Supplemental Information

Large-scale protein conformational transitions revealed by weighted ensemble simulations and EPR

Shramana Palit[‡], Darian T. Yang[‡], Xiaowei Bogetti, Anthony T. Bogetti, Olivia R. Wood,

Sunil Saxena, and Lillian T. Chong**

Department of Chemistry

University of Pittsburgh

219 Parkman Ave. Pittsburgh PA 15260

*Email: ltchong@pitt.edu; sksaxena@pitt.edu

		2LAO			1LST				
Sites A	Sites B	CA-CA (Å)	Cu-Cu (Å)	Width (Å)	CA-CA (Å)	Cu-Cu (Å)	Width (Å)	**ΔCA-CA (Å)**	**ΔCu-Cu (Å)**
124,128	5,7	39.2	39.5	5.0	33.2	33.5	5.0	6.0	6.0
	32,36	36.5	37.5	5.0	34.5	35.3	4.5	2.0	2.2
	35,39	38.9	41.0	4.0	35.7	38.0	4.0	3.2	3.0
	56,60	28.4	23.5	4.0	15.3	16.5	4.0	13.1	7.0
	204,208	36.3	35.0	6.0	33.0	31.0	5.0	3.3	4.0
	205,209	38.6	39.0	5.0	36.8	35.0	4.5	1.8	4.0
	208,212	34.8	38.0	5.0	35.2	34.5	5.0	0.4	3.5
	212,216	34.8	39.0	5.0	31.9	37.0	4.5	2.9	2.0
	215,219	35.5	37.0	5.0	36.5	35.5	5.0	1.0	1.5
	224,228	35.3	37.5	5.0	36.1	37.5	5.0	0.8	0.0
136,138	5,7	46.9	52.0	4.0	37.1	39.0	4.0	9.8	13.0
	32,36	41.3	45.0	4.0	36.3	40.0	4.5	5.0	5.0
	35,39	45.5	50.0	4.0	39.5	42.5	4.0	6.0	7.5
	56,60	37.2	38.5	4.5	26.4	26.5	5.0	10.8	12.0
	204,208	44.7	51.0	6.0	39.7	45.5	5.0	5.1	5.5
	205,209	46.6	54.0	4.0	41.2	47.0	4.0	5.4	7.0
	208,212	43.9	52.0	4.0	42.7	47.5	5.0	1.3	4.5
	212,216	43.9	51.5	5.0	44.0	48.5	5.0	0.1	3.0
	215,219	41.3	48.5	5.0	42.8	47.0	5.5	1.5	1.5
	224,228	35.2	42.0	5.0	38.9	42.0	5.0	3.7	0.0
144,148	5,7	42.2	45.0	4.0	30.0	30.0	6.0	12.2	15.0
	32,36	35.4	37.5	5.0	26.6	27.5	4.5	8.8	10.0
	35,39	39.9	42.0	5.0	30.2	29.5	4.0	9.7	12.5
	56,60	36.2	37.0	4.0	25.9	27.0	5.0	10.3	10.0
	204,208	44.9	51.0	5.0	36.9	41.5	5.0	8.0	9.5
	205,209	45.6	51.0	4.0	37.0	40.0	6.0	8.6	11.0
	208,212	43.9	51.5	5.0	36.6	41.5	5.0	7.3	10.0
	212,216	43.0	49.0	5.0	36.5	40.5	5.0	6.5	8.5
	215,219	40.9	47.0	5.0	36.1	40.0	4.0	4.9	7.0
	224,228	31.5	36.5	6.0	27.9	31.5	5.0	3.7	5.0

Table S1. List of labeling sites and the difference in Cu(II)-Cu(II) distance between the two states obtained from MMM.(1,2) Column 1 lists potential labeling sites at one domain of the protein (sites A), and column 2 lists the predicted labeling sites on the other domain (sites B). The third to fifth columns provide the Cα -Cα distances between the sites pairs computed based on the crystal structure of the open state, the predicted mean Cu(II)- Cu(II) distances and distribution width, respectively. Similarly, the sixth to eighth columns provide distance information for the closed state. Last two columns of Table S1 list the change in Cα -Cα and Cu(II)- Cu(II) distances between the two protein states. The site pairs that are colored in dark green are of our major focus, because these sites show the largest distance difference (>7 Å) between the two protein states solved experimentally. In total, there are 15 site pairs that are potentially for describing the conformational changes.

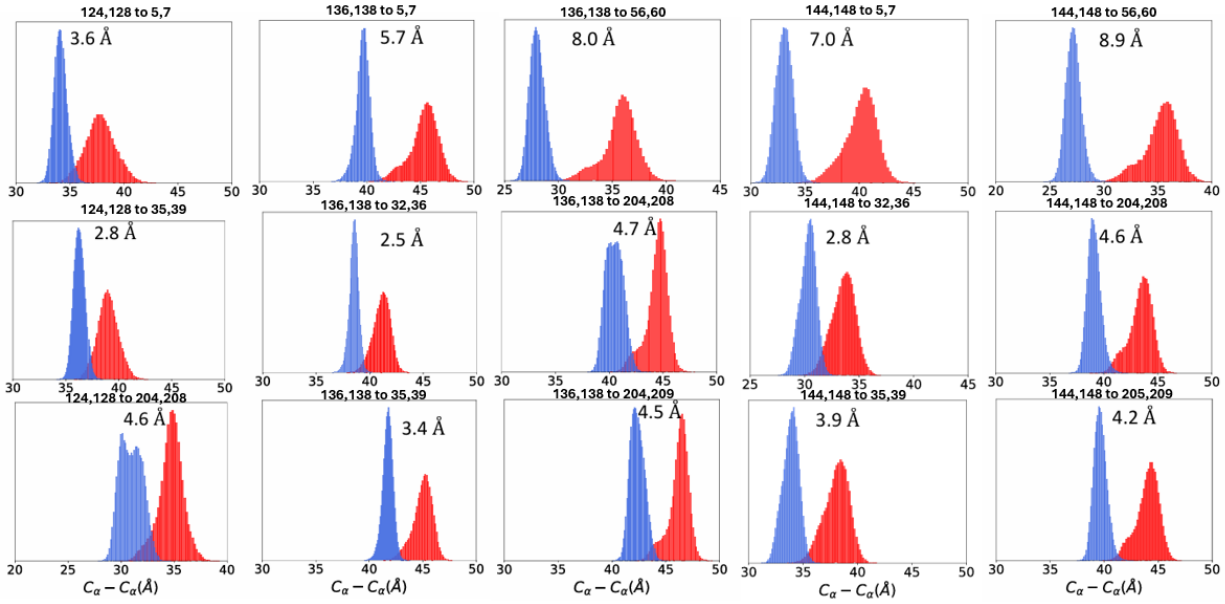

Figure S1. Distance distributions between Cα atoms of 15 different site pairs obtained from 1 µs standard molecular dynamic simulation. The distribution for the ligand free state is shown in red while the distribution for the ligand bound state is shown in blue. Site pair 56,60 to 144,148 was chosen as the protein construct as it showed the largest change in Cα- Cα distance of ca. 9 Å in the standard MD simulation.

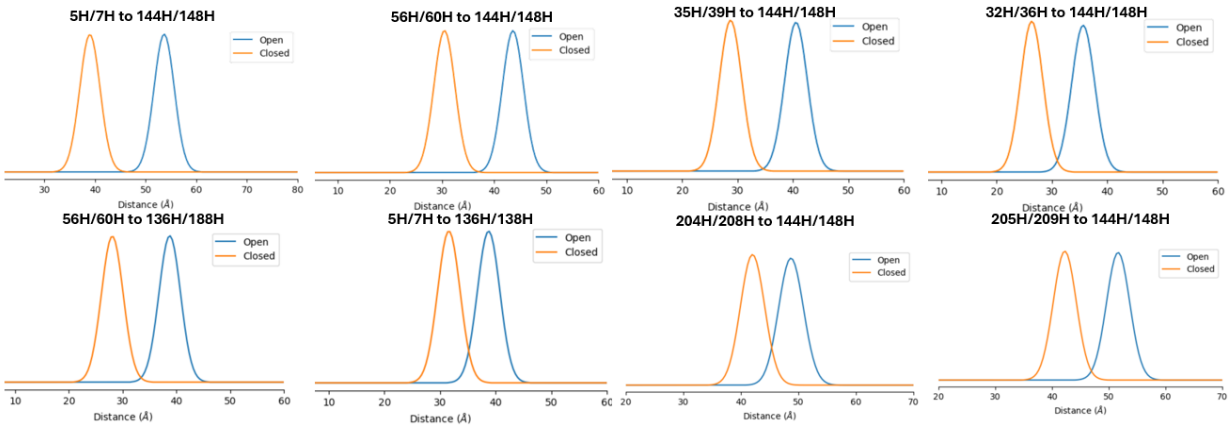

Figure S2. Distance distribution for the obtained from ChiLife (3,4), a software which can be used to predict distance distributions by modeling spin labels based on the MD rotamer libraries. Sites 56H/60H/144H/148H shows a difference of >10 Å between the ligand bound and ligand free states making it ideal for EPR experiments to capture the conformational change.

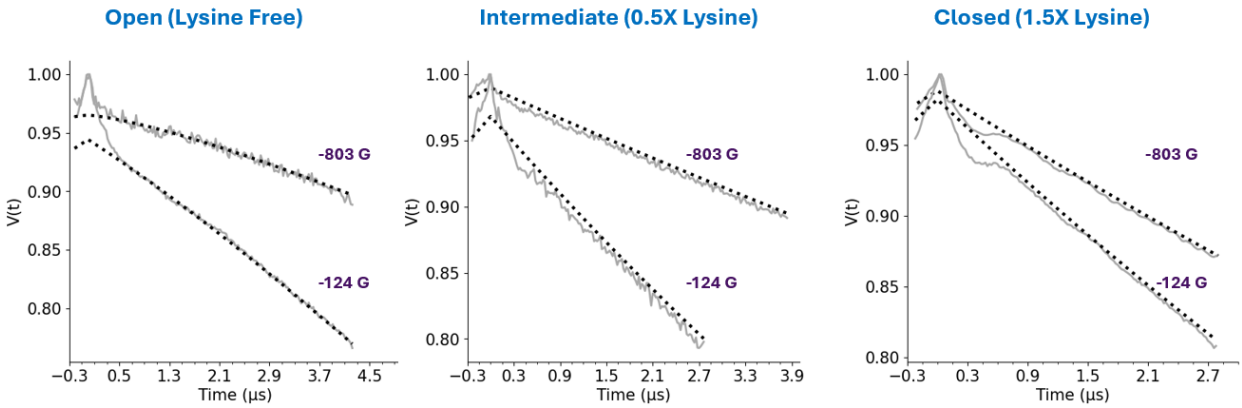

Figure S3A. Primary data at individual fields for the three LAOBP samples- Lysine free, with 0.5X Lysine and with 1.5X Lysine. The data was collected with the pump pulse at -124 G and -803 G from the position of maximum intensity of the FS-ESE of the Cu(II) spectrum to mitigate orientational selectivity effects (5).

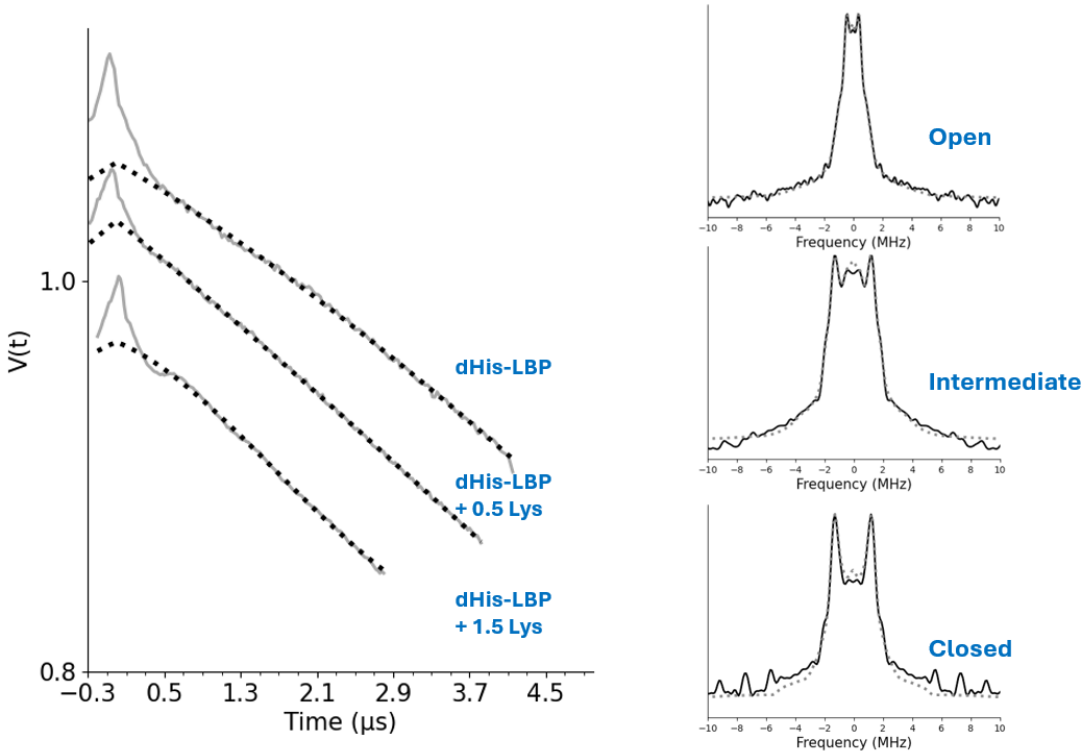

Figure S3B. Primary DEER data (left). Pake patterns or frequency domain spectra for the time domain data presented in the main text for the ligand free (no lysine), intermediate (0.5X Lysine) and ligand bound (1.5X Lysine) states of LAOBP (right).

Sample	dHis-LBP	dHis-LBP +0.5X Lys	dHis-LBP +1.5X Lys
$(\pi/2)\upsilon_A$ (ns)	8	8	8
$(\pi)\upsilon_A$ (ns)	16	16	16
$(\pi)\ \upsilon_B$	200 ns 200 MHz chirp	150 ns 200 MHz chirp	150 ns 200 MHz chirp
τ	400	400	400
T	4500	4100	3000
Δt	24	20	20
SRT (ms)	1.5	1.5	1.5
Shots per point	20	20	20
Modulation depth %	3.7	2	3.9
SNR	40	33	57

Table S2. DEER acquisition parameters, modulation depths and SNR for Q Band DEER in the main manuscript.

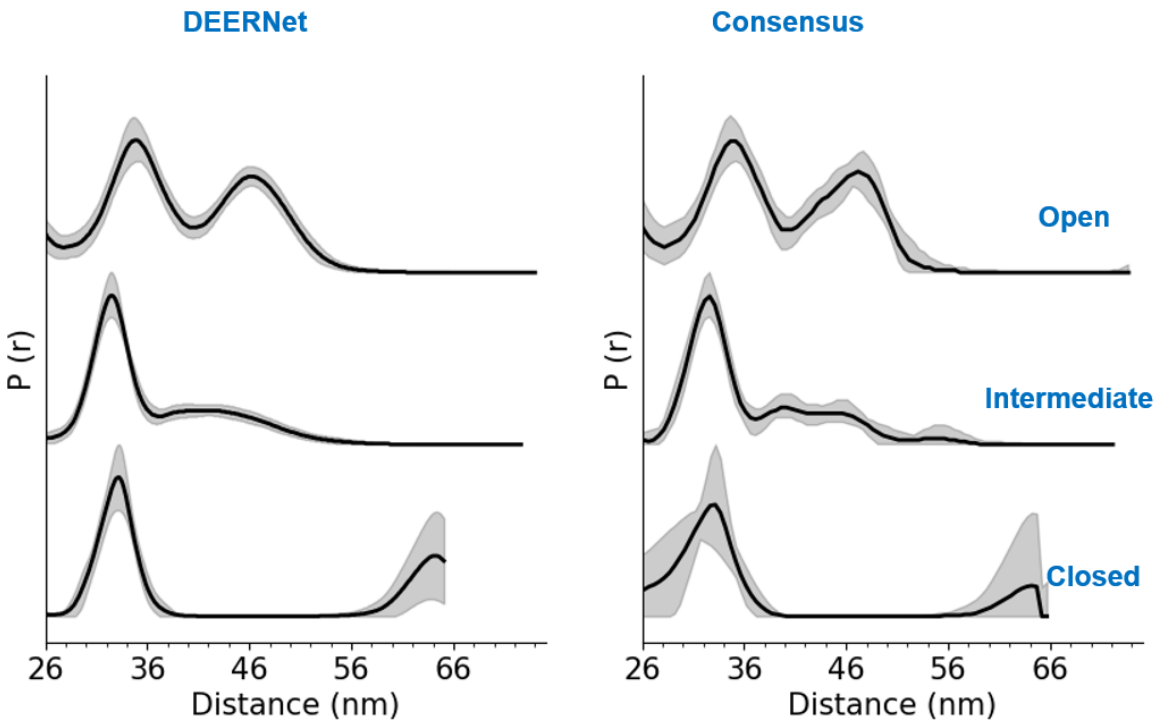

Figure S4. Distance distributions obtained from DEERNet (5) and CDA (6). The distributions are consistent with the distributions from DEERAnalysis.

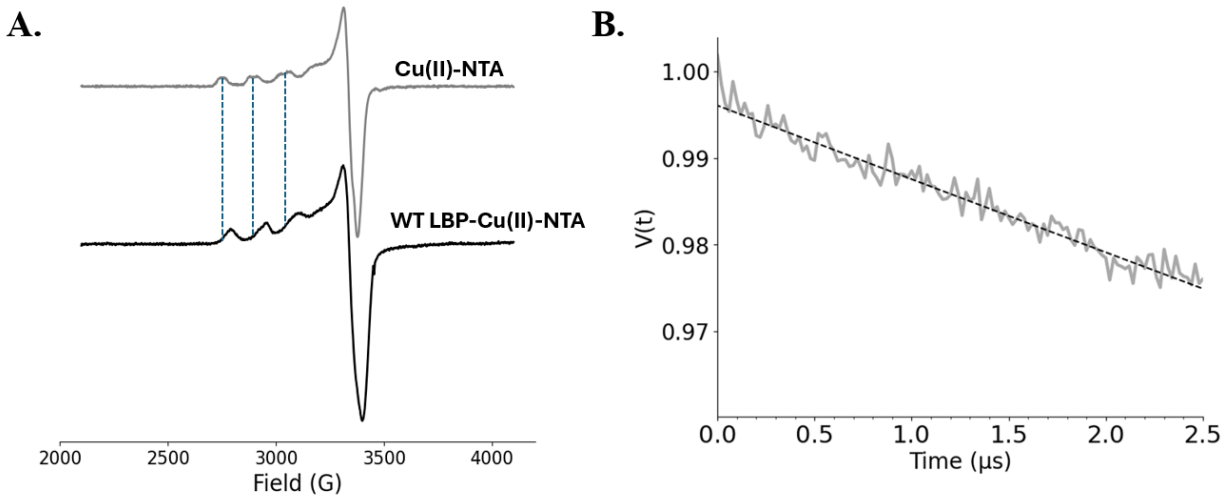

Figure S5. A) CW-EPR spectra of Cu(II)-NTA in grey and wildtype LAOBP bound to Cu(II)-NTA in black. There are distinct changes between the two spectra, indicating that Cu(II)-NTA is binding non-specifically to WT LAOBP. This is not unexpected as the protein has a native histidine site and Cu(II)-NTA can bind to single His sites but prefers a dHis coordination (7). B) Time domain signal of WT LAOBP labeled with Cu(II)-NTA. There are no distinct modulations suggesting that non-specific coordination does not contribute to any of the distances obtained on the mutated protein.

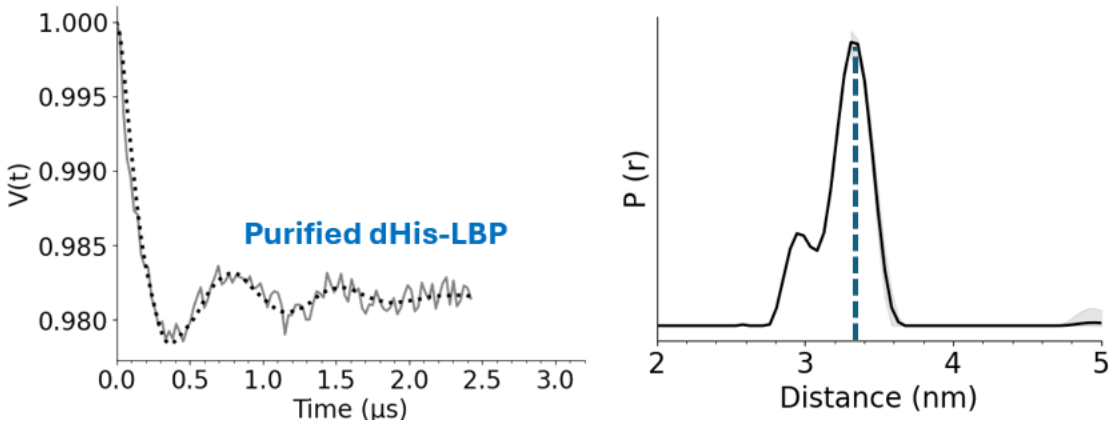

Figure S6. Background-corrected time domain data on purified LAOBP and the corresponding distribution. LAOBP adopts the closed conformation during the purification process by binding to free amino acids. Therefore, the distribution of purified LAOBP without any added lysine is consistent with the closed conformation of the protein. Hence, an unfolding and refolding technique had to be used to obtain data on the open conformation (lysine free). Details are provided in the Methods section in the main text.

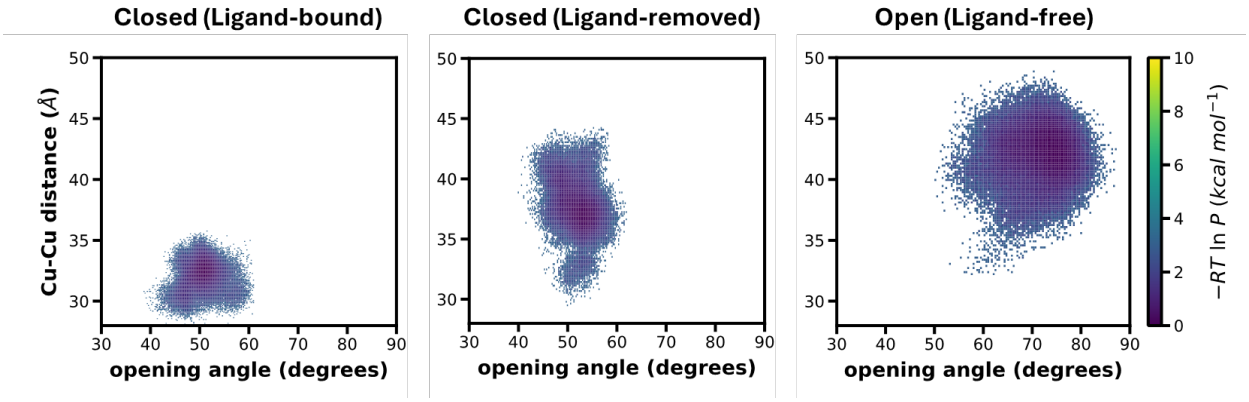

Figure S7. Standard MD simulation data (1 μs) data on LAOBP protein initialized from the closed (ligand bound), closed (ligand removed) and open (ligand free) states. Probability distributions are plotted as a function of Cu(II)-Cu(II) distance and opening angle. The simulation was unable to capture any conformational transition between the states as both distance and the opening angle do not change significantly over the course of the simulation.

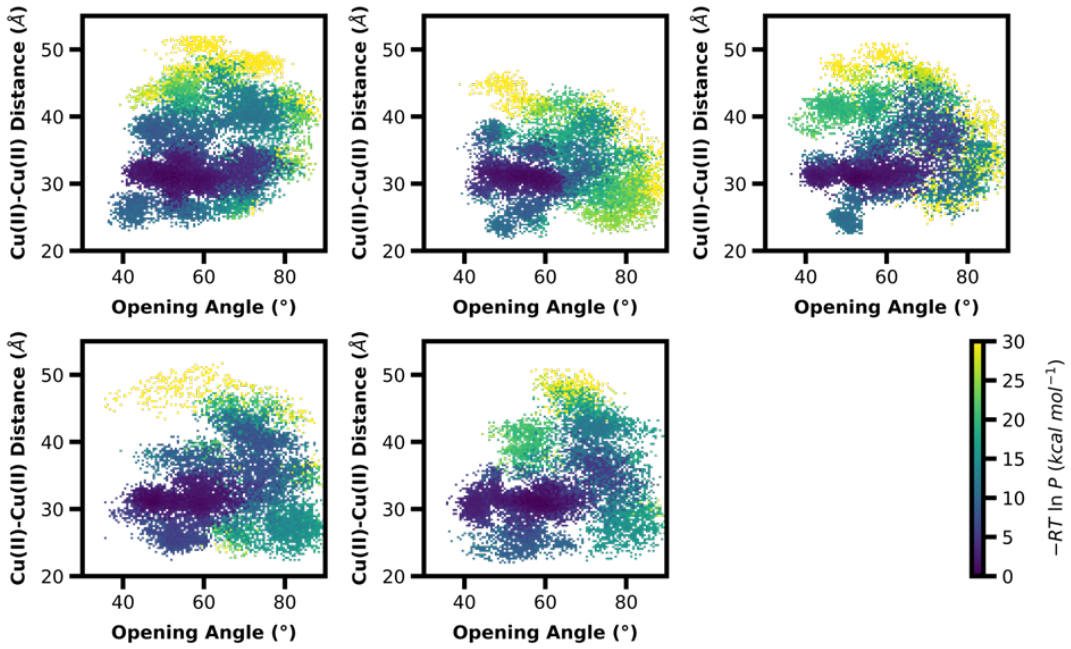

Figure S8. Data from five independent WE simulations of the LAOBP protein sampling the closed to open state transition. Probability distributions are shown as a function of the Cu(II)–Cu(II) distance and the opening angle. We were able to consistently sample the open to closed transition of LAOBP in all five simulations with the Cu(II)-Cu(II) distance changing from 32 Å to ~45 Å.

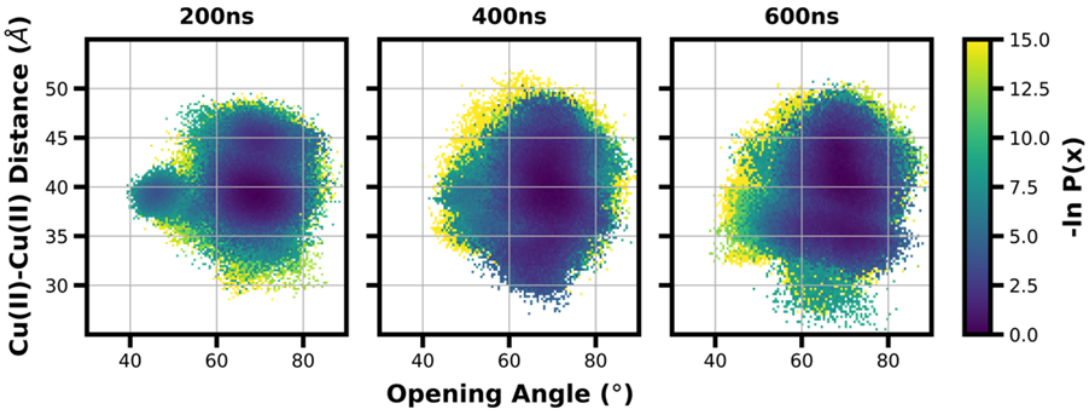

Figure S9. Probability distributions of 25 standard MD simulations (15 μs of aggregate simulation time) initialized from the ligand-free open state sampled during our WE simulations. The probability distribution evolution of the Cu(II)–Cu(II) distance and opening angle over time is plotted for all 25 simulations at 200, 400, and 600 ns of individual simulation time.

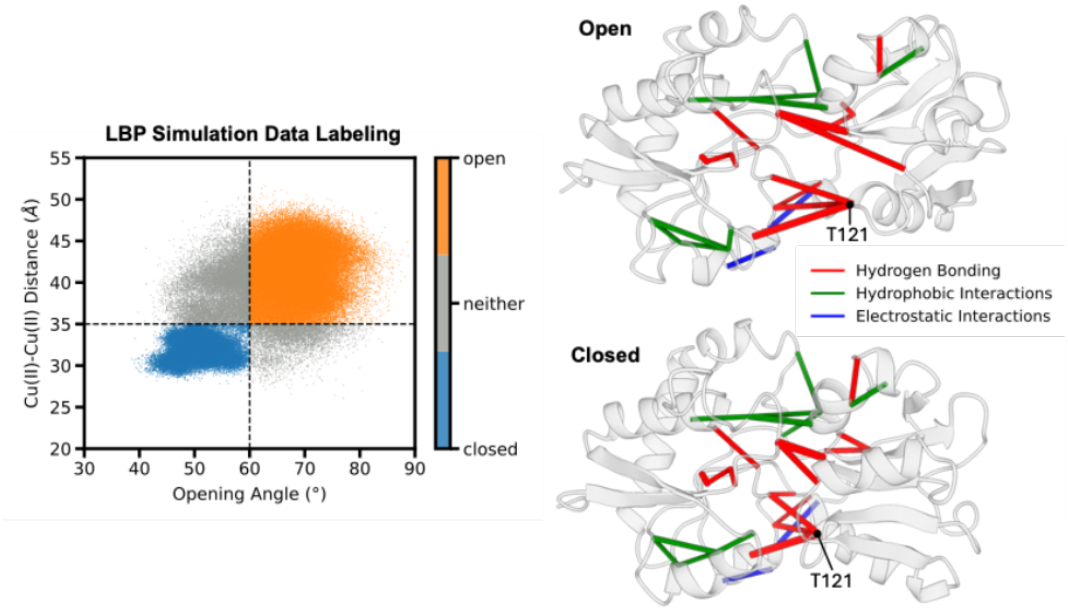

Figure S10. Labeling of the LAOBP simulation data and key interaction features overlaid on representative structures of the open and closed states.

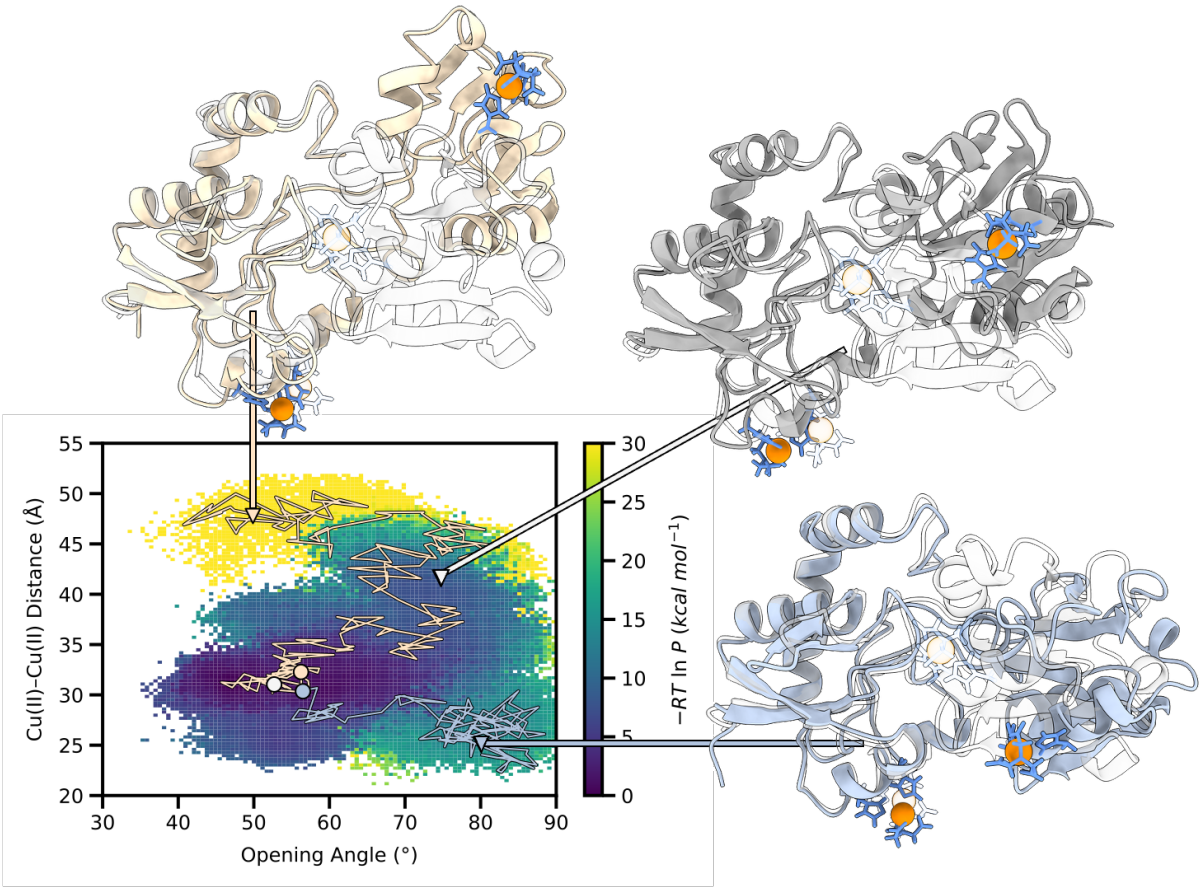

Figure S11. Representative structures from conformational extremes of LAOBP WE simulations. Starting from the closed state of LAOBP (marked by colored circles), WE simulations were able to sample the open state (marked by the gray upside down triangle) as well as lower probability states (marked by colored upside down triangles) with, e.g. a high Cu(II)–Cu(II) distance and a low opening angle (tan line, markers, and structure) and a low Cu(II)–Cu(II) distance and high opening angle (light blue line, markers, and structure). The closed state starting structure is shown as a transparent reference on each structural panel. The LAOBP backbone is shown as a ribbon cartoon while the NTA capping ligand of the spin label is shown in blue, and the Cu(II) metal ion is shown in light orange.

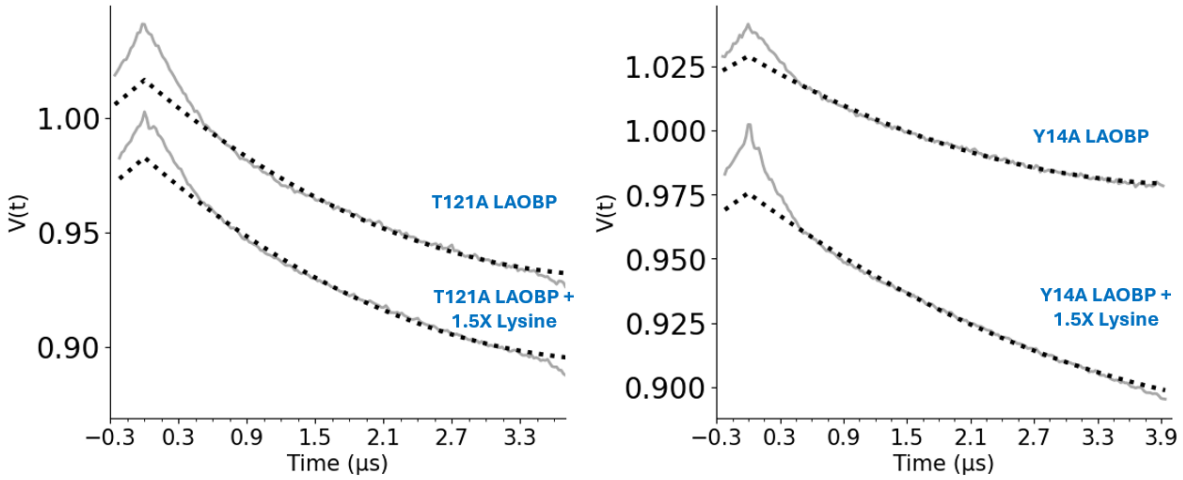

Figure S12. Primary deer data signals for the T121A mutant and Y14A mutant of LAOBP shown in Figure 7.

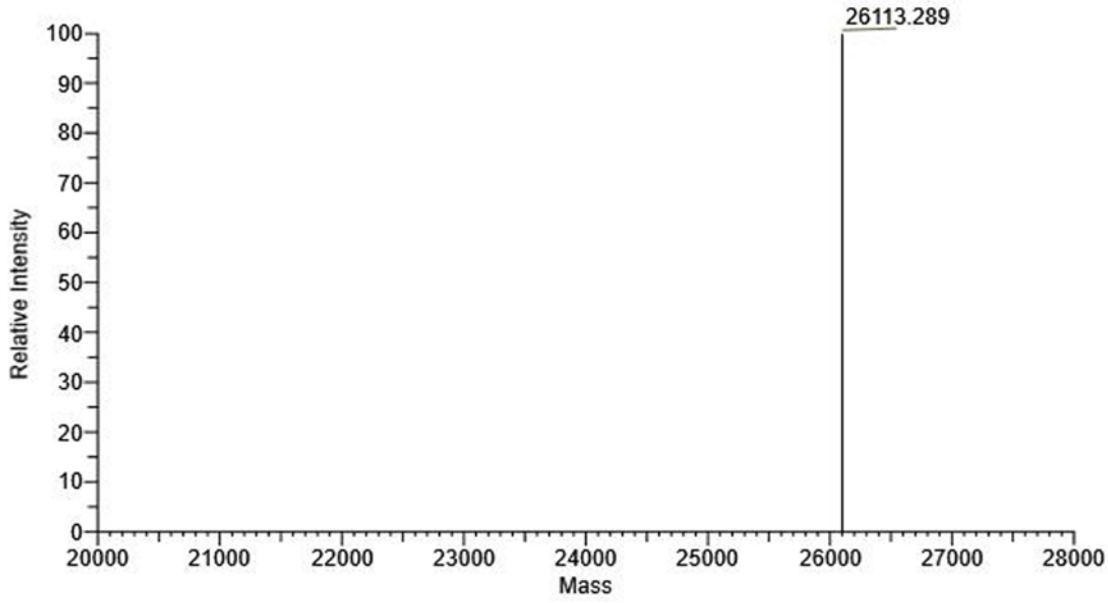

Figure S13. LC-MS data on the refolded LAOBP protein sample. There is a single peak with a molecular weight of 26113.3 kDa, which is the expected mass of apo (ligand-free) protein. There is no detectable peak broadening, suggesting lysine has been successfully removed completely from the sample.

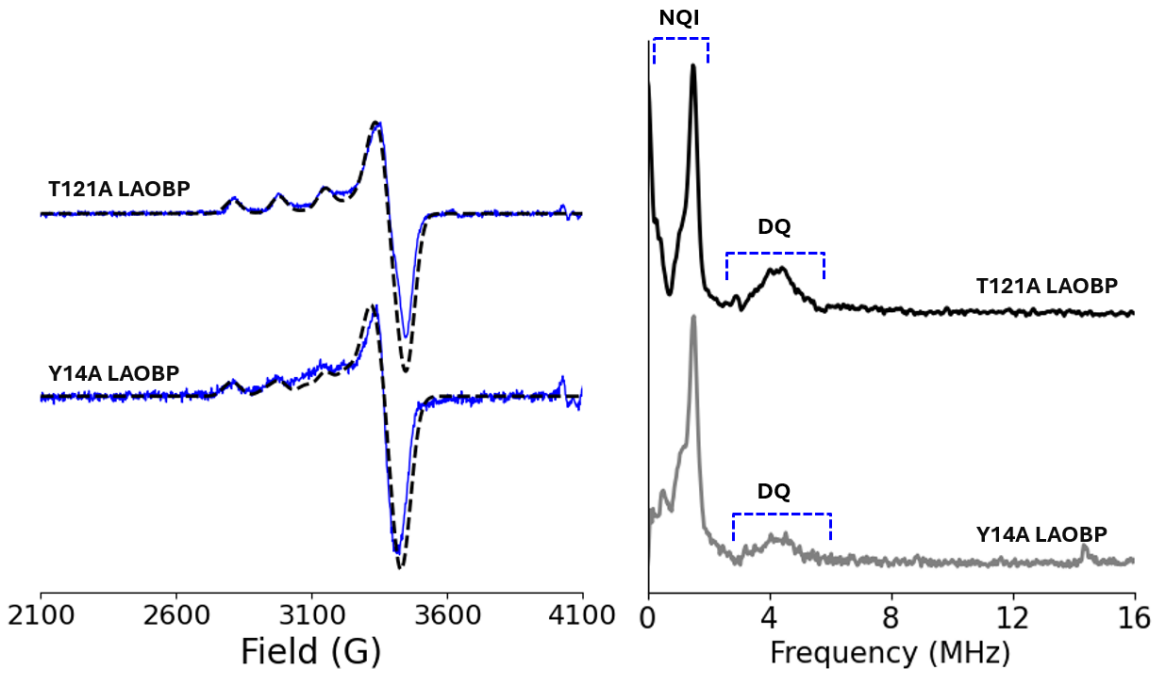

Figure S14. (Left) CW-EPR spectra and simulations of T121A LAOBP and Y14A LAOBP labeled with Cu(II)-NTA. The T121A LAOBP spectrum shows a single component with *g* and *A* tensors characteristic of dHis-bound Cu(II)-NTA, indicating complete labeling. The Y14A LAOBP spectrum was simulated using two components. Component 1 (65%) exhibits *g* and *A* tensors corresponding to protein-bound Cu(II)-NTA, while component 2 (35%) shows *g* tensors consistent with free Cu(II)-NTA, suggesting ~65% labeling efficiency in this mutant. (Right) ESEEM spectra for both mutants displaying characteristic NQI and DQ peaks indicative of imidazole coordination.

Movie S1. Movie of a continuous LAOBP closed-to-open pathway obtained from the WE simulation, highlighting the change in the Cu(II)-Cu(II) distance as it goes from the closed to the open state. (https://github.com/darianyang/wepr/raw/refs/heads/main/LAOBP-path.mp4)

1. Jeschke G. MMM: A toolbox for integrative structure modeling. Protein Sci. 2018;27(1):76–85.
2. Jeschke G. MMM: Integrative ensemble modeling and ensemble analysis. Protein Sci. 2021;30(1):125–35.
3. Tessmer MH, Stoll S. chiLife: An open-source Python package for in silico spin labeling and integrative protein modeling. PLOS Comput Biol. 2023 Mar 31;19(3):e1010834.
4. Hasanbasri Z, Tessmer MH, Stoll S, Saxena S. Modeling of Cu(II)-based protein spin labels using rotamer libraries. Phys Chem Chem Phys. 2024 Feb 22;26(8):6806–16.
5. Hasanbasri Z, Moriglioni NA, Saxena S. Efficient sampling of molecular orientations for Cu(II)-based DEER on protein labels. Phys Chem Chem Phys. 2023 May 17;25(19):13275–88.
6. Worswick SG, Spencer JA, Jeschke G, Kuprov I. Deep neural network processing of DEER data. Sci Adv. 2018 Aug 24;4(8):eaat5218.

7. Jeschke G. ComparativeDeerAnalyzer 2.0.

8. Heubach CA, Hasanbasri Z, Abdullin D, Reuter A, Korzekwa B, Saxena S, et al. Differentiating between Label and Protein Conformers in Pulsed Dipolar EPR Spectroscopy with the dHis-Cu2+(NTA) Motif. Chem – Eur J. 2023;29(72):e202302541.