

Supporting Information: Structural analysis of Sars-Cov-2 main protease H172Y mutant

Alessia Guadagnin Pattaro¹, Teresa Dalle Nogare²

This file contains supplementary material for the analysis performed on the H172Y mutation of SARS-CoV-2 main protease in the APO form. Supporting graphs and tables are reported for both the structural and functional analysis performed.

¹ Master's degree in Quantitative and Computational Biology, University of Trento

² Master's degree in Physics, University of Trento

Structural analysis

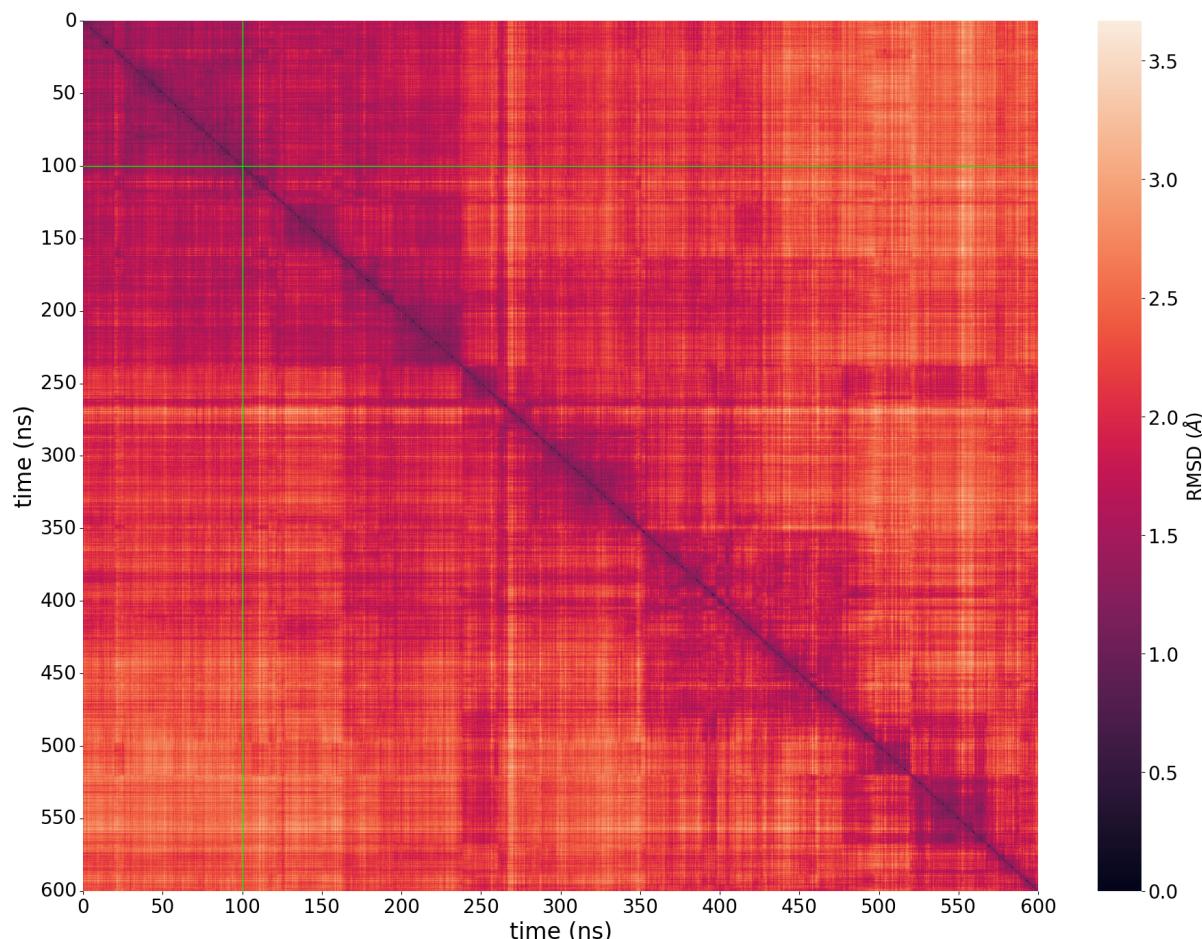


Figure S1: Pairwise RMSD of wild type 6XHU

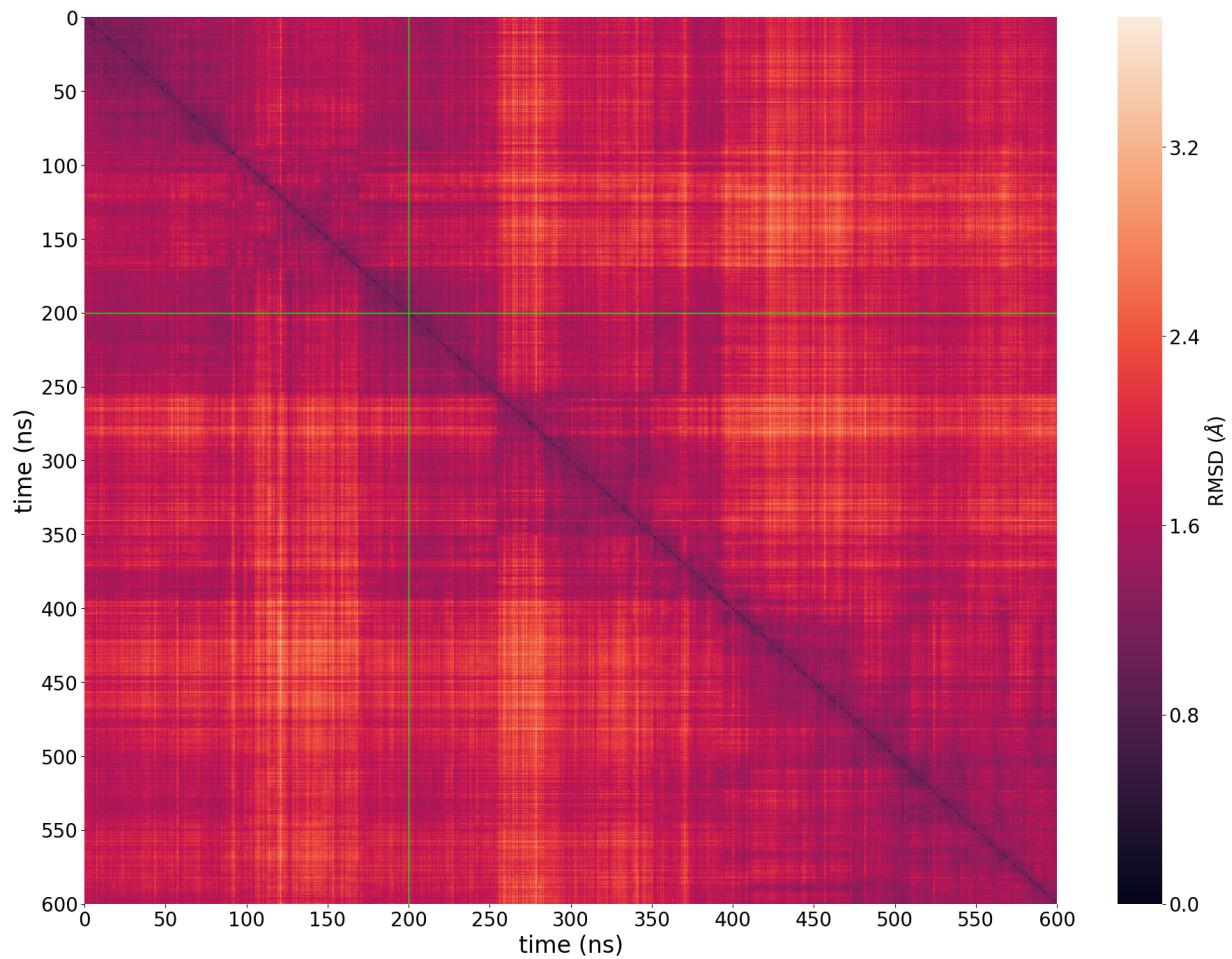
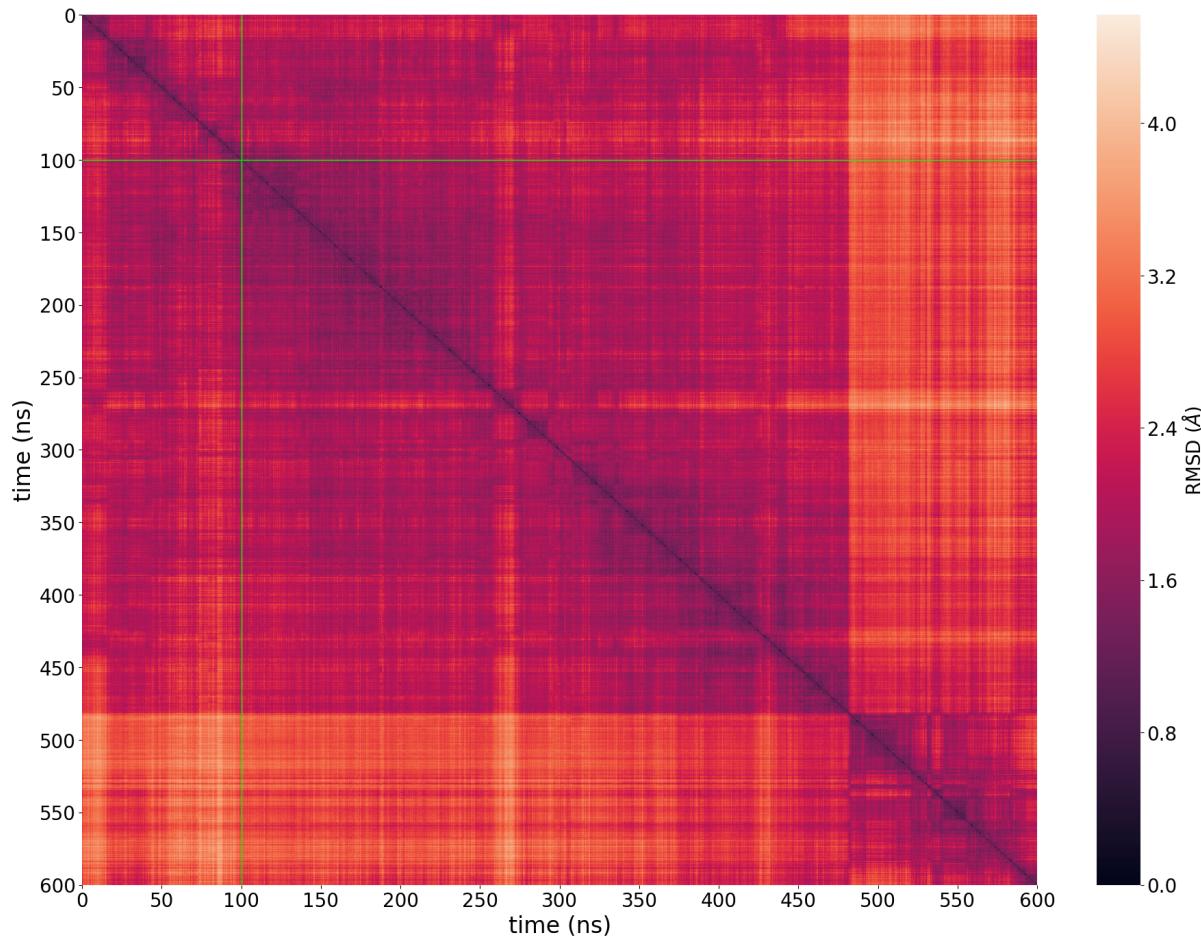
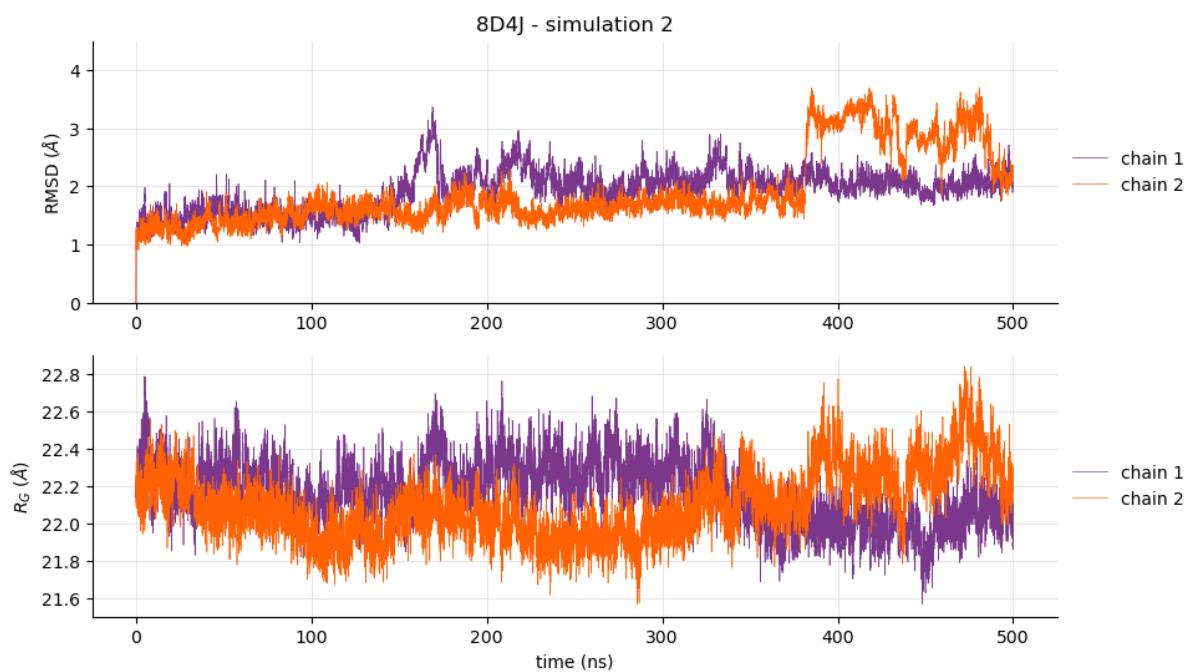
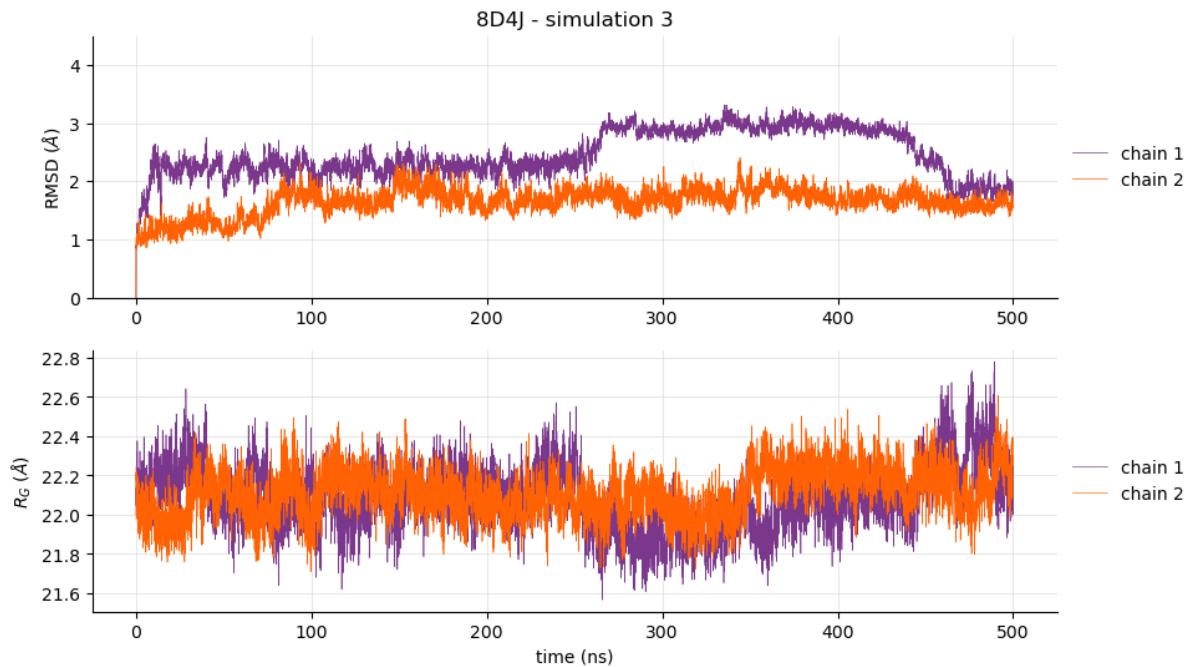
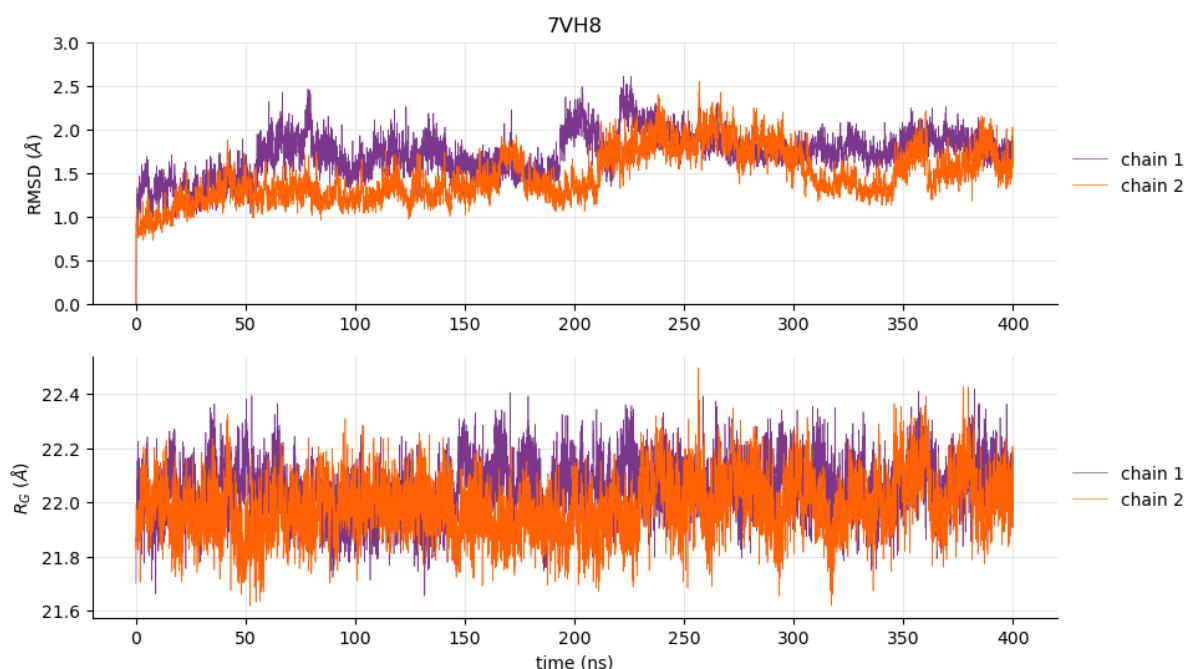


Figure S2: Pairwise RMSD of wild type 7VH8

**Figure S3:** Pairwise RMSD of 8D4J-2 simulation**Figure S4:** RMSD (top) and radius of gyration (bottom) of simulation 8D4J-2

**Figure S5:** RMSD (top) and radius of gyration (bottom) of simulation 8D4J-3**Figure S6:** RMSD (top) and radius of gyration (bottom) of simulation of WT 7VH8

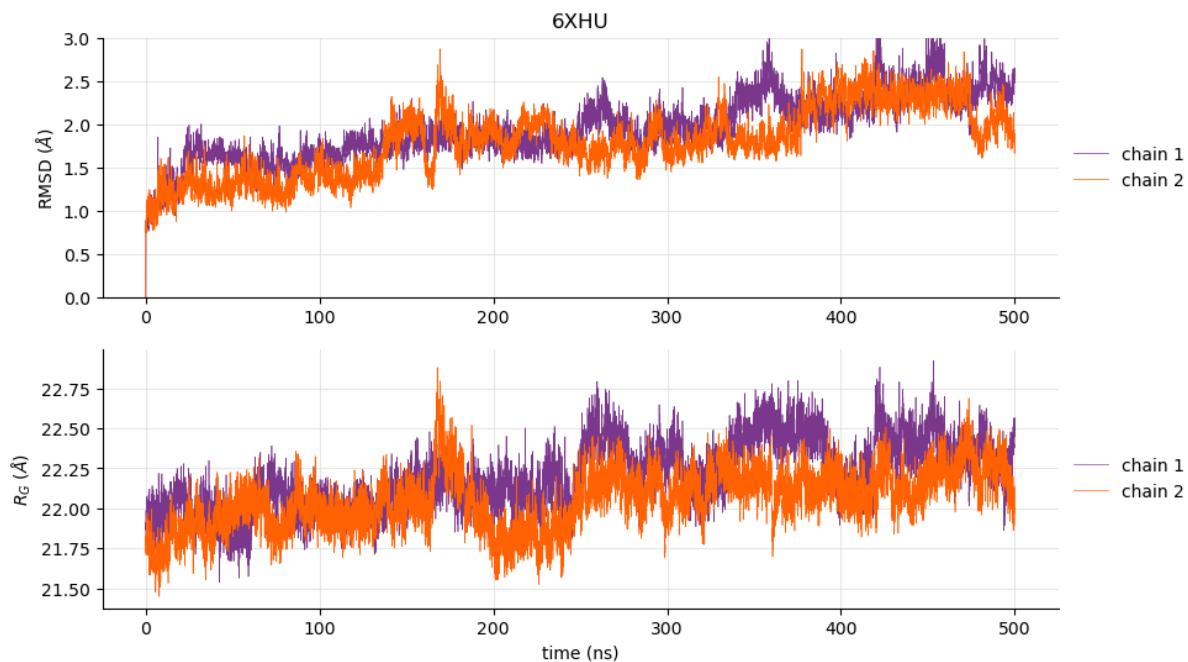


Figure S7: RMSD (top) and radius of gyration (bottom) of simulation of WT 6XHU

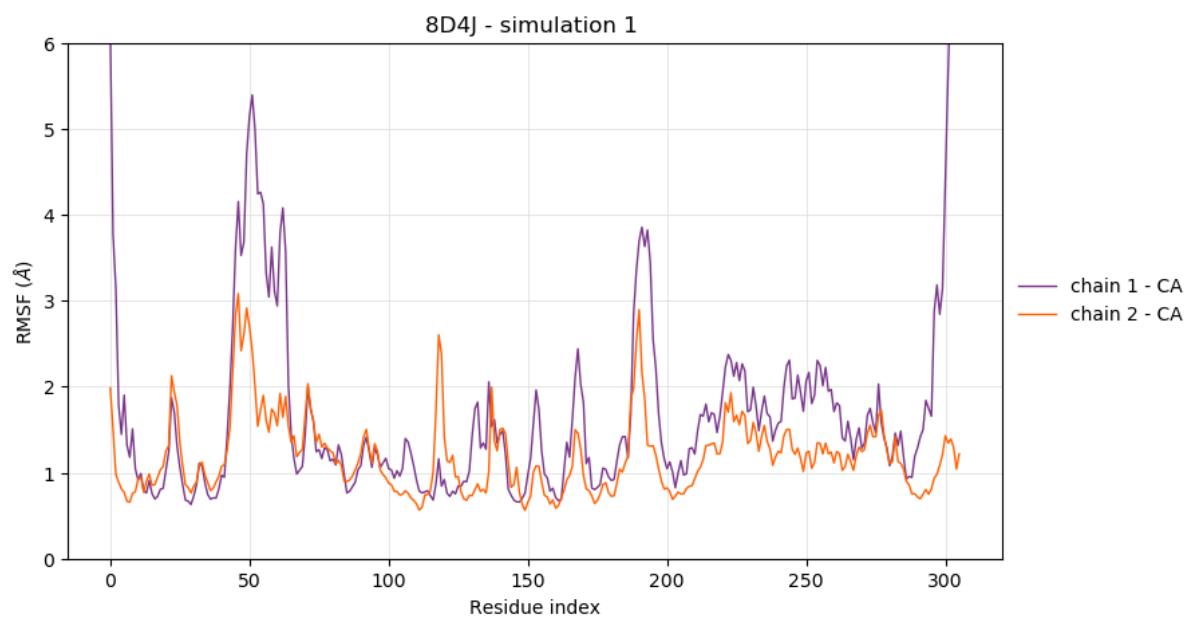


Figure S8: RMSF of chain 1 and chain 2 of 8D4J-1

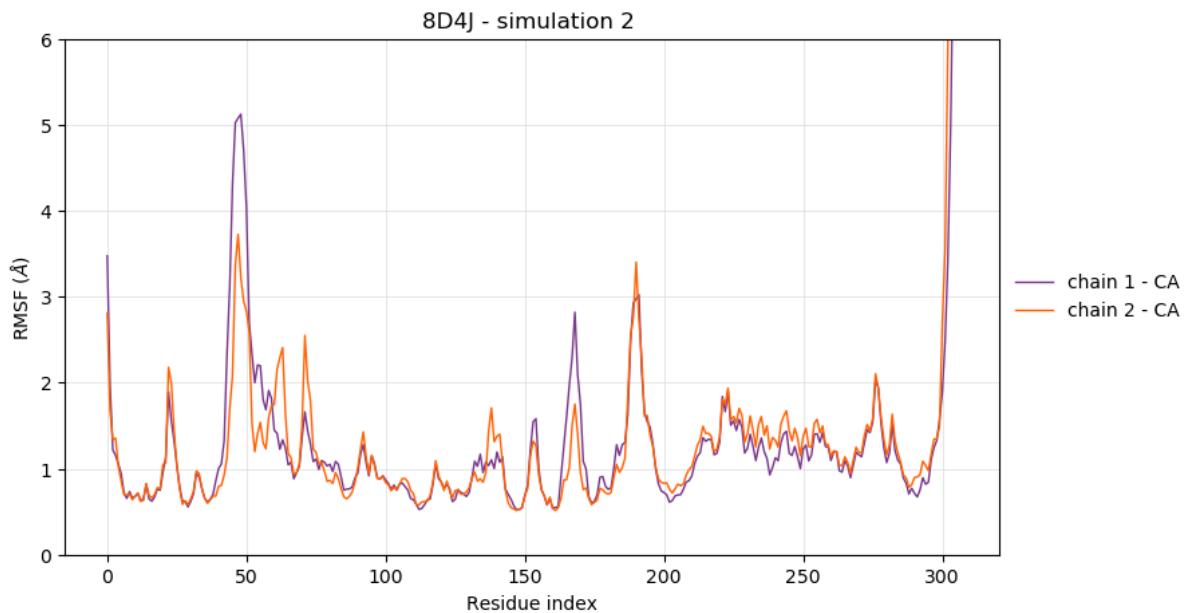


Figure S9: RMSF of chain 1 and chain 2 of 8D4J-2

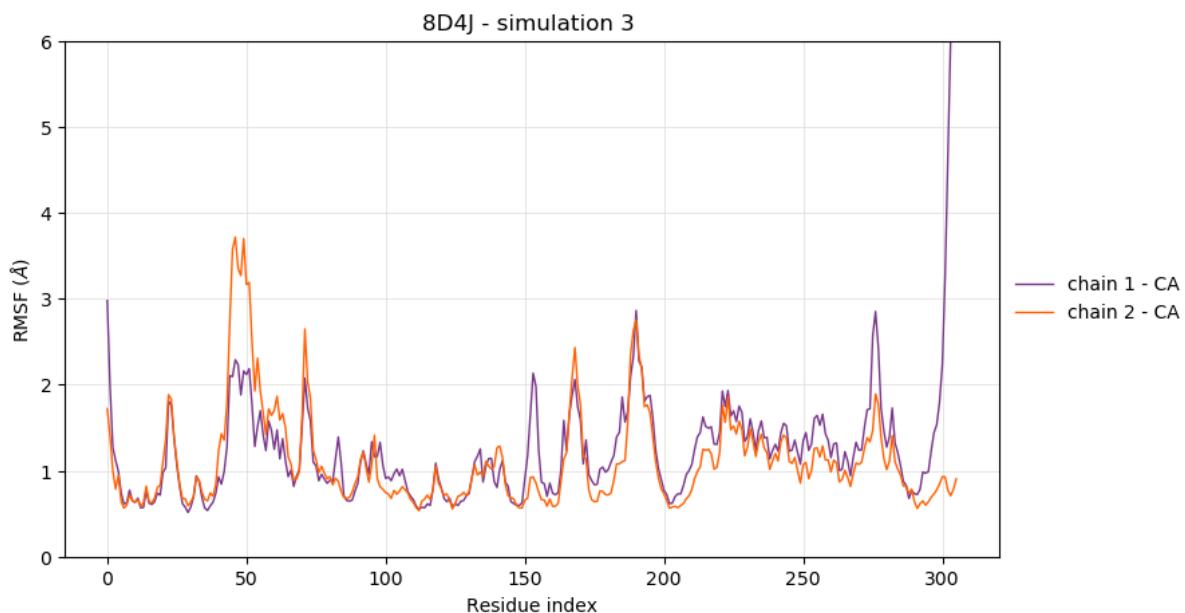


Figure S10: RMSF of chain 1 and chain 2 of 8D4J-3

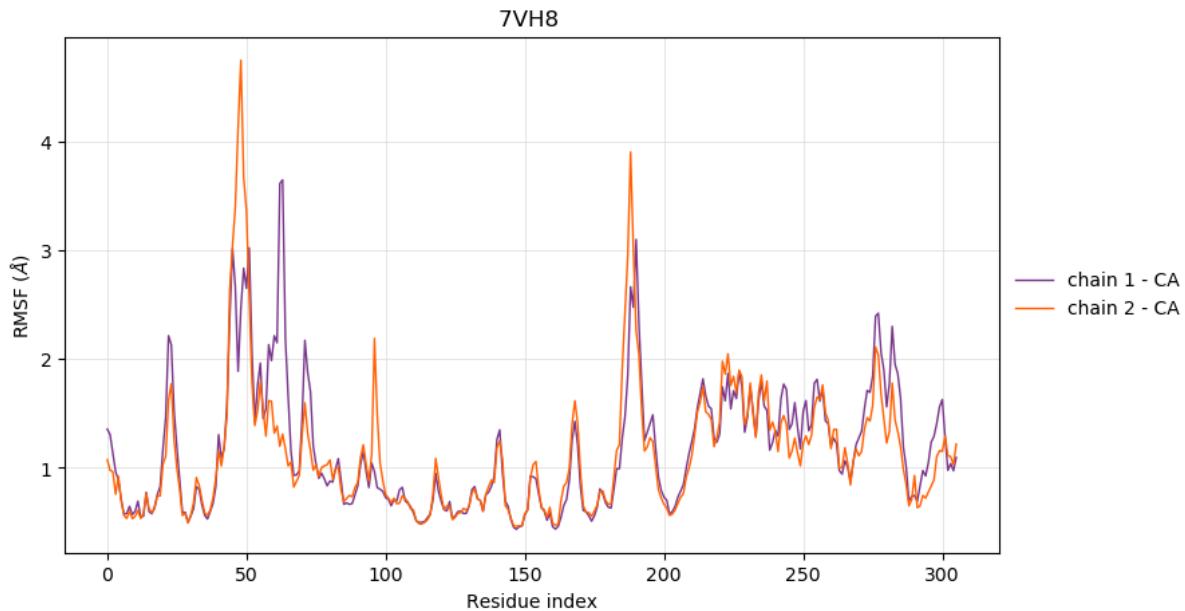


Figure S11: RMSF of chain 1 and chain 2 of simulation of wild type 7VH8

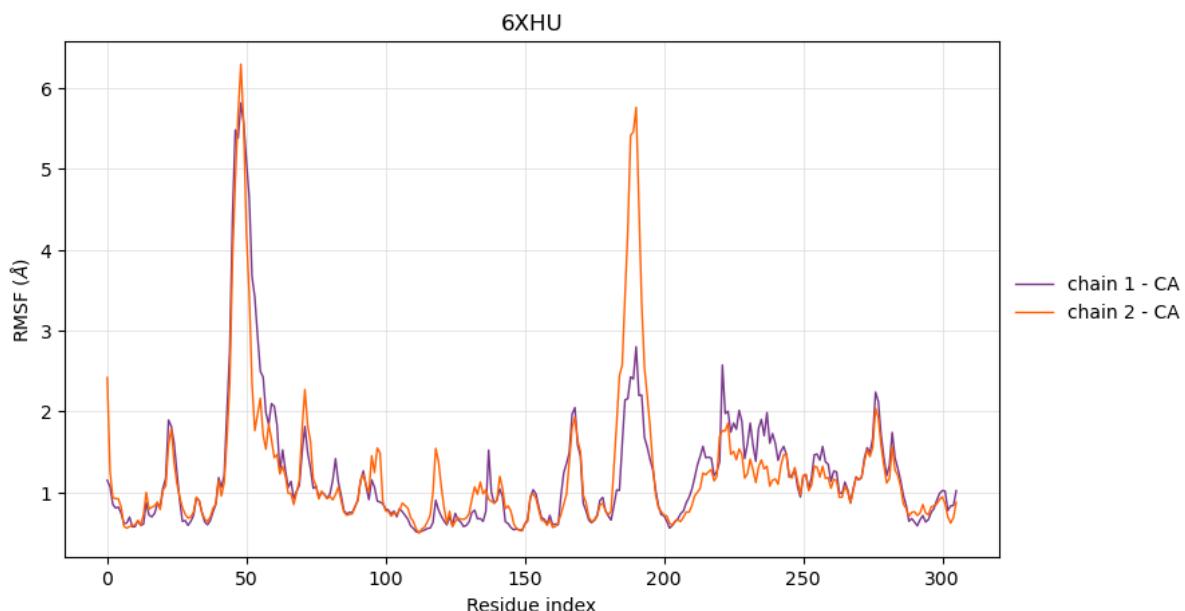


Figure S12: RMSF of chain 1 and chain 2 of simulation of wild type 6XHU

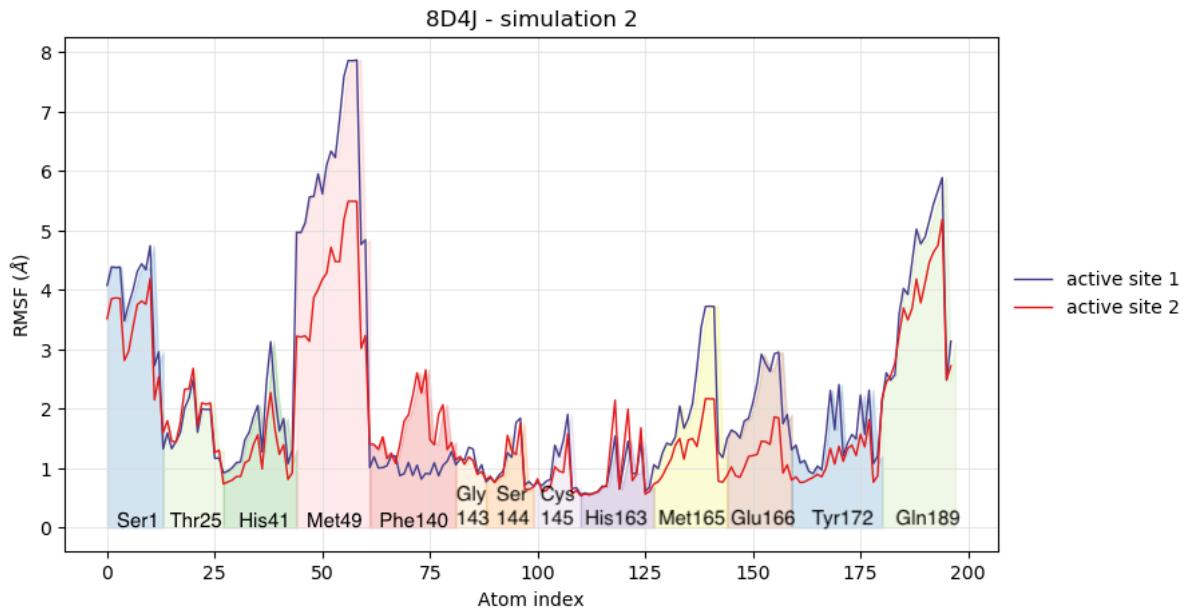


Figure S13: RMSF of residues of the active site of simulation 8D4J-2. The colours only help distinguish the different residues.

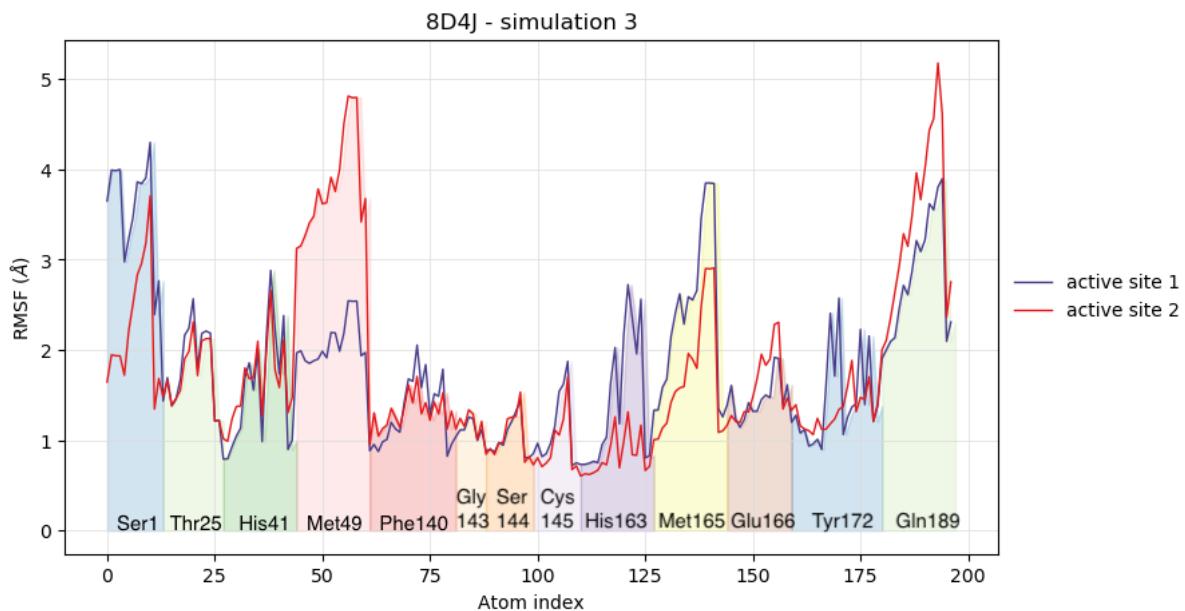


Figure S14: RMSF of residues of the active site of simulation 8D4J-3. The colours only help distinguish the different residues.

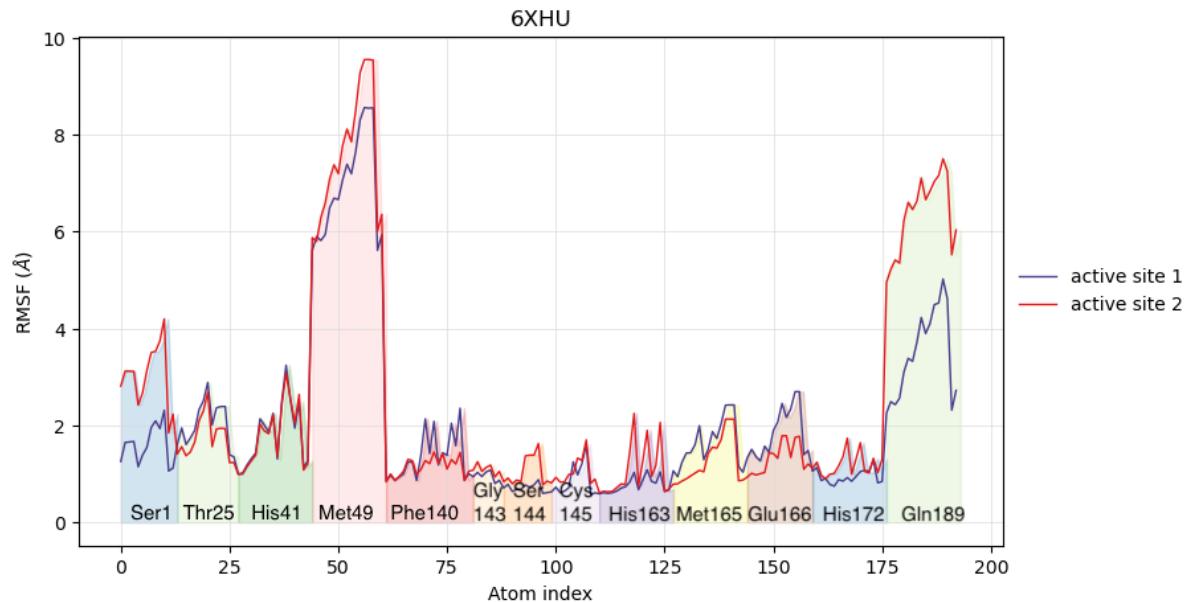


Figure S15: RMSF of residues of the active site of simulation of wild type 6XHU. The colours only help distinguish the different residues.

Functional analysis

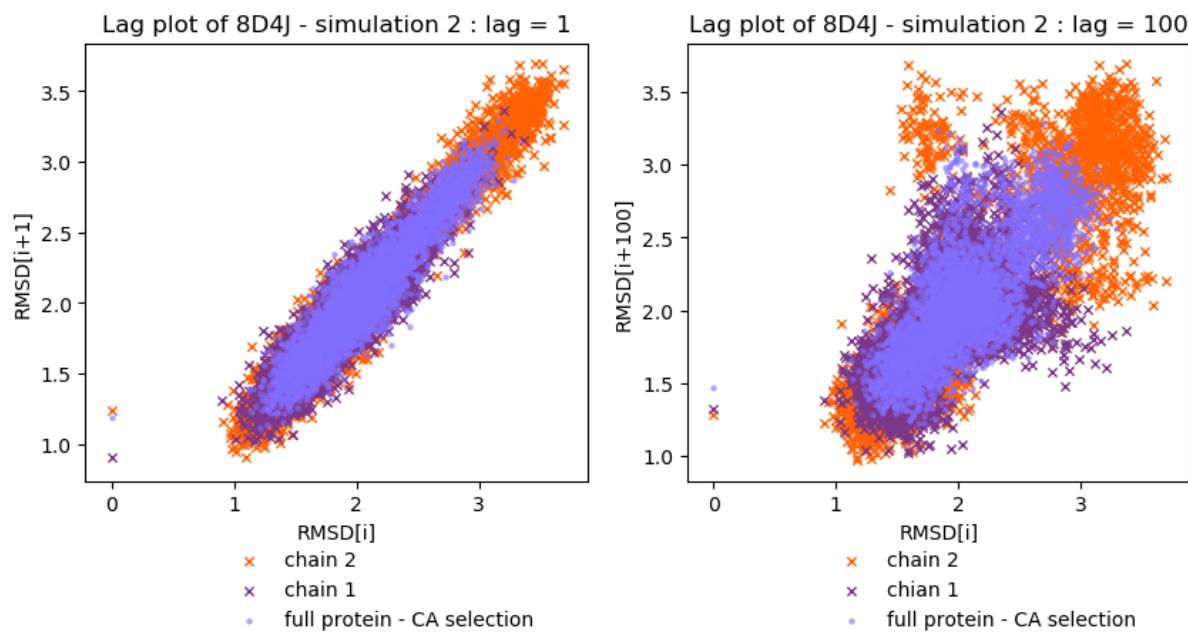


Figure S16: Lag plot for the second simulation of 8D4J. RMSD is plotted against the 1st order lag (left) and 100th order lag (right)

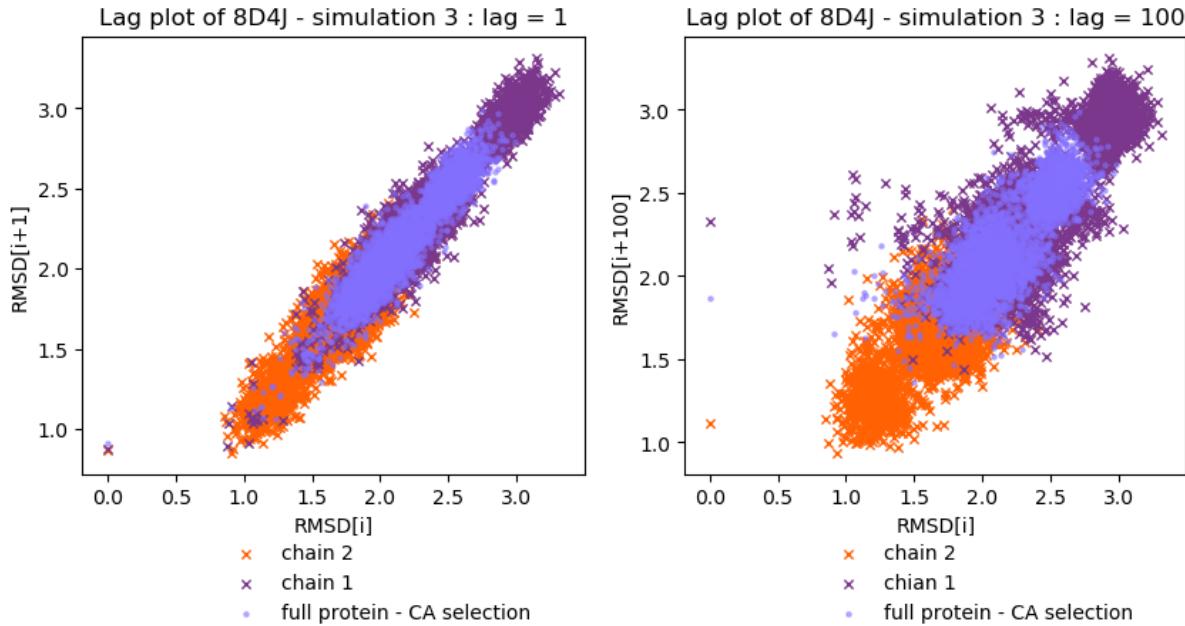


Figure S17: Lag plot for the third simulation of 8D4J. RMSD is plotted against the 1st order lag (left) and 100th order lag (right)

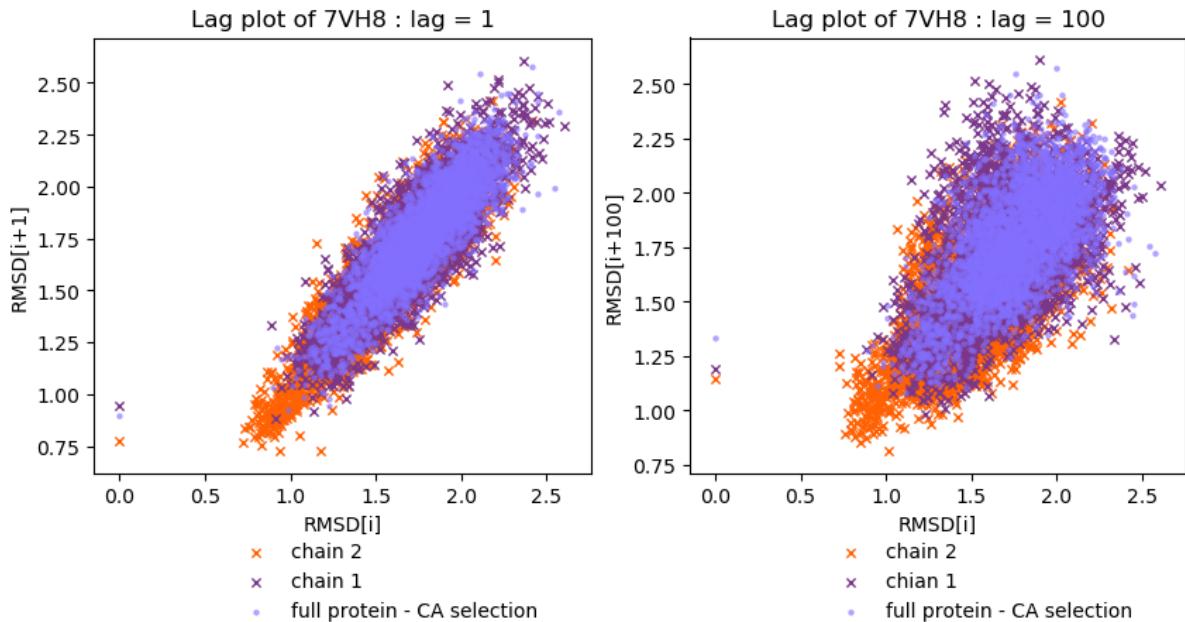


Figure S18: Lag plot for the simulation of 7VH8. RMSD is plotted against the 1st order lag (left) and 100th order lag (right)

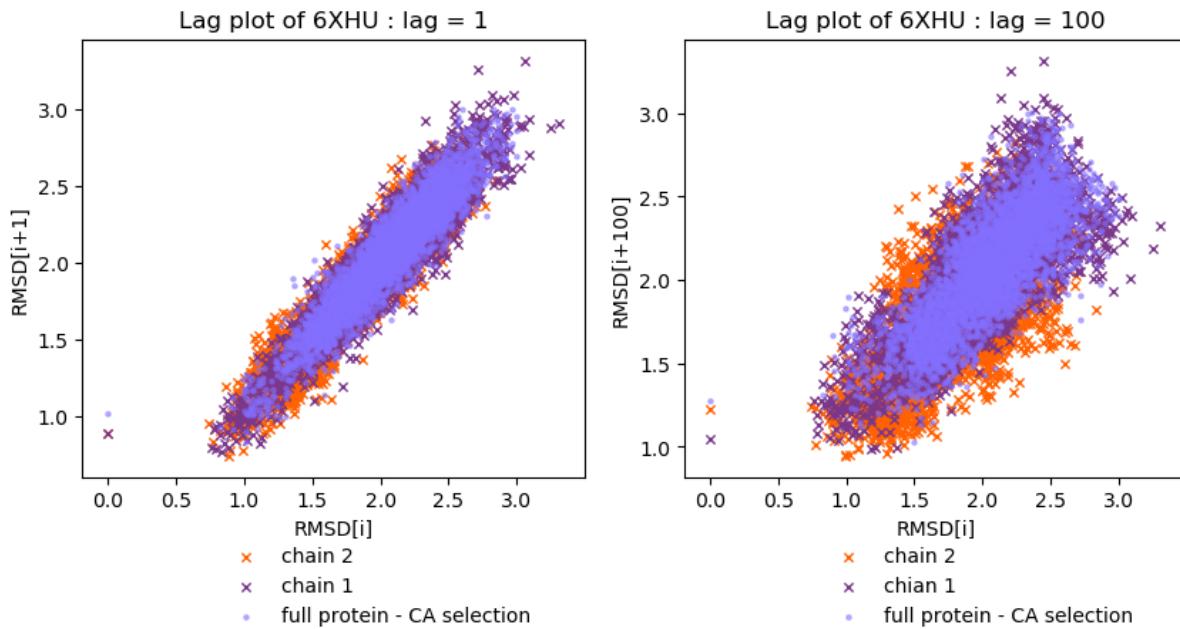


Figure S19: Lag plot for the simulation of 6XHU. RMSD is plotted against the 1st order lag (left) and 100th order lag (right)

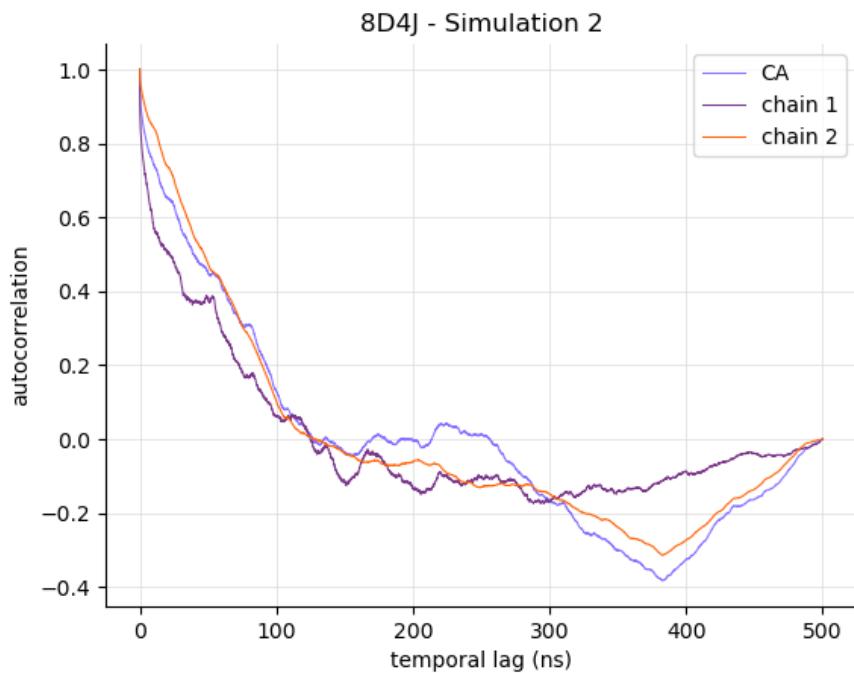


Figure S20: Autocorrelation function for the RMSD of the second simulation of 8D4J for different selections : full protein - CA, chain 1 and chain 2.

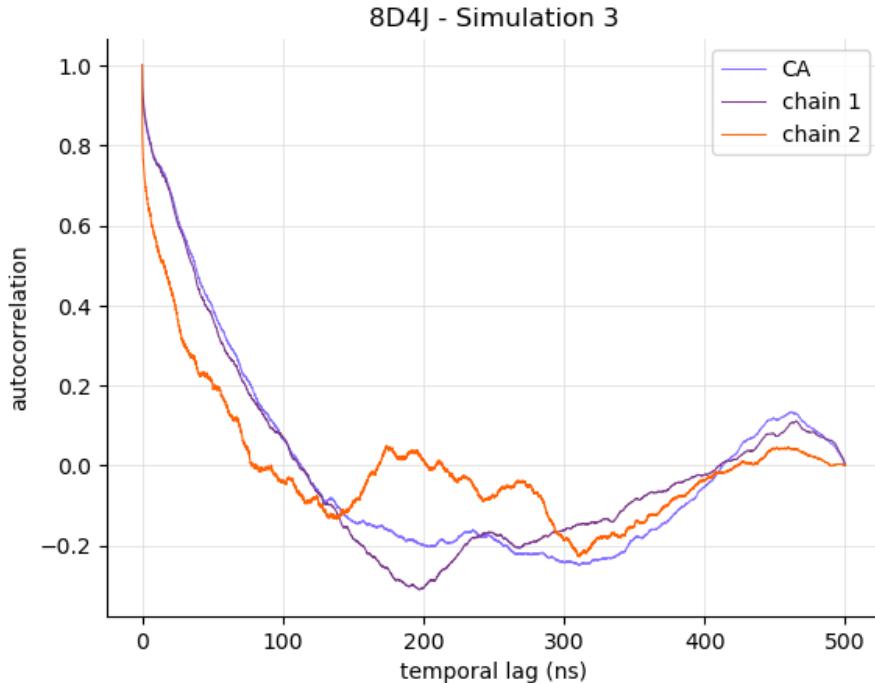


Figure S21: Autocorrelation function for the RMSD of the third simulation of 8D4J for different selections : full protein - CA, chain 1 and chain 2.

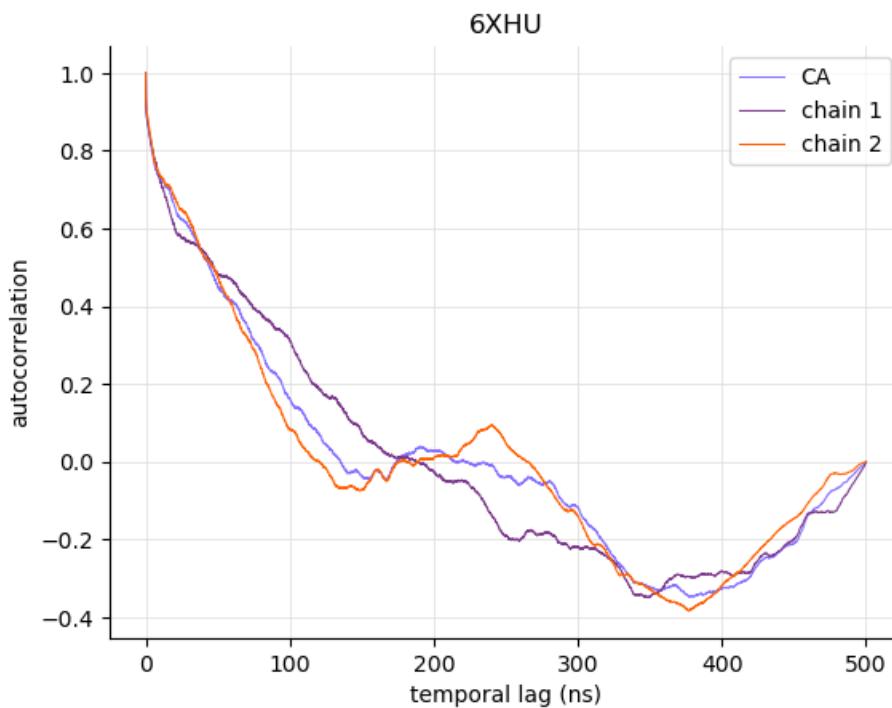


Figure S22: Autocorrelation function for the RMSD of the WT 6XHU for different selections : full protein - CA, chain 1 and chain 2.

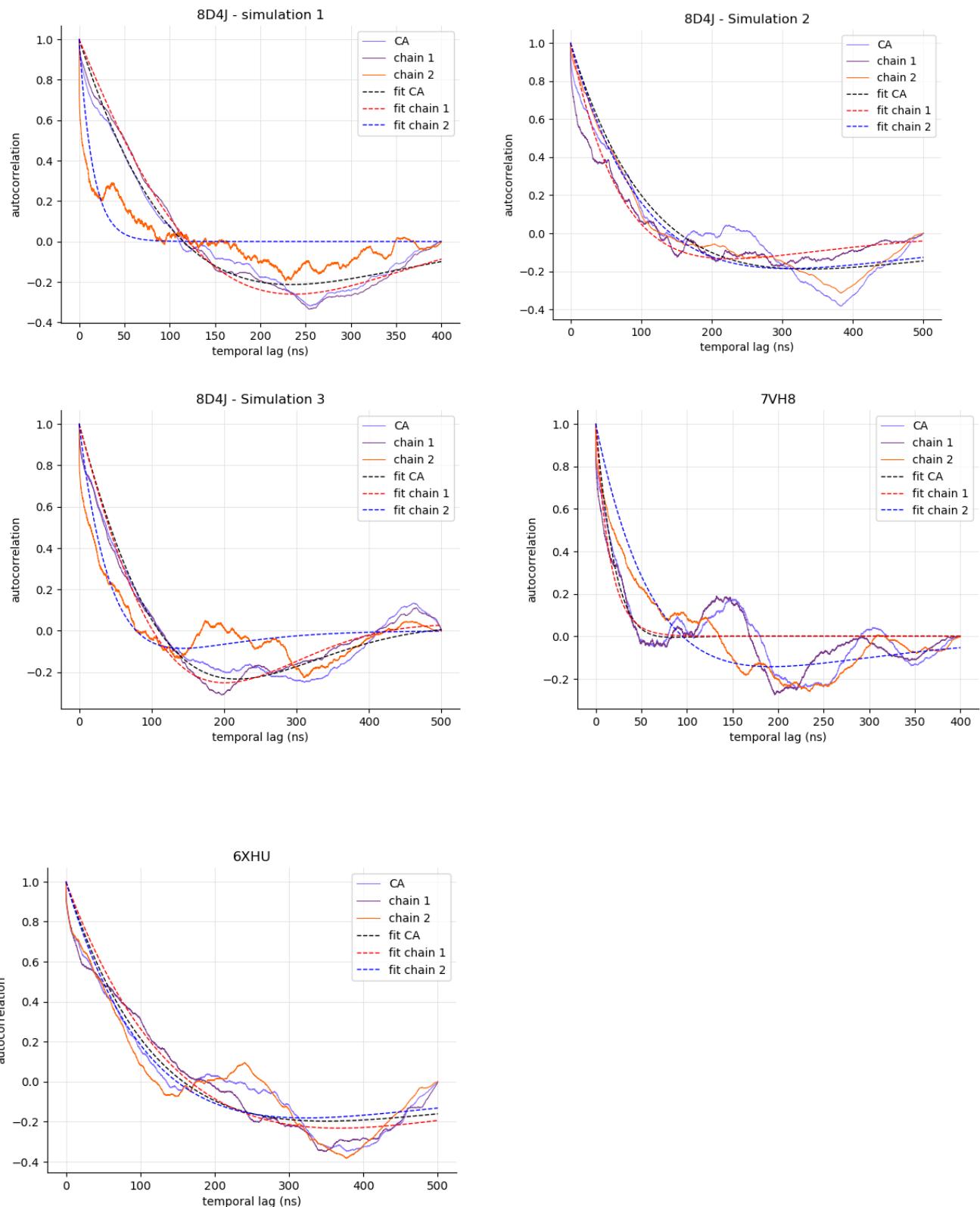


Figure S23: Autocorrelation function for the RMSD plotted together with the fit performed for each selection. From the fit, the autocorrelation time is determined.

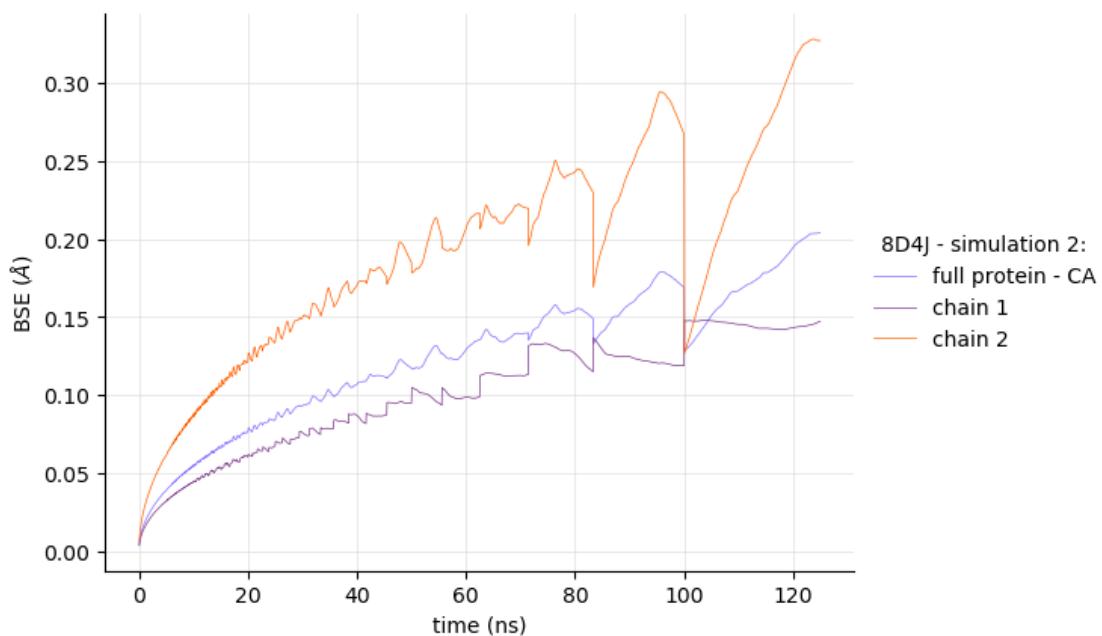


Figure S24: BSE plot calculated from the RMSD time-series of the second simulation of 8D4J.

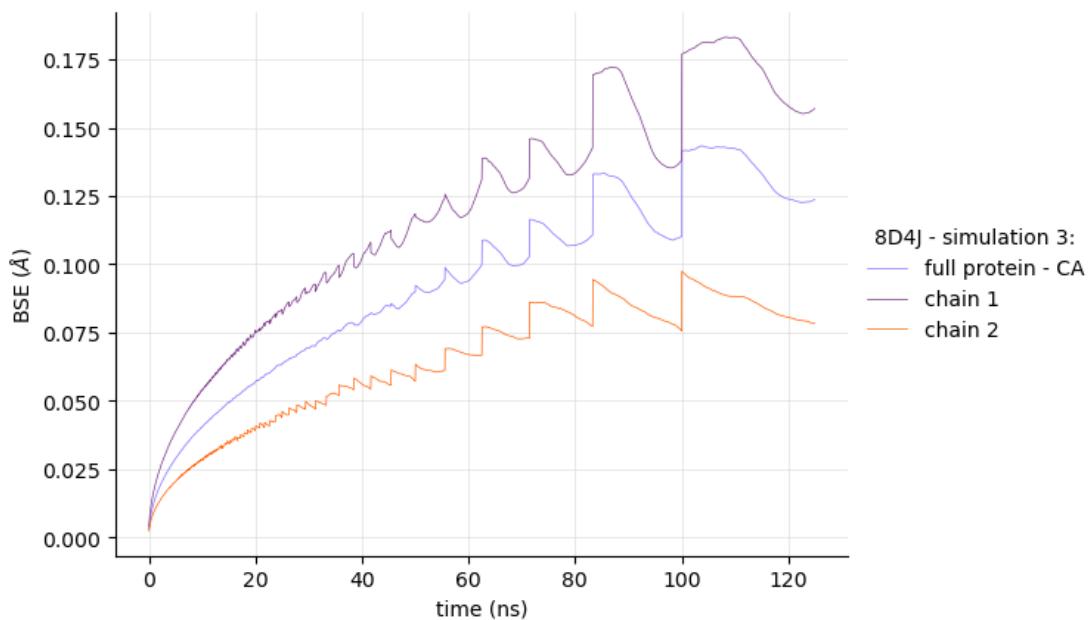


Figure S25: BSE plot calculated from the RMSD time-series of the third simulation of 8D4J.

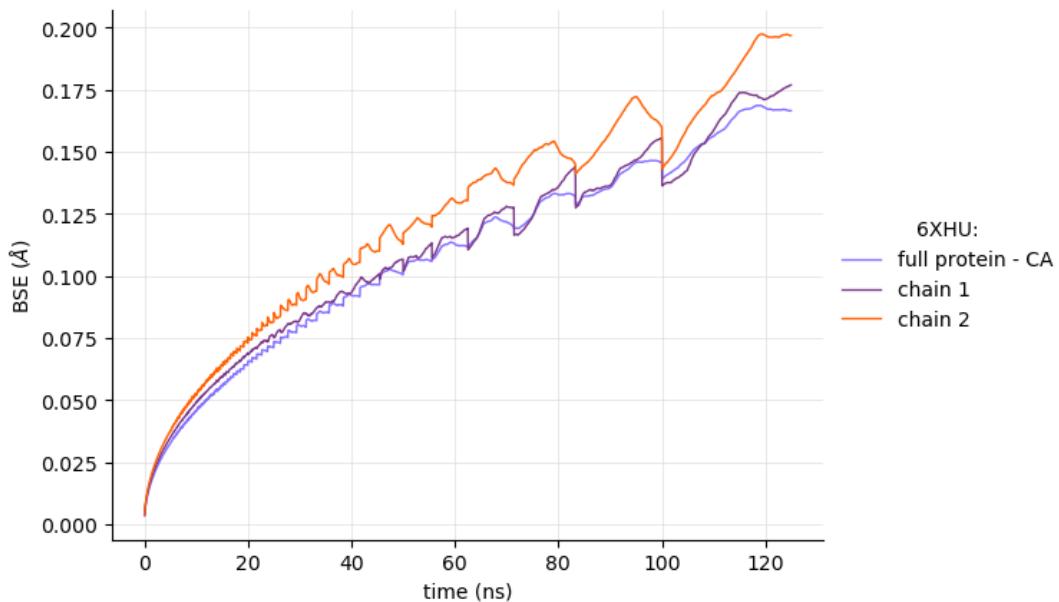


Figure S26: BSE plot calculated from the RMSD time-series of the WT 6XHU.

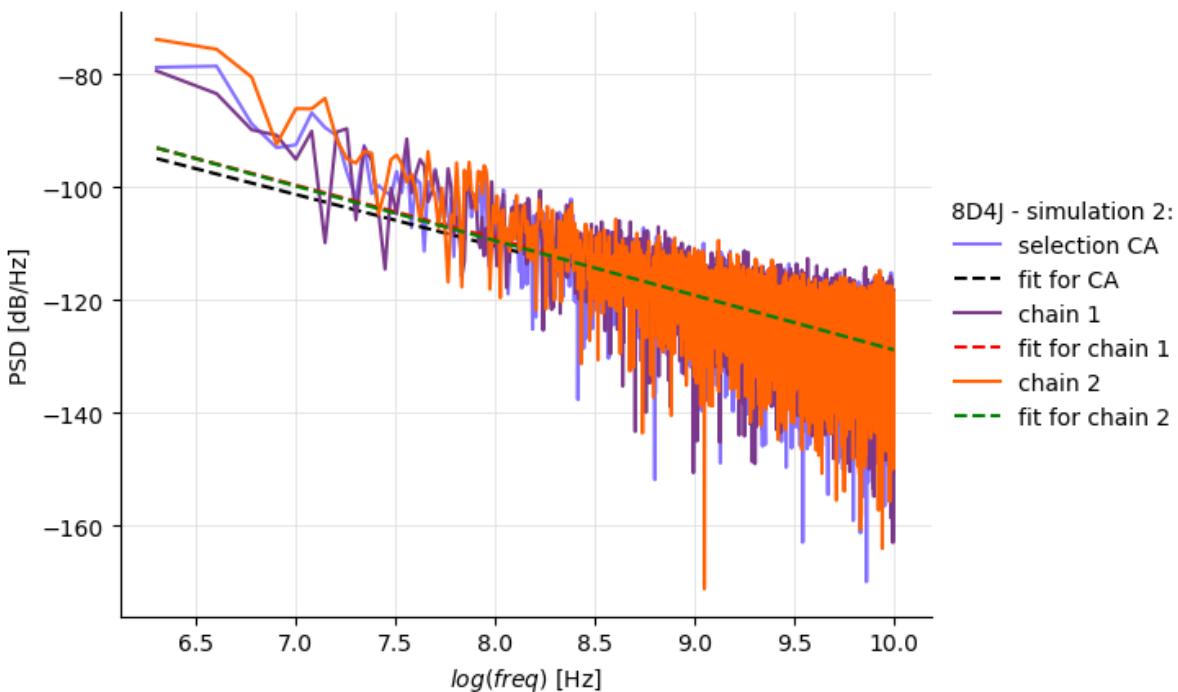


Figure S27: Power spectral density of the noisy part of the RMSD for the second simulation of 8D4J calculated for various selections : full protein, chain 1 and chain 2. The linear fit used to derive the autocorrelation time is also reported with a dashed line.

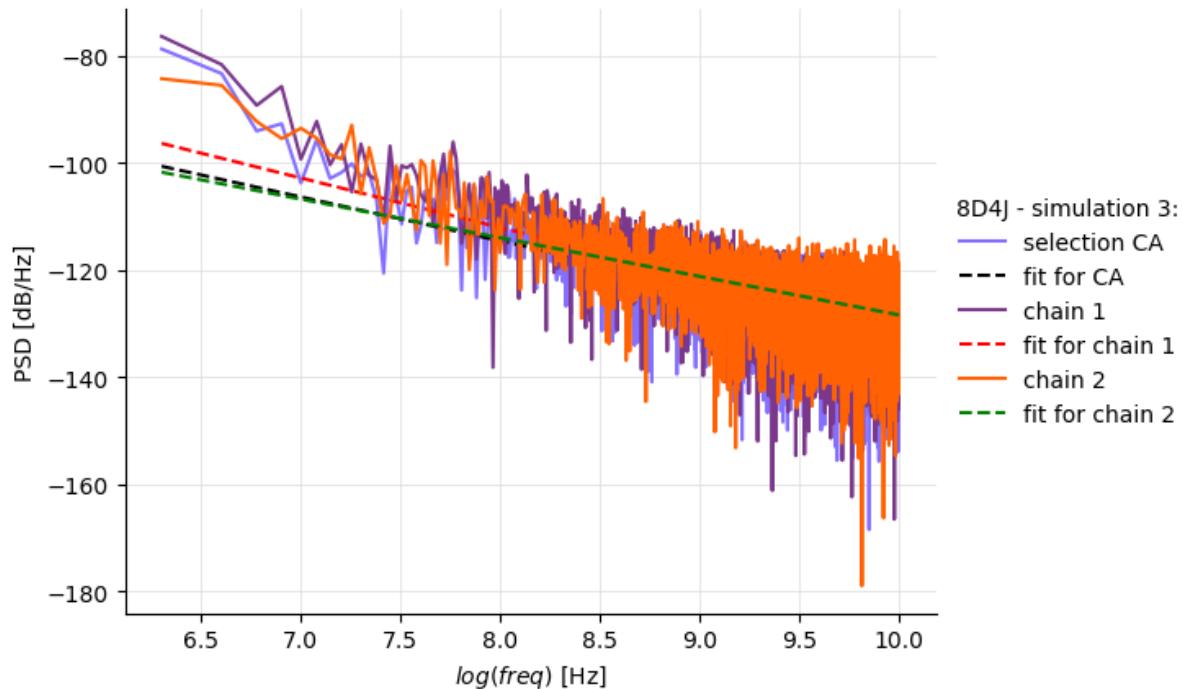


Figure S28: Power spectral density of the noisy part of the RMSD for the third simulation of 8D4J calculated for various selections : full protein, chain 1 and chain 2. The linear fit used to derive the autocorrelation time is also reported with a dashed line.

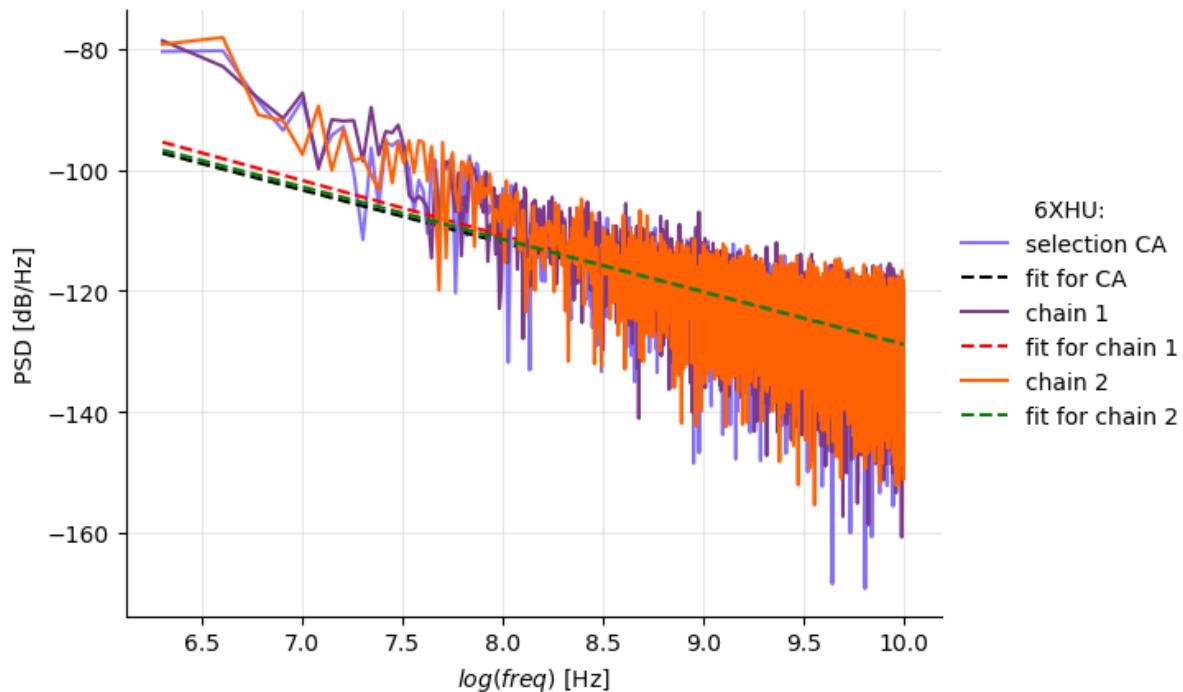
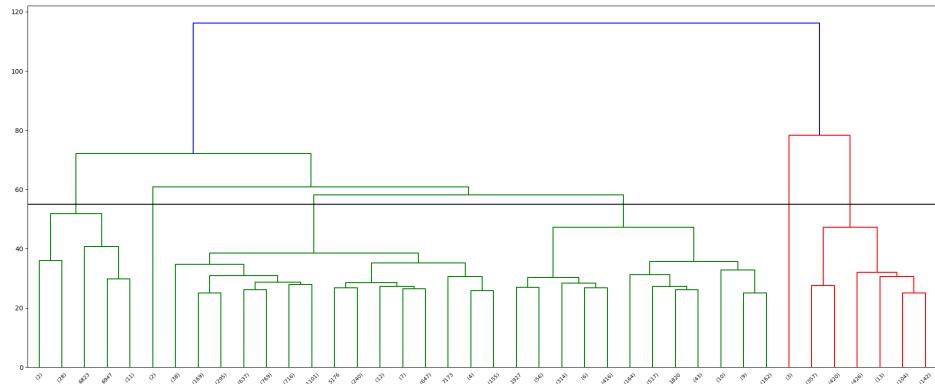
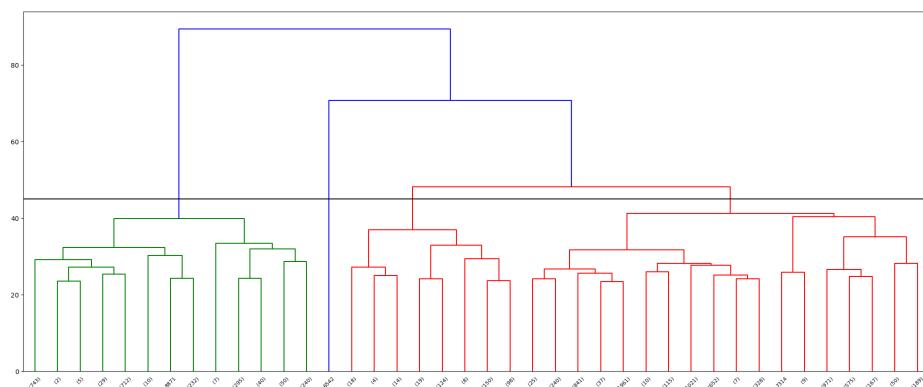


Figure S29: Power spectral density of the noisy part of the RMSD for 6XHU calculated for various selections : full protein, chain 1 and chain 2. The linear fit used to derive the autocorrelation time is also reported with a dashed line.

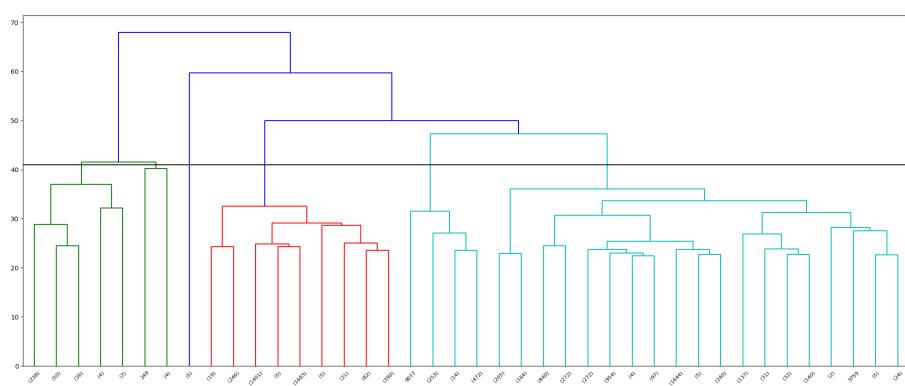
(a)



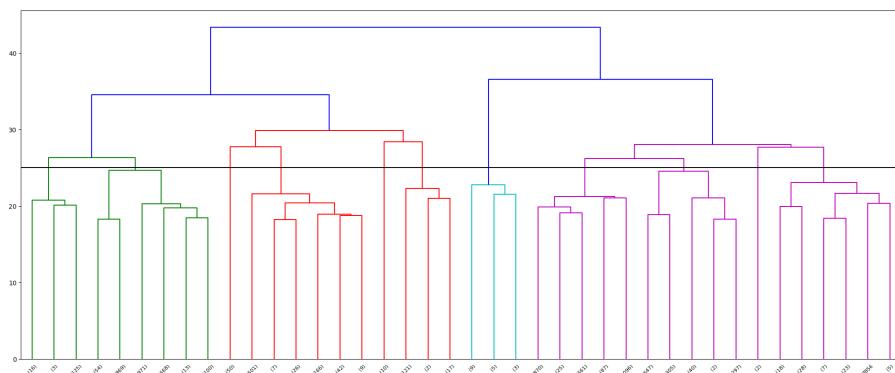
(b)



(c)



(d)



(e)

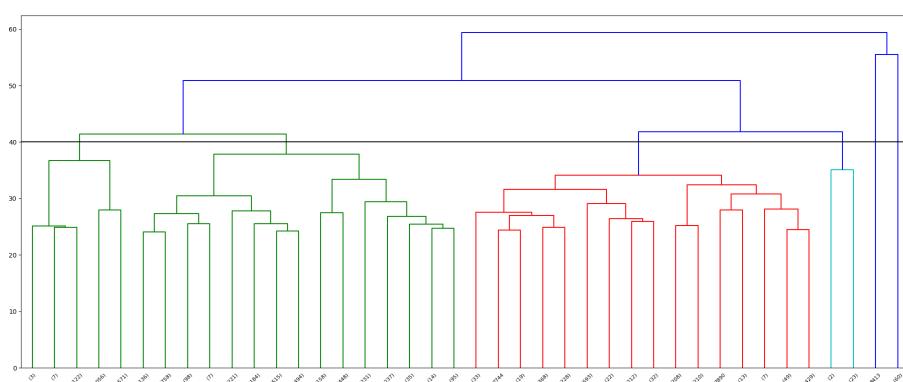


Figure S30: Dendograms constructed from the linkage matrix for : **(a)** 8D4J-1, **(b)** 8D4J-2, **(c)** 8D4J-3, **(d)** 7VH8, **(e)** 6XHU

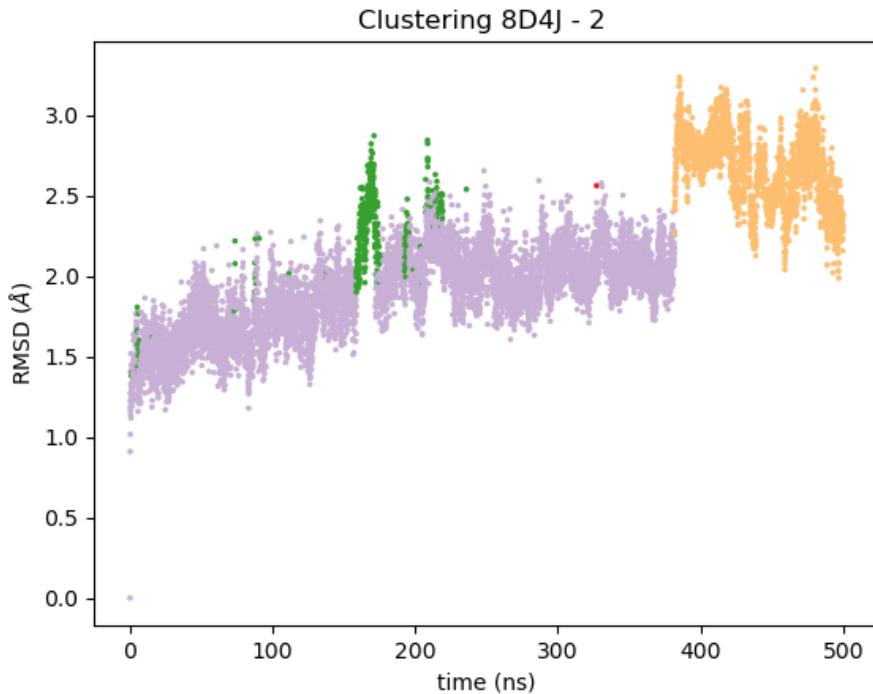


Figure S31: RMSD time series obtained from production data with highlighted groups corresponding to different protein configurations obtained from the cluster analysis for the second simulation of 8D4J

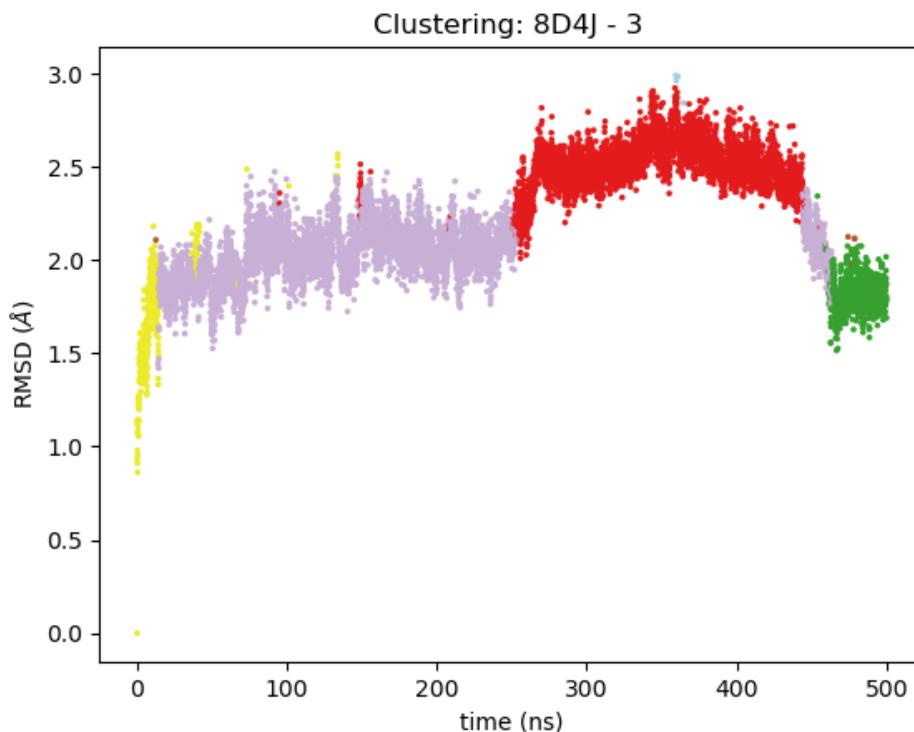


Figure S32: RMSD time series obtained from production data with highlighted groups corresponding to different protein configurations obtained from the cluster analysis for the third simulation of 8D4J

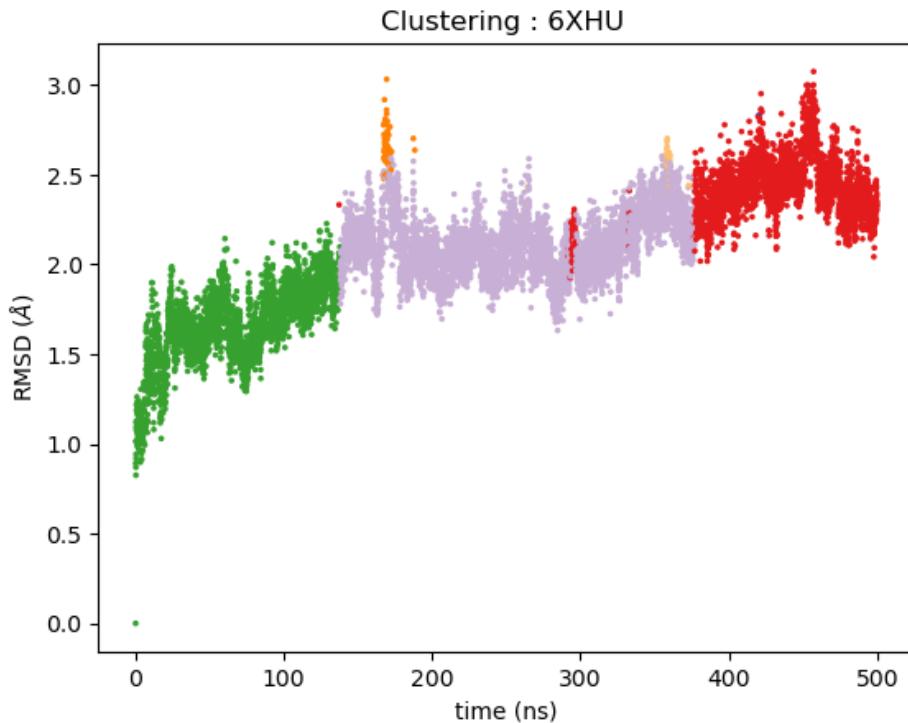


Figure S33: RMSD time series obtained from production data with highlighted groups corresponding to different protein configurations obtained from the cluster analysis for the WT 6XHU

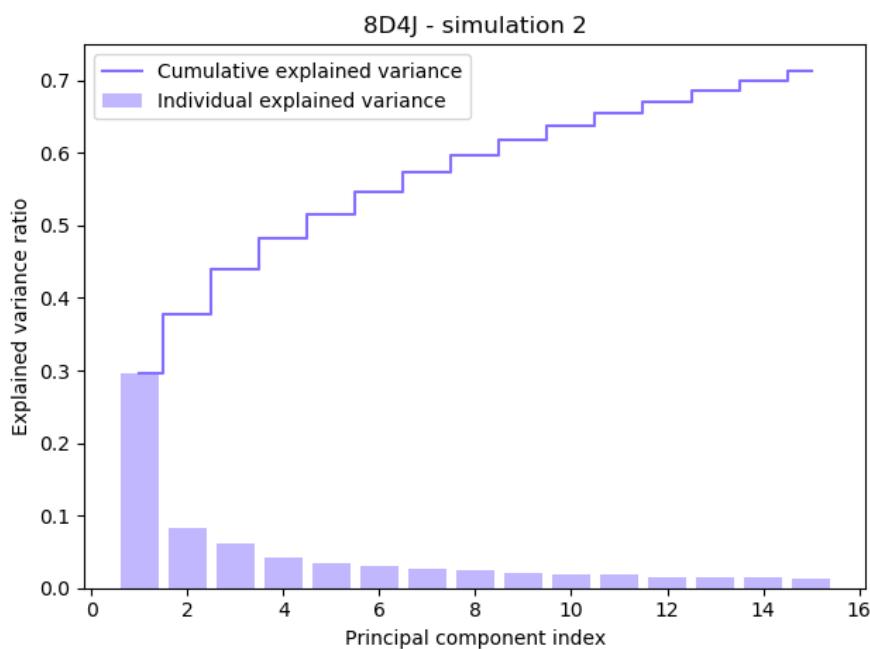


Figure S34: Individual and cumulative variance explained by the first 15 PCs for the second simulation of 8D4J

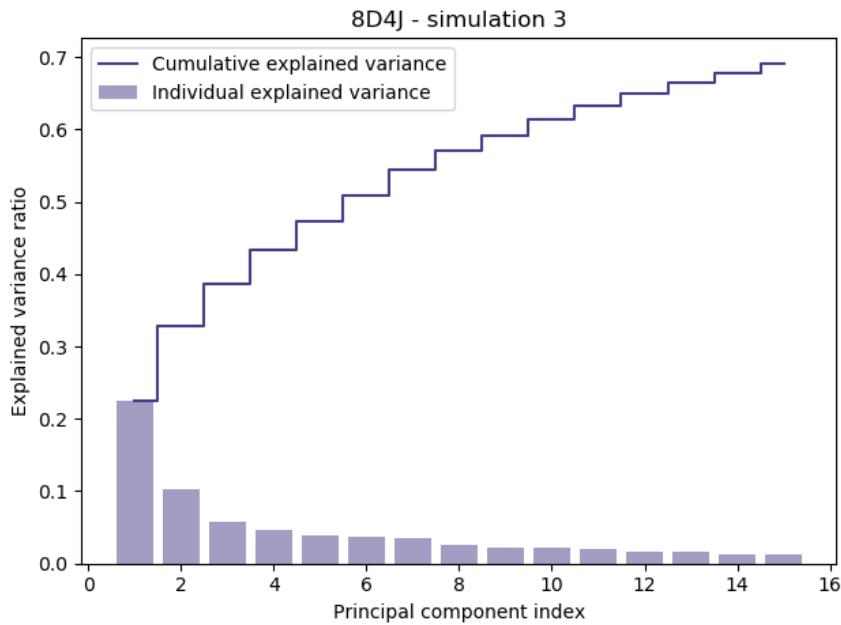


Figure S35: Individual and cumulative variance explained by the first 15 PCs for the third simulation of 8D4J

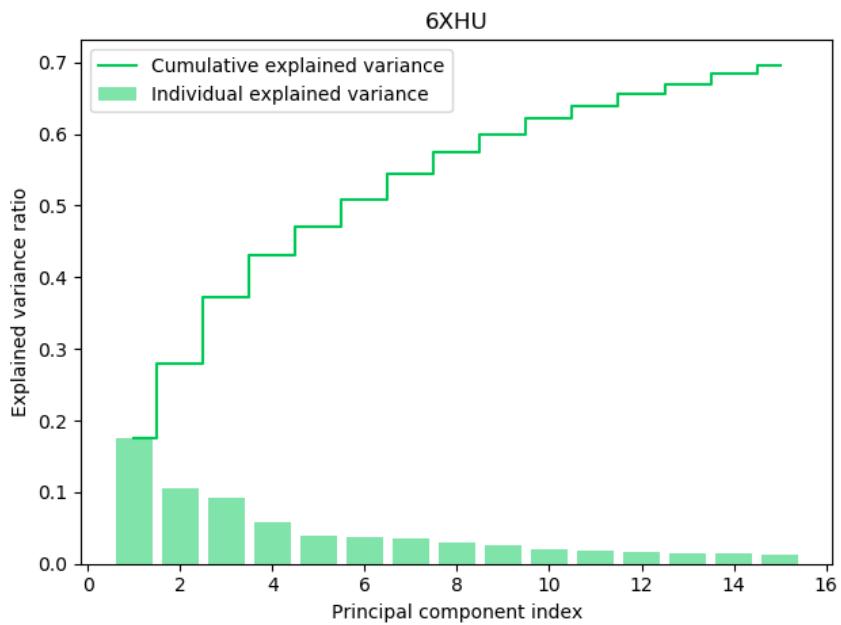


Figure S36: Individual and cumulative variance explained by the first 15 PCs for the simulation of 6XHU

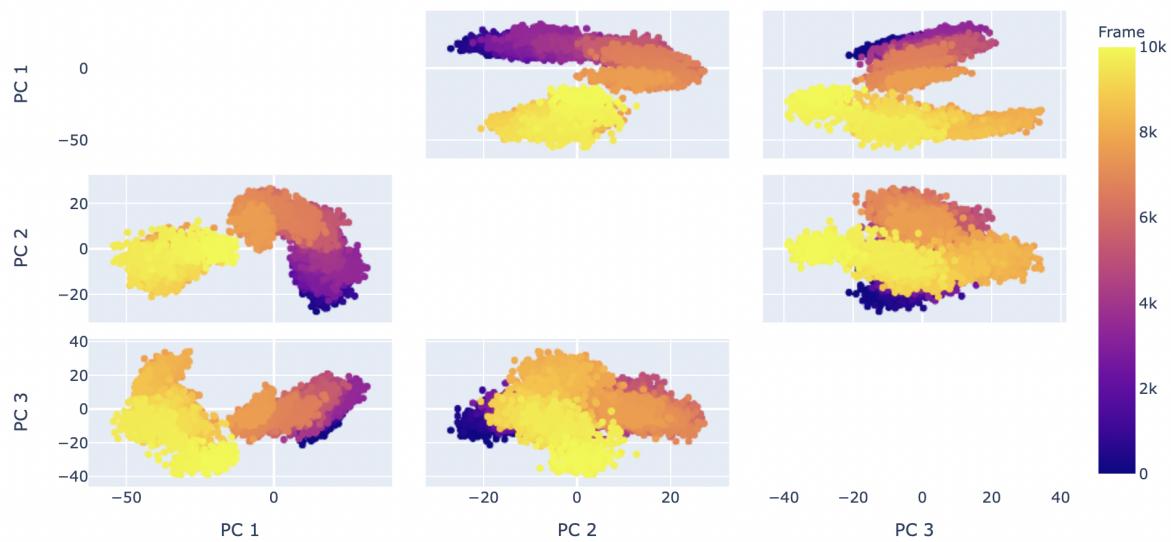


Figure S37: Trajectory of the second simulation of 8D4J projected onto the first three PCs

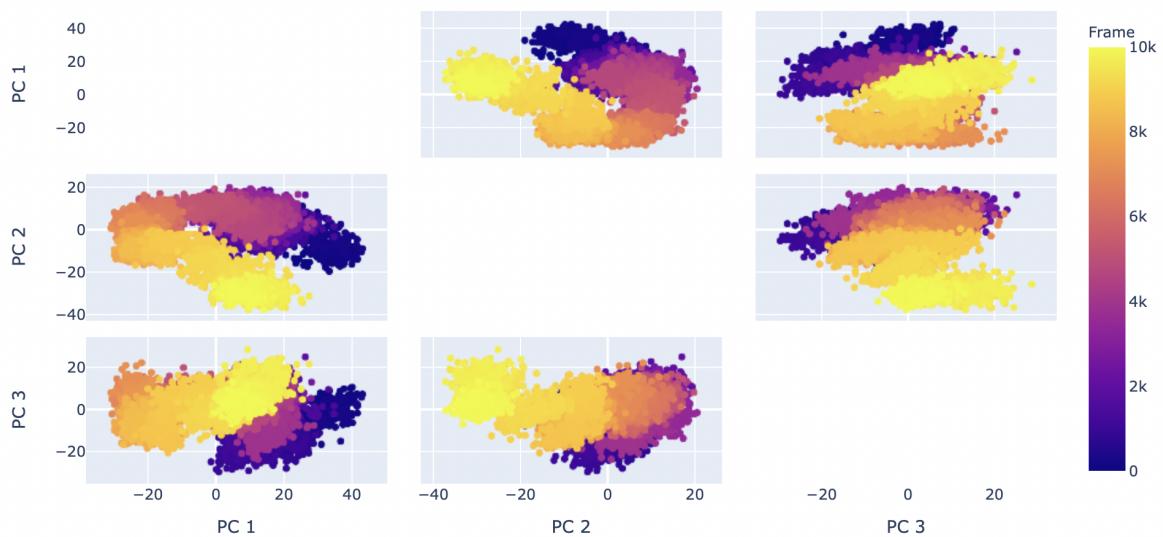


Figure S38: Trajectory of the third simulation of 8D4J projected onto the first three PCs

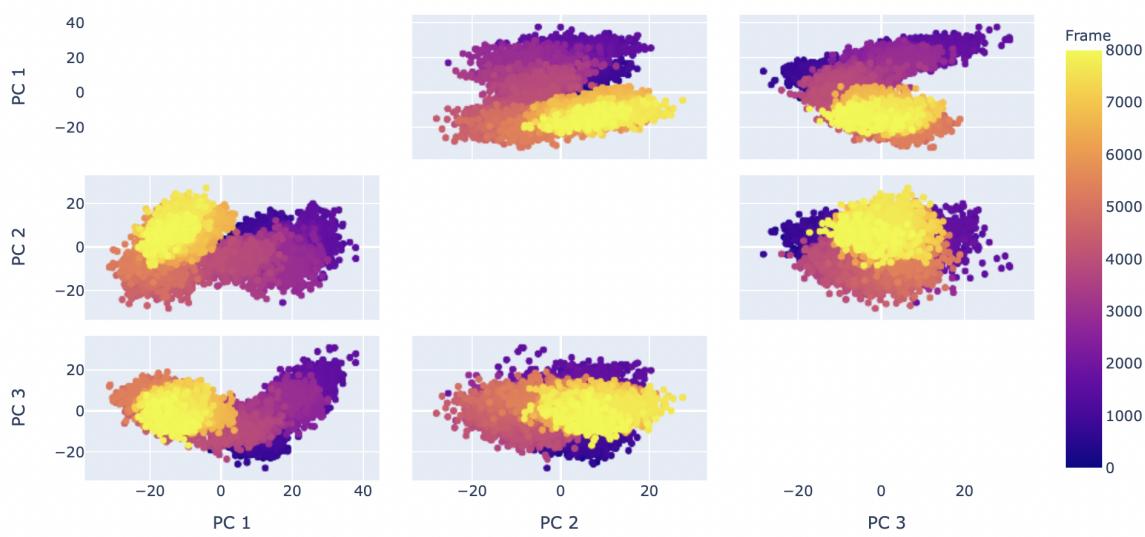


Figure S39: Trajectory of the WT 7VH8 projected onto the first three PCs

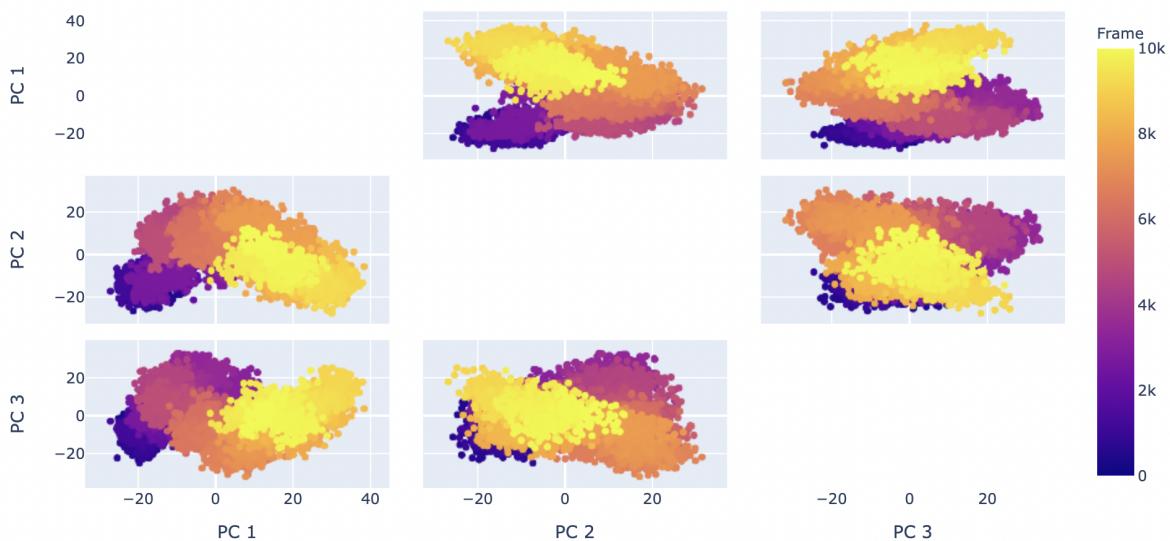


Figure S40: Trajectory of the WT 6XHU projected onto the first three PCs