Article

Interpretation of HDX Data by Maximum-Entropy Reweighting of Simulated Structural Ensembles

Richard T. Bradshaw,[1] Fabrizio Marinelli,[2] José D. Faraldo-Gómez,[2,*] and Lucy R. Forrest[1,*]
[1]Computational Structural Biology Section, National Institute of Neurological Disorders and Stroke, National Institutes of Health, Bethesda, Maryland and [2]Theoretical Molecular Biophysics Unit, National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, Maryland

ABSTRACT   Hydrogen-deuterium exchange combined with mass spectrometry (HDX-MS) is a widely applied biophysical technique that probes the structure and dynamics of biomolecules without the need for site-directed modifications or bio-orthogonal labels. The mechanistic interpretation of HDX data, however, is often qualitative and subjective, owing to a lack of quantitative methods to rigorously translate observed deuteration levels into atomistic structural information. To help address this problem, we have developed a methodology to generate structural ensembles that faithfully reproduce HDX-MS measurements. In this approach, an ensemble of protein conformations is first generated, typically using molecular dynamics simulations. A maximum-entropy bias is then applied post hoc to the resulting ensemble such that averaged peptide-deuteration levels, as predicted by an empirical model, agree with target values within a given level of uncertainty. We evaluate this approach, referred to as HDX ensemble reweighting (HDXer), for artificial target data reflecting the two major conformational states of a binding protein. We demonstrate that the information provided by HDX-MS experiments and by the model of exchange are sufficient to recover correctly weighted structural ensembles from simulations, even when the relevant conformations are rarely observed. Degrading the information content of the target data—e.g., by reducing sequence coverage, by averaging exchange levels over longer peptide segments, or by incorporating different sources of uncertainty—reduces the structural accuracy of the reweighted ensemble but still allows for useful insights into the distinctive structural features reflected by the target data. Finally, we describe a quantitative metric to rank candidate structural ensembles according to their correspondence with target data and illustrate the use of HDXer to describe changes in the conformational ensemble of the membrane protein LeuT. In summary, HDXer is designed to facilitate objective structural interpretations of HDX-MS data and to inform experimental approaches and further developments of theoretical exchange models.

SIGNIFICANCE   Hydrogen-deuterium exchange combined with mass spectrometry (HDX-MS) experiments are a powerful approach for probing the conformational dynamics and mechanisms of proteins. However, the mechanistic implications of HDX-MS observations are frequently difficult to interpret because of the limited spatial resolution of the technique, as well as the lack of quantitative tools to translate measured data into structural information. To overcome these problems, we have developed a computational approach to construct structural ensembles that are maximally diverse while reproducing target experimental HDX-MS data within a given level of uncertainty. Using both artificial and experimental test data, we demonstrate that the approach can correctly discern distinct structural ensembles reflected in the target data and thereby facilitate statistically robust evaluations of competing mechanistic interpretations of HDX-MS experiments.

INTRODUCTION

Upon exposure to a deuterated solvent such as D_2O, labile hydrogen atoms present in protein side chains and backbones will readily exchange for deuterium. The rate of this process is influenced by the chemical features of the exchanging groups and by conditions such as pD or temperature and is also critically dependent on protein conformation [1,2]. Consequently, measurements of hydrogen-deuterium exchange (HDX) rates are increasingly used as a direct probe of protein dynamics. Moreover, by combining HDX with mass spectrometry (HDX-MS), this approach has also become feasible for large complexes and membrane proteins, even at low concentrations [3].

Typically, HDX-MS is carried out using so-called bottom-up and continuous labeling strategies, in which proteins are deuterated for varying amounts of time, quenched, proteolytically fragmented, and purified in the solution phase before analysis of the individual peptide fragments by mass spectrometry. For each identified fragment, typically 5–20 residues in length, deuterium incorporation is then reported as the change in peptide mass over time. Because side-chain and terminal-amine deuterons exchange back relatively rapidly with protons during analysis, HDX-MS data reports exclusively on backbone-amide exchange. This ability to directly probe protein dynamics has led to diverse applications (4), including studies of allostery (5–7), epitope mapping for protein-protein or protein-lipid interactions (8–11), effects of ligand binding (12–15), mechanisms of membrane proteins (16–22), and dynamics of large macromolecular complexes (23–26). This progress notwithstanding, the interpretation of HDX-MS data in structural and mechanistic terms has been, generally speaking, largely qualitative and lacking objective metrics.

No matter the protein system, interpretation of HDX-MS data requires an understanding of the processes reflected in the exchange kinetics. For any given backbone amide under a given set of conditions (pH, temperature, etc.), the most rapid rate of exchange occurs when the residue is in a completely unstructured, solvent-accessible conformational state of the protein. Under these circumstances, the value of the intrinsic exchange rate constant, k_i^{int} for residue i, is determined predominantly by steric and electronic effects from neighboring side chains (27,28). In a folded conformational state, by contrast, amides will be partially or fully occluded from solvent and/or engaged in hydrogen bonding. This structural protection can diminish the intrinsic rate constant by several orders of magnitude. In this case, exchange is better described as a two-step process: first, a structural transition must occur from a so-called noncompetent exchange state to a competent one; this step is followed by the intrinsic chemical exchange reaction with rate constant k_i^{int} (2,29). If the structural transition entails only local alterations rather than complete unfolding, an equilibrium between the exchange-competent and noncompetent states may be reached rapidly, even more so than the hydrogen-deuterium substitution; this situation is referred to as occurring with "EX2" kinetics. The overall exchange rate under these conditions is thus given by the product of the equilibrium constant for the structural transition and the intrinsic rate, $k_i = K_i^{\text{eq}} k_i^{\text{int}}$. This relationship is commonly expressed as $k_i = k_i^{\text{int}}/P_i$, where P_i denotes the "protection factor" for each amide, which in turn relates to the free-energy difference between the noncompetent and competent states, $\Delta G = RT \ln P_i$. Following these concepts, HDX data is commonly interpreted in terms of the degree of protein structural flexibility and solvent accessibility for a given amide.

In practice, HDX-MS experiments measure deuteration averaged over lengthy peptide fragments rather than at the single-residue level. Even in the light of statistical analysis approaches that allow high-resolution protection factors to be derived directly from experimental data for peptides (30,31), interpretation of the observed data in structural terms is not straightforward. Oftentimes, HDX levels are color coded and mapped on known protein structures, which allows an intuitive visualization of the results and highlights dynamic or solvent-exposed protein regions. However, this kind of qualitative visual analysis can easily lead to a subjective interpretation of the experimental data (32). Moreover, HDX data reflect the properties of an ensemble of protein conformations and, in some cases, therefore might not be explained by a single structural state. To address these issues, previous studies have relied on molecular simulation methods. A typical approach is to first generate a conformational ensemble for the protein of interest with molecular dynamics (MD) or Monte Carlo simulations. The simulated data must then be translated into predicted peptide-deuteration levels that can be correlated with the experimental data (33–43). Typically, this is achieved using empirical models that predict protection factors P_i from an ensemble of protein structures. Some P_i prediction models directly estimate the free energy of exchange from the equilibrium constant obtained by defining both exchange-competent and noncompetent states from a simulated trajectory (34,36,40), which requires that both sets of states have been adequately sampled. In a second category of P_i prediction models, such as the one used here (33), the likelihood of exchange is predicted using an empirical scoring function parameterized on the basis of the characteristics of the folded state only; thus, P_i may be in principle predicted from sufficiently long (e.g., microsecond-timescale) MD simulations of one or more folded states of the protein of interest. Regardless of the specific approach, an important caveat of these kind of strategies is that for many cases of interest, a simulation may not accurately represent the conformational ensemble reflected by the experimental data, for example, because of force-field inaccuracies or incomplete sampling of alternate folded states. Thus, even if a perfectly accurate empirical model for P_i were at hand, the predicted protection factors might deviate substantially from measured data.

Here, we develop and test a methodology to resolve this problem. This approach, which we refer to as HDX ensemble reweighting (HDXer), enables us to construct conformational ensembles that faithfully reflect a given set of target HDX-MS data for a given empirical model of P_i. HDXer is based on concepts outlined in previous studies and applied to other types of biophysical data (44–54) but not yet to HDX-MS. In brief, this is a post hoc method whereby a maximum-entropy criterion is used to reassign statistical weights to each of the configurations in a structural ensemble generated computationally

(e.g., via simulations or modeling) so that calculated ensemble-averaged peptide deuterated fractions reproduce measured values within a given level of uncertainty. That is, this approach aims to adjust populations in a heterogenous conformational ensemble such that they conform ideally to the experimental data while taking into account all potential sources of uncertainty. Thus, the method can be used to rank the correspondence between a given HDX-MS data set and several candidate conformational states based on the degree of bias required to reproduce the experimental results.

To evaluate the validity of HDXer, we focus primarily on artificial HDX-MS data generated for a binding protein that undergoes a substantial conformational change (55). Specifically, we reweight a simulated structural ensemble so that calculated deuteration levels match a set of artificial HDX-MS data reflecting predefined populations of two major conformational states. The performance of the method is then assessed based on whether the conformations favored by the reweighting indeed correspond to the structural states used to generate the target data. The use of artificial data allows us first, to rigorously assess reweighting performance in a context for which the correct ensemble is known and second, to evaluate the effect of different sources of uncertainty on the ensemble reweighting. Encouragingly, the results show that the proposed approach always recovers the key features of the correct structural ensemble, even when sparse HDX data are targeted or in the presence of moderate error sources. Finally, to demonstrate the transferability of the approach to experimentally determined data, we apply HDXer to HDX-MS measurements obtained recently for the membrane transporter LeuT (16).

METHODS

Calculation of HDX residue protection factors and peptide deuterated fractions

To predict deuterium uptake based on structural snapshots (obtained from MD simulations or another molecular modeling method), we first calculate the protection factor for each residue i, P_i, using the method of Best and Vendruscolo (33). Specifically, the free-energy difference between exchange-competent and noncompetent states of a residue is approximated by a linear function of the numbers of H-bonds and heavy-atom contacts of the corresponding backbone amide, denoted as $N_{H,i}$ and $N_{C,i}$, respectively:

$$\ln P_i = \langle \beta_C N_{C,i} + \beta_H N_{H,i} \rangle. \quad (1)$$

The notation $\langle \ldots \rangle$ signifies an ensemble average over all available snapshots. $N_{C,i}$ is calculated as the number of nonhydrogen atoms within 6.5 Å of the amide N atom of residue i, excluding atoms in residues $i - 2$ to $i + 2$; $N_{H,i}$ is the number of O or N atoms within 2.4 Å of the amide hydrogen atom. In the original formulation by Best and Vendruscolo, the scaling factors β_C and β_H are set to 0.35 and 2.0, respectively. These values reflect an empirical optimization with respect to experimental HDX data for several water-soluble proteins (33); however, their optimal value depends on the

protein or experimental conditions (43), and therefore, we will treat them as optimizable parameters.

In addition to P_i, we consider the intrinsic exchange rate constant for each residue type, k_i^{int}, from Bai and co-workers, updated for acidic residues and glycine (27,28). Deuterated fractions for peptide segments of the protein, $D_{j,t}^{sim}$, can then be calculated for any given time point of exchange, t, using the exchange rate constants of each individual residue and according to first-order kinetics. That is,

$$D_{j,t}^{sim} = \frac{\sum_{i=m_j+1}^{i=n_j} 1 - \exp\left(-\frac{k_i^{int}}{P_i}t\right)}{n_j - m_j}, \quad (2)$$

where m_j and n_j are the first and last residue numbers of the j-th protein fragment respectively. Note that proline residues do not have an exchangeable amide proton and were therefore excluded from the deuterated fraction calculation. The first residue (m_j) in each peptide segment was also omitted from the average because hydrogens in the amine N-terminus are labile after proteolytic fragmentation and are assumed to have fully exchanged back to protons during the HDX-MS purification and analysis step. It should also be noted that in direct comparisons of experimental and predicted data, the measured deuterated fractions should be corrected for the fraction of D_2O/H_2O in the reaction buffer and for back exchange during the analysis process. Both corrections can be achieved by normalizing to deuterated fractions observed in identical control experiments performed under maximal deuteration conditions (32).

Maximum-entropy ensemble reweighting with HDX data

In this section, we describe the basic formulation for calculating corrections to the statistical weight of the individual structural snapshots in an ensemble, each denoted by X_k, such that the predicted deuteration fractions reproduce a set of HDX experimental data. Our approach is related to that of Marinelli and Fiorin (46), in which the only bias applied is that strictly required to conform to the experiments, following the so-called maximum-entropy principle (44,45,53,54,56). In general terms, the minimal bias needed to correct the mean value of one or more observables of interest is provided by a linear function of those observables, added as a perturbation term to the molecular force field or energy function, $U(X)$ (44). In this case, the target observables are P_i (or functions thereof) (Eqs. 1 and 2), and therefore the corrected force field is defined as

$$U_{corr}(X) = U(X) - k_B T \sum_i \lambda_i [\beta_C N_{C,i}(X) + \beta_H N_{H,i}(X)]. \quad (3)$$

In the initial sample, the statistical weight of each configuration X_k is proportional to $exp\{-U(X_k)/k_BT\}$. Similarly, in the corrected ensemble, these weights are proportional to $exp\{-U_{corr}(X_k)/k_BT\}$. The set of weight adjustments we seek, $\Omega(X_k)$, are therefore simply a Boltzmann factor of the linear term of Eq. 3:

$$\Omega(X_k) = \frac{exp\left\{\sum_i \lambda_i [\beta_C N_{C,i}(X_k) + \beta_H N_{H,i}(X_k)]\right\}}{\sum_{k'} exp\left\{\sum_i \lambda_i [\beta_C N_{C,i}(X_{k'}) + \beta_H N_{H,i}(X_{k'})]\right\}}, \quad (4)$$

in which the denominator is a normalization term calculated by summing over all simulation configurations.

The scaling factors λ_i in Eqs. 3 and 4 are the key adjustable parameters in this methodology. These parameters will be uniquely determined so that deuteration fractions deduced from the reweighted ensemble fit the

experimental data within a defined error distribution, ρ_{err}, and with the smallest possible bias. To quantify this bias, we report the amount of apparent work, W_{app}, required to reweight the ensemble. In formal terms, the optimal value of λ_i is at the global minimum of the following (Kullback-Leibler) likelihood function (46,57):

$$L = \frac{W_{app}}{k_B T} - \ln \rho_{err}. \quad (5)$$

The apparent work, W_{app}, depends on the correction to the potential applied in Eq. 3 as follows:

$$W_{app} = k_B T \ln \left\langle \exp \left\{ -\sum_i \lambda_i [\beta_C N_{C,i}(\mathbf{X}) + \beta_H N_{H,i}(\mathbf{X}) - \ln P_i] \right\} \right\rangle, \quad (6)$$

where $\langle \ldots \rangle$ denotes a mean value over the corrected ensemble or, in other words, a weighted average according to the weights of Eq. 4. Note that W_{app} is related to the Kullback-Leibler divergence between the initial and corrected ensembles (46,57,58), $d_{KL} = W_{app}/k_B T = \sum_k \Omega(\mathbf{X}_k) \ln \Omega(\mathbf{X}_k) + \ln N$, where N is the number of simulation frames.

The function ρ_{err} is an error distribution that, for simplicity, we assume to be Gaussian and uncorrelated across all target data points:

$$\rho_{err}(D^{sim}) \propto \exp \left\{ -\sum_t \sum_j \gamma \frac{\left(D_{j,t}^{sim} - D_{j,t}^{exp}\right)^2}{2\eta^2} \right\}, \quad (7)$$

where the parameter γ controls the final level of agreement with the target experimental data (see below), η is an estimate of the uncertainty (here set to 1, such that γ instead imposes equal uncertainty for all target data points), and $D_{j,t}^{exp}$ and $D_{j,t}^{sim}$ are the experimental and predicted deuterated fractions, respectively. $D_{j,t}^{sim}$ is calculated according to Eq. 2 using the protection factors for each amide, but after adjusting for reweighting, $\ln P_i = \langle \beta_C N_{C,i} + \beta_H N_{H,i} \rangle = \sum_k [\beta_C N_{C,i}(\mathbf{X}_k) + \beta_H N_{H,i}(\mathbf{X}_k)] \Omega(\mathbf{X}_k)$.

In practice, we use a gradient-based minimization of the likelihood function L in Eq. 5, in which the parameters λ_i are calculated iteratively according to the derivative of L:

$$\lambda_i^{n+1} = \lambda_i^n (1 - \varepsilon) + \varepsilon \frac{\partial \ln \rho_{err}(D^{sim})}{\partial \ln P_i}, \quad (8)$$

where ε is an update rate selected to ensure convergence. The corrected (reweighted) protection factors entered into Eq. 8 depend on λ_i (Eq. 4) and thus are also updated at each iteration. The model parameters β_C and β_H are optimized at each step using a Monte Carlo procedure to reduce the discrepancy between simulated and experimental data, measured by the mean squared deviation, $MSD = \chi^2/N_D$, where N_D is the number of data points and $\chi^2 = \sum_t \sum_j (D_{j,t}^{sim} - D_{j,t}^{exp})^2/\eta^2$. Optimization was performed with 100 Monte Carlo trials at each step, and new values of β_C and β_H were accepted only if MSD decreased. The maximum step sizes for the trials ($\Delta\beta_{C,max} = 0.15$, $\Delta\beta_{H,max} = 1.6$) were chosen to correspond to 10% of the maximum range of β_C and β_H we proposed to explore.

We note that although HDX-MS measures the total deuterated fraction for protein fragments, our approach uses the minimal bias condition to spread such experimental information across individual residues (Eqs. 3, 4, 5, 6, 7, and 8). Nevertheless, if multiple experimental data points incorporating a single amide are available or if deuteration is otherwise correlated between amides, the contribution of each amide to the ensemble correction is constrained by a simultaneous fit to all the experimental data. Therefore, in practical applications of reweighting, the inclusion of HDX-MS measurements for overlapping peptide segments will ultimately lead to enhanced resolution.

Reweighting parameters and metrics of robustness

In the reweighting procedure, the presence of unknown errors in predicted and experimental data is implicitly considered by setting a parameter γ (Eq. 7) that regulates the variance in the error distribution and that can be tuned to achieve a compromise between the applied bias and the level of agreement with experiments (46,57). To identify a reasonable value of γ, a decision plot of W_{app} vs. MSD can be constructed for different values of γ. Typically, the presence of undetermined, systematic errors such as forward-model uncertainty or sampling inefficiency induces a rapid increase of the work value below a certain value of MSD, resulting in an L-shaped decision plot (see Fig. 6 A for an example). In this case, a reasonable value of γ can be found at the kink of the L-curve, provided that the associated value of W_{app} is within say, two or three $k_B T$.

TeaA simulation data and generation of the artificial target HDX data

The simulation data used for ensemble reweighting were taken from the unbiased replica ($\sim$45 ns, with frames at 1 ps intervals) of bias-exchange metadynamics simulations performed previously (55) for the periplasmic binding protein TeaA from *Halomonas elongata* (UniProt: E1VBK1), including both "closed" and "open" states of TeaA (Fig. 1). The artificial HDX-MS data used as a target for the reweighting were created from this trajectory so as
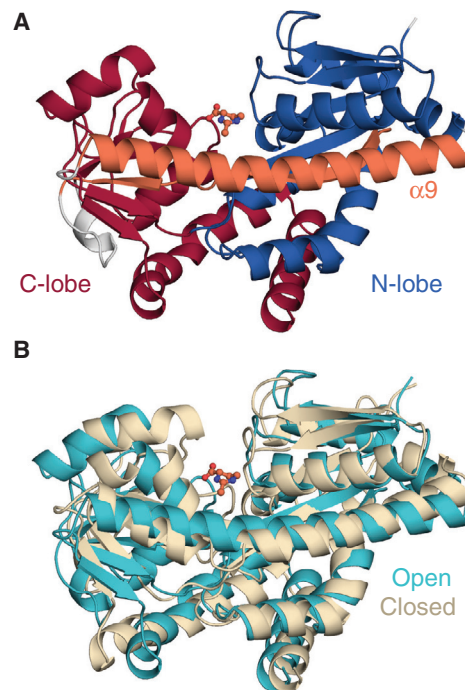
FIGURE 1 Structures of ectoine-bound TeaA in open and closed conformations. (A) A representative open structure is shown as cartoon helices, highlighting the N-lobe (*blue*), the C-lobe (*red*), and the β4/α9 segments (*orange*) that span both lobes. The ectoine ligand bound to the central binding cleft is shown in ball-and-stick representation. (B) An overlay of representative open (*cyan*) and closed (*wheat*) conformations is given. The C_α RMSD between the two conformations is 3.2 Å.

to represent a rapidly interconverting conformational ensemble comprising 60% closed and 40% open states. Specifically, two reference configurations were chosen to represent typical "closed" and "open" states based on their structural similarity to available structures (Fig. 1 B), and two subensembles of closed and open configurations (corresponding to 37.2 and 1.6% of the initial frames, respectively) were then obtained by extracting highly related frames, defined as those in which the root mean-square deviation (RMSD) of the C_α atoms was <1.0 Å from those in the closed or open reference structures. The remaining 61.2% of frames remained unassigned.

Artificial target HDX-MS data sets were then derived from the closed and open subensembles according to Eq. 2. Residue protection factors for the mixed target ensemble were calculated as $\ln P_i^{\text{mix}} = 0.6 \ln P_i^{\text{closed}} + 0.4 \ln P_i^{\text{open}}$, in which P_i^{closed} and P_i^{open} represent protection factors calculated across the subensemble of the closed and open conformations, respectively. Protection factors were calculated using Eq. 1, with $\beta_\text{H} = 2.0$ and $\beta_\text{C} = 0.35$. To assess the ability of our reweighting method to extract subensembles of different relative size, additional target ensemble mixtures were created with 95:5, 80:20, 40:60, 20:80, and 5:95 ratios of closed/open conformations. Artificial HDX-MS data were also constructed for the open and closed ensembles separately using the same values of P_i^{closed} and P_i^{open}. All artificial data were calculated at time points of 0.167, 1.0, 10.0, 60.0, and 120.0 min. These time points reflect typical HDX-MS experiments and capture both short- and long-timescale EX2 exchange. A scheme of the generation and use of the artificial mixed-ensemble HDX data for ensemble refinement is provided in Fig. S1.

To assess the impact of segment averaging and sequence coverage upon reweighting, multiple TeaA HDX-MS data sets were generated. The largest set of artificial HDX measurements, obtained at residue-level resolution and with full sequence coverage, comprised 294 residues at five time points, giving a total of 1470 individual predicted observables to be refined against. To evaluate the effect of segment averaging, five other target data sets were generated, in which TeaA was divided into fragments of size 5, 10, 15, 20, or 50 residues, including prolines. Note that because deuteration of the N-terminal amine is excluded from HDX-MS data, neighboring protein segments were defined with a one-residue overlap (e.g., 1–10, 10–19, etc.). The final peptide segment in each data set was extended up to the C-terminal residue 310. Analysis of the effect of sequence coverage was based on the 10-residue segment target HDX data set, which comprises 34 peptides, from which coverage was reduced in five cumulative steps from 100 to 20% of the sequence (six to seven peptides at each step; Fig. S2). Assuming that buried peptides are less likely to be proteolytically hydrolyzed, we preferentially excluded peptides with lower solvent accessibility.

To assess the effect of experimental noise on the reweighting, we added an error term to each target HDX data point; the magnitude of this error was randomly obtained from Gaussian distributions of standard deviation $\sigma = 0.01$ or 0.1 (in units of deuterated fraction). The target HDX-MS data in this case were those generated to reflect 100% protein coverage and 10-residue segments.

To evaluate the impact of sampling errors, we removed from the simulation data all frames with C_α RMSD < 1.5 Å with respect to the reference closed-state structure.

To assess the effect of the accuracy of the model used to predict protection factors (Eq. 1) on the ensemble reweighting, we modified the value of the β parameters in different protein regions (as opposed to uniform values of β). Two different target HDX data sets were generated, both containing variations in the model parameters for residues 225–261, which comprise the $\alpha 9$ helix. Specifically, the β parameters were selected to produce "low-error" ($\beta_\text{H} = 7.0$, $\beta_\text{C} = 0.2$) and "high-error" ($\beta_\text{H} = 2.0$, $\beta_\text{C} = 0.2$) target data sets using values that are either consistent or inconsistent, respectively, with their observed inverse relationship (see Fig. S8).

Trajectory clustering

To interrogate the results of the reweighting procedure without the requirement for a reference protein configuration, the structures ("samples") in the final ensembles were clustered based on their pairwise similarity (RMSD of the C_α traces) using the density-based algorithm DBSCAN as implemented in scikit-learn v0.21.2 (59). The minimum size of a cluster, n, was set to 10% of the total ensemble size, but the contribution of each frame to n corresponded to the weight assigned after ensemble reweighting (Eq. 4) and normalized to the number of structures in the entire ensemble. The maximum radius, ε, which defines the neighborhood of an individual sample, was chosen by evaluating cluster quality for the ensemble obtained after reweighting to the residue-level data set with $\gamma = 10^3$. Scanning a range of values of ε from 10.546 to 105.46 Å (equivalent to pairwise RMSD values of 0.05 or 0.50 Å, respectively) on this test set revealed well-defined clusters with high silhouette scores (60) at an ε-value of 42.187 Å (a pairwise RMSD of 0.20 Å).

Application of HDXer to LeuT HDX-MS data

We considered the experimental HDX-MS data for the amino acid transporter LeuT (UniProt: O67854) described in a previous study by Adhikary et al. (16). Briefly, these data had been obtained at 0.167, 1.0, 10.0, and 120.0 min time points for wild-type (WT) and Y268A mutant LeuT reconstituted into 60/40 1-palmitoyl-2-oleoylphosphocholine (POPC)/1-palmitoyl-2-oleoylphosphoglycerol (POPG) lipid nanodiscs in 20 mM Tris-HCl (pH 7.4), 100 mM NaCl, 0.5 mM EDTA buffer. Of the 21 peptides previously identified and compared (16), 17 peptides were used for HDXer analysis. The other four peptides, in either the N-terminus (1–12, 1–14, 1–16) or the C-terminus (505–517), were not used because they are found in regions not resolved in the crystal structures used in our MD simulations.

Reference structural ensembles of WT and Y268A LeuT were generated by extending the atomistic MD simulations performed by Adhikary et al. (16). Briefly, structures of either outward-facing WT LeuT (starting from Protein Data Bank, PDB: 3TT1) or inward-facing Y268A LeuT (starting from PDB: 3TT3) were embedded in a 1,2-dimyristoylphosphocholine (DMPC) bilayer, and three independent, 2-μs-long simulations were performed for each configuration. Structures were saved at 100 ps intervals for a total of 20,000 frames per 2 μs simulation, and hence, a total of 120,000 frames in the reference ensemble were used as input to HDXer.

Data availability

All underlying data used in this study are made freely available (https://doi.org/10.5281/zenodo.3385168), including the initial simulation trajectories, target HDX data sets, and analysis code for extracting contacts and H-bonds, generating artificial target data sets, reweighting ensembles, and clustering. The code and underlying data used to create figures are also available in this repository.

RESULTS

The TeaA test system undergoes a substantial conformational change

In the proposed computational approach, we seek to be able to reweight a heterogenous structural ensemble so that it optimally reflects a given set of HDX-MS data. The success of such a method requires that it be able to detect and upweight the protein configurations that are most consistent with the data but also to detect and downweight those that are not. To meaningfully test this method, therefore, one must begin with a sample that is sufficiently heterogeneous for a system with several states of known structure. To this end, we considered the ectoine-binding protein TeaA

and extracted a broad sample of configurations from enhanced-sampling MD simulations carried out in a previous study (55). The structure of TeaA consists of two distinct lobes interconnected by a single β-strand (β4) and a single α-helix (α9) (Fig. 1 A). Ectoine binding at a central cleft between the lobes fosters a clamshell-like structural change (Video S1), whereby the distance between lobes changes by up to ~10 Å. We refer to the two endpoints of this conformational change as the "open" and "closed" states (Fig. 1 B). These states have nearly identical secondary structure, except that closure requires local unwinding and kinking of the α9 helix at residues K247-L249.

The existing simulations, based on bias-exchange metadynamics, capture the full range of this structural change and explain how the affinity for ectoine is modulated by the conformational state of the protein (55). These data demonstrated that the closed state of TeaA is most favored when ectoine is bound; however, partial opening of this bound form was also observed and found to entail a free-energy penalty of only ~2 kcal mol^{-1} (55). Accordingly, the unbiased replica in these simulations samples open, closed, and intermediate configurations of the protein (Fig. S3). This structural heterogeneity makes these data an ideal choice as a reference set on which to test the performance of our reweighting method.

Artificial HDX-MS data for open and closed TeaA

To test the protocol proposed here, we also need target HDX data sets for each of the conformation states of the protein of interest. To our knowledge, however, no experimental HDX-MS data exist for TeaA. We therefore decided to generate artificial, high-resolution HDX data for the two major states of TeaA (open and closed) to evaluate whether a hypothetical experiment would yield a measurable contrast. To this

end, we extracted separate ensembles of open and closed conformations from the simulation data and compared the predicted deuterated fractions at the single-residue level for each set (calculated using Eqs. 1 and 2). The HDX data were generated at single-residue resolution and across five time points to capture both spatial and temporal differences in deuterium uptake at high resolution (see Methods).

We observed substantial differences between the predicted deuterated fractions of closed and open ensembles (Fig. 2 A), confirming that these artificial data sets are well suited for our purpose. As might be expected, this contrast is most pronounced at the binding site interface and in the α9 helix (Fig. 2 B). Interestingly, though, subtle differences are also observed across almost the entire protein and vary from one time point to another. These complex patterns cannot be easily interpreted visually, e.g., by mapping the data onto the representative structures (Fig. 2 B) because they reflect the dynamical nature of the simulated ensembles. For the same reasons, such comparisons based on single structures also offer limited insights into experimentally determined HDX-MS data, as has been noted elsewhere (32). The striking variability of the idealized artificial data for TeaA further illustrates the need for an ensemble perspective to rigorously interpret HDX measurements at the structural level.

Ensemble reweighting with idealized single-residue HDX target data

To begin to evaluate the HDXer method, we next produced artificial HDX-MS data for a hypothetical measurement in which TeaA spontaneously interconverts between closed and open states, populating these states in a 60:40 ratio. The sample derived from the unbiased metadynamics replica (hereafter referred to as the reference ensemble)
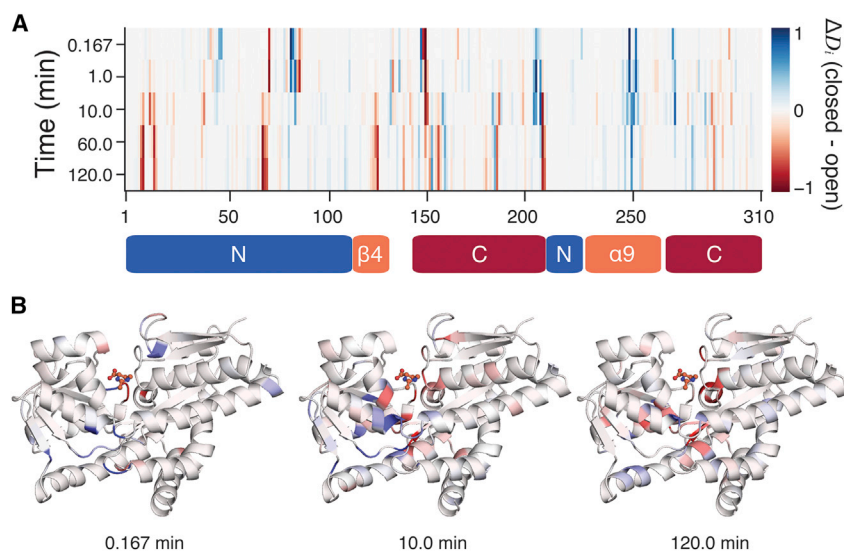

FIGURE 2 Difference in predicted deuterated fractions between closed and open ensembles of TeaA. (A) By-residue $\Delta D_i = D_{i,\text{closed}} - D_{i,\text{open}}$ for each time point is shown, where red indicates that a residue is more deuterated in the open conformation than in the closed, and blue indicates the opposite. Domain definitions are indicated using bars beneath the plot. (B) A representative closed structure of TeaA, colored by residue ΔD_i at the 0.167, 10, and 120 min time points, is shown. The largest ΔD_i-values are observed for residues either lining the central binding cleft or involved in the partial unfolding of helix α9 but are clearly not uniform across time points.

comprises a heterogeneous set of conformations that were either assigned to closed and open states or unassigned (decoys), with a ratio of 37.2:1.6:61.2 (see Methods). Note that some of the decoy structures do share structural similarities with either open or closed states, as demonstrated by the continuity of the RMSD distributions (Fig. 3, *A* and *B*, *cyan*). The challenge for the HDXer method, therefore, is to identify to the appropriate weights for each and all of the configurations in the reference ensemble so that ensemble-averaged HDX levels calculated for the reweighted sample exactly reflect the 60:40 ratio of open/closed conformations in the target data set.

As expected, without reweighting, the predicted HDX levels for the reference ensemble were in poor agreement with the target HDX data ($MSD = 2.2 \times 10^{-3}$, equivalent to a root mean-square error of 0.047 Da in mass for every residue), owing to the mismatch in populations. In reweighting applications, acceptable levels of MSD are highly system and application dependent. With artificial data, because the target ensemble is exactly present in the reference ensemble, we considered $MSD \leq 10^{-6}$, which is equivalent to a root mean-square error of 0.001 Da in mass for every residue, to be an acceptable agreement in this case. Ensemble reweighting with HDXer succeeds in matching the target data (Fig. 3; Fig. S4) to an extremely high level of precision. By increasing the value of the parameter γ in Eq. 7, an increasingly tighter agreement with the target HDX data was achieved (Fig. S4 *A*), requiring a larger apparent work, W_{app}, to be applied (Fig. S4 *B*) and resulting in an increasing deviation from the initial reference ensemble, whereas the optimized values of β_C and β_H remain close to their target values (Fig. S4, *C* and *D*). Enforcing closer agreement with the target HDX (by

increasing γ) resulted in the gradual development of a RMSD distribution profile containing two distinct peaks, corresponding to the closed and open states of TeaA, as the decoy trajectory frames became downweighted (*dark brown*, *orange*, and *yellow*; Fig. 3, *A* and *B*). Notably, the bimodal features of the target distribution could already be detected with only a small applied bias of $W_{app} = 0.9$ kJ mol^{-1} relative to the 2.6 kJ mol^{-1} bias applied at $\gamma = 10^3$. After reaching an $MSD \leq 10^{-7}$ (reweighting with $\gamma \approx 10^2$ or larger), no further substantial changes in the ensemble were observed (Fig. 3, *A* and *B*), and W_{app} reached a plateau (Fig. S4 *B*).

To more quantitatively characterize the outcome of the reweighting, we applied a clustering algorithm to the configurations in the reweighted ensemble obtained using $\gamma = 10^3$ (see Methods). Two clusters were found: the largest cluster clearly represented a closed conformation (Fig. 3 *C*) and comprised 59.3% of the final ensemble by weight, whereas the second cluster comprised 35.5% of the final ensemble and reflected an open conformation (Fig. 3 *D*). The remaining 5.2% of the ensemble consisted of outliers that, owing to structural dissimilarities and/or low weight after reweighting, could not be assigned to either of the clusters.

From the RMSD distributions of the reweighted ensembles, it was clear that the final ensemble still contained a non-negligible fraction of frames >1.0 Å RMSD to either the closed or open state. Moreover, these decoy frames were included in the extracted clusters alongside the "correct" frames (i.e., those assigned to the closed or open reference ensembles). These observations raise concerns about the fidelity that can be achieved with ensemble-averaged observables such as these. We therefore asked how similar
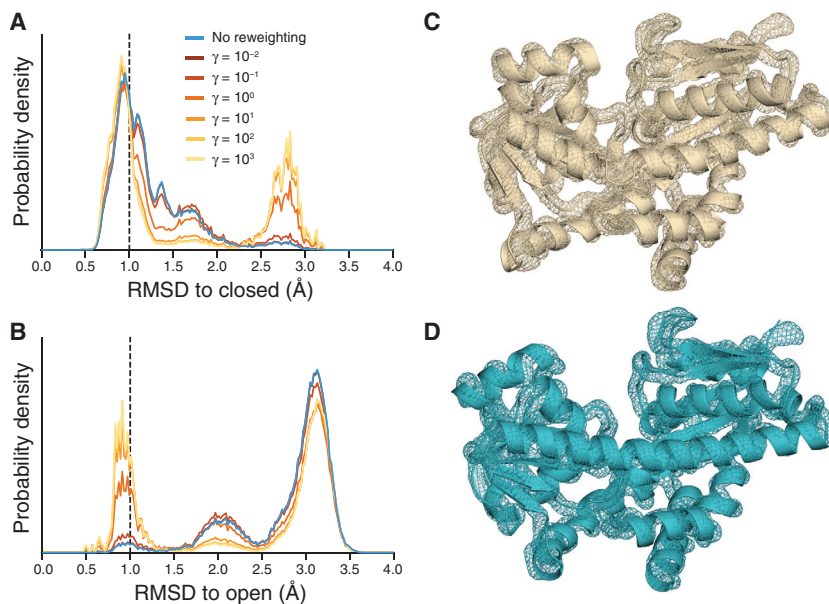
FIGURE 3 Results of HDX ensemble reweighting at single-residue resolution. (*A* and *B*) Probability distributions of the RMSD are shown with respect to the closed (*A*) or open (*B*) reference structure of TeaA for the initial reference ensemble (*cyan*) and for ensembles obtained after reweighting with progressively higher γ-values (*dark brown* to *orange* to *yellow*). The dashed line indicates the 1.0 Å RMSD cutoff used to assign frames to the reference closed (*A*) or open (*B*) ensemble. (*C* and *D*) Ensemble density maps of the closed (*C*) or open (*D*) clusters, extracted by structural clustering after reweighting with $\gamma = 10^3$, are given. The mesh reflects the density of backbone N, CA, and C atoms overlaid onto the representative closed (*C*) or open (*D*) structure of TeaA. Maps were created using the AtomProb (75) feature of Xplor-NIH v2.51 and are shown at 0.25 σ.

these decoy structures are to those used to generate the target protection factor data. According to the root mean-square fluctuation of the backbone atoms, both clusters exhibited minimal structural variance, with a maximum root mean-square fluctuation of 1.2 Å, excluding the N-terminal residue (Fig. S5), and well-defined backbone density when calculated across all structures in each cluster (Fig. 3, C and D). Reassuringly, then, the inclusion of decoy frames reflected conformationally correlated frames, indicating that the reweighting identified key structural features of the target data and, based on those features, created populations of the two conformational states in good agreement with the target ratio of 60:40.

HDX ensemble reweighting with realistic peptide segments and sequence coverage

The results so far demonstrate that HDX reweighting can successfully extract key structural features of artificial target HDX data for an ensemble of conformations defined at the residue level and with 100% sequence coverage. However, this level of information content is not representative of typical HDX-MS experiments, which report deuterated fractions for proteolytic fragments of a protein, whereas complete sequence coverage requires extensive optimization of experimental conditions. To evaluate the extent to which lower-resolution HDX-MS data can be meaningfully interpreted with a quantitative method such as HDXer, we systematically degraded the information content of the artificial target data produced at single-residue resolution while maintaining the 60:40 ratio of closed/open state data. First, the deuterated fraction values were averaged over peptides of increasing length, from 5 to 50 residues, while maintaining full sequence coverage. Second, using fragment lengths typical for HDX-MS, sequence coverage was reduced by removing peptide segments from the target data. To compare ensembles obtained with different target data sets, for which γ-values are not directly comparable, we instead fixed the level of agreement with the target data at $MSD = 10^{-6}$.

Averaging the deuterated fractions over peptide segment lengths from 5 to 50 residues represents a loss of spatial resolution in the HDX-MS signal and increases the degeneracy of the structural information present in the data. When reweighting the reference ensemble, increasing the length of the segments progressively reduced the value of W_{app} required to achieve the same level of agreement with the target data, which is also increasingly less resolved and thus more easily reproduced (Fig. 4 A). That is, the smaller values of W_{app} reflect a greater similarity between the initial and reweighted ensembles. However, this degradation of the target data translates into a reduced ability to discern between conformational states. Specifically, both the RMSD probability distributions and the structure-based clustering after reweighting (Fig. 4 B; Table 1) show that increasing
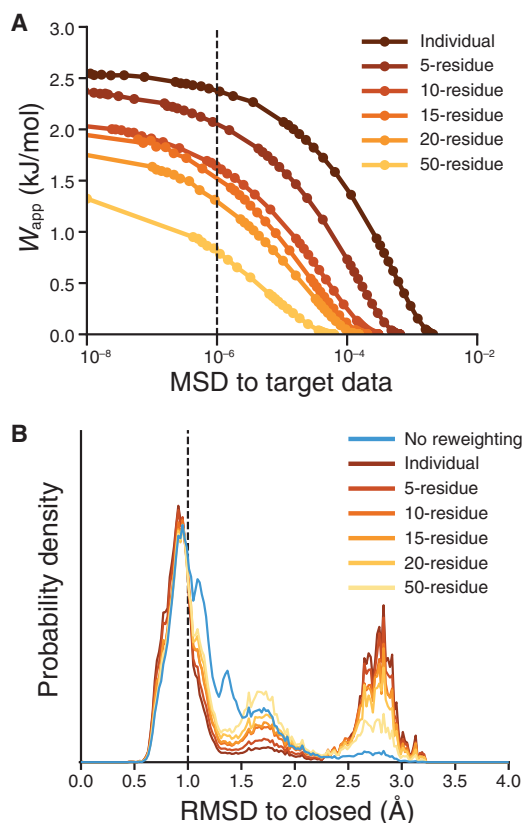

FIGURE 4 Effects of segment averaging on ensemble reweighting. (A) A decision plot is given showing the work applied during reweighting against the MSD of the reweighted ensemble to target HDX data. Circles indicate independent reweighting experiments. (B) RMSD probability distributions, with respect to the closed TeaA structure, are shown before (*cyan*) and after ensemble refinement to $MSD = 10^{-6}$. In both panels, data from reweighting performed with individual residue deuterated fractions are shown in dark brown, and the data obtained by increasing the peptide segment lengths are shown using gradual color variation from light brown to orange to yellow.

the fragment length reduced the ability of HDXer to discriminate between open and semiopen (RMSD ~1.7 Å) protein structures. Indeed, using $\geq$20-residue-long segments, the semiopen state was still highly populated and was identified as a separate, unique cluster (Table 1). The

TABLE 1 Cluster Populations after Ensemble Reweighting with Segment-Averaged Target Data

Segment Length	Closed (%)	Open (%)	Semiopen (%)	Outliers (%)
1	59.1	34.4	–	6.4
5	59.5	31.0	–	9.5
10	58.9	26.6	–	14.6
15	59.1	24.9	–	16.0
20	59.5	21.0	15.6	3.9
50	61.2	10.3	22.8	5.7

Populations are measured as percentage by weight of the total ensemble. Predicted deuterated fractions from the reweighted ensembles fit the target data with $MSD = 10^{-6}$. The data for segment length = 1 represent a reweighting with residue-resolved target data.

decreasing structural fidelity of the results was correlated with larger deviations of the HDXer-optimized β_H and β_C parameters. However, optimized parameters remained close to their target values for segments ≤ 20 residues, in which maximal deviations of $\beta_C \pm 2.2\%$ and $\beta_H \pm 8.5\%$ were observed. For longer fragments, therefore, a quantitative interpretation is not possible, unless information from overlapping (redundant) peptides is available (Fig. S6). Overall, however, for the peptide lengths typical of HDX-MS experiments (5–20 residues), HDX reweighting correctly identified the trends in closed and open state populations present in the target data.

The data obtained so far assume relatively similar (60:40) populations of the open and closed states, but cases in which the ratio of states is more imbalanced can be easily envisaged. We therefore carried out additional tests with varying populations of the two states for target data consisting of 10-residue peptides. Encouragingly, HDXer was able to correctly identify trends in the target ensemble, even with low (5%) populations of either closed or open states (Fig. S7).

Even at low levels of amide resolution, the target data up to this point covered the entire length of the protein. Loss of sequence coverage increases the degeneracy of the structural information present in HDX-MS data. We therefore investigated the effects of reducing coverage using the 10-residue-long segment data set analyzed earlier, for which reweighting at 100% sequence coverage resulted in cluster populations of 58.9 and 26.6% for the closed and open states, respectively (Fig. 4 B; Table 1).

As expected, gradual degradation of the sequence coverage also reduced the value of W_{app} required to match the target data, e.g., with $MSD = 10^{-6}$ (Fig. 5 A), for the same reasons discussed above for increasing peptide lengths. The effect in terms of structural interpretation was also similar: reducing coverage incorrectly increased the contribution of semiopen states relative to 100% coverage (Fig. 5 B) and was again correlated with increasing changes to the optimized model parameters. These effects were particularly marked when the coverage was $\leq 40\%$ (Table 2). Nevertheless, parameter changes remained small, up to $\beta_C \pm 1.4\%$ and $\beta_H \pm 7.8\%$, after reweighting with $\geq 40\%$ sequence coverage.

It is perhaps surprising that even at 20% coverage, HDXer produced a 10-fold enrichment of the population of the open state, i.e., in qualitative agreement with the target data. Inspection of the peptides included in this set (Fig. S2) shows that at least one peptide spanning the $\alpha 9$ helix was included at all coverage levels. Because the conformational change in helix $\alpha 9$ correlates strongly with the open-to-closed transition, peptides in this helix likely include crucial target observables that allow our method to correctly discern between states of TeaA. In actual HDX experiments, this correlation might not exist for any one peptide fragment among those available, in which case 20% coverage would not likely be sufficient to derive a clear interpretation. Over-
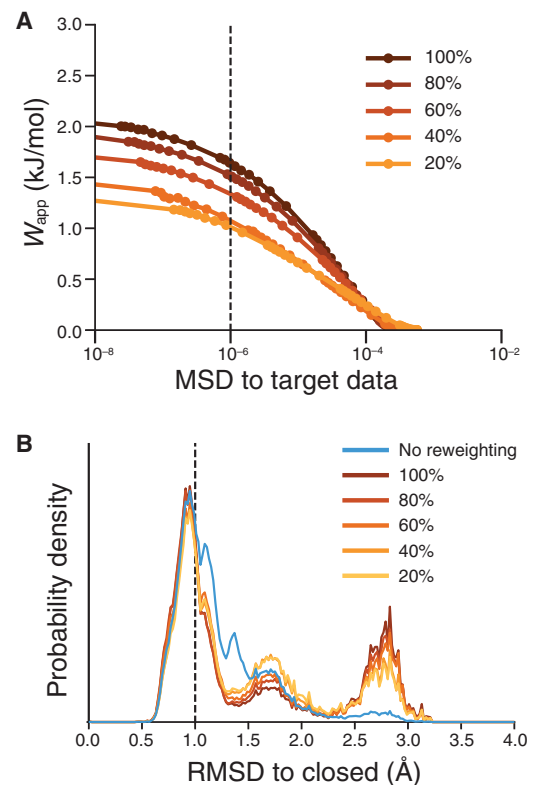

FIGURE 5 Effects of reduced sequence coverage on ensemble reweighting. (A) Decision plot and (B) RMSD probability distributions after reweighting with reduced sequence coverage in the target HDX-MS data, using target data with 10-residue segments, are shown. See legend to Fig. 4 for more details. In both panels, data from reweighting performed with full coverage are shown in dark brown, and the data obtained by decreasing sequence coverage lengths are shown using gradual color variation from light brown to orange to yellow.

all, therefore, our results suggest that although low sequence coverage does not preclude ensemble reweighting, HDX-MS data at high coverage are likely to be substantially advantageous for HDXer applications.

HDXer with noisy target data

The data so far suggest that the ability of our method to identify open and closed states from the initial sample of TeaA, based

TABLE 2 Cluster Populations after Ensemble Refinement with Target Data Covering Smaller Proportions of the Protein

Coverage (%)	Closed (%)	Open (%)	Semiopen (%)	Outliers (%)
100	58.9	26.6	–	14.6
80	58.8	23.9	13.5	3.8
60	56.6	22.2	17.5	3.7
40	55.7	17.4	22.4	4.5
20	55.5	16.2	23.6	4.8

Populations are measured as percentage by weight of the total ensemble. Predicted deuterated fractions from the final ensembles fit the target data with $MSD = 10^{-6}$. Peptide segments were 10 residues long, so the data presented for 100% coverage represent the same final ensemble as segment length = 10 in Table 1.
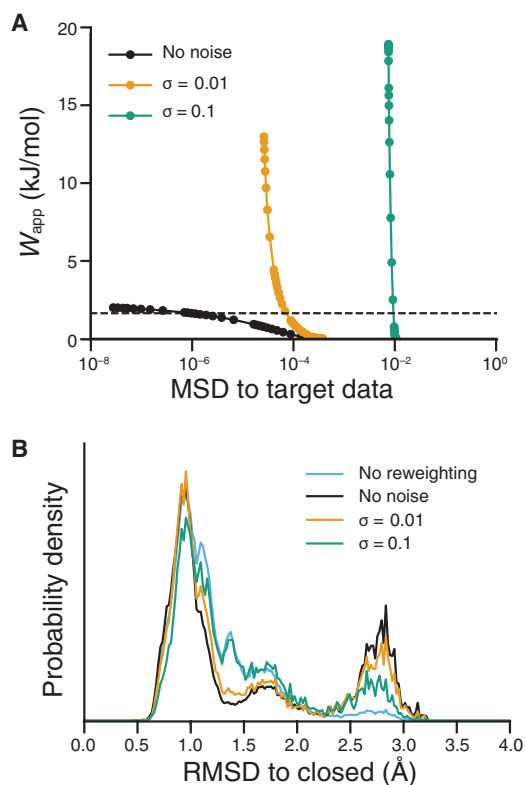
FIGURE 6 Effects of experimental noise on ensemble reweighting. (A) Decision plot and (B) RMSD probability distributions after reweighting with noise added to target data are shown. In both panels, data from reweighting performed without noise added to the target data are shown in black, while orange and teal represent results from reweighting with added noise of increasing magnitude. RMSD distributions are compared after reweighting with $W_{app} = 1.64$ kJ mol^{-1}, indicated by the dashed line in (A). The unbiased RMSD distribution is shown in cyan.

only on similarity with target HDX data, does not critically depend on peptide segment length, nor does it require complete coverage. However, all test cases so far have assumed that the target HDX data are perfectly precise, with zero random uncertainty, which is obviously not reflective of experimentally determined data. To evaluate the ability of HDXer to reweight noisy data, we added random noise of standard deviation σ to each target data point in the set corresponding to 100% protein coverage and 10-residue peptide segments. As expected, the accuracy of the predicted HDX-MS data after ensemble reweighting depended on the level of noise incorporated in the target data (Fig. 6 A). When targeting data containing Gaussian random error with $\sigma = 0.01$ (deuterated fraction units), a larger apparent bias (W_{app}) was required to fit the final ensemble with $MSD \leq 10^{-4}$, compared to the experiments in which the target data were noise-free. Target data generated with $\sigma = 0.1$ could not be fitted with $MSD \leq 10^{-2}$, even after applying high levels of apparent bias, clearly demonstrating overfitting. Nevertheless, the final structural ensembles, evaluated at an equivalent level of W_{app} after reweighting with each data set, were still substantially modified from their initial distributions (Fig. 6 B).

In fact, with a small value of noise in the target data ($\sigma = 0.01$) the final ensemble showed very similar structural features to the ensemble obtained with noise-free data. When targeting data with higher noise levels ($\sigma = 0.1$), the final ensemble deviated more substantially from the ensemble obtained with noise-free data. In particular, HDXer was incapable of substantially downweighting semi-open structural frames (Fig. 6 B) when targeting the data with the most noise. For comparison, uncertainties from technical HDX-MS replicates have been estimated to be well below $\sigma = 0.01$, which would be equivalent to 0.1 Da error per time point per 10-residue peptide (61). Larger errors may arise from differences in experimental protocol and biological replicates (61), but our results suggest such errors may only impact the structural insights provided by HDXer if they are at least an order of magnitude larger than those typically measured between replicates. This finding provides reassurance that HDXer can provide structurally useful interpretations of HDX-MS data even with realistic levels of experimental uncertainty.

HDXer of ensembles with insufficient conformational sampling

In addition to potentially uninformative or noisy data, it is entirely possible that the initial ensemble being reweighted does not contain any of the structural states reflected by the HDX-MS measurements being targeted. In the case of MD sampling, this situation might arise because of sampling-time inadequacies or force-field discrepancies. For the TeaA system, this situation can be exemplified by removing all the closed-state conformations from the reference ensemble before applying HDXer exactly as before (see Methods). The resultant decision plot (Fig. 7 A) clearly shows that the fit could not be improved beyond an agreement of $MSD \approx 10^{-4}$, in contrast to the fits with $MSD < 10^{-7}$ attained when closed-state structures were present in the reference ensemble. Concomitant with the decrease in MSD was a rapid increase in the apparent work required ($W_{app} > 10$ kJ mol^{-1}), indicating poor overlap between the reference and reweighted ensembles, i.e., only a handful of configurations have predicted HDX values in agreement with the target data. This interpretation is supported by the distribution of structures after reweighting, which predominantly consists of semiopen states that must be only partially representative of the closed-state HDX data (Fig. 7 B). Encountering such a characteristic decision plot and structural distribution when using experimental data would motivate the use of enhanced-sampling methods to improve the reference pool of structures (62–65).

HDXer with inaccuracies in the empirical model

In the aforementioned data, the phenomenological model of Best and Vendruscolo used to translate structural frames into
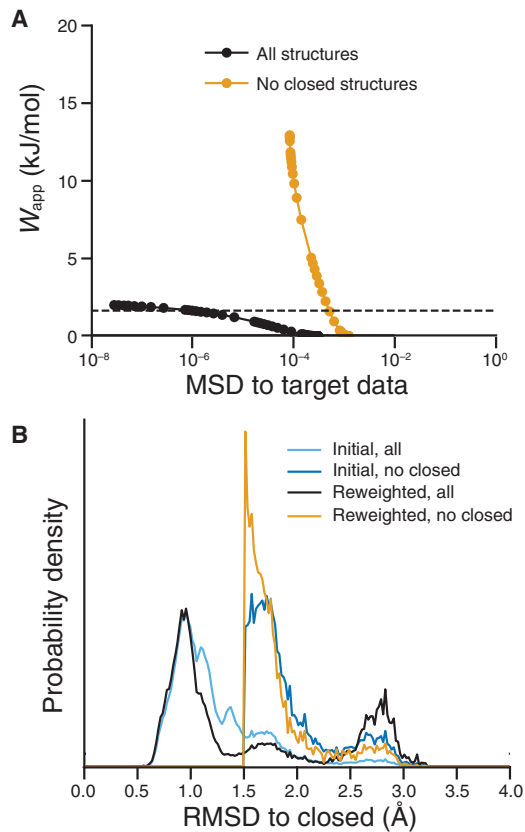
A

B

FIGURE 7 Detection of critically insufficient conformational sampling in the input ensemble to be reweighted. (*A*) Decision plot and (*B*) RMSD probability distributions after reweighting are shown. In both panels, data were obtained from reweighting performed with an initial ensemble lacking closed conformations (*orange*) or with the complete initial ensemble (*black*). For comparison, the distribution of the complete reference ensemble without reweighting is shown for the full ensemble containing all structures (*cyan*) and for the ensemble without structures of the closed conformation (*blue*). RMSD distributions are compared after reweighting with $W_{app} = 1.64$ kJ mol^{-1}, denoted by the dashed line in (*A*).

protection factors is assumed to be perfectly accurate, with constant values of the empirical scaling factors β_H and β_C used throughout the protein. In particular, the target HDX data were obtained using the values $\beta_H = 2.0$ and $\beta_C = 0.35$ that were identified based on the original parameterization (33). However, the optimum values of these scaling factors might vary between proteins or within the protein environment. Revisiting the target HDX data set and systematically varying the values of β_H and β_C indicated that the predicted HDX levels themselves were not sensitive to small changes in β parameters (Fig. S8). However, this analysis also illustrated that β_H and β_C should be inversely related: any perturbations away from an inverse relationship resulted in large discrepancies relative to the initial predicted HDX calculated with $\beta_H = 2.0$ and $\beta_C = 0.35$. These findings are broadly consistent with prior observations (33).

By treating β_H and β_C as additional ("nuisance") parameters during the reweighting of λ_i, as has been done here, any systematic inaccuracy of the model parameters across the

whole protein should be automatically reduced. However, if deuteration in different regions of the protein is best described by different model parameters, the nuisance parameter optimization will result in an imprecise, "averaged" model that may be inappropriate in individual protein regions. To analyze the potential effects of model inaccuracy in a structural context, we perturbed the target data set by generating artificial HDX-MS data using different β_H and β_C parameters for residues in the $\alpha 9$ helix. Specifically, values of β_H and β_C that were either consistent or inconsistent with their observed inverse relationship (Fig. S8) were used to generate "low-error" and "high-error" target data sets, respectively.

The introduction of errors in the β parameters was detrimental to the ability of HDXer to fit a final ensemble with moderate agreement ($MSD < 10^{-4}$) to either target data set (Fig. 8 *A*), demonstrating the importance of these parameters for reweighting, at least quantitatively. Nevertheless, the structural distributions obtained from this optimization differed significantly depending on the magnitude of the perturbation to the parameters. Specifically, compared at equal values of W_{app}, reweighting to the low-error data set resulted in a final
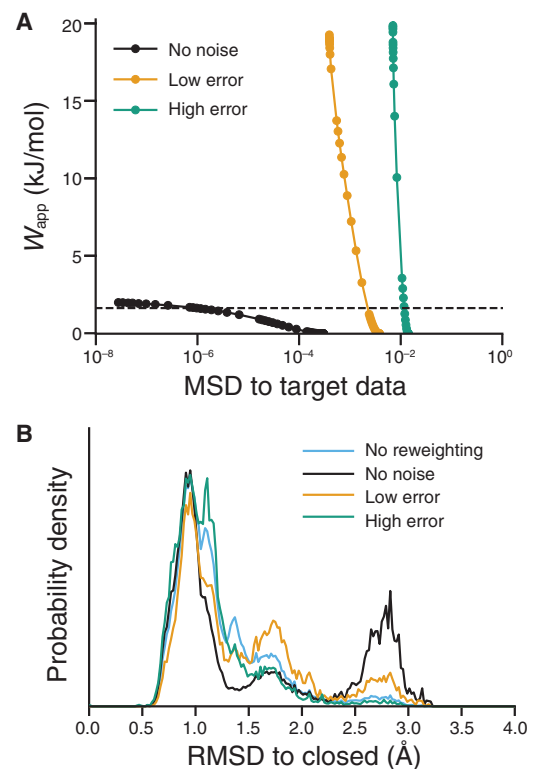
A

B

FIGURE 8 Effect of variations in the HDX model on ensemble reweighting. (*A*) Decision plot and (*B*) RMSD probability distributions after reweighting are shown. In both panels, data were obtained from reweighting to target data sets created using mixtures of β_H and β_C parameters designed to correspond with low-error (*orange*) and high-error (*teal*) conditions, alongside target sets obtained either with the default parameters (*black*) or without reweighting (*cyan*). RMSD distributions are compared after reweighting with $W_{app} = 1.64$ kJ mol^{-1}, denoted by the dashed line in (*A*).

ensemble with upweighted populations of both semiopen and open frames, whereas reweighting to the high-error target data was unable to recreate the target TeaA structural distribution (Fig. 8 *B*). The effects of the low-error parameter set are similar to the trends observed when reweighting to data sets degraded by increasing peptide length (Fig. 4 *B*) or by reducing sequence coverage (Fig. 5 *B*). In both cases, the final β_H and β_C parameters, after optimization during the HDXer reweighting procedure, differed from the values used for either the $\alpha 9$ helix or the remainder of the protein (Table S1). These results suggest that if different regions of the protein are best described by fundamentally different models, perhaps reflective of distinct exchange mechanisms, the use of a single set of averaged β_H and β_C parameters may not produce informative structural ensembles. On the other hand, HDXer can be expected to provide structurally useful information when applied to experimental HDX-MS data that are uniformly well described by a given model or in which the HDX prediction model is only slightly incorrect, for example, in which exchange is well described by β_H and β_C parameters that follow the inverse relationship of the original Best and Vendruscolo model.

Conformational ensembles of a membrane transport protein identified by HDXer

The tests presented so far have been based on artificial HDX data, which allowed us to evaluate how the different kinds of plausible uncertainty in the input data impact the outcome of our ensemble reweighting for an assumed model of the protection factor. It is, however, reasonable to ask how HDXer might perform when applied to true experimental data. To evaluate HDXer in the context of a real-life biological question, we applied this method to experimental HDX-MS data obtained for the bacterial amino acid transporter LeuT. The structure and function of LeuT has been well studied because it serves as a prototype of a wide variety of membrane transport proteins that share its fold. The functional mechanisms of these proteins require that they alternate between states in which substrate binding sites are exposed to either the outside or the inside of the cell; these conformations are referred to as outward- and inward-facing states, respectively. Transport proteins like LeuT are therefore inherently dynamic; however, their conformational preferences can be biased by specific environmental conditions, ligands, and/or point mutations. For example, an earlier study on LeuT reconstituted into lipid nanodiscs suggested that the population of inward-facing conformations is greatly amplified by the Y268A mutation as compared to the WT protein (16). To assess whether application of HDXer would validate or refute this conclusion, we used MD simulations to generate a conformational ensemble containing a mixture of inward- and outward-facing LeuT in equal populations. This ensemble was then reweighted separately so as to reproduce HDX-MS data reported for either WT or Y268A.

After reweighting, the relative population of inward- and outward-facing LeuT conformations was calculated for each set of target HDX-MS data and quantified as the excess population of inward-facing states (i.e., inward population − outward population). Application of HDXer to the Y268A target experimental data resulted in a greater excess of inward-facing conformations in the final structural ensemble, consistent with the conclusions of the original HDX-MS study (Fig. 9). Importantly, the proportion of inward-facing frames increased with W_{app}, suggesting that improved agreement with Y268A HDX-MS data was predominantly driven by the selection of a larger population of inward-facing conformations in the final ensemble.

Available crystal structures and biophysical studies of LeuT with Förster resonance energy transfer (FRET) and double electron electron resonance (DEER) also support the observation that the Y268A mutation shifts the transporter ensemble toward inward-facing populations (66–68). The absolute populations of inward- and outward-facing states have not been reliably determined, however, and are liable to be affected by experimental differences such as the lipid or detergent environment or ionic conditions. Our interpretation of the LeuT HDX-MS data is therefore limited to identifying the trend in population shift upon Y268A mutation in lipid nanodiscs. In this context, we have established that HDXer correctly discriminates the shift toward more inward-facing states associated with the mutation.

Overall, our findings underscore the potential of the proposed method to generate structure-based interpretations of experimentally determined HDX-MS data that are not only quantitative and objective but also mechanistically informative, even when the target data are imperfect.

DISCUSSION

Broad applicability and label-free sample preparation have made HDX-MS an increasingly attractive biophysical
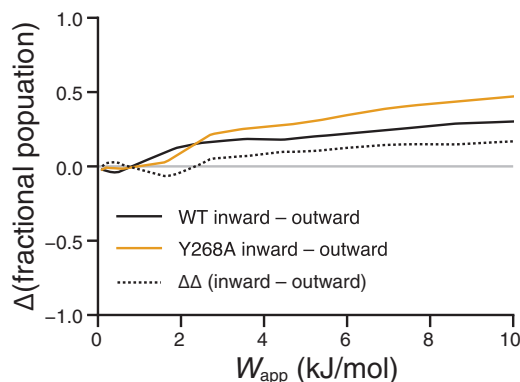
FIGURE 9 Application of HDXer to experimentally determined HDX-MS data for LeuT. A mixed reference ensemble consisting of both outward- and inward-facing LeuT structures was fitted separately to either WT LeuT (*solid black line*) or Y268A mutant HDX-MS data (*solid orange line*). The difference between the WT and mutant data is shown as a dotted line.

technique to study global biomolecular structure and dynamics under native conditions, as demonstrated by the variety of reported applications on both globular and membrane proteins, as well as frequently updated reviews (3,4,69). The major challenge, however, has been to objectively translate the HDX data into structural information so as to be able to formulate conclusive mechanistic insights. The methodology introduced here, named HDXer, is intended to facilitate this structural interpretation. In this approach, a distribution of conformations in a pre-existing ensemble is reweighted post hoc so that calculated ensemble-averaged deuteration levels match a given set of target data. Further analysis of the resulting reweighted ensemble (for example, through clustering) thus provides the desired structural interpretation of the inputted HDX data.

As noted, the overall performance of the HDXer method was assessed on artificially-generated target data. Two factors motivated this deliberate choice. First, we aimed to focus our evaluation on the reweighting method itself, leaving aside other factors that contribute to the prediction of HDX data. By using the same empirical model of P_i (Eq. 1) and EX2-like kinetics, both in the generation of the artificial HDX data and in the calculation of weights (Eq. 4), we ensured that potential inaccuracies in this empirical model did not influence our assessment. Similarly, by using a pre-existing configurational ensemble with a predefined population of states to generate the artificial data, we ensured that there was a correct answer against which our methodology could be evaluated.

The second advantage of artificial data is that it can be systematically degraded in ways that reflect the limitations of actual measurements so as to judge the usability of the technique for structure determination. Indeed, HDX-MS studies vary greatly in terms of the level of peptide coverage and redundancy, and a priori there is no guarantee that an observed set of peptides will contain sufficient information to allow a clear structural interpretation. Our method performs optimally the better the coverage and resolution of the data, as one should expect. However, it is worth noting and is also very promising that even with incomplete sequence coverage or lengthy peptide segments, well beyond those typically attained in well-optimized HDX-MS experiments, our reweighting method can qualitatively identify the major conformational states contributing to the target set (Tables 1 and 2). Furthermore, the identification of relevant conformational states is consistent upon addition of moderate levels of artificial noise to the target data set. These observations, together with the encouraging results obtained for LeuT, lead us to conclude that HDXer will successfully provide structural insights when used to interpret experimentally determined data exhibiting typical coverage and noise.

Notwithstanding these reasons for optimism, it should be noted that the ability of this or any other computational method to facilitate the interpretation of measured HDX data ultimately depends on the fidelity of the empirical model used to calculate the residue protection factors, P_i, for a given structural snapshot. Indeed, to date no HDX prediction model has yet been shown to be uniformly accurate across different biomolecular systems (70–72). In the current HDXer implementation, which predicts P_i using the Best and Vendruscolo forward model (Eq. 1), the structural correlate of the data is the folded protein rather than exchange-competent protein conformations. Therefore, this model is well suited to applications with ensembles generated by, for example, microsecond-timescale MD simulations. Note, however, that with sufficient sampling, HDXer may also be applied with alternative forward models that explicitly define and explore exchange-competent states (34).

Our results indicate that the Best and Vendruscolo model is sensitive to large conformational changes and assigns similar deuterium exchange levels to structurally correlated frames (Fig. 3), which are positive features that are well suited to ensemble reweighting. Moreover, our controlled evaluations of HDXer with model errors incorporated into the target data suggest that reweighting with this model can provide structural insights even when applied to experimentally determined data. On the other hand, our evaluation also makes it clear, reassuringly, that a reweighting method will not be practically useful if the reference ensemble does not include the conformational states present in the target HDX-MS data.

Given the different sources of potential error, it is crucial to be able to assess the reweighting process in absolute terms, i.e., to discern when the optimal solution is less than realistic. The HDXer method is equipped to do so, specifically through the calculation of the W_{app} required to achieve a given MSD. W_{app} and other metrics of reweighting robustness such as the Kish effective sample size (51,73) may be used to identify situations in which the reweighting results are liable to overfitting. Along a similar vein, W_{app} and MSD may also be used as metrics to rank results obtained using alternative empirical protection factor models or alternative reference ensembles so as to evaluate and improve their accuracy. The framework provided by HDXer to assess these potential sources of error is a key advantage in applications with experimental data.

Finally, the HDXer method could be straightforwardly applied to cross validate the HDX data itself (Supporting Materials and Methods). Deuteration levels measured at different time points could be separated into training and validation sets, and inconsistencies in the resultant reweighted ensembles may reveal sources of experimental error. However, as has been extensively discussed for other ensemble refinement methods (50,51,57,58,73), disentangling the exact sources of error in a given set of reweighting results is a challenging proposition and likely to require comparison and cross validation with multiple reference ensembles and experimental data sets.

On a technical note, it is worth underscoring that in contrast to the canonical maximum-entropy reweighting

approach, which enforces exact agreement with an experimental observable, we use a parameter γ to control the degree of fitness to the target to account for potential uncertainties in the measurements. Consequently, HDXer shares some of the theoretical underpinnings of Bayesian approaches used to optimally recreate experimental observables, either through ensemble reweighting or on-the-fly biased sampling (50,57,74). We would argue, however, that biased sampling might not be an appropriate strategy to interpret HDX data, given the empirical nature of HDX prediction models and their imperfect correlation with experiment (70–72) and, more generally, our incomplete understanding of the structural determinants of exchange across different biomolecular systems. Thus, post hoc reweighting seems the most effective approach at this time.

In conclusion, we have developed an effective maximum-entropy-based method to derive a structural-level interpretation of HDX-MS experiments via reweighting of conformational ensembles. We anticipate that HDXer will contribute to more systematic, quantitative analyses of HDX prediction methodologies and aid studies of individual proteins and their functional mechanisms via objective structural interpretation of experimental HDX-MS measurements.

SUPPORTING MATERIAL

Supporting Material can be found online at https://doi.org/10.1016/j.bpj.2020.02.005.

AUTHOR CONTRIBUTIONS

Conceptualization, all authors; methodology development, programming, analysis, and writing of original drafts, R.T.B. and F.M.; supervision and resource and funding acquisition, J.D.F.-G. and L.R.F.; visualization of data, R.T.B., F.M., and L.R.F.; writing—review and editing, all authors; project administration, L.R.F.

ACKNOWLEDGMENTS

REFERENCES

1. Hvidt, A., and K. Linderstrøm-Lang. 1954. Exchange of hydrogen atoms in insulin with deuterium atoms in aqueous solutions. *Biochim. Biophys. Acta.* 14:574–575.

2. Englander, S. W., and N. R. Kallenbach. 1983. Hydrogen exchange and structural dynamics of proteins and nucleic acids. *Q. Rev. Biophys.* 16:521–655.

3. Oganesyan, I., C. Lento, and D. J. Wilson. 2018. Contemporary hydrogen deuterium exchange mass spectrometry. *Methods.* 144:27–42.

4. Trabjerg, E., Z. E. Nazari, and K. D. Rand. 2018. Conformational analysis of complex protein states by hydrogen/deuterium exchange mass spectrometry (HDX-MS): challenges and emerging solutions. *Trends Analyt. Chem.* 106:125–138.

5. Englander, J. J., C. Del Mar, …, V. L. Woods, Jr. 2003. Protein structure change studied by hydrogen-deuterium exchange, functional labeling, and mass spectrometry. *Proc. Natl. Acad. Sci. USA.* 100:7057–7062.

6. Klontz, E. H., A. D. Tomich, …, E. J. Sundberg. 2017. Structure and dynamics of FosA-mediated fosfomycin resistance in *Klebsiella pneumoniae* and *Escherichia coli*. *Antimicrob. Agents Chemother.* 61:e01572-17.

7. Ramirez-Sarmiento, C. A., and E. A. Komives. 2018. Hydrogen-deuterium exchange mass spectrometry reveals folding and allostery in protein-protein interactions. *Methods.* 144:43–52.

8. Zhang, Q., L. N. Willison, …, A. G. Marshall. 2011. Epitope mapping of a 95 kDa antigen in complex with antibody by solution-phase amide backbone hydrogen/deuterium exchange monitored by Fourier transform ion cyclotron resonance mass spectrometry. *Anal. Chem.* 83:7129–7136.

9. Li, J., H. Wei, …, L. A. Schneeweis. 2017. Mapping the energetic epitope of an antibody/interleukin-23 interaction with hydrogen/deuterium exchange, fast photochemical oxidation of proteins mass spectrometry, and alanine shave mutagenesis. *Anal. Chem.* 89:2250–2258.

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

11. Martens, C., M. Shekhar, …, A. Politis. 2018. Direct protein-lipid interactions shape the conformational landscape of secondary transporters. *Nat. Commun.* 9:4151.

12. Zhu, M. M., D. L. Rempel, …, M. L. Gross. 2003. Quantification of protein-ligand interactions by mass spectrometry, titration, and H/D exchange: PLIMSTEX. *J. Am. Chem. Soc.* 125:5252–5253.

13. Deredge, D. J., W. Huang, …, A. Wilks. 2017. Ligand-induced allostery in the interaction of the *Pseudomonas aeruginosa* heme binding protein with heme oxygenase. *Proc. Natl. Acad. Sci. USA.* 114:3421–3426.

14. Sowole, M. A., and L. Konermann. 2014. Effects of protein-ligand interactions on hydrogen/deuterium exchange kinetics: canonical and noncanonical scenarios. *Anal. Chem.* 86:6715–6722.

15. Masson, G. R., S. L. Maslen, and R. L. Williams. 2017. Analysis of phosphoinositide 3-kinase inhibitors by bottom-up electron-transfer dissociation hydrogen/deuterium exchange mass spectrometry. *Biochem. J.* 474:1867–1877.

16. Adhikary, S., D. J. Deredge, …, S. K. Singh. 2017. Conformational dynamics of a neurotransmitter:sodium symporter in a lipid bilayer. *Proc. Natl. Acad. Sci. USA.* 114:E1786–E1795.

17. Merkle, P. S., K. Gotfryd, …, K. D. Rand. 2018. Substrate-modulated unwinding of transmembrane helices in the NSS transporter LeuT. *Sci. Adv.* 4:eaar6179.

18. Möller, I. R., M. Slivacka, …, K. D. Rand. 2019. Conformational dynamics of the human serotonin transporter during substrate and drug binding. *Nat. Commun.* 10:1687.

19. Nielsen, A. K., I. R. Möller, …, C. J. Loland. 2019. Substrate-induced conformational dynamics of the dopamine transporter. *Nat. Commun.* 10:2714.

20. Eisinger, M. L., A. R. Dörrbaum, …, J. D. Langer. 2017. Ligand-induced conformational dynamics of the *Escherichia coli* Na^+/H^+ antiporter NhaA revealed by hydrogen/deuterium exchange mass spectrometry. *Proc. Natl. Acad. Sci. USA.* 114:11691–11696.

21. Giladi, M., L. Almagor, …, D. Khananshvili. 2016. Asymmetric preorganization of inverted pair residues in the sodium-calcium exchanger. *Sci. Rep.* 6:20753.

22. Giladi, M., L. van Dijk, …, D. Khananshvili. 2017. Dynamic distinctions in the Na$^+$/Ca^{2+} exchanger adopting the inward- and outward-facing conformational states. *J. Biol. Chem.* 292:12311–12323.

23. Rostislavleva, K., N. Soler, …, R. L. Williams. 2015. Structure and flexibility of the endosomal Vps34 complex reveals the basis of its function on membranes. *Science.* 350:aac7365.

24. Lim, X.-X., A. Chandramohan, …, G. S. Anand. 2017. Epitope and paratope mapping reveals temperature-dependent alterations in the dengue-antibody interface. *Structure.* 25:1391–1402.e3.

25. van de Waterbeemd, M., A. Llauró, …, A. J. R. Heck. 2017. Structural analysis of a temperature-induced transition in a viral capsid probed by HDX-MS. *Biophys. J.* 112:1157–1165.

26. Radou, G., F. N. Dreyer, …, E. Paci. 2014. Functional dynamics of hexameric helicase probed by hydrogen exchange and simulation. *Biophys. J.* 107:983–990.

27. Bai, Y., J. S. Milne, …, S. W. Englander. 1993. Primary structure effects on peptide group hydrogen exchange. *Proteins.* 17:75–86.

28. Nguyen, D., L. Mayne, …, S. Walter Englander. 2018. Reference parameters for protein hydrogen exchange rates. *J. Am. Soc. Mass Spectrom.* 29:1936–1939.

29. Hvidt, A., and S. O. Nielsen. 1966. Hydrogen exchange in proteins. *Adv. Protein Chem.* 21:287–386.

30. Saltzberg, D. J., H. B. Broughton, …, A. Sali. 2017. A residue-resolved bayesian approach to quantitative interpretation of hydrogen-deuterium exchange from mass spectrometry: application to characterizing protein-ligand interactions. *J. Phys. Chem. B.* 121:3493–3501.

31. Skinner, S. P., G. Radou, …, E. Paci. 2019. Estimating constraints for protection factors from HDX-MS data. *Biophys. J.* 116:1194–1203.

32. Masson, G. R., J. E. Burke, …, K. D. Rand. 2019. Recommendations for performing, interpreting and reporting hydrogen deuterium exchange mass spectrometry (HDX-MS) experiments. *Nat. Methods.* 16:595–602.

33. Best, R. B., and M. Vendruscolo. 2006. Structural interpretation of hydrogen exchange protection factors in proteins: characterization of the native state fluctuations of CI2. *Structure.* 14:97–106.

34. Persson, F., and B. Halle. 2015. How amide hydrogens exchange in native proteins. *Proc. Natl. Acad. Sci. USA.* 112:10383–10388.

35. Wan, H., Y. Ge, …, V. A. Voelz. 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.

36. Craig, P. O., J. Lätzer, …, P. G. Wolynes. 2011. Prediction of native-state hydrogen exchange from perfectly funneled energy landscapes. *J. Am. Chem. Soc.* 133:17463–17472.

37. Kieseritzky, G., G. Morra, and E.-W. Knapp. 2006. Stability and fluctuations of amide hydrogen bonds in a bacterial cytochrome c: a molecular dynamics study. *J. Biol. Inorg. Chem.* 11:26–40.

38. Ma, B., and R. Nussinov. 2011. Polymorphic triple β-sheet structures contribute to amide hydrogen/deuterium (H/D) exchange protection in the Alzheimer amyloid β42 peptide. *J. Biol. Chem.* 286:34244–34253.

39. Borysik, A. J. 2017. Simulated isotope exchange patterns enable protein structure determination. *Angew. Chem. Int.Engl.* 56:9396–9399.

40. Liu, T., D. Pantazatos, …, V. L. Woods, 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.

41. Petruk, A. A., L. A. Defelipe, …, A. G. Turjanski. 2013. Molecular dynamics simulations provide atomistic insight into hydrogen exchange mass spectrometry experiments. *J. Chem. Theory Comput.* 9:658–669.

42. Claesen, J., and A. Politis. 2019. POPPeT: a new method to predict the protection factor of backbone amide hydrogens. *J. Am. Soc. Mass Spectrom.* 30:67–76.

43. Park, I.-H., J. D. Venable, …, A. Brock. 2015. Estimation of hydrogen-exchange protection factors from MD simulation based on amide hydrogen bonding analysis. *J. Chem. Inf. Model.* 55:1914–1925.

44. Pitera, J. W., and J. D. Chodera. 2012. On the use of experimental observations to bias simulated ensembles. *J. Chem. Theory Comput.* 8:3445–3451.

45. Boomsma, W., J. Ferkinghoff-Borg, and K. Lindorff-Larsen. 2014. Combining experiments and simulations using the maximum entropy principle. *PLoS Comput. Biol.* 10:e1003406.

46. Marinelli, F., and G. Fiorin. 2019. Structural characterization of biomolecules through atomistic simulations guided by DEER measurements. *Structure.* 27:359–370.e12.

47. Różycki, B., Y. C. Kim, and G. Hummer. 2011. SAXS ensemble refinement of ESCRT-III CHMP3 conformational transitions. *Structure.* 19:109–116.

48. Cesari, A., A. Gil-Ley, and G. Bussi. 2016. Combining simulations and solution experiments as a paradigm for RNA force field refinement. *J. Chem. Theory Comput.* 12:6192–6200.

49. Hermann, M. R., and J. S. Hub. 2019. SAXS-restrained ensemble simulations of intrinsically disordered proteins with commitment to the principle of maximum entropy. *J. Chem. Theory Comput.* 15:5103–5115.

50. Bottaro, S., T. Bengtsen, and K. Lindorff-Larsen. 2020. Integrating molecular simulation and experimental data: a Bayesian/maximum entropy reweighting approach. *Methods Mol. Biol.* 2112:219–240, Published online October 31, 2018.

51. Cesari, A., S. Reißer, and G. Bussi. 2018. Using the maximum entropy principle to combine simulations and solution experiments. *Computation.* 6:15.

52. Bonomi, M., G. T. Heller, …, M. Vendruscolo. 2017. Principles of protein structural ensemble determination. *Curr. Opin. Struct. Biol.* 42:106–116.

53. Marinelli, F., and J. D. Faraldo-Gómez. 2015. Ensemble-biased metadynamics: a molecular simulation method to sample experimental distributions. *Biophys. J.* 108:2779–2782.

54. Hustedt, E. J., F. Marinelli, …, H. S. Mchaourab. 2018. Confidence analysis of DEER data and its structural interpretation with ensemble-biased metadynamics. *Biophys. J.* 115:1200–1216.

55. Marinelli, F., S. I. Kuhlmann, …, J. D. Faraldo-Gómez. 2011. Evidence for an allosteric mechanism of substrate release from membrane-transporter accessory binding proteins. *Proc. Natl. Acad. Sci. USA.* 108:E1285–E1292.

56. Olsson, S., H. Wu, …, F. Noé. 2017. Combining experimental and simulation data of molecular processes via augmented Markov models. *Proc. Natl. Acad. Sci. USA.* 114:8265–8270.

57. Hummer, G., and J. Köfinger. 2015. Bayesian ensemble refinement by replica simulations and reweighting. *J. Chem. Phys.* 143:243150.

58. Köfinger, J., L. S. Stelzl, …, G. Hummer. 2019. Efficient ensemble refinement by reweighting. *J. Chem. Theory Comput.* 15:3390–3401.

59. Pedregosa, F., G. Varoquaux, …, É. Duchesnay. 2011. Scikit-learn: machine learning in Python. *J. Mach. Learn. Res.* 12:2825–2830.

60. Rousseeuw, P. J. 1987. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *J. Comput. Appl. Math.* 20:53–65.

61. Hudgens, J. W., E. S. Gallagher, …, S. Urata. 2019. Interlaboratory comparison of hydrogen-deuterium exchange mass spectrometry measurements of the fab fragment of NISTmAb. *Anal. Chem.* 91:7336–7345.

62. Marinelli, F. 2013. Following easy slope paths on a free energy landscape: the case study of the Trp-cage folding mechanism. *Biophys. J.* 105:1236–1247.

63. Peacock, R. B., J. R. Davis, …, E. A. Komives. 2018. Dynamic consequences of mutation of tryptophan 215 in thrombin. *Biochemistry.* 57:2694–2703.

64. Camilloni, C., and F. Pietrucci. 2018. Advanced simulation techniques for the thermodynamic and kinetic characterization of biological systems. *Adv. Phys. X.* 3:1477531.

65. Markwick, P. R. L., R. B. Peacock, and E. A. Komives. 2019. Accurate prediction of amide exchange in the fast limit reveals thrombin allostery. *Biophys. J.* 116:49–56.

66. Zhao, Y., D. Terry, …, J. A. Javitch. 2010. Single-molecule dynamics of gating in a neurotransmitter transporter homologue. *Nature.* 465:188–193.

67. Kazmier, K., S. Sharma, …, H. S. McHaourab. 2014. Conformational dynamics of ligand-dependent alternating access in LeuT. *Nat. Struct. Mol. Biol.* 21:472–479.

68. Krishnamurthy, H., and E. Gouaux. 2012. X-ray structures of LeuT in substrate-free outward-open and apo inward-open states. *Nature.* 481:469–474.

69. Masson, G. R., M. L. Jenkins, and J. E. Burke. 2017. An overview of hydrogen deuterium exchange mass spectrometry (HDX-MS) in drug discovery. *Expert Opin. Drug Discov.* 12:981–994.

70. Skinner, J. J., W. K. Lim, …, S. W. Englander. 2012. Protein hydrogen exchange: testing current models. *Protein Sci.* 21:987–995.

71. McAllister, R. G., and L. Konermann. 2015. Challenges in the interpretation of protein h/d exchange data: a molecular dynamics simulation perspective. *Biochemistry.* 54:2683–2692.

72. Mohammadiarani, H., V. S. Shaw, …, H. Vashisth. 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.

73. Rangan, R., M. Bonomi, …, M. Vendruscolo. 2018. Determination of structural ensembles of proteins: restraining vs reweighting. *J. Chem. Theory Comput.* 14:6632–6641.

74. Bonomi, M., C. Camilloni, …, M. Vendruscolo. 2016. Metainference: a Bayesian inference method for heterogeneous systems. *Sci. Adv.* 2:e1501177.

75. Schwieters, C. D., and G. M. Clore. 2002. Reweighted atomic densities to represent ensembles of NMR structures. *J. Biomol. NMR.* 23:221–225.
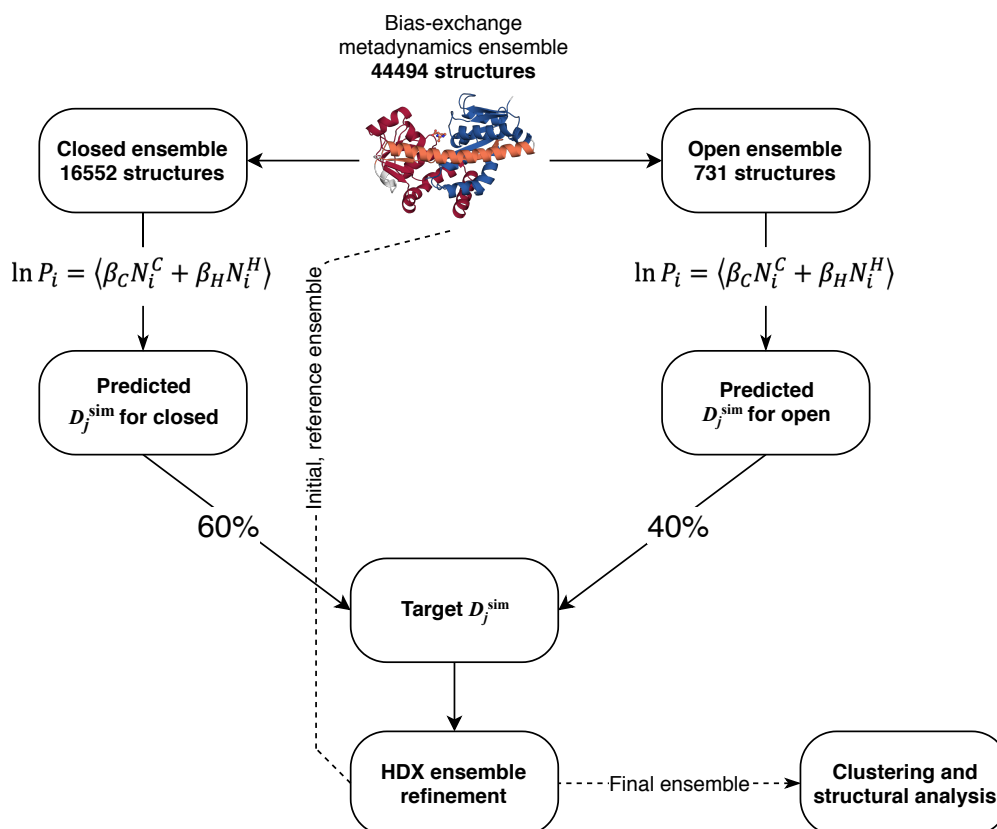
Supplemental Information

**Interpretation of HDX Data by Maximum-Entropy Reweighting of Simulated Structural Ensembles**

Richard T. Bradshaw, Fabrizio Marinelli, José D. Faraldo-Gómez, and Lucy R. Forrest

Figure S1 – Structural ensembles and target data used for HDX ensemble refinement testing. Artificial HDX data (*solid arrows*) were generated for each exchangeable amide for a mixed ensemble corresponding to 60% closed and 40% open TeaA. These HDX data were used as the target for ensemble refinement of the complete test ensemble, which includes closed, open, and decoy semi-open frames (*dashed lines*).

1   10   20   30   40   50   60   70   80

DNWRYAHEEYEGDVQDVFAQAFKGYVEDNSDHTVQVYRFGELGESDDIMEQTQNGILQFVNQSPGFTGSLIPSAQIFFIP

81   90   100   110   120   130   140   150   160

YLMPTDMDTVLEFFDESKAINEMFPKLYAEHGLELLKMYPEGEMVVTADEPITSPEDFDNKKIRTMTNPLLAETYKAFGA

161   170   180   190   200   210   220   230   240

TPTPLPWGEVYGGLQTGIIDGQENPIFWIESGGLYEVSPNLTFTSHGWFTTAMMANQDFYEGLSEEDQQLVQDAADAAYD

241   250   260   270   280   290   300   310

HTIEHIKGLSEESLEKIKAASDEVTVTRLNDEQIQAFKERAPQVEEKFIEMTGEQGQELLDQFKADLKAV

Coverage: 100% 80% 60% 40% 20%   Domains: N C β4/ α9

Figure S2 – Residues included in target HDX data for TeaA with reduced levels of sequence coverage. The sequence is colored according to the protein domain region as defined in Fig. 1A. Bars underneath the sequence indicate the residues included in the target data at 100% coverage (*dark brown*) and at approximately 80, 60, 40, or 20% coverage (from *light brown to orange to yellow*). See Methods section in the main text for more details.
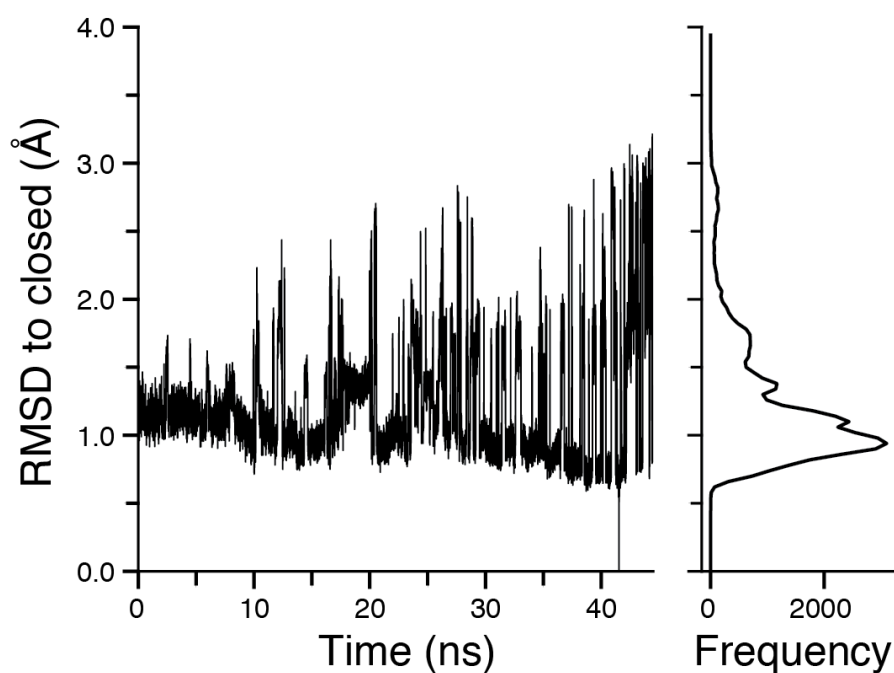
Figure S3 – Structural variability in the reference simulation of TeaA and selection of reference structures. The root mean square deviation (RMSD) of the C_α atoms from the reference closed conformation is plotted against time for the unbiased trajectory from bias-exchange metadynamics simulations performed by Marinelli & coworkers (1). Trajectory frames were taken at 1 ps intervals. The frame at 41.546 ns was arbitrarily selected as the reference closed configuration, among many frames with similarly-low RMSD from the holo structure. Visual analysis confirmed that this configuration adequately represented the two key features of closed structures, namely a small inter-lobe distance and a kinked $\alpha 9$ helix. The frame at 44.428 ns has the highest RMSD to the closed conformation and was therefore selected as the reference open configuration. The distribution of the RMSD values to the closed structure is shown in the panel on the right.
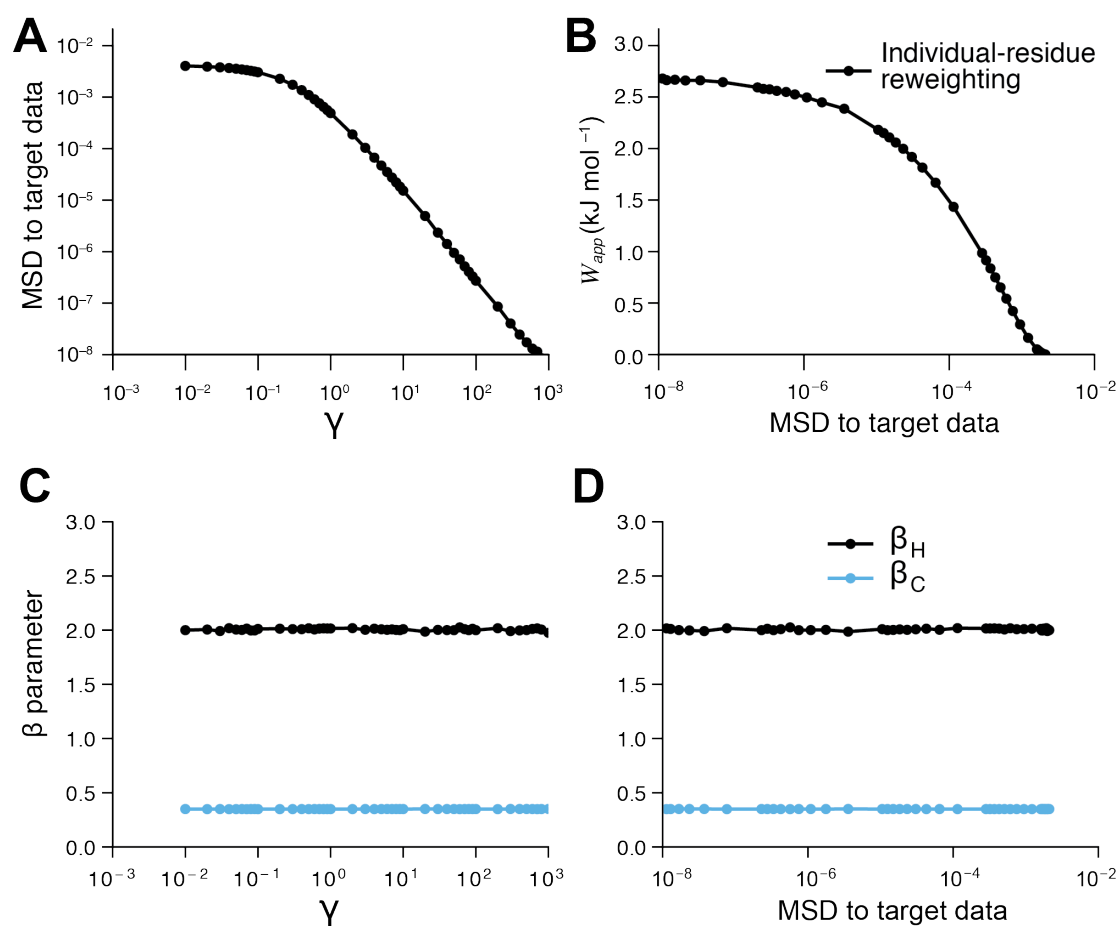
Figure S4 – Relationship between γ, W_{app}, and the agreement with the target data during HDX reweighting analyses in which the target TeaA HDX data is resolved at the individual residue level. Circles indicate independent reweighting experiments. **(A)** The mean square deviation (MSD) between the predicted and target HDX values increases as the value of γ decreases. Little improvement in MSD is observed for values of γ below ca. 10^{-1}, suggesting that, beyond this point, the initial HDX data lies within the uncertainty distribution ρ_{err} that is defined by γ. **(B)** For the same reweighting analyses, the reduction in MSD is coupled to an increase in W_{app} until a plateau is reached for MSD values below ca. 10^{-7}. **(C)** and **(D)** Optimized values of the β_H (*black*) and β_C (*blue*) parameters after reweighting do not change as a function of γ or MSD. The target HDX-MS data was generated using β_H = 2.0 and β_C = 0.35.
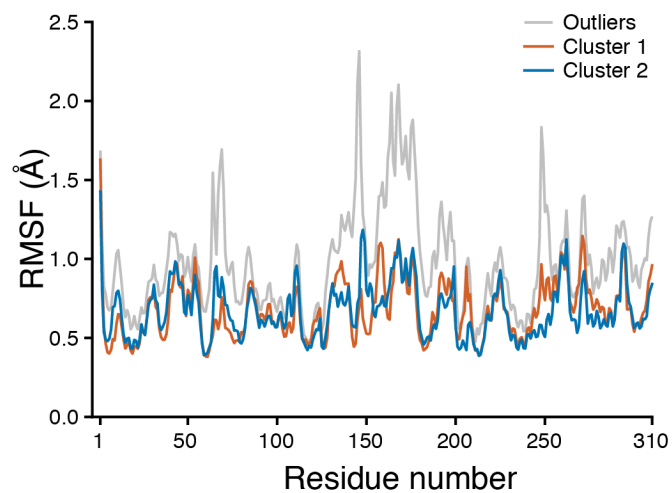
Figure S5 – Structural variance in the clustered ensembles of TeaA obtained after ensemble reweighting by HDX data resolved at the individual residue level. The backbone root mean squared fluctuation (RMSF) was averaged over each residue for all conformations in the two main clusters (*orange and cyan*), or in the outliers (*gray*).
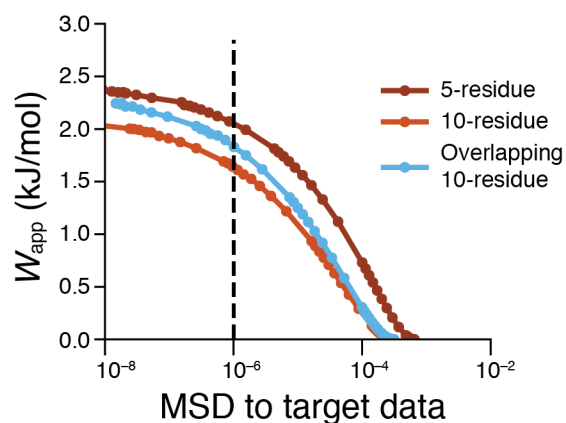
Figure S6 – Effect of overlap (i.e., redundancy) in the peptide segment residue ranges on reweighting. The decision plot shows the work applied during reweighting against the MSD of the reweighted ensemble to the target HDX data. Circles indicate independent reweighting experiments. Target data with overlapping 10-residue segments were defined such that peptides maintained 100% sequence coverage, but overlapped in approximately 5-residue intervals (e.g., residues 1-10, 6-15, 10-19, 15-24…). Reweighting using overlapping segments (*blue*) initially shows similar performance to reweighting with non-overlapping 10-residue segment data (*light brown*), but trends towards the results obtained with shorter, 5-residue, peptide segments (*dark brown*), owing to the additional information content of target data with peptide redundancy.
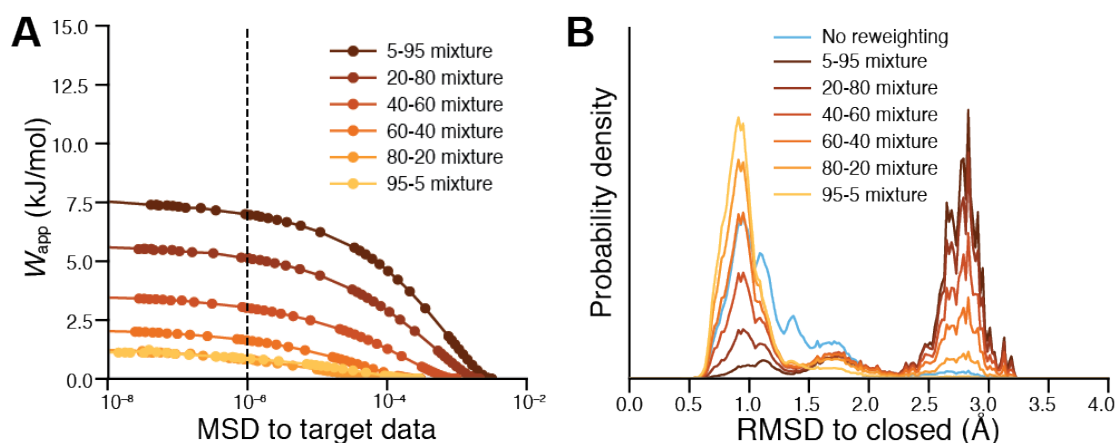
Figure S7 – Effect of targeting the HDX reweighting to ensembles with varying populations of the two major states. Target ensembles are defined as percentage mixtures of closed and open states. The 60-40 mixture is the target ensemble used in the main text. All data assume 10-residue long peptide segments. **(A)** Decision plot showing the apparent work applied during reweighting, W_{app}, against the MSD of the reweighted ensemble to the target HDX data. As the target population of closed states decreases from 80% to 5%, greater values of W_{app} are required to achieve an agreement of MSD = 10^{-6} (*dashed line*). Reweighting to mixtures with 80% and 95% closed states required approximately equal W_{app}, as the large increase in closed state population compared to the reference ensemble was balanced by a comparatively smaller increase in open state population. In all cases the reference ensemble populations were 37.2 : 1.6 : 61.2 for closed : open : decoy states. **(B)** RMSD distributions, with reference to the closed TeaA conformation, after reweighting. Trends in the target ensembles are at least qualitatively recreated across all datasets.
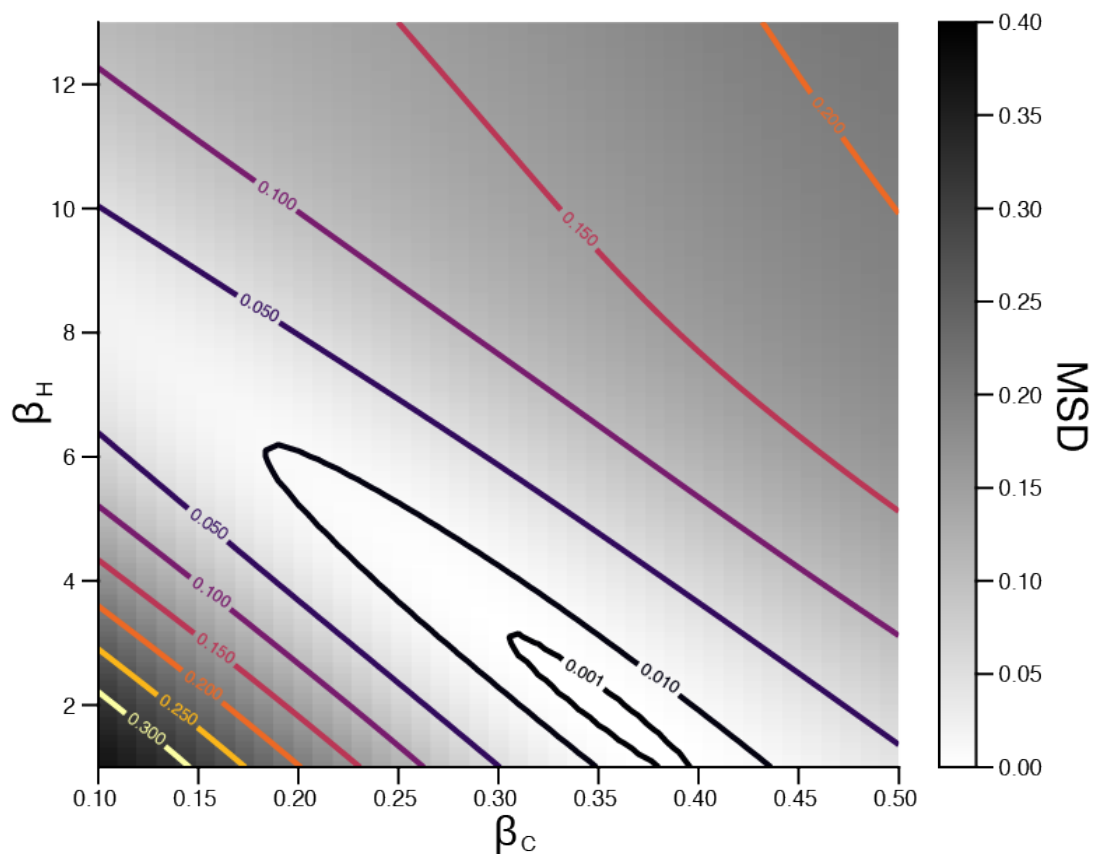
Figure S8 – Effects of varying the β parameters on predicted HDX values obtained using the Best and Vendruscolo forward model for TeaA. Contours indicate the mean square deviation (MSD) of the predicted HDX to the reference HDX values calculated with β_H = 2.0 and β_C = 0.35. The artificial TeaA HDX data generated with 100% sequence coverage in 10-residue peptide segments were used as reference HDX values.

Table S1 – HDX model parameters, β_C and β_H, after reweighting to original, low-error and high-error target datasets, with W_{app} = 1.64 kJ mol^{-1}

Dataset	*Target β_H	*Target β_C	†Observed β_H	†Observed β_C
Original	2.0	0.35	2.02	0.349
Low error	7.0	0.20	2.61	0.336
High error	2.0	0.20	1.67	0.343

* Values of β_H and β_C used to calculate target HDX-MS data for the α9 helix of TeaA; HDX-MS data for the remaining residues were calculated using the default parameters, β_H = 2.0 and β_C = 0.35. † Values of β_H and β_C obtained by optimization during the reweighting procedure.

Table S2 – Agreement of reweighted ensembles and target data for noisy and zero-noise datasets, measured as MSD. Each row reflects a cross-validation of a reweighted ensemble (trained model) against multiple target HDX data (test datasets). The accuracy of each reweighted ensemble to its own training data is shown in **bold**.

Training dataset for reweighting	Test dataset for MSD calculation					
	No noise		σ = 0.1		σ = 0.01	
	1 min	60 min	1 min	60 min	1 min	60 min
No noise	**1.02 x 10^{-6}**	**8.06 x 10^{-7}**	1.34 x 10^{-2}	7.68 x 10^{-3}	1.30 x 10^{-4}	7.71 x 10^{-5}
σ = 0.1	2.11 x 10^{-4}	3.40 x 10^{-4}	**1.25 x 10^{-2}**	**7.12 x 10^{-3}**	4.02 x 10^{-4}	4.33 x 10^{-4}
σ = 0.01	3.32 x 10^{-5}	3.32 x 10^{-5}	1.33 x 10^{-2}	7.97 x 10^{-3}	**8.73 x 10^{-5}**	**5.52 x 10^{-5}**

Text S1 – Cross-validation experiments with HDXer reweighted ensembles

Overfitting is an unwanted, but common, pitfall of methodologies designed to construct or refine structural ensembles based on limited-resolution data. We therefore evaluated the potential for overfitting in the context of HDXer. In the article enclosed we highlighted how the apparent work (W_{app}) as a function of the MSD provides an effective descriptor of the reweighting process; when this apparent work increases in magnitude with no significant improvement in MSD, overfitting is likely the cause. Nevertheless, it is unlikely that monitoring of a single metric will unambiguously preclude overfitting, and thus complementary approaches, such as cross validation experiments, are desirable. To illustrate possible applications of HDXer for cross-validation, we first defined training and test HDX-MS datasets. Specifically, we treated the reweighted ensemble as if it were a trained model with a given level of accuracy (e.g., MSD = 1 x 10^{-6}) to the target HDX data that were used for the training (reweighting) of the ensemble. In the absence of overfitting, a statistically-robust reweighted ensemble should have similar accuracy when compared to the training set data as to test data that were *not* included in the reweighting process, but were sampled from the same underlying distribution. Our first cross-validation test investigated the reweighting performed with the data comprised of 10-residue segments and with 100% coverage (**Fig. 4**). We generated new test HDX-MS data at two timepoints (5 and 30 min) that were *not* used in the HDXer reweighting, which was instead "trained" with HDX data at 0.167, 1, 10, 60, and 120 min timepoints. The reweighted ensemble had similar agreement with the test HDX dataset (MSD = 3.01 x 10^{-6} and 1.04 x 10^{-6} for the 5- and 30-minute timepoints, respectively) as with the training dataset (MSD = 1 x 10^{-6}). This result reassuringly suggests that the final ensemble had not been overfitted in this case.

With HDX-MS experiments, multiple test datasets can be easily acquired, for example from multiple independent repeat samples and multiple independent deuteration timepoints. A robustly-reweighted ensemble should agree similarly to all test data that are sampled from the same protein structural distribution, provided that the level of uncertainty of the test data is similar to that of the training data. Conversely, if test and training data feature different levels of uncertainty, in a robustly-reweighted ensemble the accuracy in reproducing each data set must be comparable with the relative error level. Our second cross-validation therefore involved test HDX datasets with a *different level of experimental noise* to the training sets included in the reweighting. Specifically, using the final ensembles obtained after reweighting to noisy target data (**Fig. 6B**), we calculated the MSD to the zero-noise target data at the 1- and 60-min timepoints. *Vice-versa*, using the final ensemble after reweighting to

11

zero-noise target data, we also calculated the MSD to the noisy target data at the 1- and 60-min timepoints. In each cross-validation case, the reweighted ensembles exhibit different levels of agreement to the test datasets than to the training dataset included in the HDXer reweighting. Nonetheless, these results are fully compatible with the noise level in each data set, underscoring that the reweighted ensembles were generated to a level of precision consistent with the uncertainty present in the target data (**Table S2**).

These tests demonstrate how cross-validation may be used to assess whether reweighted ensembles suffer from overfitting. Of course, this application is easily characterized here, in part owing to our use of artificial datasets with controlled sources of error. In practice, identifying potential sources of overfitting (e.g., experimental datapoints with differing uncertainty levels, simulated ensembles that do not adequately describe the variance of the experimental data, etc.) might require extensive investigation. Here, we limit ourselves to highlighting this potential use of HDX-MS data, and we propose that similar approaches will be advantageous when applied to real experimental target data that require thorough cross-validation of reweighted structural ensembles.

Movie S1 – Artificially-generated morph between the closed and open representative structures of TeaA. TeaA is shown in cartoon representation (*wheat*), and the ectoine substrate from the closed configuration is shown in ball and stick representation (*peach*).

References for Supporting Materials

1. Marinelli, F., S.I. Kuhlmann, E. Grell, H.-J. Kunte, C. Ziegler, and J.D. Faraldo-Gómez. 2011. Evidence for an allosteric mechanism of substrate release from membrane-transporter accessory binding proteins. Proc. Natl. Acad. Sci. U. S. A. 108: E1285–E1292.