Biophysical Journal, Volume *120*

Supplemental information

Modeling the native ensemble of PhuS using enhanced sampling MD and HDX-ensemble reweighting

Kyle C. Kihn, Tyree Wilson, Ally K. Smith, Richard T. Bradshaw, Patrick L. Wintrode, Lucy R. Forrest, Angela Wilks, and Daniel J. Deredge

Supporting Material:

Modeling the native ensemble of PhuS using enhanced-sampling MD and HDX-ensemble reweighting

K. Kihn[1], T. Wilson[1], A. Smith[1], R. T. Bradshaw[2], P. L. Wintrode[1], L. R. Forrest[3], A. Wilks[1], D. J. Deredge[1]*

[1]Department of Pharmaceutical Sciences, School of Pharmacy, University of Maryland, Baltimore, MD, USA, [2]Department of Chemistry, King's College London, London, UK, and [3]Computational Structural Biology Section, National Institute of Neurological Disorders and Stroke, National Institute of Health, Bethesda, MD, USA

*For correspondence: dderedge@rx.umaryland.edu (DJD)

Running title:

Enhanced sampling MD and HDXer of PhuS

Supplemental Methods

Ensemble reweighting adjusts the relative populations ('weights') of frames within an initial probability distribution such that the final ensemble average of a given property conforms to a target value. A maximum entropy criterion enforces that the minimal possible bias is applied to the initial distribution during the reweighting process. In the case of HDXer, conformity with the target value (the experimental HDX-MS data) can be modified by means of an uncertainty distribution of adjustable width, controlled by the parameter γ. Overall, frames from the initial structural ensemble (in our case from simulations) which most conform with the experimental data will be upweighted, and vice versa.

Briefly and methodologically, HDXer adjusts the original potential energy function of the simulation $U(X)$ to create a new potential $U_{corr}(X)$, with populations that better reflect the experimental data. The correction itself takes the form of the predicted experimental data (here at the protection factor level, using the phenomenological equation of Best & Vendruscolo), and a per-residue variable λ_i as in Eq. 3.

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

The values λ_i must be solved iteratively during the reweighting process. Optimal values of λ_i minimize a cost function that balances the error to the target data (including the uncertainty controlled by g), with the magnitude of the bias applied to the initial distribution. After optimization, the final λ_i factors are then normalized over the entire simulation resulting in an adjusted weight value for each frame as follows in Eq. 4.

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

The denominator is a normalization term calculated by summing over all simulation configurations.

The details of the statistical methodology to determine the optimal λ_i factors such that the adjusted deuteration fraction from the reweighted ensemble fits the experimental deuteration fraction within defined statistical error has been described (1).

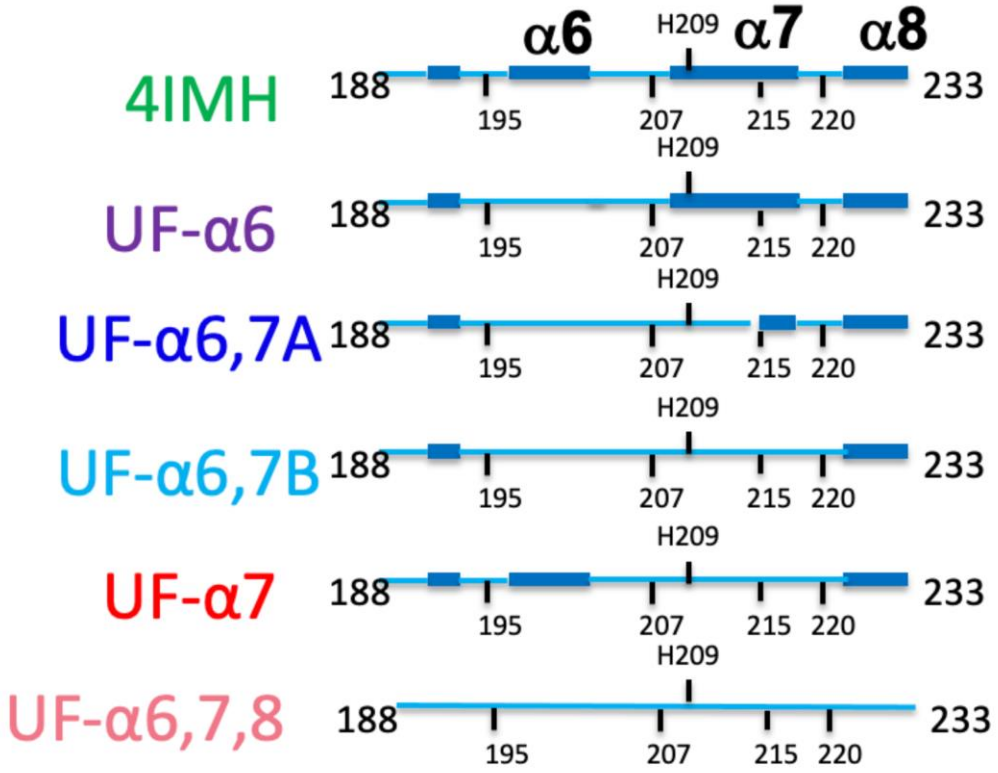
Supplemental Results:

TICA and maximum entropy reweighting using locally-unfolded models. For a thorough validation of the use of TICA in conjunction of HDXer, we employed the same strategy as in UF-α6,7B (Fig. 2D,E and F) with all UF models. The ensemble from a given UF model was combined with the ensemble from the apo-PhuS crystal structure configuration to generate 5 mixed ensembles. These mixed ensembles were then subjected to TICA which consistently separated the frames based on the simulation of origin (Fig. S9A) in the first dimension of separation (TICA1). The separation is also seen following randomization of the frames in the mixed ensemble. In the second TICA dimension, distinct secondary clustering was consistently observed

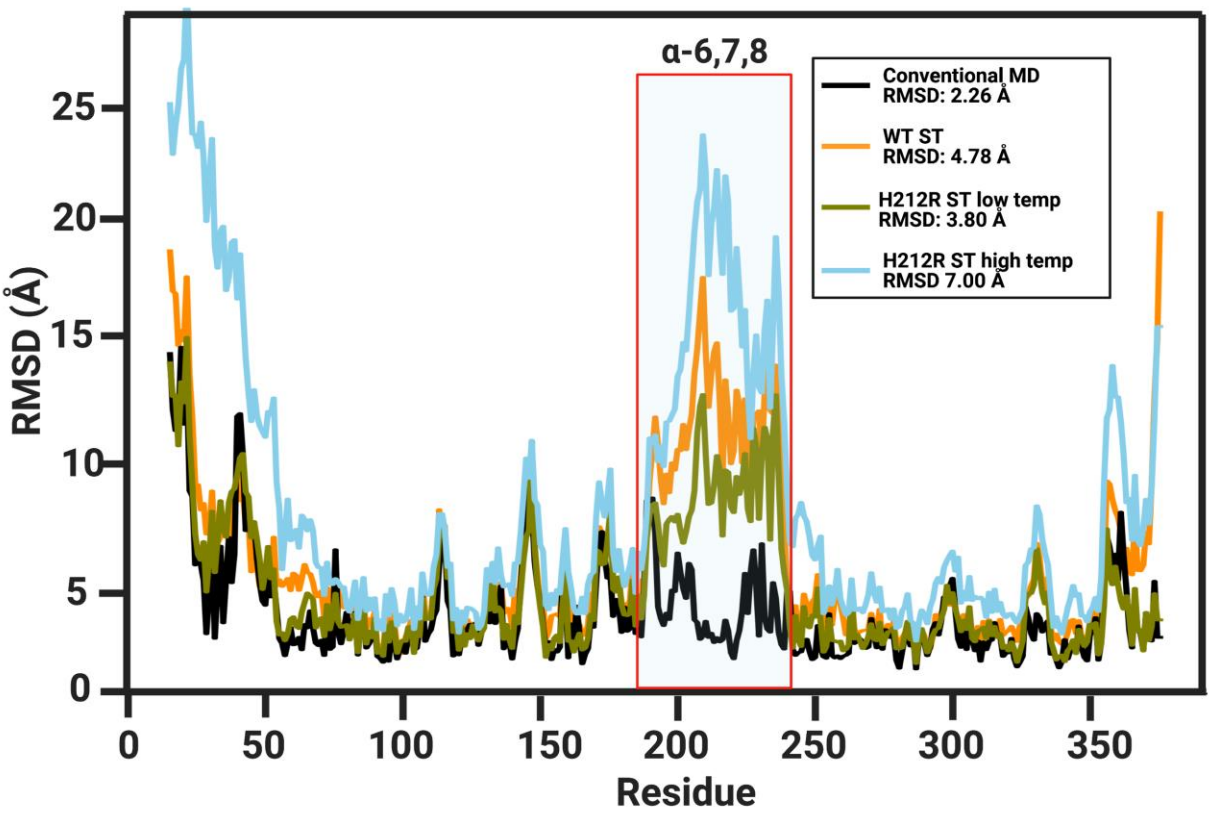
for the frames originating from UF model simulation (Fig. S9A), possibly reflecting the more dynamic behavior of the UF ensembles.

Separately, HDXer was performed on the mixed ensembles using experimental HDX-MS data from apo and holo-PhuS. A range of γ values from 0.1 to 9 were used for the reweighting and a decision plot, an apparent work W_{app} *vs* Mean Squared Deviation plot, was constructed to determine a suitable value. An examination of the decision plots for the reweighting of mixed ensembles vs apo-PhuS HDX-MS as a target data (Fig. S9C) shows that less W_{app} is required to achieve closer agreement with the experimental data (smaller MSD values), in ensembles containing frames from models UF-α6,7A, UF-α6,7B and UF-α6,7,8 than those containing frames from UF-α6 and UF-α7. However, when reweighting to holo-PhuS HDX-MS data, lower MSD values were achieved with less W_{app} than when reweighted to apo-PhuS HDX-MS. Moreover, the decision plots were observed to be independent from the make-up of the mixed ensemble when reweighted to holo-PhuS. This is likely because, in the case of reweighting to holo-PhuS HDX-MS data, most of the effective W_{app} is being applied to the upweighting of frames originating from the simulation with the crystallographic configuration (4IMH), a constant component of the mixed ensembles.

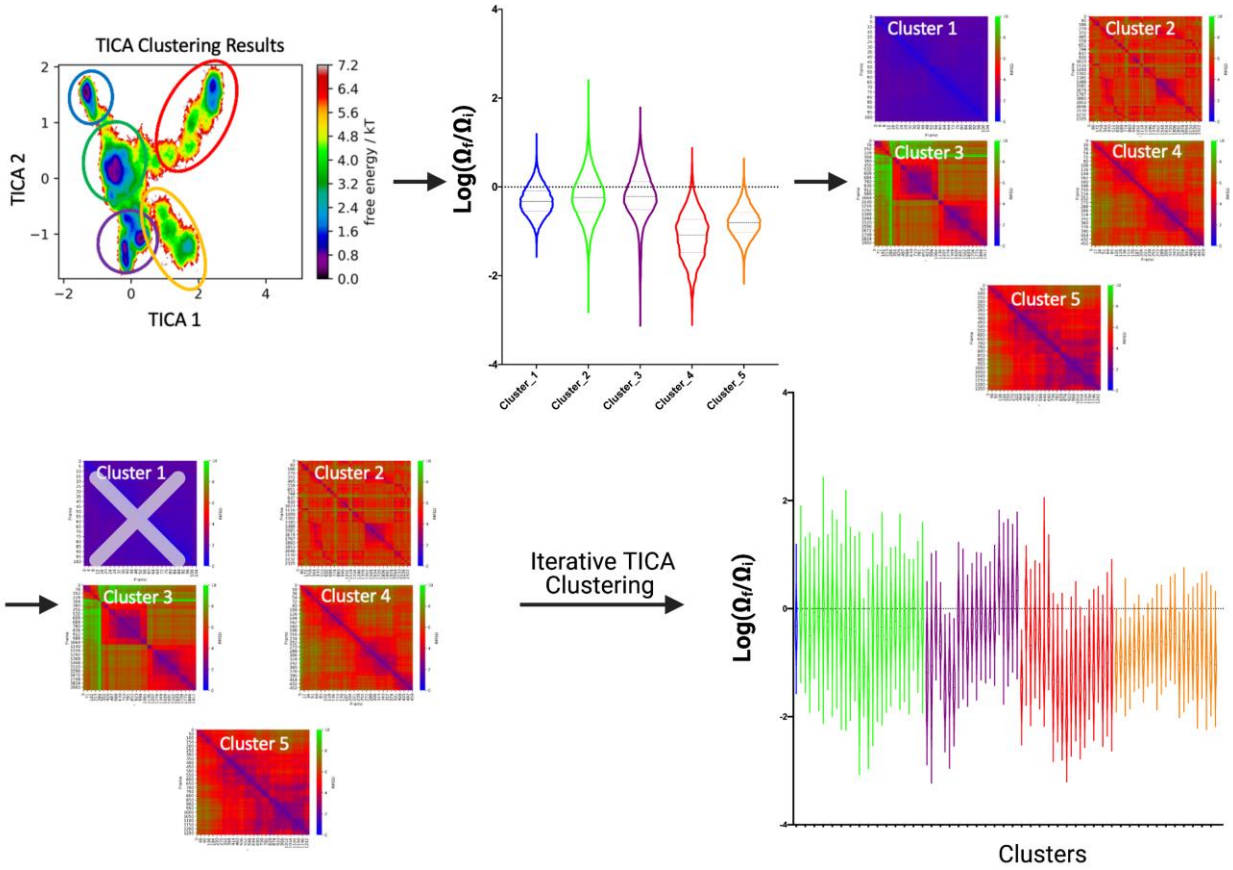
The ideal γ value selected from these plots represented a W_{app} of ~2 kJ/mol and was below the inflection point of the curve, i.e., where the best fit is achieved without overfitting the data (Fig. S9C). This value is then used to determine the final weights (Ω_f) for each frame of the ensemble in the reweighting process. Using this value, the resulting fold change in weight ($\log(\Omega_f/\Omega_i)$ where Ω_i is the initial weight of the frame prior to reweighting) of each individual frame was mapped to the respective clusters resulting from TICA dimensional reduction (Fig. S9B) to determine weight redistribution within and across clusters. When HDXer was performed against apo-PhuS HDX-MS data, the average $\log(\Omega_f/\Omega_{\mathrm{i}})$ of the cluster originating from the apo crystal structure was consistently downweighted (Fig. S9B top). On the other hand, clusters originating from the UF models displayed mixed behavior. One subcluster is typically seen to be heavily downweighted, while the other consistently displays the highest average $\log(\Omega_f/\Omega_i)$ (Fig. S9B top). Here again, the mixed ensemble containing model UF-α6,7B with α6 and α7 partially unfolded is observed to have the most upweighted subcluster (subcluster 2) (Fig. S9B top). In contrast, when reweighting was performed against holo-PhuS HDX-MS (Fig. 2C bottom left, Fig. S9B bottom), the clusters containing frames generated from the apo crystal structure simulations consistently displayed the greatest average $\log(\Omega_f/\Omega_i)$, except for model UF-α7.



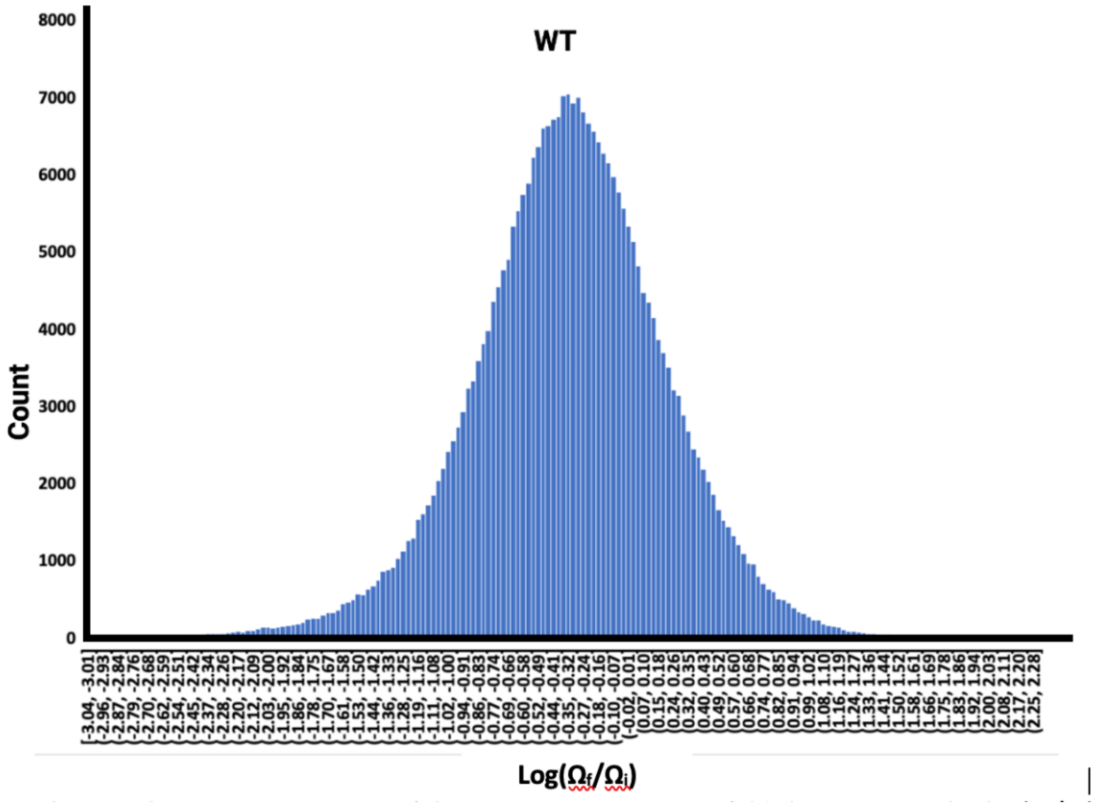
Supplemental Figure 1: Schematic diagram illustrating the secondary structure elements that were modeled as unstructured loops in the α-6,7,8 region for the various locally-unfolded models.



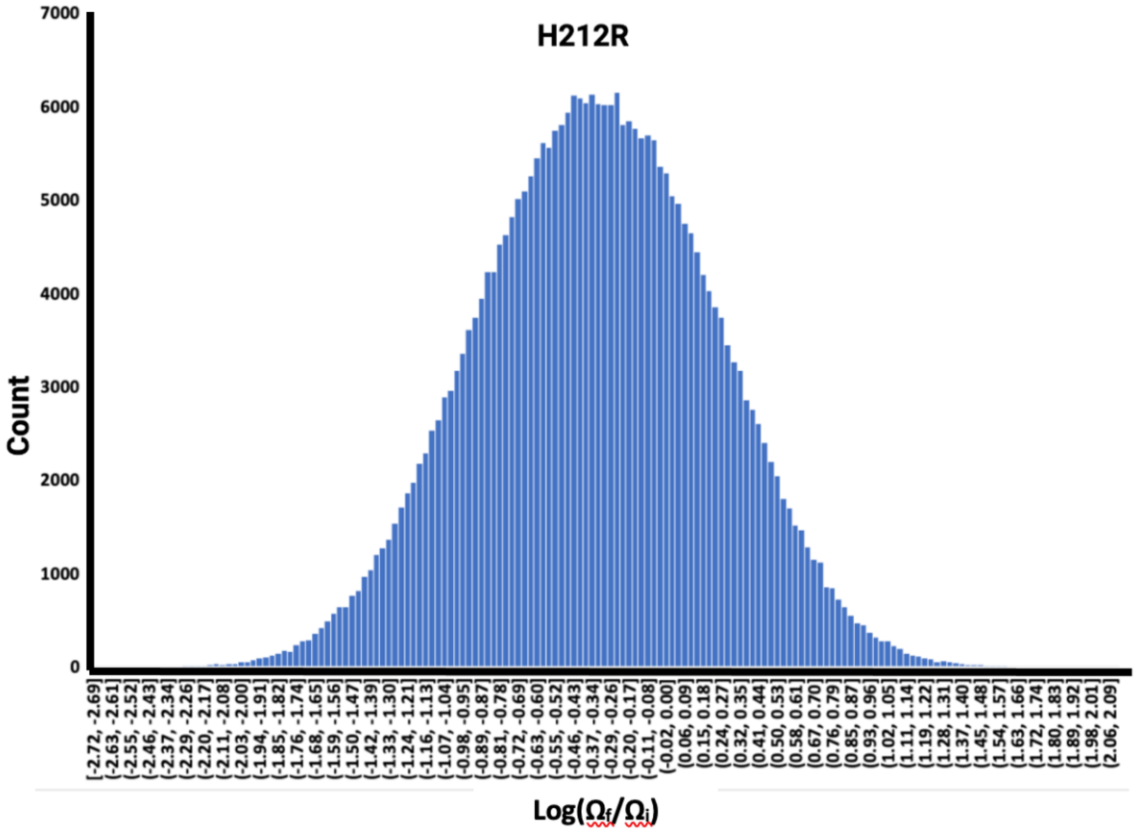
Supplemental Figure 2: Trajectory-averaged local C_α-RMSD (Å) for each residue in PhuS for the combined 6 ST runs of WT apo-PhuS (orange), the 5 μs conventional MD simulation of WT apo-PhuS (black), the combined 420K and 440K ST runs of H212R apo-PhuS (green), and the combined 520K ST runs of H212R (blue). The average overall C_α RMSD to the starting crystal structure is given in the legend. RMSD values for the wild type ST and the conventional MD were calculated to the 4IMH structure and the RMSD values for the H212R ST low temperature and H212R ST high temperature runs were calculated with respect to the *in silico* H212R structure derived from 4IMH. The region comprising the α-6,7,8 motif is highlighted with a red box.



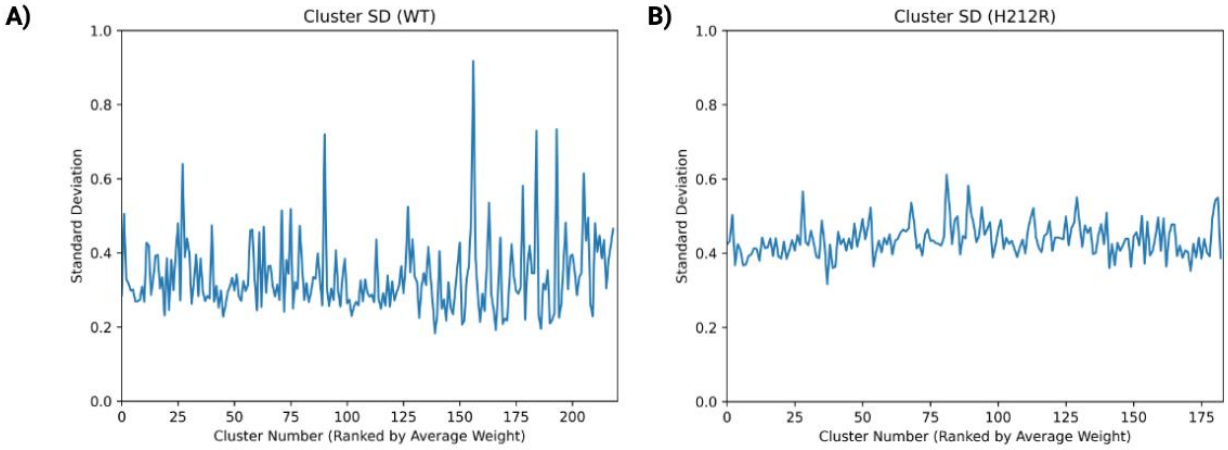
Supplemental Figure 3: Workflow incorporating HDXer with iterative clustering through TICA. Initial TICA clustering is performed (top left) and clusters are extracted (colored ellipses). HDXer reassigned weights are applied to each cluster (top middle). A frame-to-frame pairwise RMSD is calculated for each cluster and plotted (top right). When all pairwise RMSD values are < 3 Å, no further TICA is performed on that cluster (cluster 1, bottom left). The remaining clusters are then further analyzed with iterative TICA clustering until the endpoints are achieved (bottom right).



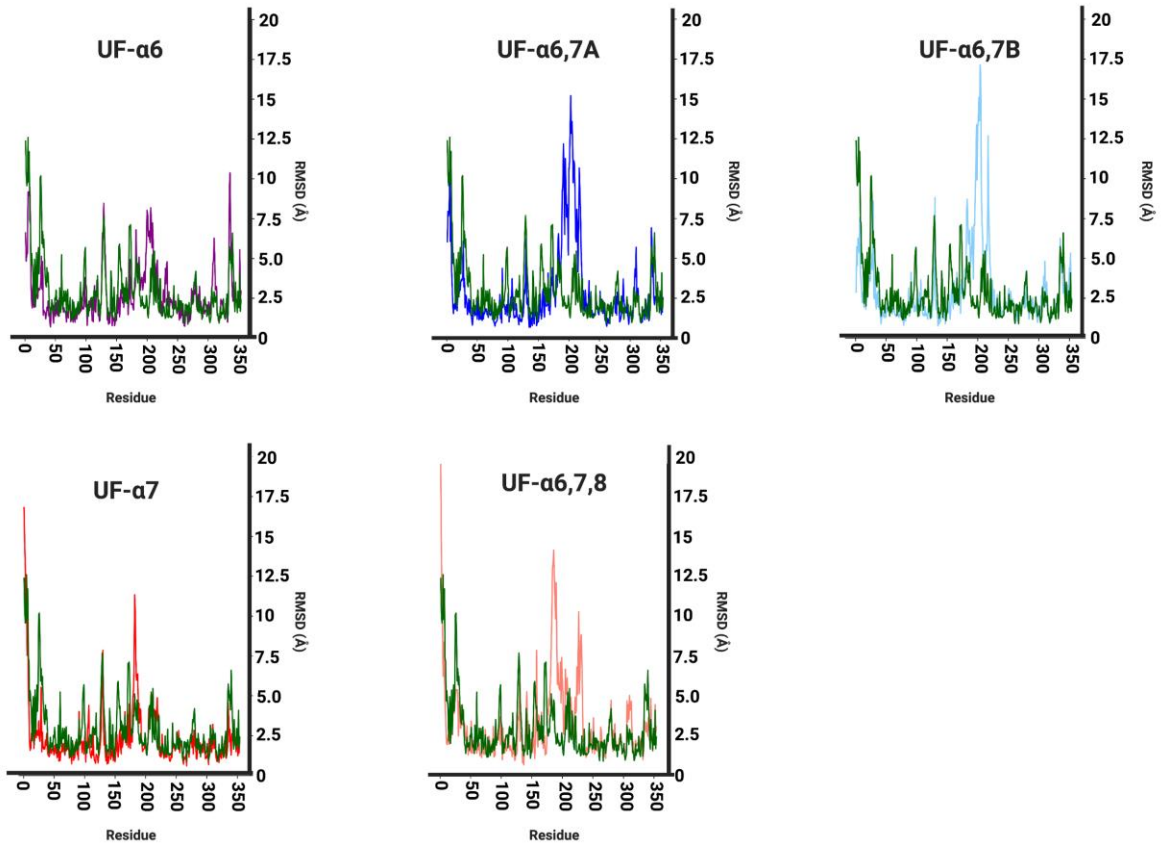
Supplemental Figure 4: Histogram of the distribution of the log fold change in weight, $\log(\Omega_f/\Omega_i)$ for each of the 300,000 reweighted frames resulting from ST of WT apo-PhuS.



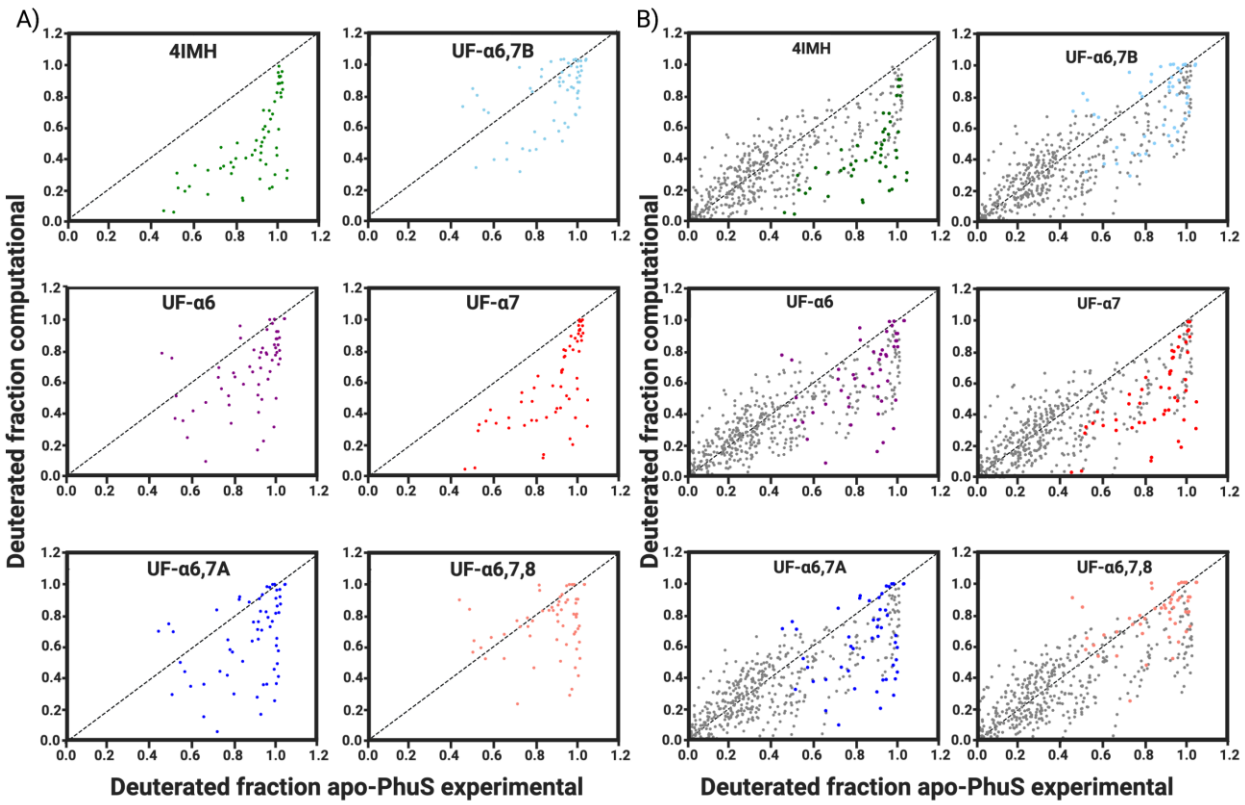
Supplemental Figure 5: Histogram of the distribution of the log fold change in weight, $\log(\Omega_f/\Omega_i)$ for each of the 300,000 reweighted frames resulting from ST of H212R apo-PhuS.



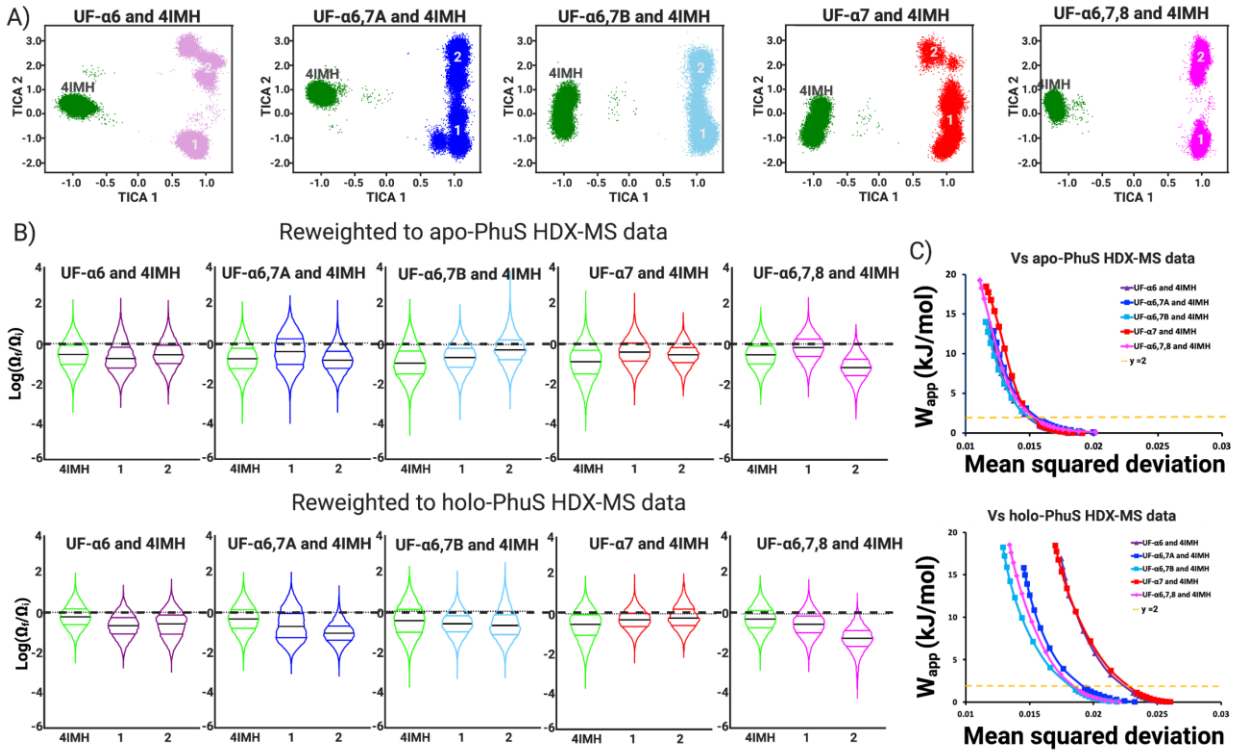
Supplemental Figure 6: A) Standard deviation in cluster average $\log(\Omega_f/\Omega_i)$ plotted for each cluster of WT apo-PhuS. B) Standard deviation in cluster average $\log(\Omega_f/\Omega_i)$ plotted for each cluster of H212R apo-PhuS.



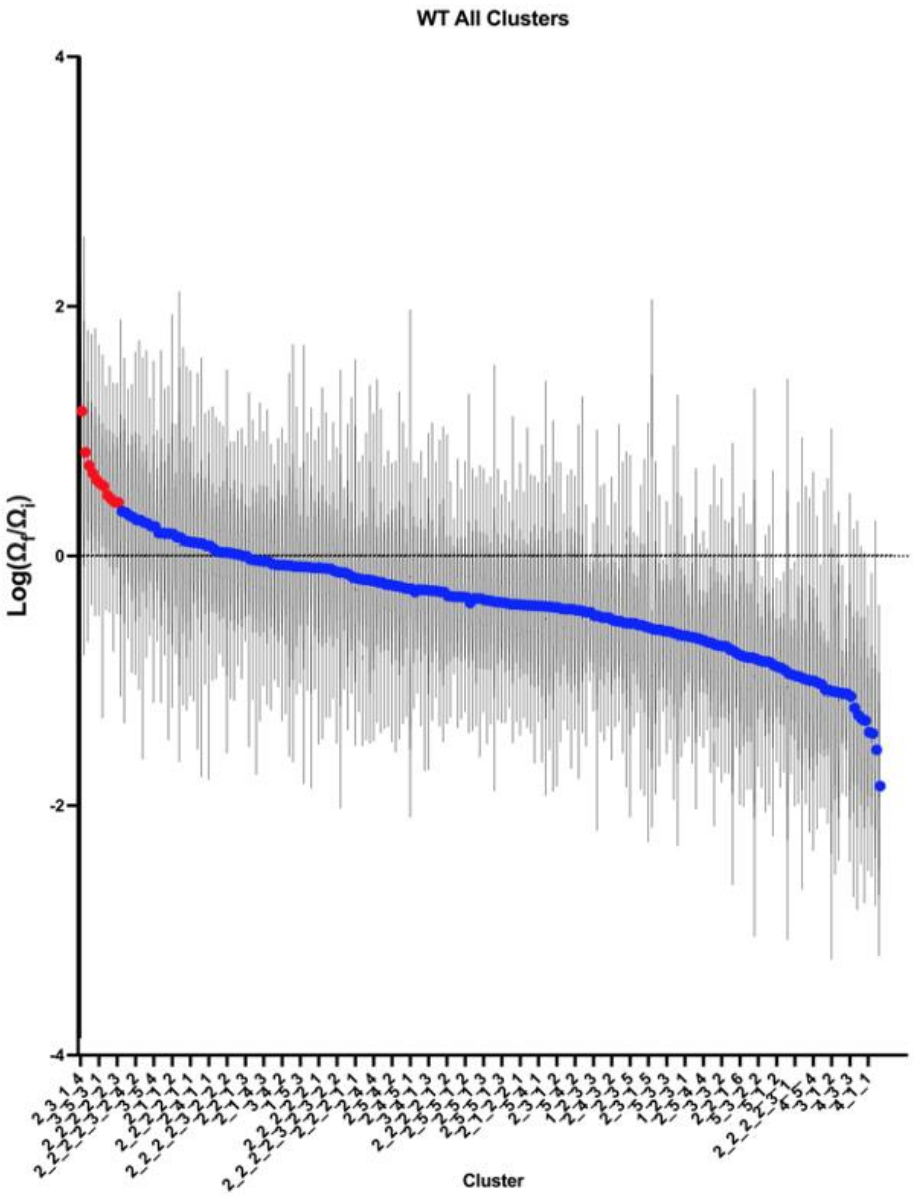
Supplemental Figure 7: Trajectory-averaged local Cα RMSD (Å) for the apo-PhuS 5 μs simulation (green) versus the trajectory-averaged local Cα RMSD (Å) of the 50 ns unfolded model simulations (UF-α6: purple, UF-α6,7A: blue: UF-α6,7B cyan, UF-α7: red, UF-α6,7,8: pink).



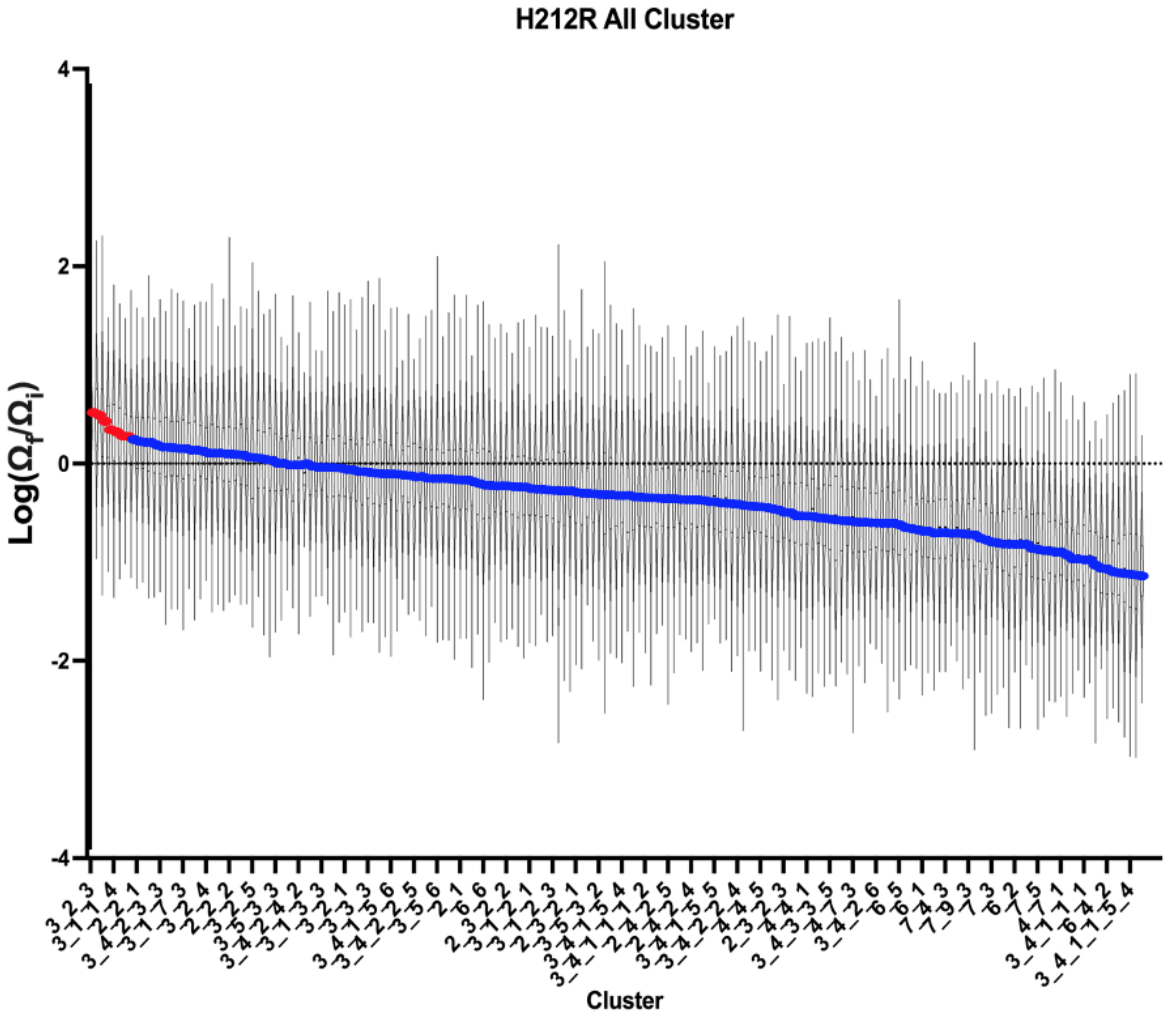
Supplemental Figure 8: Correlation between experimental and calculated HDX. A) Plots of the calculated deuterium uptake for each unfolded model and 4IMH (4IMH: green, UF-α6: purple, UF-α6,7A: blue: UF-α6,7B cyan, UF-α7: red, UF-α6,7,8: pink) vs. the experimental deuterium uptake for apo-PhuS for all deuterium exposure time points for peptides encompassing the α-6,7,8 region. B) Plots of calculated deuterium uptake for each unfolded model and 4IMH (4IMH: green, UF-α6: purple, UF-α6,7A: blue: UF-α6,7B cyan, UF-α7: red, UF-α6,7,8: pink) vs. the experimental deuterium uptake for apo-PhuS for all deuterium exposure time points for all peptides (gray). Peptides in the α-6,7,8 region are highlighted in the corresponding model color.



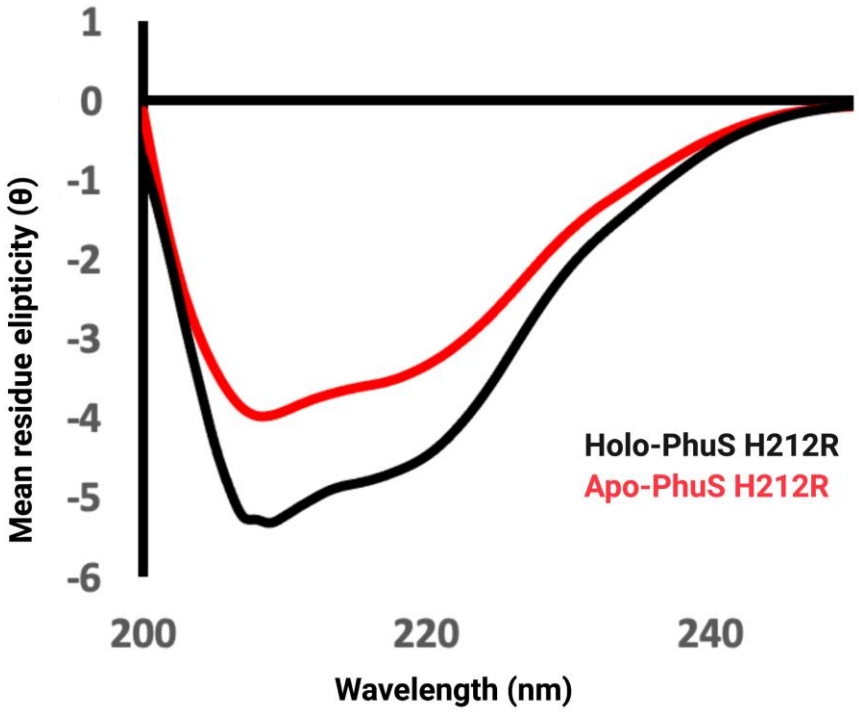
Supplemental Figure 9: TICA clustering and HDXer of mixed ensembles of apo-PhuS and locally-unfolded models. A) TICA separation of the mixed ensembles of 4IMH (green) and the locally-unfolded models (UF-α6: pink, UF-α6,7A: blue: UF-α6,7B cyan, UF-α7: red, UF-α6,7,8: magenta). B) Violin plots of HDXer weights mapped to the TICA clusters of mixed ensembles reweighted to apo-PhuS experimental data (top) and to holo-PhuS experimental data (bottom). C) Work vs MSD decision plots for the mixed ensembles reweighted to apo-PhuS experimental data (top) and holo-PhuS experimental data (bottom).



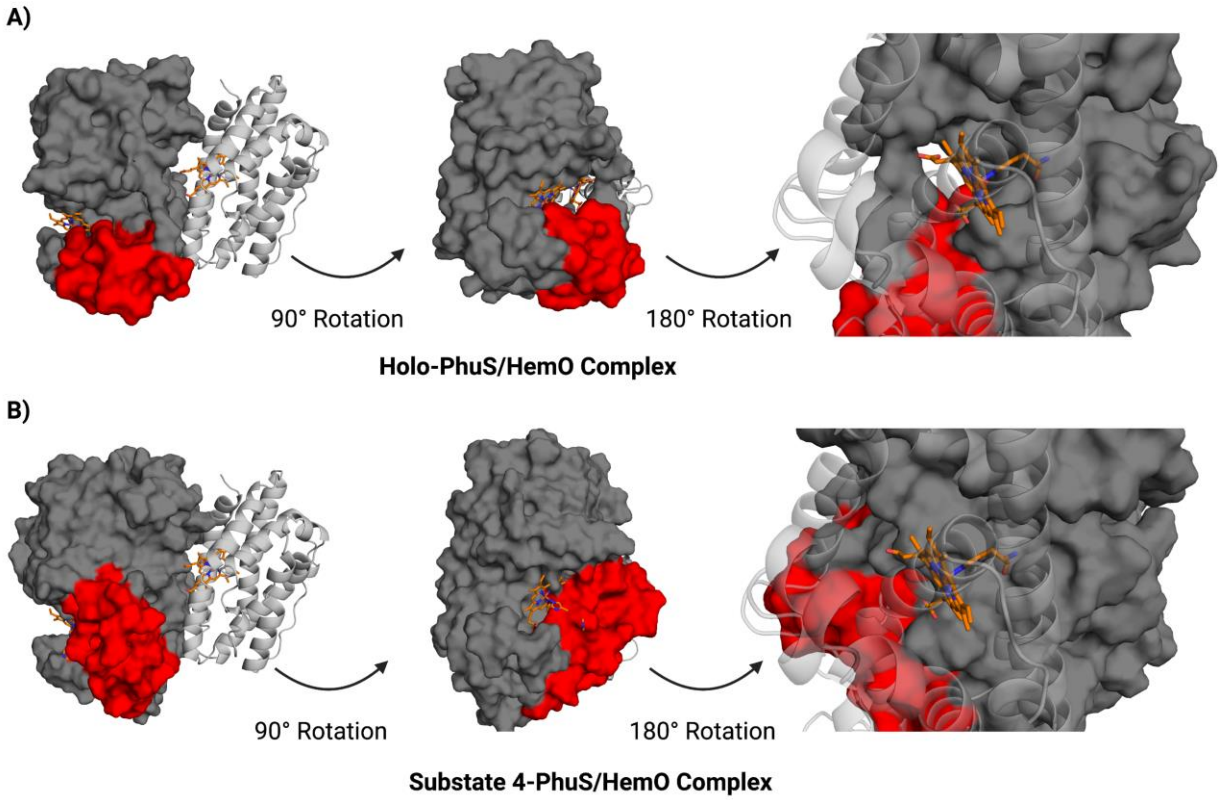
Supplemental Figure 10: Violin plots of all 221 endpoint TICA clusters of the WT apo-PhuS ST simulations with HDXer weight reassignment mapped as $\log(\Omega_f/\Omega_i)$. The average $\log(\Omega_f/\Omega_i)$ for each cluster are represented by colored dots (blue or red) with red dots representing the 11 significantly upweighted clusters.



Supplemental Figure 11: Violin plots of all 187 endpoint TICA clusters of the H212R apo-PhuS ST simulations with HDXer weight reassignments mapped as $\log(\Omega_f/\Omega_i)$. The average $\log(\Omega_f/\Omega_i)$ for each cluster are represented by colored dots (blue or red) with red dots representing the 7 significantly upweighted clusters.



Supplemental Figure 12: Experimental CD spectra (mean residue ellipticity) for apo-PhuS H212R (red) and holo-PhuS H212R (black) vs wavelength (nm).



Supplemental Figure 13: A) Holo-PhuS/HemO modeled complex with PhuS in surface representation and HemO in cartoon representation, viewing the dimer interface (left), viewed from the PhuS binding pocket (middle, rotated by 90° from initial view) and viewed from the HemO binding pocket (right, rotated by 180° from middle figure) with HemO in transparent rendering. α-6,7,8 region residues are colored in red. B) Representative frame from substate 4 aligned to holo-PhuS within the holo-PhuS/HemO modeled complex, indicating the effect of the rearrangement of the α-6,7,8 region (red) on the pathway between the two binding sites. The complex is viewed from the same perspective as in panel A.

Model name	RMS error over all peptides	RMS error for peptides in α-6,7,8 region
4imh	0.208	0.419
UF-α6	0.182	0.281
UF-α6,7A	0.191	0.308
UF-α6,7B	0.177	0.208
UF-α7	0.204	0.345
UF-α6,7,8	0.183	0.245

Supplemental Table 1: RMS error of the scatter plot of experimental apo-PhuS deuterium uptake vs. calculated deuterium uptake relative to the line of identity ($y=x$) for all peptides (left) or only peptides in the α-6,7,8 region (right) at all timepoints for each locally-unfolded model.

Cluster	Log(Ω_f/Ω_i)	Number of Frames
2_3_1_4	1.159685171	333
2_2_2_1_2	0.828684706	900
3_5_5_7	0.722032256	1266
2_2_2_2_3_2_3_3_3	0.658970357	575
3_5_1	0.605141263	1926
3_5_3_1	0.576709561	974
3_5_5_6	0.555782211	1142
3_5_5_4	0.482374827	279
3_5_5_1	0.452328138	1001
2_2_2_2_3_2_3_3_2_1	0.425696715	541
2_2_2_2_2_3	0.425121385	940
2_2_2_2_3_3	0.352027071	1117
2_2_2_2_3_2_	0.343332014	1022
3_5_5_3	0.319061236	865
2_2_2_2_3_2_3_3_2_3	0.304776502	740
2_2_2_2_3_2_4_2	0.283623325	1081
2_2_2_2_3_2_3_4	0.282589695	555
2_5_5_2	0.263253536	2198
2_2_2_2_3_2_2_	0.255366161	929
3_5_3_3_4	0.232936303	2403
3_5_4	0.23085644	6846
3_4_1_1	0.178388473	558
2_2_2_2_3_2_4	0.177578272	3585
2_2_2_2_3_2_4_1	0.176003191	692
2_1_2_3	0.173951671	1503
2_1_1_2	0.169558871	1802
2_2_2_2_3_2_3_3_1	0.143741463	682
2_3_1_3	0.141387786	1181
2_2_2_2_3_2_4_3	0.114692749	1811
2_1_1_1	0.110739791	519
2_2_2_2_1_1	0.105040578	359
2_2_2_2_3_2_3_2_1	0.102201496	1005
2_2_2_2_3_2_3_3_2_2	0.096141048	973
3_5_2	0.092359268	2544
3_5_3_3_2	0.076146278	2963
2_4_1_1	0.074570785	894
3_5_5_2	0.052272988	842
3_5_5_5	0.028855101	1339

2_2_2_2_1_3	0.023049533	827
2_5_1_2	0.021637988	701
2_2_2_2_3_2_2_2	0.020924281	856
2_5_3	0.01467007	1303
2_4_5_3	0.009117217	948
3_5_3_3_1	0.003914447	772
2_2_2_2_3_2_2_3	-0.0064953	949
2_2_1_3	-0.0077613	472
2_2_1_2	-0.0342666	1837
1_3_6	-0.0387576	2060
2_2_2_3_1	-0.0424393	1745
2_2_1_7	-0.0475926	1175
2_1_4_3	-0.0499468	802
3_3_1_3	-0.0522987	674
3_5_5_8	-0.0702254	598
1_3_7	-0.0777378	2032
2_2_2_2_3_2_3_1	-0.0791489	613
3_4_2	-0.0793532	5560
2_1_2_1_3	-0.0829196	446
2_2_2_3_3	-0.084026	1274
2_2_2_2_2_2	-0.0915298	486
1_2_1	-0.092003	1745
1_5_3	-0.0924104	1591
2_2_2_3_2	-0.0944764	1947
2_1_2_1_2	-0.0955465	491
2_4_3_4	-0.1012385	2140
2_5_2_3	-0.101383	550
2_2_2_2_2_1	-0.1018841	558
2_1_2_1_1	-0.1061136	1063
2_1_4_2	-0.1080687	952
2_1_1_4	-0.1098821	368
3_5_3_3_3	-0.1265235	2714
2_2_2_2_3_2_3_2_2	-0.1366772	435
2_5_5_3	-0.1406756	1719
1_5_1	-0.1424804	279
2_3_2_2	-0.1577683	2589
2_2_2_2_3_2_3_2_3	-0.1772616	576
2_2_2_1_1	-0.1852846	787
3_3_1_1	-0.1911668	2789

2_2_2_2_4_4	-0.1974706	1908
2_2_1_1	-0.1981641	2763
2_4_1_6	-0.2022959	779
2_4_4	-0.2085109	1945
2_1_1_3	-0.2181652	749
3_4_5	-0.2189228	6052
1_4_5	-0.2355828	2697
3_3_3_2	-0.2374482	1084
2_5_4_2	-0.2425773	605
2_1_2_2_2	-0.2487356	807
1_1_2	-0.2533086	601
2_2_1_4	-0.2637666	761
3_4_4	-0.2680286	9252
2_4_5_1	-0.2700271	500
1_1_1	-0.277364	211
1_5_4	-0.2790205	442
3_3_3	-0.2794899	1938
2_5_2_2	-0.2836817	1740
3_4_1_3	-0.2865318	1402
3_1_3	-0.289236	3386
1_4_1	-0.2949457	2337
1_3_3	-0.2967328	4668
2_5_1_1	-0.2980379	2080
2_2_2_2_1_2	-0.3283641	1555
1_2_2	-0.3310032	983
1_5_2	-0.3343294	66
3_4_3	-0.3347847	4944
2_1_4_1	-0.3353318	511
2_5_5_1_2	-0.3385499	1915
3_5_3_2	-0.3447794	10256
1_3_1	-0.3478895	3214
1_2_4_2	-0.3482753	1665
2_4_3_3	-0.348995	1931
2_5_1_3	-0.361528	2976
5_4_5	-0.363708	893
2_5_5_1_1	-0.3702563	1104
2_3_2_3	-0.3767906	925
3_3_3_1	-0.3788784	449
2_5_5_1_3	-0.3823608	1188

1_2_4_1	-0.3856093	2947
1_3_4	-0.3943613	617
1_3_5	-0.395284	1393
2_4_5_4	-0.3962294	1227
2_1_2_2_1	-0.3997725	945
1_4_2	-0.402876	1045
2_5_4_1	-0.4035032	866
2_2_2_2_4_1	-0.4050916	1386
5_5_5	-0.4071344	411
5_4_1	-0.4080798	455
2_4_1_3	-0.4102279	853
2_1_3	-0.414754	522
2_5_1_4	-0.415422	433
2_2_2_2_3_2_3_5	-0.4202175	1347
2_3_1_2	-0.4209303	1122
2_2_1_5	-0.4324502	376
1_2_4_3	-0.4340968	1282
2_3_1_1	-0.4349568	827
3_3_1_2	-0.4349842	2483
5_4_2	-0.4437877	580
2_2_3_1	-0.4444573	863
2_5_4_3	-0.4517748	1289
3_4_1_2	-0.4597299	798
1_3_2	-0.4611876	38
1_2_3_3	-0.487132	417
3_1_2_2	-0.4897209	1236
1_4_3	-0.5007569	2283
2_4_3_1	-0.5021891	2634
5_3_2_5	-0.5038424	1121
2_4_3_2	-0.5234033	487
1_2_3_2	-0.5310625	1464
3_5_3_3_5	-0.5316096	2088
2_2_3_4	-0.5437825	3820
2_4_1_4	-0.5484894	677
2_3_5	-0.5486208	2153
5_3_1	-0.549904	394
1_2_3_4	-0.5607341	723
2_4_5_2	-0.5649412	633
2_4_2	-0.5762641	1853

2_3_1_5	-0.5832398	1628
2_4_1_5	-0.5958404	268
5_5_1	-0.5992203	627
3_1_4	-0.5999336	2834
1_4_4	-0.6119515	106
5_3_3	-0.6130941	2616
5_2_4	-0.6183952	413
4_5_1	-0.6307067	1183
2_2_3_7	-0.6377364	2342
5_3_2_6	-0.6473835	1061
1_2_3_1	-0.6478129	465
5_3_4	-0.6570648	950
5_3_2_4	-0.6619375	1142
2_2_3_5	-0.6672568	577
5_2_6	-0.6771615	1440
2_5_4_4	-0.6884291	1741
5_3_2_1	-0.7008723	836
4_5_2	-0.7057492	803
2_4_1_2	-0.7213624	378
5_5_3	-0.7279607	148
3_3_2	-0.7323867	2956
2_5_1_5	-0.7351061	1131
5_5_2	-0.7597704	451
4_5_5	-0.7704541	733
5_2_3	-0.793405	1239
2_2_1_6	-0.8077498	442
2_5_2_1	-0.8176149	934
2_2_2_2_3_1_3	-0.8237756	350
2_2_2_2_3_1_2	-0.8237756	350
3_2	-0.8336905	1143
5_3_2_2	-0.8472769	317
5_1_4	-0.8543315	1405
4_4_3	-0.8568503	439
2_2_2_2_4_2	-0.8638472	2427
5_6_1	-0.8874862	1370
5_1_2	-0.8981713	2889
5_2_2	-0.9081168	588
5_2_5	-0.9250898	898
2_3_4	-0.9537962	1031

2_2_2_2_4_3	-0.9621251	260
2_2_2_2_3_1_2	-0.9720378	593
5_6_2	-0.976235	494
3_1_2_1	-0.9930389	1970
5_1_3	-1.0018327	1094
5_5_4	-1.0093478	598
4_5_4	-1.0128579	836
5_6_3	-1.0281646	1429
5_4_3	-1.0405811	715
2_3_2_1	-1.0821042	1159
2_2_2_1_3	-1.0857006	755
3_1_2	-1.0967875	7455
4_5_3	-1.1003722	1455
4_4_4	-1.1084928	273
5_3_2_3	-1.1118242	363
5_2_1	-1.1157271	1525
4_3_3	-1.1359068	328
4_3_2	-1.2288698	411
4_1_2	-1.2865568	570
4_1_3	-1.317961	232
4_4_1	-1.3324284	648
4_1_1	-1.4244904	251
4_4_2	-1.4356977	506
4_2	-1.5671043	702
4_3_1	-1.857777	511

Supplemental Table 2: The results of iterative TICA clustering of the wild type apo-PhuS ST simulations ranked by $\log(\Omega_f/\Omega_i)$ following HDXer ensemble reweighting with the 11 significantly upweighted clusters highlighted in yellow. Clusters are named so that the first number is the parent cluster and the last number is the final sub cluster. Weights reassigned by HDXer are used to calculate $\log(\Omega_f/\Omega_i)$ and are averaged over all the frames of a given cluster. The number of frames that comprise each cluster is listed.

Cluster	Log(Ω_f/Ω_i)	Number of Frames
3_2_3	0.49807073	608
2_2_3	0.472860251	1308
2_2_4	0.413082691	910
3_4_2_2_1	0.330872922	509
3_1_1_4	0.310030939	2081
3_4_2_3_4	0.271193672	933
3_1_1_5	0.266516819	1731
3_6_1	0.236449334	4501
2_2_1	0.216716694	860
3_6_2	0.207983887	1118
3_4_2_3_2	0.207141341	1260
3_4_2_2_2	0.180919524	956
3_4_2_3_3	0.163398618	827
3_1_7_2	0.161408648	1424
3_6_4	0.154579639	1775
3_6_3_3	0.148749265	1744
3_1_7_3	0.146892755	1151
2_2_6	0.133448491	500
3_4_2_4_1	0.132715101	1619
2_2_8	0.122732095	1506
3_2_4	0.10367551	1450
3_1_1_3	0.103485418	425
3_1_7_1	0.102351675	690
3_4_2_3_1	0.096880772	1692
2_2_2	0.093804023	1617
3_1_3_4	0.088185887	870
3_4_2_3_5	0.083296565	1170
3_1_5_3	0.063828727	1739
3_2_5	0.05841897	861
3_6_3_1	0.050437836	737
2_2_5	0.039794885	1557
3_1_8_3	0.034610694	1526
3_5_3_3	0.004816866	4736
3_4_2_4_3	0.002121369	491
3_1_2_6	-0.007166977	177
3_2_2	-0.010684868	853
3_4_2_4_2	-0.012374399	90

2_2_7	-0.012494348	16
3_1_1_2	-0.025810114	1472
3_1_2_7	-0.035292873	1256
3_1_3_3	-0.036658668	1389
3_1_5_1	-0.037478289	768
3_1_8_2	-0.038631354	1727
3_4_2_2_3	-0.04237942	871
3_2_1	-0.054477818	839
3_5_3_4	-0.058213049	5165
3_4_1_1_1_5	-0.076809011	582
3_4_1_6	-0.077739738	1443
3_3_3	-0.085978894	2452
3_1_2_5	-0.091543582	1132
3_4_2_3_6	-0.096842251	1638
3_1_5_2	-0.097842583	2109
3_4_1_5_6	-0.102633046	615
3_4_1_5_3	-0.107123278	212
3_1_8_1	-0.112868552	789
3_4_1_1_5_2_2_4	-0.11703231	534
3_4_2_2_5	-0.120555647	64
3_1_1_1	-0.125621361	2216
3_5_1	-0.140820683	956
2_4_2_1	-0.14338153	498
3_5_6	-0.143655494	2686
3_4_1_1_1_4	-0.144739821	659
3_4_1_4	-0.148573546	3884
3_4_2_5	-0.155771241	2636
2_1	-0.158174509	1670
3_4_1_1_1_3	-0.160938024	855
2_4_2_2	-0.176517375	290
3_1_3_2	-0.194169132	1480
6_6	-0.208541297	1634
3_3_5	-0.21041343	1465
3_1_3_1	-0.216279695	870
3_4_1_3	-0.216722665	2065
2_3_2_2	-0.217741668	1257
3_4_3_3_3	-0.224308587	252
3_5_2_4	-0.226485166	1736
3_5_4	-0.227650709	4612

3_1_2_1	-0.245069448	754
2_4_1	-0.248642723	1694
3_4_1_5_2	-0.250279714	450
3_4_1_1_5_2_2_1	-0.256902955	907
3_1_2_3	-0.258803873	789
3_4_1_1_5_1	-0.263818909	1765
3_3_4	-0.265024114	696
3_3_2	-0.265716879	1557
2_3_2_1	-0.280180354	645
3_4_1_1_5_2_2_2	-0.287551183	1805
3_4_3_3_4	-0.289713334	349
3_5_2_1	-0.293471951	1228
3_5_3_2	-0.299380085	1589
3_4_3_1	-0.302832301	8474
3_4_3_2_3	-0.303408531	1068
3_5_3_1	-0.309877529	2437
3_4_1_5_4	-0.310411116	2370
4_5	-0.317435085	38
3_4_3_2_1	-0.323098297	1223
3_4_3_2_2	-0.323276291	1200
3_4_1_1_1_2	-0.332233792	824
3_4_1_1_5_2_4	-0.33396777	392
3_3_1	-0.335297225	2725
3_4_2_1	-0.339790206	918
2_4_2_5	-0.340088883	381
2_4_3	-0.344030246	289
3_1_2_4	-0.346938089	643
3_1_2_2	-0.348864414	1050
2_4_2_4	-0.348944836	435
2_3_1	-0.353872712	606
3_5_5	-0.357904136	8110
2_3_4	-0.366410936	169
3_4_1_5_5	-0.372238863	245
3_4_1_1_5_2_2_3	-0.38136456	397
2_4_2_3	-0.389141735	389
3_4_1_1_1_1	-0.391314385	1244
3_4_2_2_4	-0.396614016	621
3_4_3_3_1	-0.408621065	1574
3_4_1_1_1_6	-0.411817324	672

3_4_1_1_6	-0.418441904	1650
4_4_5	-0.420833345	1956
4_4_6	-0.426406809	321
2_3_3	-0.440385602	5924
1	-0.451943098	33485
2_3_2_3	-0.4745489	762
3_6_3_2	-0.477580061	2166
3_4_1_5_1	-0.506476834	1243
7_9_1	-0.508221244	631
4_4_1	-0.510851413	1774
3_4_3_3_2	-0.514355331	331
3_4_1_2	-0.527064037	11674
3_4_1_1_5_2_1	-0.533090984	1682
3_4_3_3_5	-0.54094937	1220
4_6	-0.54950379	1506
3_4_1_1_5_2_3	-0.553754386	1153
2_3_2_4	-0.557475801	1385
4_7_3	-0.565141908	3573
6_4_1	-0.571939696	680
3_5_2_2	-0.572051286	2134
4_4_3	-0.576220944	616
3_4_2_6	-0.577199471	462
3_5_2_3	-0.579537651	683
3_4_1_1_5_3	-0.579560346	889
7_9_4	-0.580737211	497
6_5	-0.597405836	448
4_4_7	-0.623585612	3707
4_4_4	-0.63200861	1028
4_4_8	-0.644112762	1713
6_1	-0.65605816	5807
4_4_2	-0.658088036	553
7_9_2	-0.671873297	365
7_4_4	-0.672962628	466
7_4_3	-0.672962628	466
4_1_2	-0.680308267	1338
7_8	-0.681981762	239
6_4_3	-0.683880725	507
7_9_3	-0.690399771	1164
3_4_3_3_6	-0.695362055	1854

4_3	-0.726242142	2650
6_4	-0.743527733	1944
7_3	-0.764027385	572
7_1	-0.772557573	3294
7_2	-0.782634763	334
7_4_2	-0.785363813	617
6_2	-0.785539983	706
7_6	-0.787230643	538
4_1_1	-0.787512281	762
7_4_1	-0.82481582	545
7_5	-0.835327443	868
4_1_4	-0.850575196	736
4_8	-0.861562342	1205
4_7_1	-0.862381542	765
4_9_2	-0.888350156	539
4_9_3	-0.92581328	1062
4_1_3	-0.931040438	1071
3_4_1_1_1	-0.935215004	633
4_9_4	-0.939680691	598
3_4_1_1_2	-0.989277132	790
4_9_1	-1.02013386	443
6_4_2	-1.026938744	498
4_2	-1.052067352	2939
4_7_2	-1.063832626	3940
7_7	-1.06730495	246
3_4_1_1_5_4	-1.077814072	1106
3_4_1_1_3	-1.085304573	963
4_10	-1.09379047	1752

Supplemental Table 3: The results of iterative TICA clustering of the H212R apo PhuS ST simulations ranked by $\log(\Omega_f/\Omega_i)$ following HDXer ensemble reweighting with the 7 significantly upweighted clusters highlighted in yellow. Clusters are named so that the first number is the parent cluster and the last number is the final sub cluster. Weights reassigned by HDXer are used to calculate $\log(\Omega_f/\Omega_i)$ and are averaged over all the frames of a given cluster. The number of frames that comprise each cluster is listed.

Supporting references:

1. Bradshaw, R. T., F. Marinelli, J. D. Faraldo-Gomez, and L. R. Forrest. 2020. Interpretation of HDX Data by Maximum-Entropy Reweighting of Simulated Structural Ensembles. Biophys J 118(7):1649-1664, 10.1016/j.bpj.2020.02.005.