

SUPPLEMENTARY INFORMATION

Substrate Inhibition by the Blockage of Product Release and Its Control by Tunnel Engineering

Piia Kokkonen^{*a}, Andy Beier^{*a,b}, Stanislav Mazurenko^a, Jiri Damborsky^{a,b}, David Bednar^{a,b†}, Zbynek Prokop^{a,b†}

^a Loschmidt Laboratories, Department of Experimental Biology and RECETOX, Faculty of Science, Masaryk University, Kamenice 5/A13, 625 00 Brno, Czech Republic

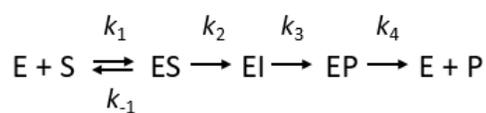
^b International Clinical Research Center, St. Anne's University Hospital Brno, Pekarska 53, 656 91 Brno, Czech Republic

Content

Section I: Kinetic models and global data analysis	2
Section II: Molecular dynamics simulations and Markov state models	6
Section III: Mutagenesis of p3 tunnel residues	23
Section IV: Multivariate data analysis	25
References	27

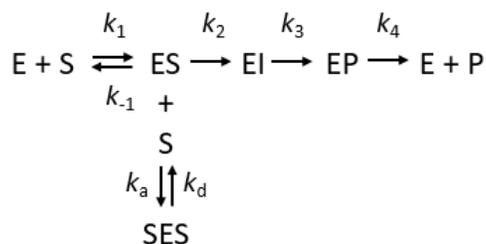
Section I: Kinetic models and global data analysis

We applied the minimal kinetic model for HLDs reaction (**Supplementary Scheme S1**) which has been proposed earlier based on X-ray crystallographic (Verschueren, 1993) and transient kinetic studies (Schanstra, 1996; Bosma, 2003; Prokop, 2003). The kinetic pathway includes four individual steps: (i) initial binding of a substrate (S) to the enzyme (E) forming enzyme-substrate complex (ES), (ii) cleavage of carbon-halogen bond leading to the formation of covalent alkyl-enzyme intermediate (EI), (iii) hydrolyses of alkyl-enzyme intermediate and, (iv) final release of reaction products. The rate constant for association and dissociation of ES complex is indicated as k_1 and k_{-1} , respectively, k_2 is the rate constant for cleavage of carbon-halogen bond, k_3 is the rate constant for the hydrolysis of the alkyl-enzyme intermediate and k_4 is the rate constant for product release.

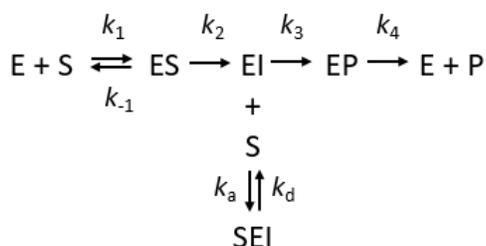


Supplementary Scheme S1

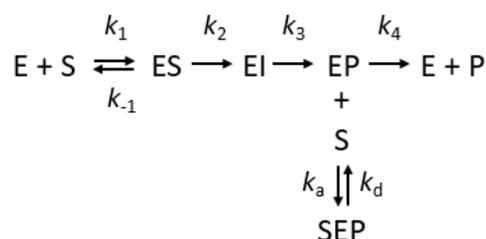
Even for the simple kinetic pathway including one intermediate, there are three possible scenarios of how substrate can inhibit the catalytic cycle (**Supplementary Scheme S2-4**). The rate constant for association and dissociation of substrate inhibitory complex is indicated as k_a and k_d , respectively (the equilibrium dissociation constant for substrate inhibitory complex $K_{SI} = k_d/k_a$).



Supplementary Scheme S2



Supplementary Scheme S3

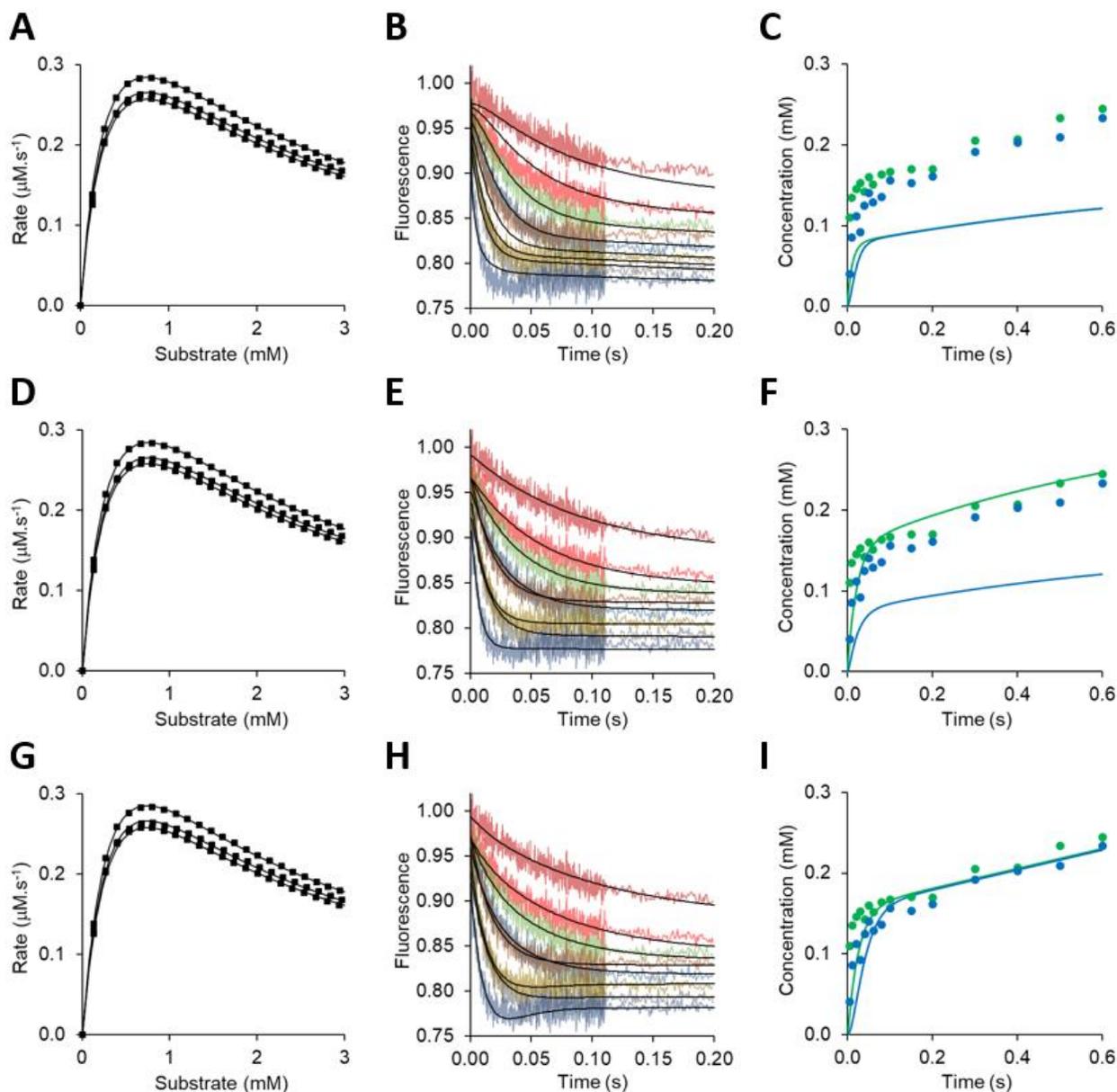


Supplementary Scheme S4

To address which model accounts for the kinetic observations we fit the steady-state and transient kinetic data globally to all three expected models (**Supplementary Figure S1**) and examine the errors in the parameters and evaluate the goodness of fit both visually and computationally. Before the global data analysis, each kinetic trace was individually fit by conventional methods to an analytical function (e.g., double exponential, single exponential + linear phase) that mimics the data sufficiently to allow an estimate of the standard deviation of each measurement based upon the residuals using the fitted curve. This calculated standard deviation value was then used as a basis to normalize residuals for global fitting. The normalized standard deviation χ^2 provided a metric to estimate the goodness of fit. For a good global fit, an average standard deviation is comparable to that obtained based upon the best fit to individual traces and the ratio χ^2/DoF (degrees of freedom) should approach unity. In the case of three tested kinetic models, Supplementary Scheme S2, S3 and S4, the χ^2/DoF was 3.50, 2.29 and 1.27, respectively (**Supplementary Table S1**). This suggests that the model where the substrate binds to the enzyme-product complex is more likely than the two other considered scenarios.

The visual inspection of the fit indicates, that while steady-state data can achieve a good correspondence to the simulations for all three models (**Supplementary Figure S1, panels A, D and G**) there is increasing discrepancy of the experimental and simulated values for transient-state data. The significant indications of the quality of the fit were observed for both multiple turnover data (**Supplementary Figure S1 panels B, E and H**) and burst data (**Supplementary Figure S1 panels C, F and I**), the latter strongly determine the pattern of inhibition. In the case of the traditional model including enzyme complex with two substrates (SES), the fraction of enzyme-substrate complex that can undergo the following chemical step (k_2) is reduced by the formation of SES complex and burst amplitudes are decreased accordingly for both, bromide and alcohol product which is in disagreement with observed experimental values (**Supplementary Figure S1 C**). By testing the second model including substrate binding to alkyl-enzyme intermediate (EI), the formation of the reaction intermediate is not compromised and the amplitude of halide burst reached the expected value. However, the fraction of EI that can undergo following hydrolytic step (k_3) is reduced by the formation of the substrate-enzyme-intermediate complex (SEI) and the simulated amplitude of alcohol product formation is again significantly reduced in comparison to experimental observation (**Supplementary Figure S1 F**). The only model satisfying the rates and amplitudes of all observed phases of the burst and multiple turnover data is the model described in **Supplementary Scheme S4**, where the substrate is bound to enzyme-product complex. The kinetic and equilibrium constant obtained from individual global data analysis are summarized in **Supplementary Table S1**. Interestingly, the equilibrium dissociation constant for substrate inhibitory complex K_{SI} strongly

different for individual models, when it acquires sub-micromolar values for models including the formation of SES and SEI. Such values would suggest tight binding of the substrate to inhibitory complex, which is very unlikely when compared to the dissociation constant describing the productive binding of the substrate to the active site of LinB ($K_S = 37$ mM). The model including inhibitory binding of substrate to enzyme-product complex provided reasonable value in the millimolar range ($K_{SI} = 1.21 \pm 0.01$ mM).



Supplementary Figure S1. Steady-state initial rates recorded at 0 - 3 mM of 1,2-dibromoethane and 49, 68, and 83 nM LinB wild type (A, D, G). Stopped-flow fluorescence traces recorded upon rapid mixing of 4.8 mM LinB wild type with 0 – 4 mM 1,2-dibromoethane, each trace shows the average of seven individual experiments (B, E, H). Reaction burst of halide (green circles) and alcohol (blue circles) product monitored upon rapid mixing 6.8 mM 1,2-dibromoethane with 160 μM LinB wild type (C, F, I) and 7.8 mM 1,2-dibromoethane with 160 μM LinB wild type (D). Solid lines represent the global fit of the model described in Supplementary Scheme S2 (A, B, C), S3 (D, E, F) and S4 (G, H, I) to the kinetic data.

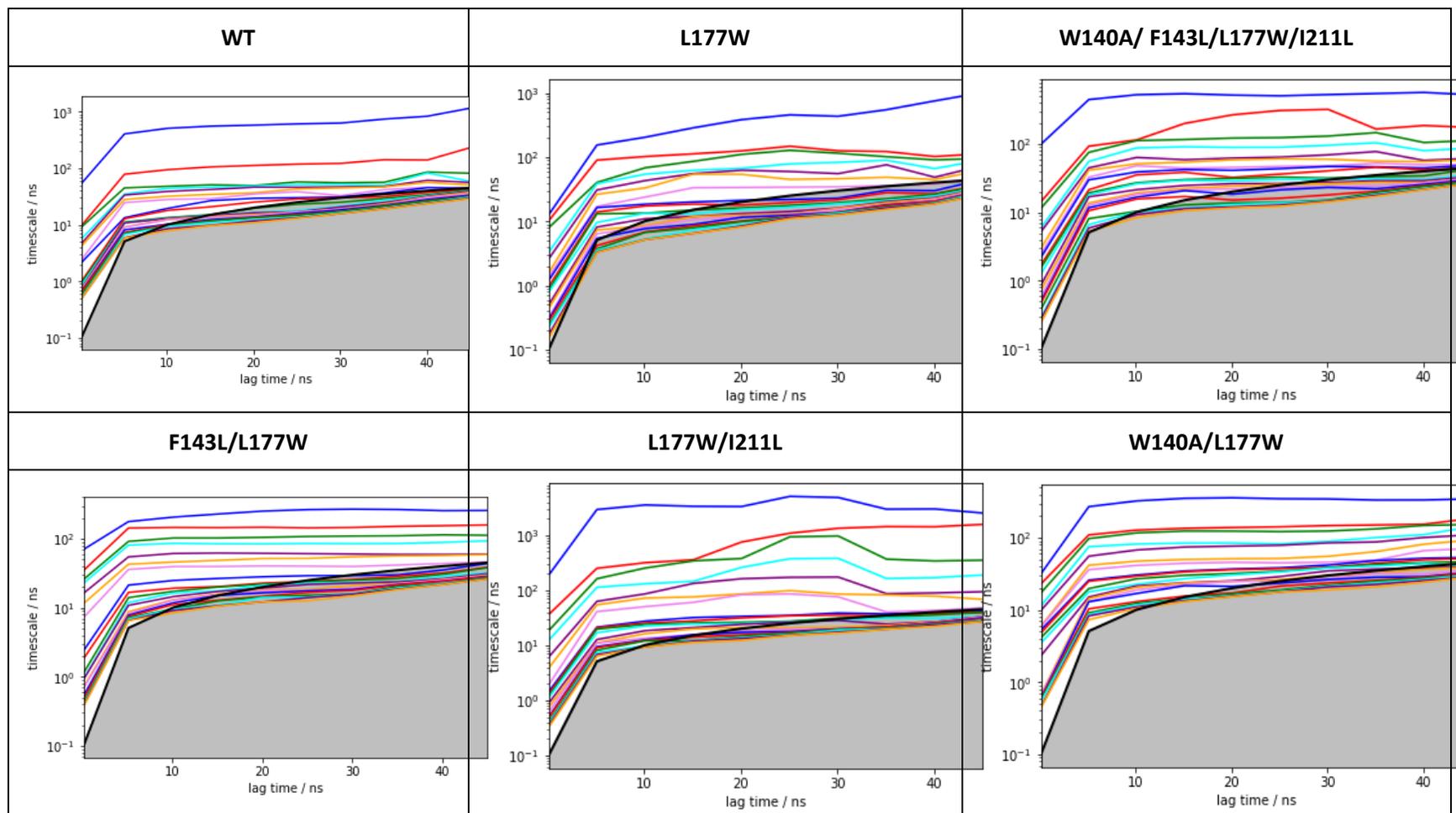
Supplementary Table S1: Kinetic constants and scaling factors obtained by global data analysis. The parameters were derived by simultaneous fitting steady-state and transient-state kinetic data using numerical integration of rate equations derived from Model I, II and III, described in Supplementary Scheme S2-4. The fluorescence signal F recorded by stopped-flow was defined as the sum of fluorescence intensities from all contributing species $F = f \cdot (E + a \cdot ES + b \cdot EI + c \cdot EP)$ where f scales the fluorescence signal to the free enzyme, scaling factors a , b and c reflect the change of fluorescence intensity corresponding to the formation of the enzyme-substrate complex (ES), reaction intermediate (EI) and enzyme-product complex (EP), respectively. The rate of substrate binding k_1 was set to a diffusion-limited rate $1000 \mu\text{M}^{-1} \text{s}^{-1}$ to assume a rapid-equilibrium (the equilibrium dissociation constant for enzyme-substrate complex $K_S = k_{-1}/k_1$ could be derived from the analysis using k_{-1} as a fitted parameter).

	Model I	Model II	Model III
Goodness of Fit (χ^2/DoF)	3.50	2.29	1.27
K_{SI} (mM)	< 0.0001	< 0.001	1.21 ± 0.01
K_S (mM)	36 ± 2	38 ± 2	37 ± 2
k_2 (s^{-1})	360 ± 20	380 ± 10	330 ± 10
k_3 (s^{-1})	100 ^a	130 ± 20	109 ± 4
k_4 (s^{-1})	2.4 ± 0.1	2.7 ± 0.1	3.2 ± 0.1
f	232.8 ± 0.2	234.1 ± 0.1	235.0 ± 0.1
a	0.77 ± 0.01	1.07 ± 0.01	0.99 ± 0.01
b	1.05 ± 0.01	0.85 ± 0.01	0.75 ± 0.01
c	0.89 ± 0.01	0.86 ± 0.01	0.85 ± 0.01

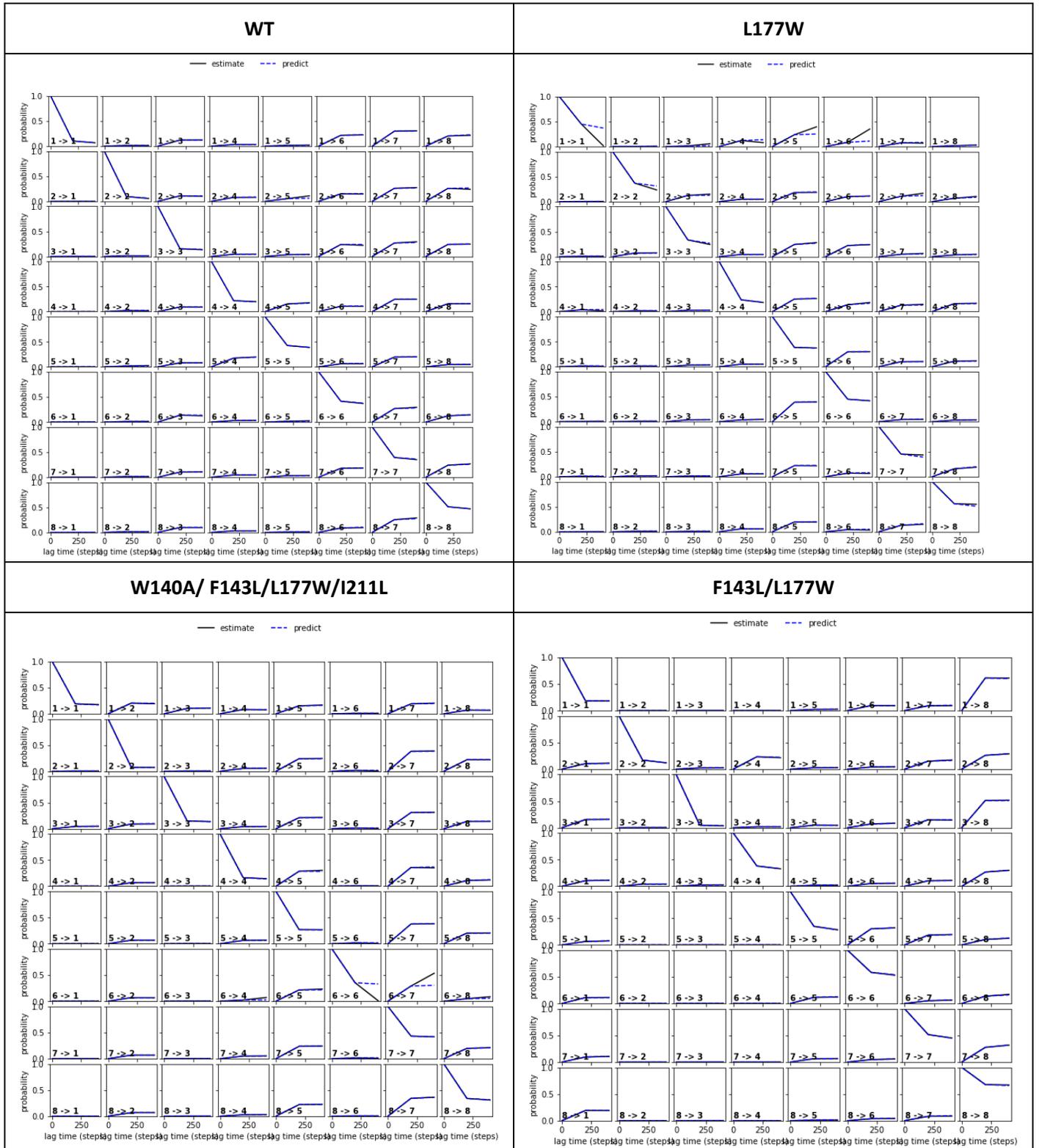
^a In fitting kinetic data to Model I, rate constant k_3 was fixed at 100 s^{-1} .

Section II: Molecular dynamics simulations and Markov state models

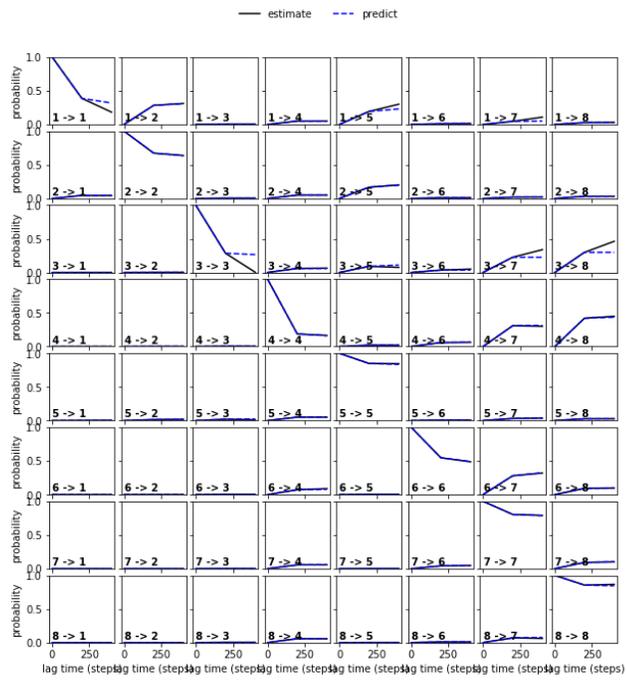
Supplementary Figure S2. The implied time-scales plots of the simulations used for the Markov state modeling.



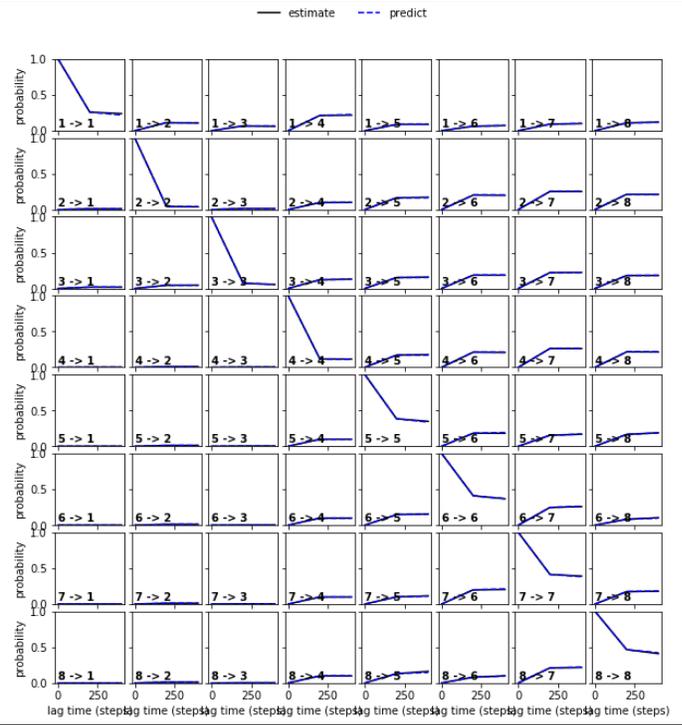
Supplementary Figure S3. The Chapman-Kolmogorov tests for the built Markov models.



L177W/I211L



W140A/L177W



Supplementary Table S4. All Markov States of the different enzyme variants organized by their equilibrium probability.

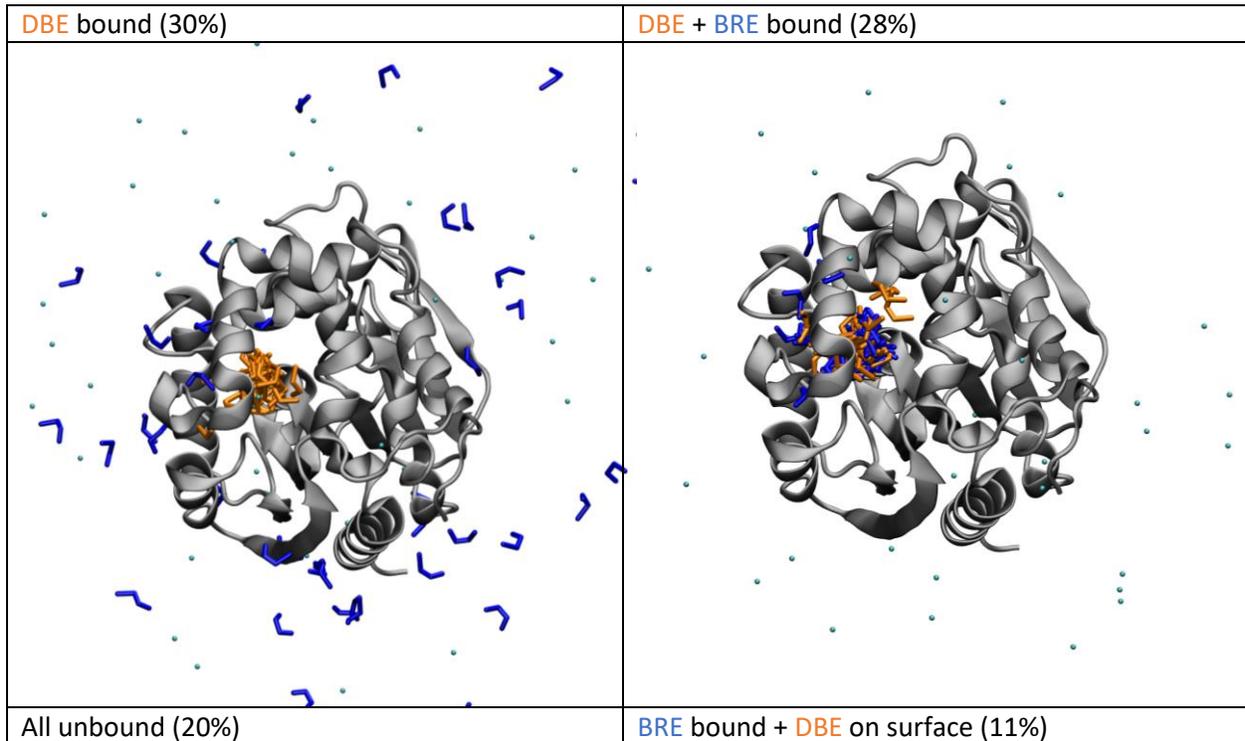
	WT	L177W	W140A/ F143L/L177W/I211L	F143L/L177W	L177W/I211L	W140A/L177W
Common states (>20%)	DBE bound (30%)	DBE on surface + BR bound (31%)	DBE + BRE bound (39%)	DBE bound (50%)	DBE bound (48%)	DBE + BRE bound (27%)
	DBE + BRE bound (28%)	BR bound (24%)	All unbound (24%)		All unbound (37%)	DBE bound (22%)
	All unbound (20%)		DBE bound (22%)			BRE bound (22%)
Uncommon states (5-20%)	BRE bound + DBE on surface (11%)	DBE + BR bound (17%)	BRE + BR bound, DBE on surface (6.8%)	BRE + BR bound (16%)	DBE + BRE bound (6.7%)	All unbound (17%)
	DBE + BRE bound (5.3%)	DBE + BR bound (14%)	BRE + DBE bound (5.3%)	DBE + BRE bound (15%)	DBE on surface + BRE bound (5.6%)	DBE + BR bound (10%)
		All unbound (6.1%)		All unbound (13%)		
Rare states (<5%)	BRE bound (4.7%)	BRE + BR bound + DBE on surface (3.9%)	DBE + BRE bound (1.8%)	BR bound (0.1%)	BR bound (2.6%)	BRE + BR bound + DBE on surface (1.2%)
	DBE + BRE bound (1.4%)	All bound (2.2%)	All bound (0.3%)	DBE + BR bound (0.1%)	DBE + BR bound (0.6%)	BR bound (0.4%)
	All bound (0.2%)	DBE on surface (1.8%)	All bound (0.1%)	All bound (0.02%)	BRE + BR bound (0.2%)	All bound (0.2%)
				All bound (0.04%)		

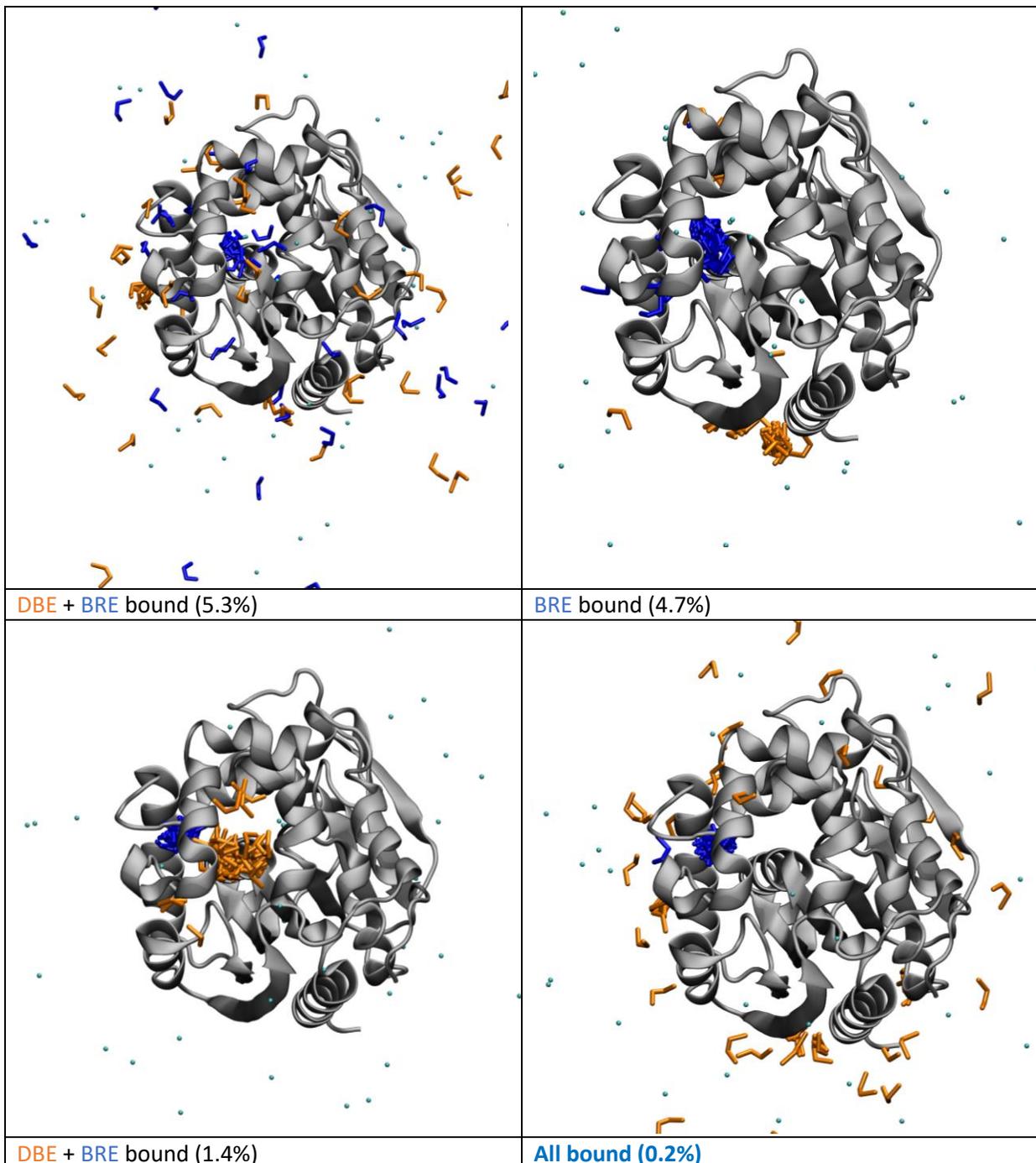
Ila: Figures of the Markov states

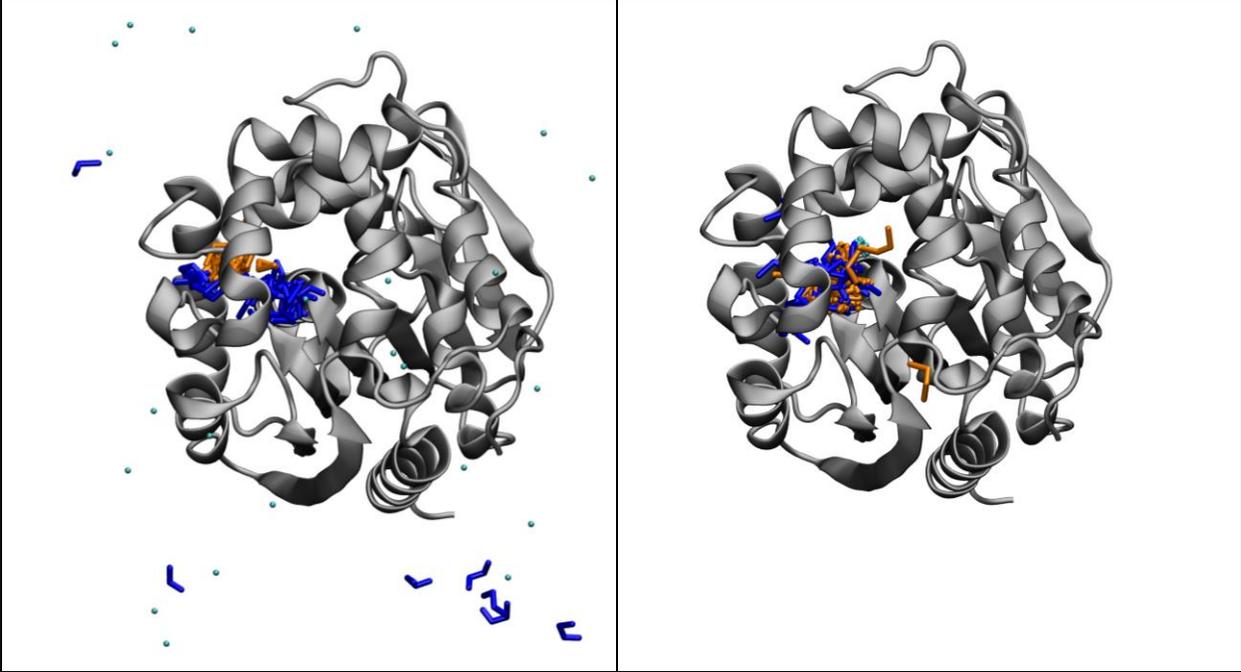
Figures were made from 50 random snapshots belonging to the states overlaid on top of the starting structure shown as cartoon.

Legend: Bromide (cyan sphere), DBE (orange sticks), BRE (blue sticks)

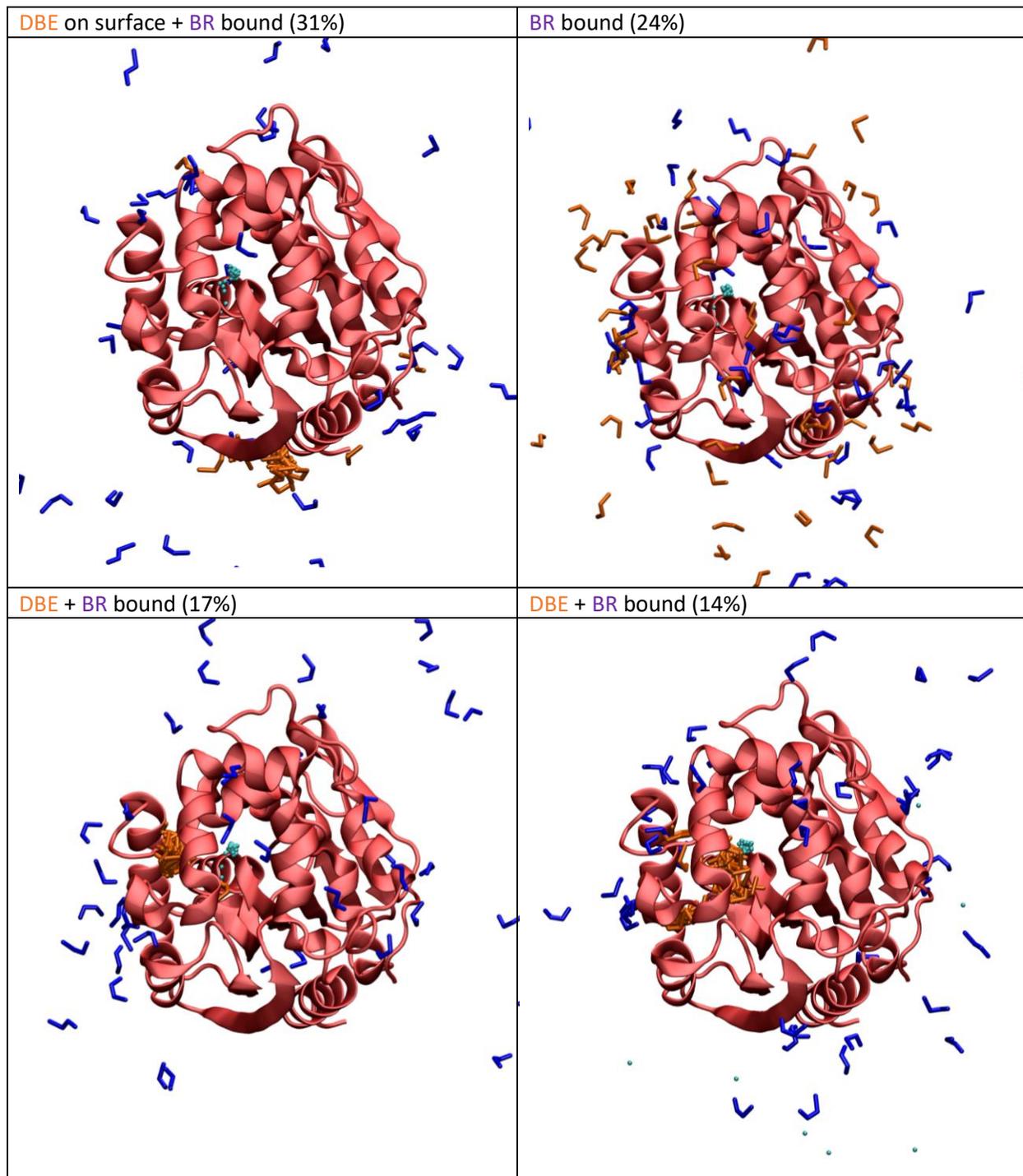
Supplementary Figure S4. The Markov states and their equilibrium probabilities obtained for WT.

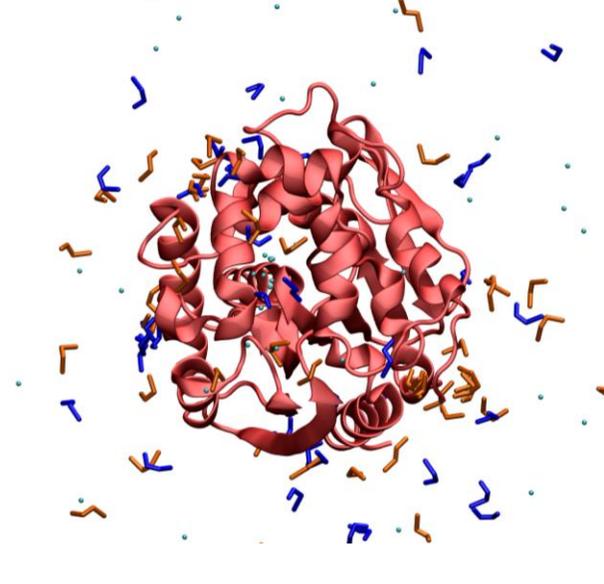
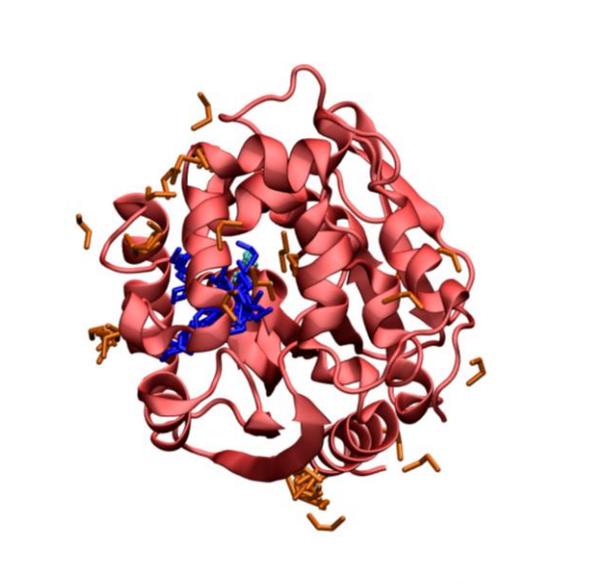
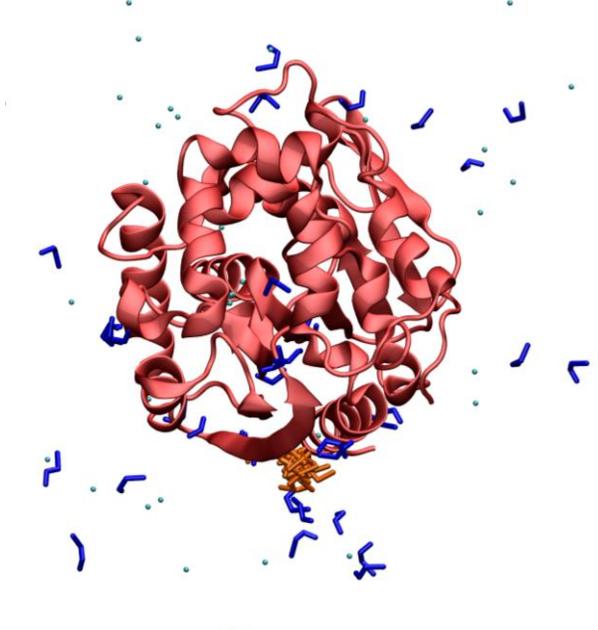




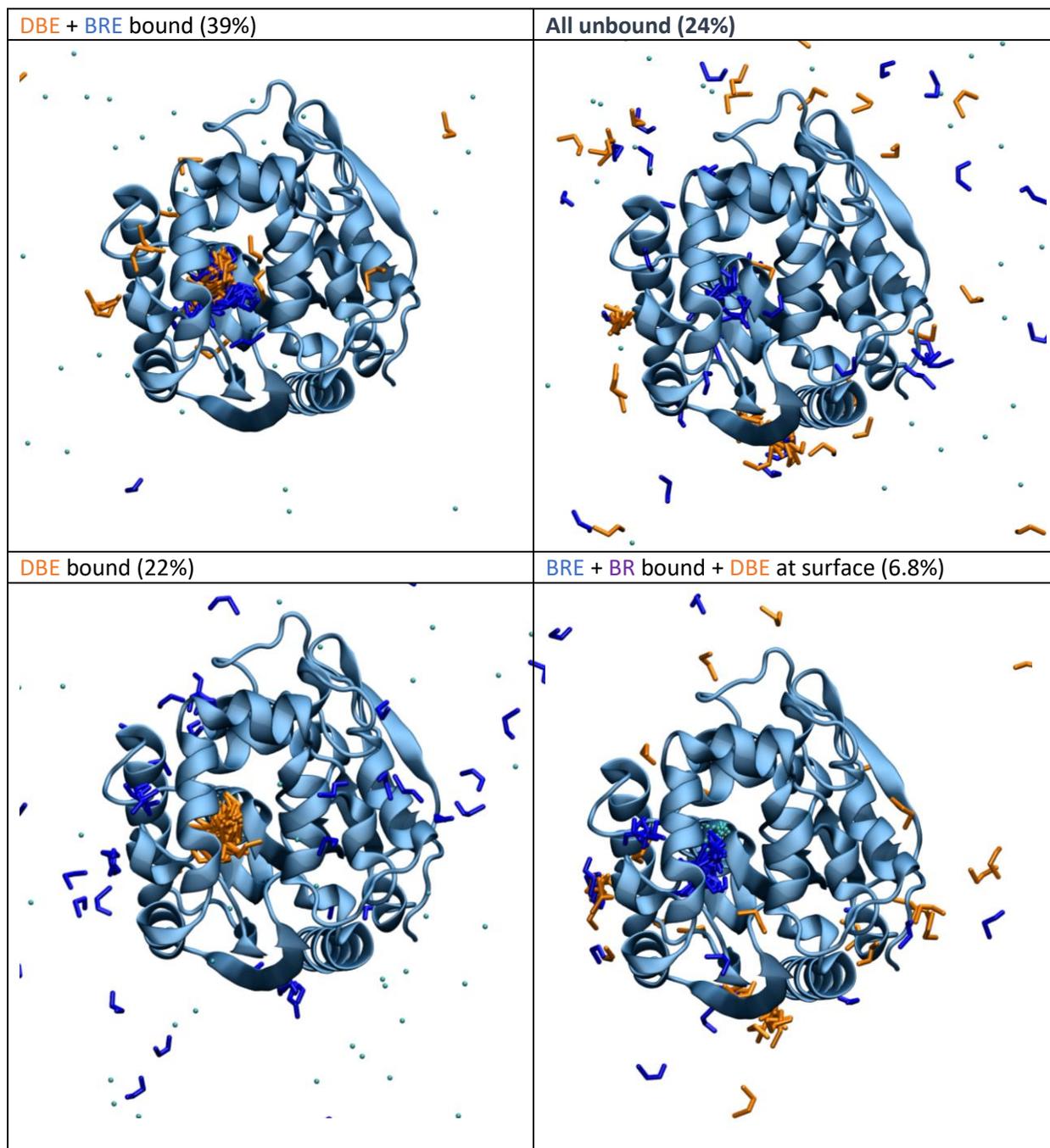


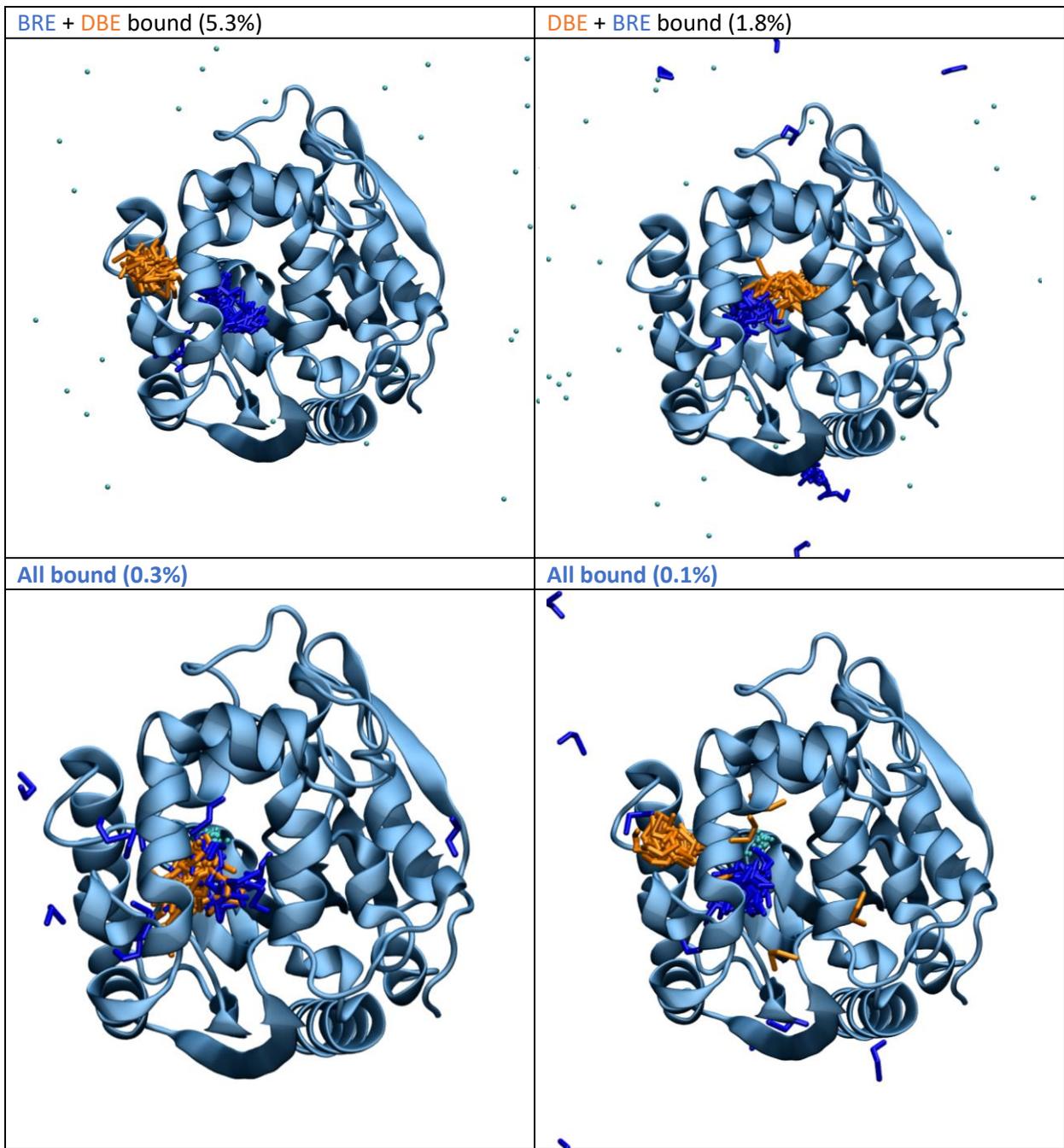
Supplementary Figure S5. The Markov states and their equilibrium probabilities obtained for L177W.



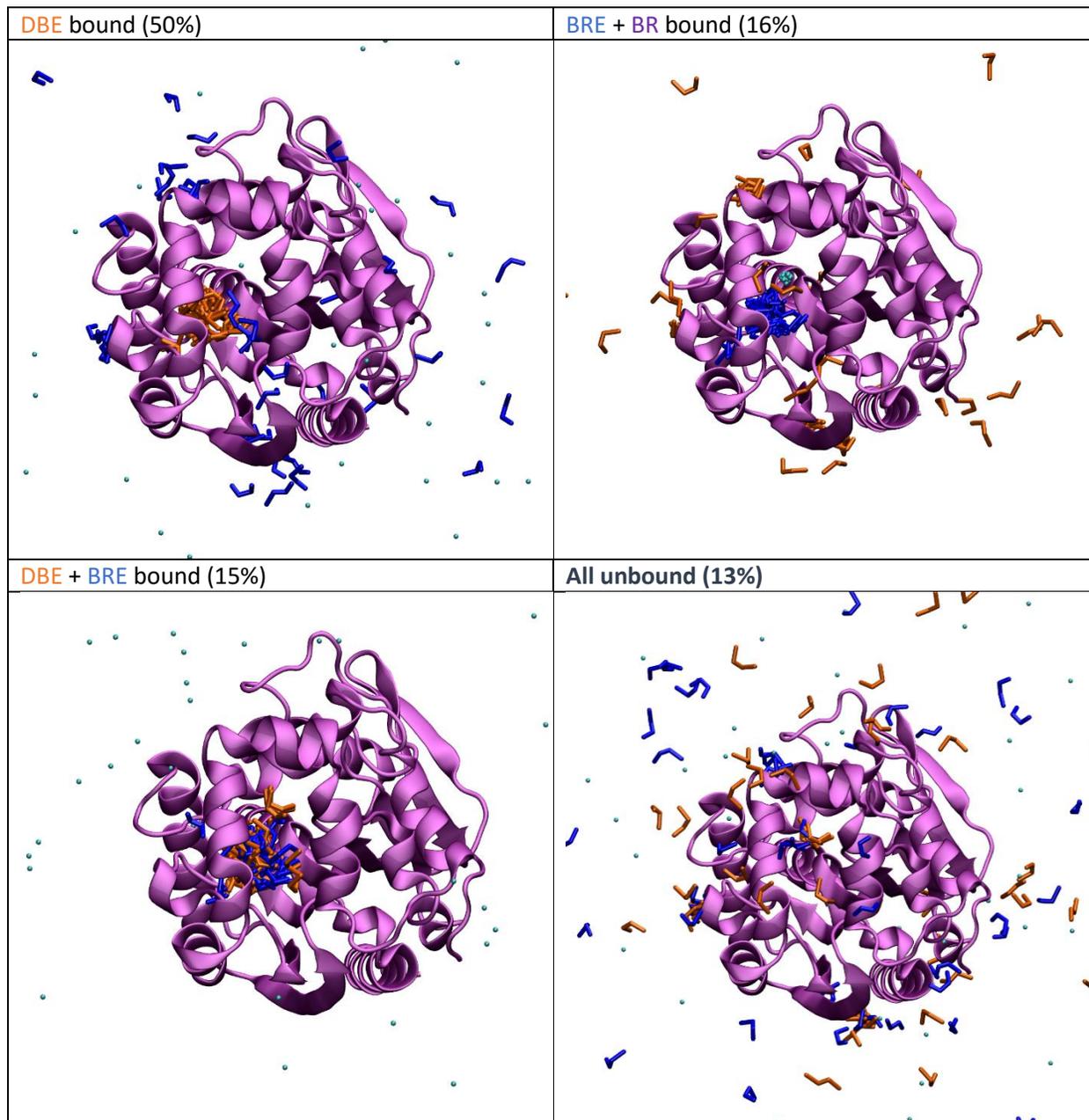
All unbound (6.1%)	BRE + BR bound + DBE on the surface (3.9%)
	
All bound (2.2%)	DBE on the surface (1.8%)
	

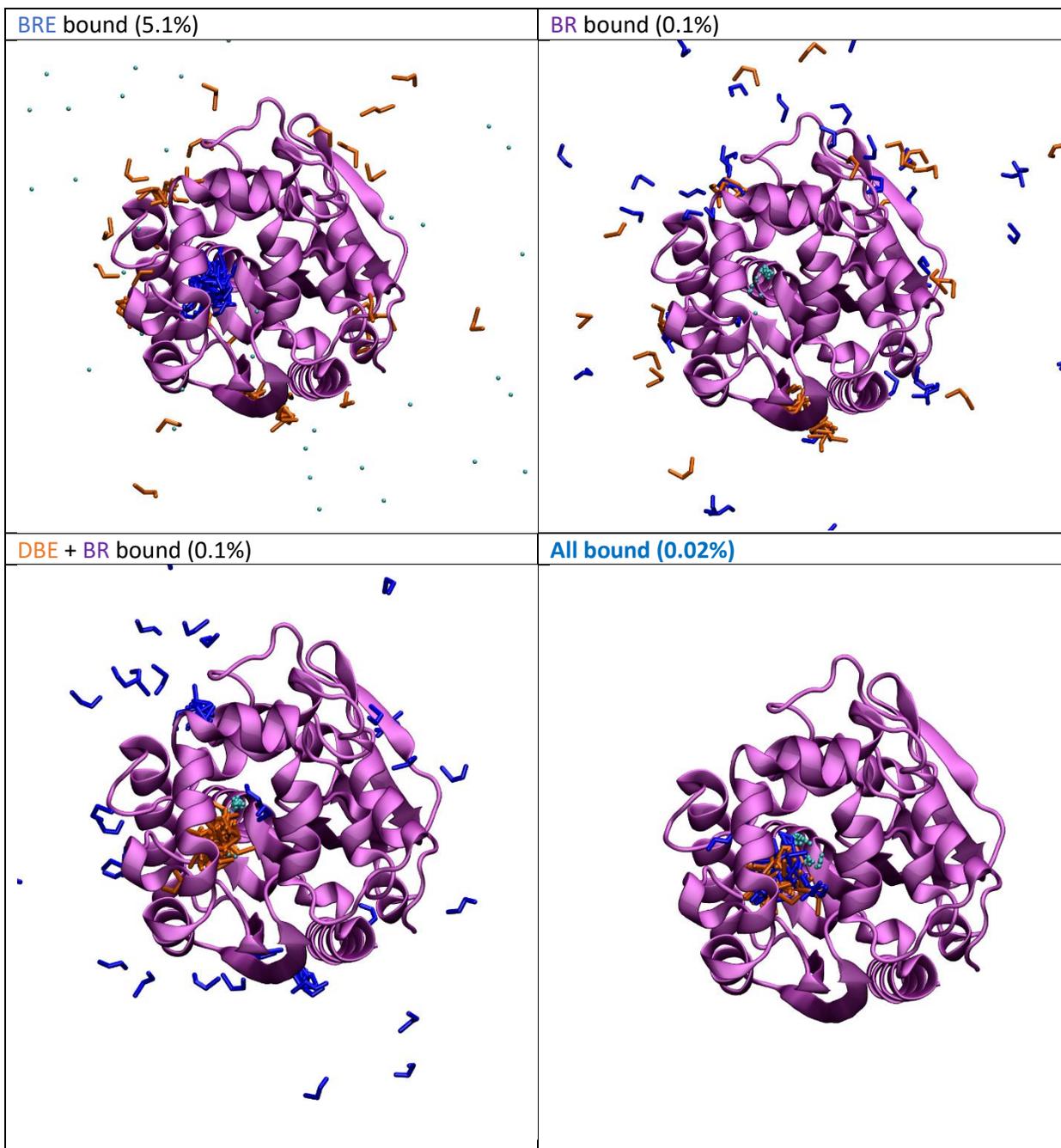
Supplementary Figure S6. The Markov states and their equilibrium probabilities obtained for W140A/F143L/L177W/I211L.



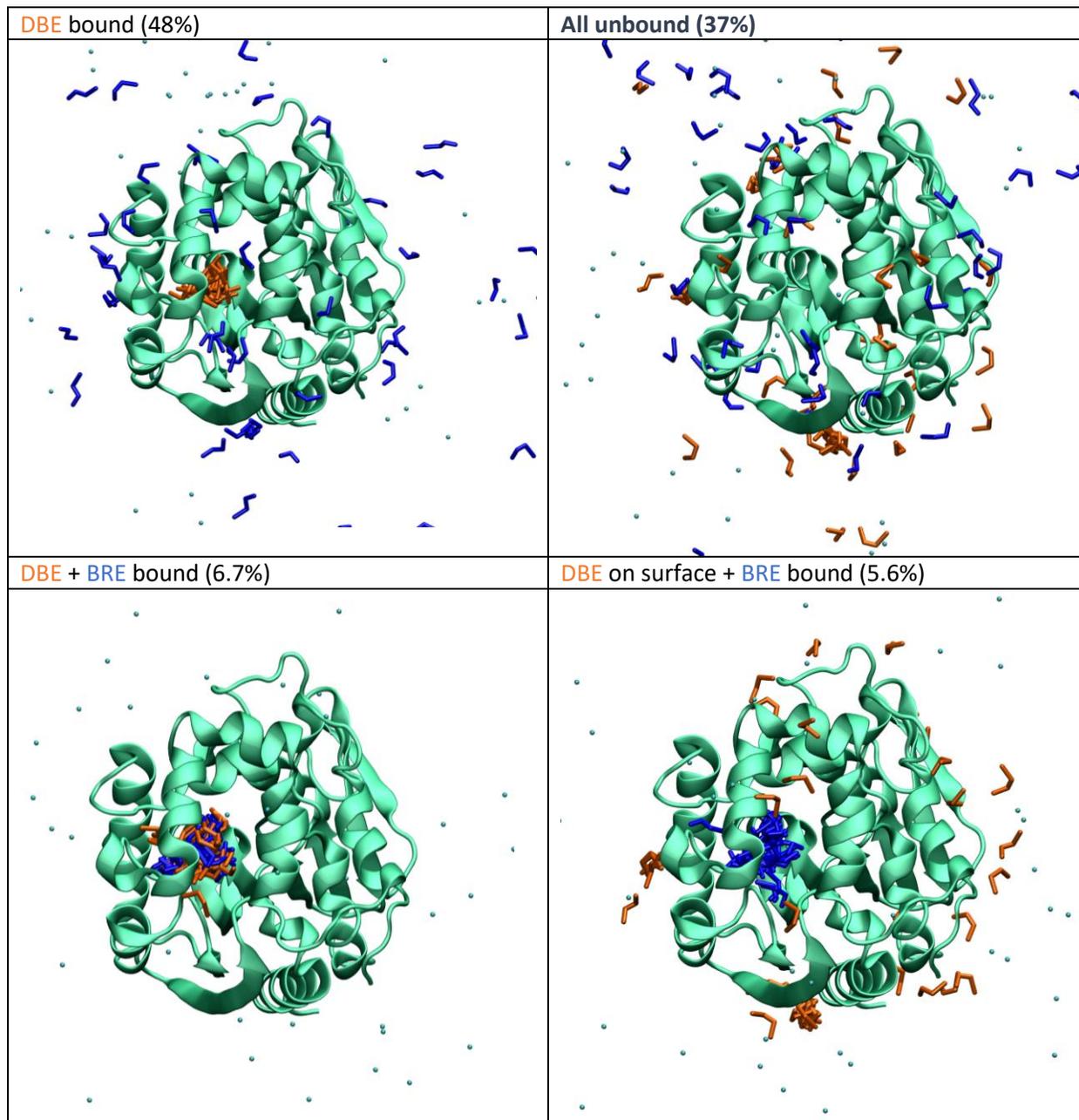


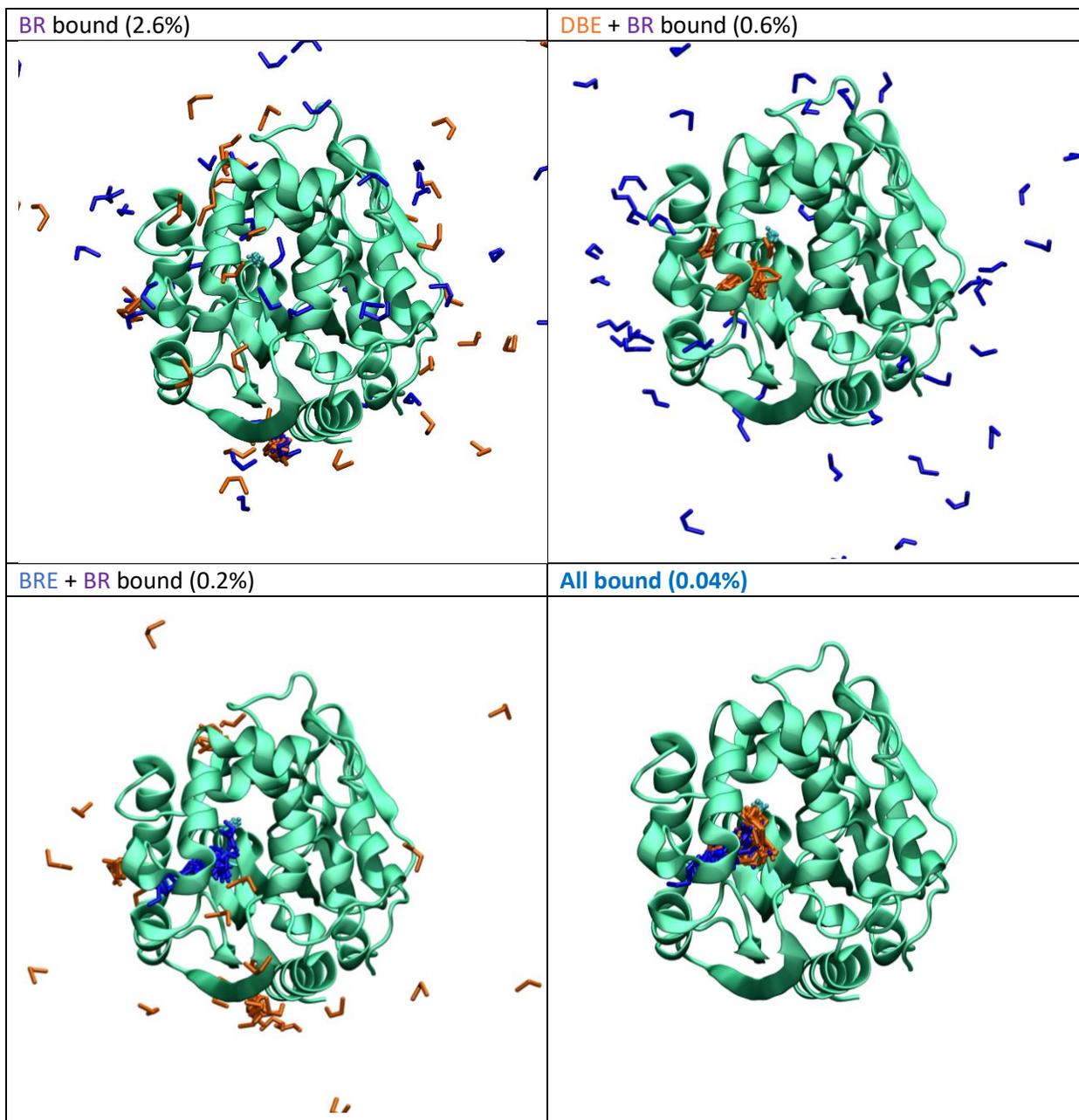
Supplementary Figure S7. The Markov states and their equilibrium probabilities obtained for F143L/L177W.



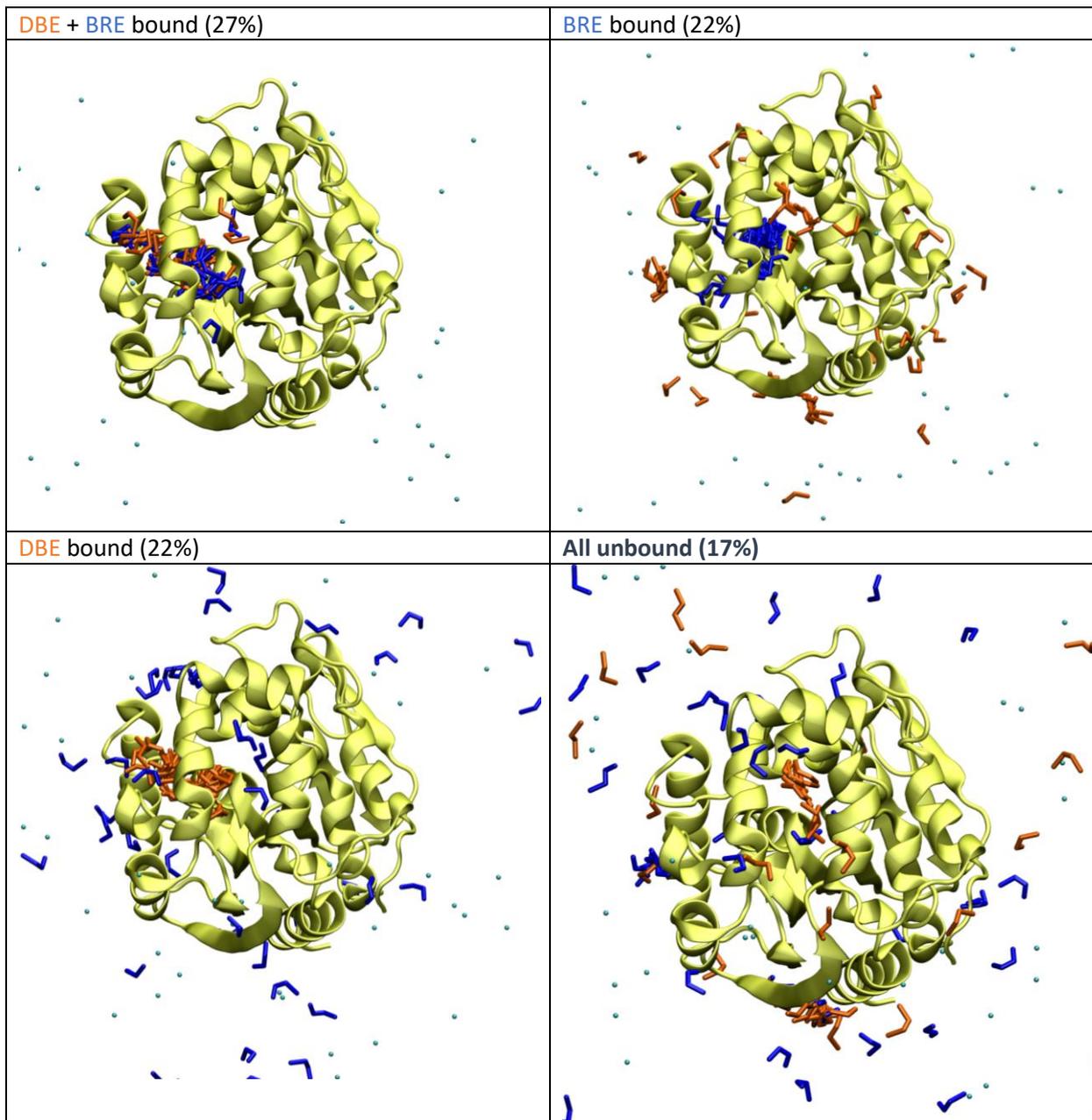


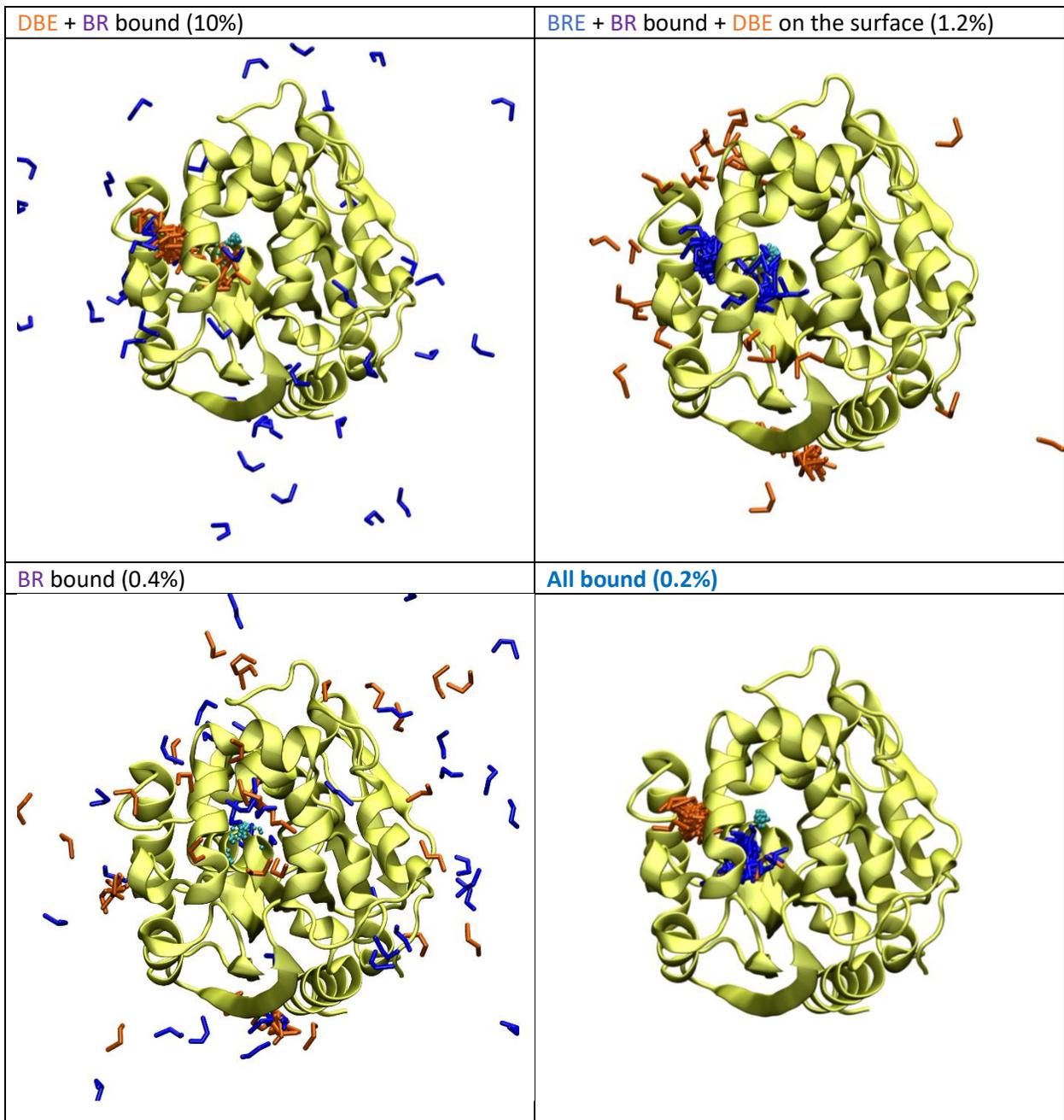
Supplementary Figure S8. The Markov states and their equilibrium probabilities obtained for L177W/I211L.





Supplementary Figure S9. The Markov states and their equilibrium probabilities obtained for W140A/L177W.

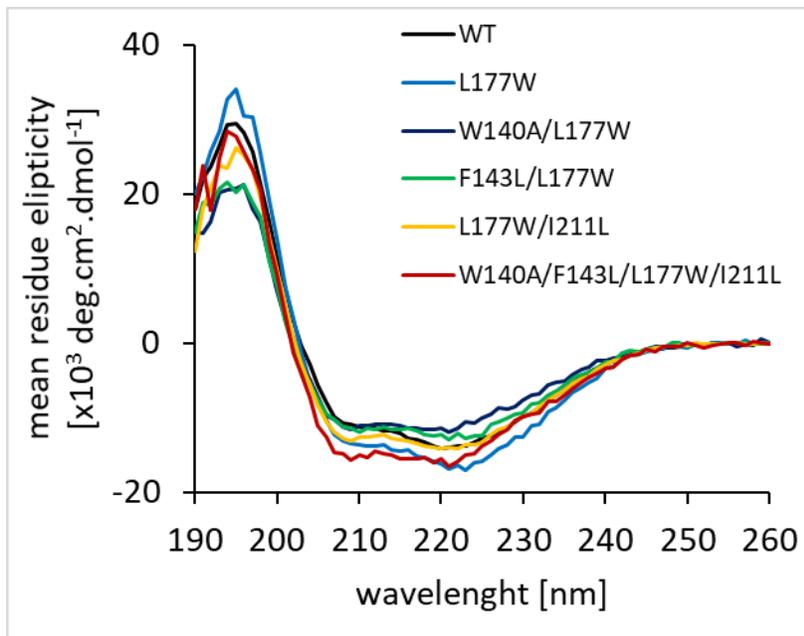




Section III: Mutagenesis of p3 tunnel residues

Supplementary Table S5. Sequences of mutagenic forward primers and used reverse primer to introduce the desired nucleotide exchanges. The bold bases indicate the mutation sites.

W140A	5'-GCGATTGCGATGCCGATCGAG GCG GCGGACTTCCGGAGCAGGAC-3'
I211L	5'-ACTCTGAG CTGGCCGCGTCAG CTG CCGATTGCCGGCAGCCTGCT-3'
F143L	5'-ATGCCGATCGAGTGGGCGGAC CTG CCGGAGCAGGACCGCGATTG-3'
pET vector reverse primer	5'-GCTAGTTATTGCTCAGCGG-3'



Supplementary Figure S10. Comparison of CD spectra of the constructed LinB variants in their native states. The spectra indicate a functional fold for all enzymes.

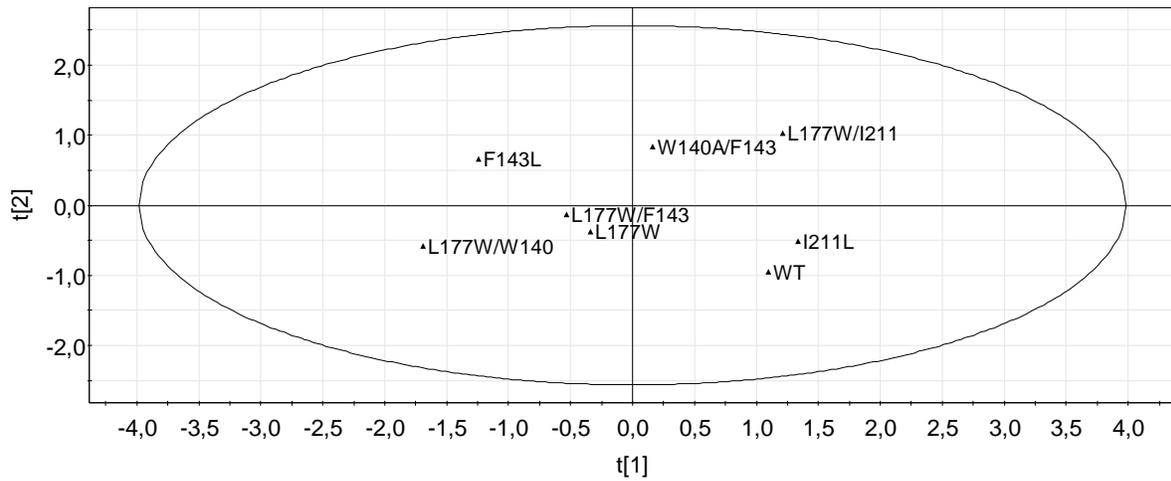
Supplementary Table S6. Comparison of the melting temperatures (T_m) obtained by CD for the constructed LinB variants.

LinB variant	T_m (°C)
WT	48.03 ± 0.52
L177W	49.13 ± 0.52
W140A/L177W	37.21 ± 0.30
F143L/L177W	47.18 ± 0.21
L177W/I211L	54.76 ± 0.31
W140A/ F143L/L177W/I211L	42.09 ± 0.26

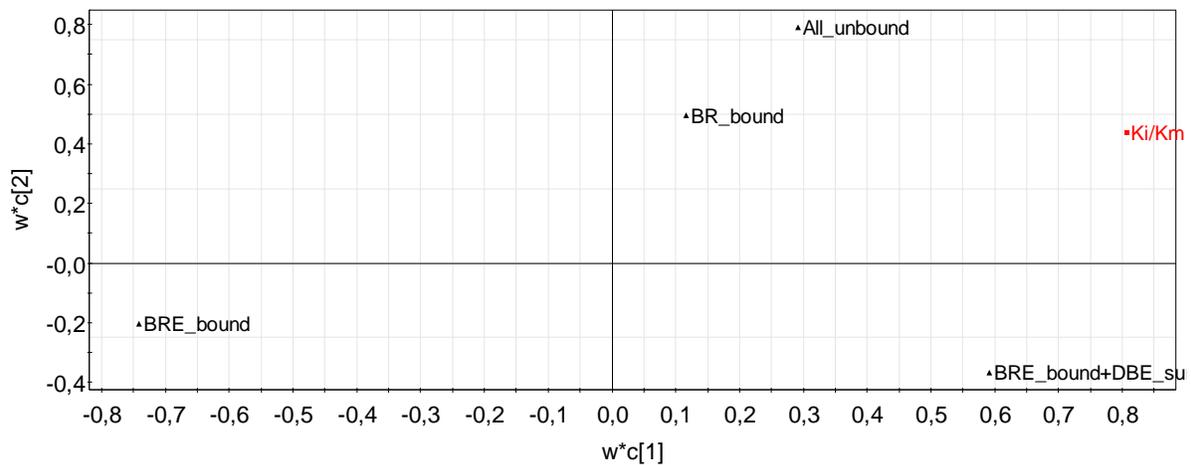
Supplementary Table S7. Comparison of the specific activities obtained by the colorimetric method by Iwasaki *et al* of investigated LinB variants with DBE as substrate.

LinB variant	Specific activity [$\mu\text{mol}/\text{mg}/\text{s}$]
WT	0.13
L177W	0.10
W140A/L177W	0.14
F143L/L177W	0.31
L177W/I211L	0.04
W140A/ F143L/L177W/I211L	0.65

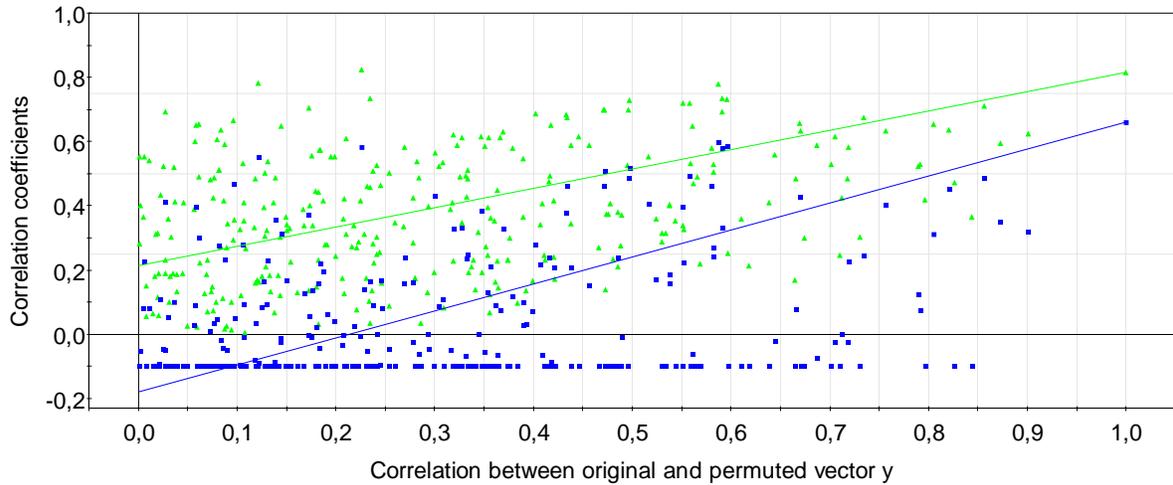
Section IV: Multivariate data analysis



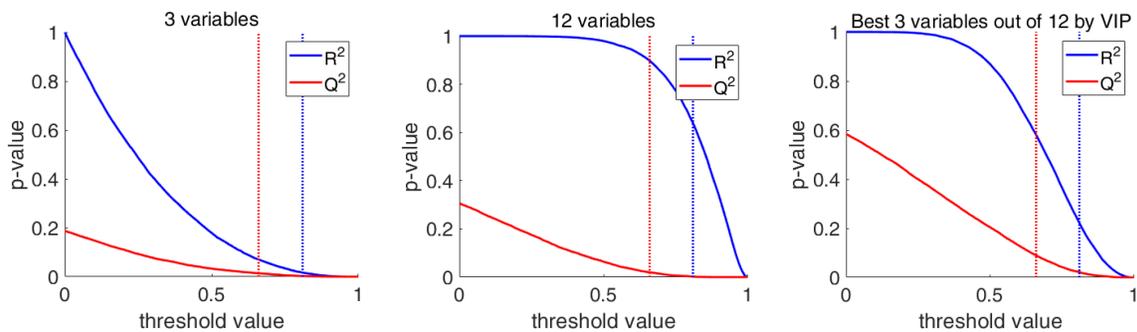
Supplementary Figure S11. The scores plot t_1 versus t_2 from the PLS analysis. The plot shows the distribution of eight protein variants along two principal components. Position of the variants in the space directly corresponds to their scores in the latent space. Variants with similar properties are located close to each other, while those with different properties are far away.



Supplementary Figure S12. Loadings plot wc_1 versus wc_2 from PLS analysis. The plot shows the distribution of three variables and the y-vector along two principal components. Position of the variables in the space corresponds to the loadings of the variables for the two components.



Supplementary Figure S13. Validation plot from permutation testing of the PLS model. The plot presents 300 permutations of original vector y . R^2 values are shown in green, Q^2 values are shown in blue. As the correlation between permuted and original y vectors decreases, the coefficients of determination for the permuted data show significantly lower values than those obtained for the model (the rightmost points), confirming that the results are unlikely to be based on the chance correlation.



Supplementary Figure S14. Monte Carlo simulations for significance testing of the PLS model. The plot presents p-values for R^2 and Q^2 coefficients based on 10 000 simulations: the proportion of simulated values above corresponding thresholds. The probability matrix of the same size as X , i.e. 3 columns (left) or 12 columns (middle and right), and vector y were drawn independently and randomly from a uniform distribution. The values for the original data are shown in dotted lines. The p-values for $Q^2 = 0.66$ are 0.015 (left), 0.021 (middle), and 0.083 (right), confirming the conclusion that the observed PLS results are unlikely to be based on chance alone even given the small sample size.

References

Verschuere KHG, Seljée F, Rozeboom HJ, Kalk KH & Dijkstra BW (1993) Crystallographic analysis of the catalytic mechanism of haloalkane dehalogenase. *Nature* **363**, 693–698.

Schanstra JP, Kingma J & Janssen DB (1996) Specificity and Kinetics of Haloalkane Dehalogenase. *J. Biol. Chem.* **271**, 14747–14753.

Bosma T, Pikkemaat MG, Kingma J, Dijk J & Janssen DB (2003) Steady-State and Pre-Steady-State Kinetic Analysis of Halopropane Conversion by a Rhodococcus Haloalkane Dehalogenase. *Biochemistry* **42**, 8047–8053.

Prokop Z, Monincová M, Chaloupková R, Klvaňa M, Nagata Y, Janssen DB & Damborský J (2003) Catalytic Mechanism of the Haloalkane Dehalogenase LinB from *Sphingomonas paucimobilis* UT26. *J. Biol. Chem.* **278**, 45094–45100.