe from homology modeling combined with hydrogen-deuterium exchange mass spectrometry (HDX-MS). *Protein Sci* 27, 1624-1635. 10.1002/pro.3459.

Roberts, V.A., Pique, M.E., Hsu, S., and Li, S. (2017). Combining H/D Exchange Mass Spectrometry and Computational Docking To Derive the Structure of Protein-Protein Complexes. *Biochemistry* 56, 6329-6342. 10.1021/acs.biochem.7b00643.

Rogov, V.V., Bernhard, F., Löhr, F., and Dötsch, V. (2004). Solution structure of the *Escherichia coli* YojN histidine-phosphotransferase domain and its interaction with cognate phosphoryl receiver domains. *J Mol Biol* 343, 1035-1048. 10.1016/j.jmb.2004.08.096.

Rosa, J.J., and Richards, F.M. (1979). An experimental procedure for increasing the structural resolution of chemical hydrogen-exchange measurements on proteins: application to ribonuclease S peptide. *J Mol Biol* 133, 399-416. 10.1016/0022-2836(79)90400-5.

Schulman, B.A., Redfield, C., Peng, Z.Y., Dobson, C.M., and Kim, P.S. (1995). Different subdomains are most protected from hydrogen exchange in the molten globule and native states of human alpha-lactalbumin. *J Mol Biol* 253, 651-657. 10.1006/jmbi.1995.0579.

Seffernick, J.T., Harvey, S.R., Wysocki, V.H., and Lindert, S. (2019a). Predicting Protein Complex Structure from Surface-Induced Dissociation Mass Spectrometry Data. *ACS Cent Sci* 5, 1330-1341. 10.1021/acscentsci.8b00912.

Seffernick, J.T., and Lindert, S. (2020). Hybrid methods for combined experimental and computational determination of protein structure. *J Chem Phys* 153, 240901. 10.1063/5.0026025.

Seffernick, J.T., Ren, H., Kim, S.S., and Lindert, S. (2019b). Measuring Intrinsic Disorder and Tracking Conformational Transitions Using Rosetta ResidueDisorder. *J Phys Chem B* 123, 7103-7112. 10.1021/acs.jpcb.9b04333.

Sievers, F., and Higgins, D.G. (2018). Clustal Omega for making accurate alignments of many protein sequences. *Protein Sci* 27, 135-145. 10.1002/pro.3290.

Simons, K.T., Kooperberg, C., Huang, E., and Baker, D. (1997). Assembly of protein tertiary structures from fragments with similar local sequences using simulated annealing and Bayesian scoring functions. *J Mol Biol* 268, 209-225. 10.1006/jmbi.1997.0959.

Skinner, J.J., Lim, W.K., Bédard, S., Black, B.E., and Englander, S.W. (2012). Protein hydrogen exchange: testing current models. *Protein Sci* 21, 987-995. 10.1002/pro.2082.

Song, Y., DiMaio, F., Wang, R.Y., Kim, D., Miles, C., Brunette, T., Thompson, J., and Baker, D. (2013). High-resolution comparative modeling with RosettaCM. *Structure* 21, 1735-1742. 10.1016/j.str.2013.08.005.

Srivastava, A., Nagai, T., Miyashita, O., and Tama, F. (2018). Role of Computational Methods in Going beyond X-ray Crystallography to Explore Protein Structure and Dynamics. *Int J Mol Sci* 19. 10.3390/ijms19113401.

Vadas, O., and Burke, J.E. (2015). Probing the dynamic regulation of peripheral membrane proteins using hydrogen deuterium exchange-MS (HDX-MS). *Biochem Soc Trans* 43, 773-786. 10.1042/bst20150065.

van Zundert, G.C.P., Melquiond, A.S.J., and Bonvin, A. (2015). Integrative Modeling of Biomolecular Complexes: HADDOCKing with Cryo-Electron Microscopy Data. *Structure* 23, 949-960. 10.1016/j.str.2015.03.014.

Vendruscolo, M., Paci, E., Dobson, C.M., and Karplus, M. (2003). Rare fluctuations of native proteins sampled by equilibrium hydrogen exchange. *J Am Chem Soc* 125, 15686-15687. 10.1021/ja036523z.

Wan, H., Ge, Y., Razavi, A., and Voelz, V.A. (2020). Reconciling Simulated Ensembles of Apomyoglobin with Experimental Hydrogen/Deuterium Exchange Data Using Bayesian Inference and Multiensemble Markov State Models. *J Chem Theory Comput* 16, 1333-1348. 10.1021/acs.jctc.9b01240.

Wang, A., Robertson, A.D., and Bolen, D.W. (1995). Effects of a naturally occurring compatible osmolyte on the internal dynamics of ribonuclease A. *Biochemistry* 34, 15096-15104. 10.1021/bi00046a016.

Zhang, M.M., Beno, B.R., Huang, R.Y., Adhikari, J., Deyanova, E.G., Li, J., Chen, G., and Gross, M.L. (2019). An Integrated Approach for Determining a Protein-Protein Binding Interface in Solution and an Evaluation of Hydrogen-Deuterium Exchange Kinetics for Adjudicating Candidate Docking Models. *Anal Chem* 91, 15709-15717. 10.1021/acs.analchem.9b03879.

Zhang, Y., Majumder, E.L., Yue, H., Blankenship, R.E., and Gross, M.L. (2014). Structural analysis of diheme cytochrome c by hydrogen-deuterium exchange mass spectrometry and homology modeling. *Biochemistry* 53, 5619-5630. 10.1021/bi500420y.

