Supporting Information for

Exploring the Temperature Stability of CRISPR–Cas12b using Molecular Dynamics Simulations

Yinhao Jia[1], Katelynn Horvath[2], Santosh R. Rananaware[1], Piyush K. Jain[1,3,4], Janani Sampath[1,*]

[1]Department of Chemical Engineering, University of Florida, Gainesville, FL, USA

[2] Department of Chemical and Biomolecular Engineering, University of Connecticut, Storrs, CT, USA

[3] Department of Molecular Genetics and Microbiology, College of Medicine, University of Florida, Gainesville, FL, USA

[4] Health Cancer Center, University of Florida, Gainesville, FL, USA

* jsampath@ufl.edu

Table of content

Principal Component Analysis

Structure and Stability

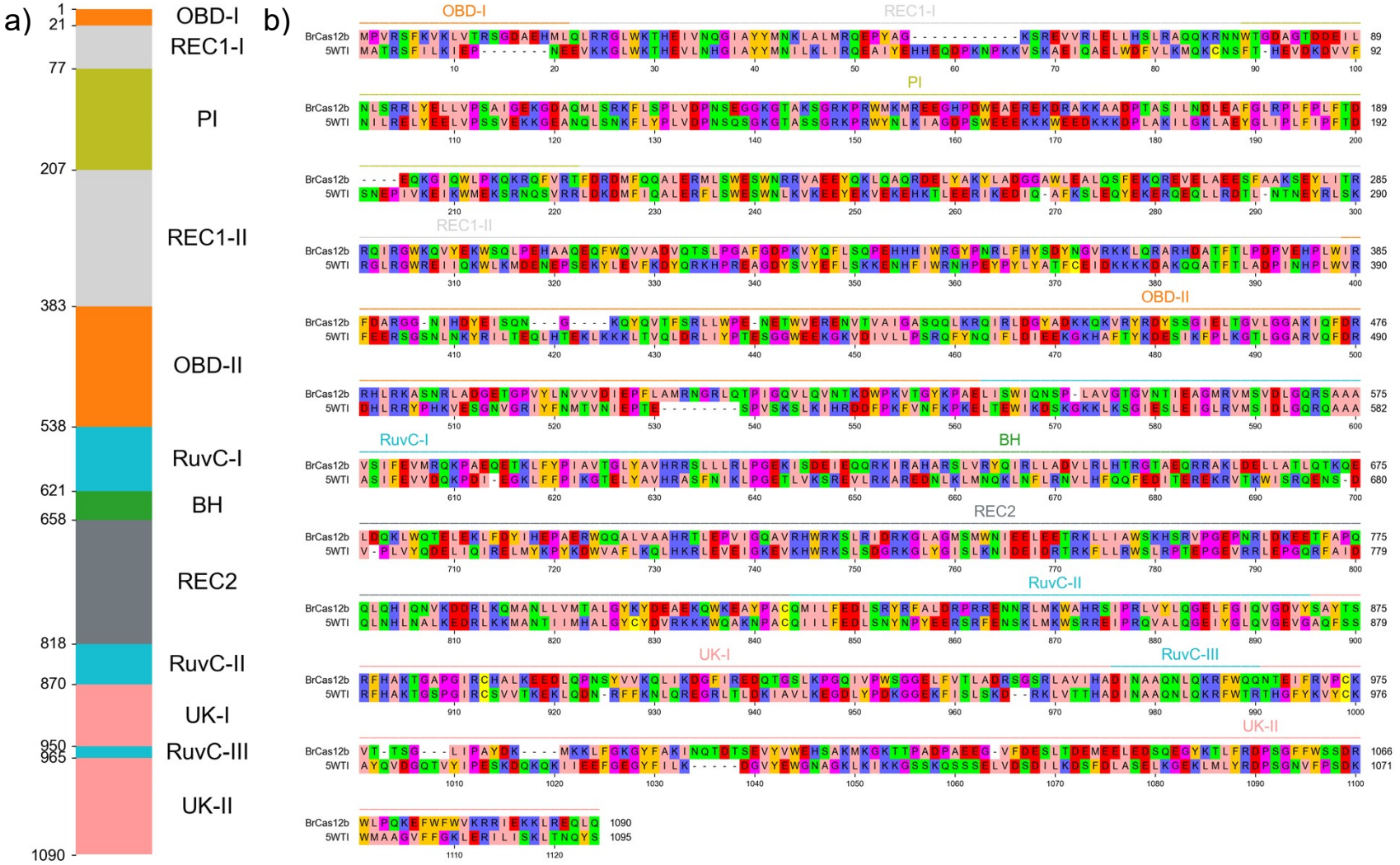

Figure S1. BrCas12b (a)domain assignment and (b)corresponding sequence alignment plotted using *pyMSAviz* package with the existing crystal structure of BthC2c1/BthCas12b (PDB: 5WTI) .

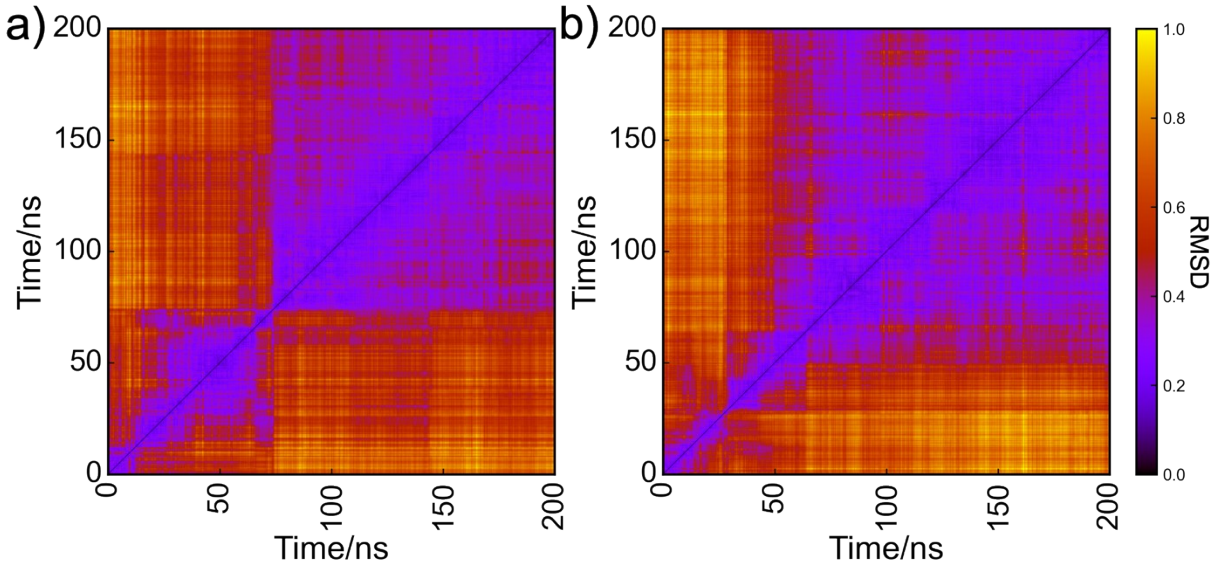

Figure S2. All-to-all RMSD matrix of (a) wild-type BrCas12b and (b) mutant-type BrCas12b at 300 K.

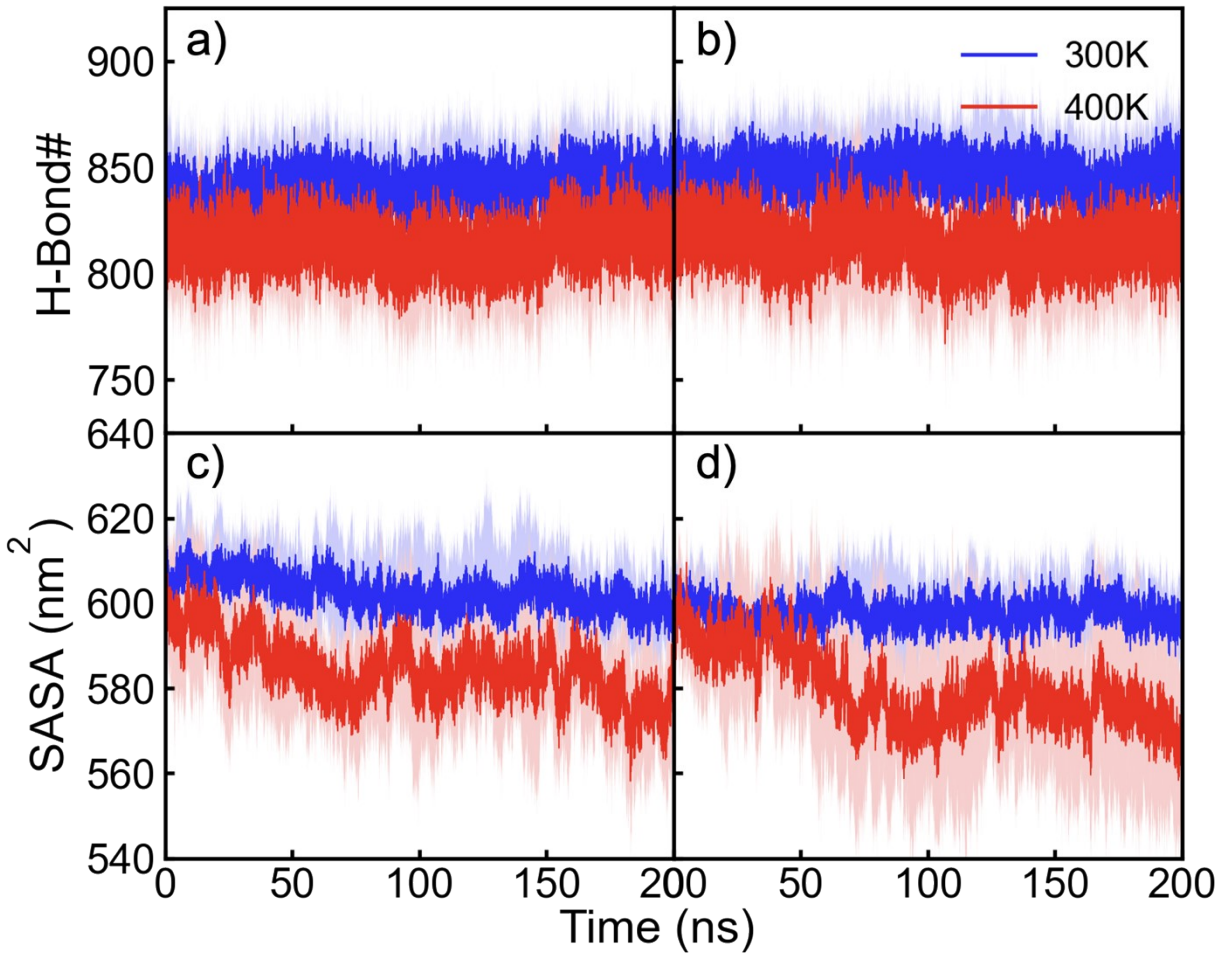

Figure S3. Intra protein hydrogen bond number (a, b) and total solvent accessible surface area (SASA) (c, d) of BrCas12b (a, c) wild type and (b, d) mutant type as a function of simulation time. The solid lines indicate values averaged over three independent replicas, while the shaded region indicates the error calculated from the four replicas.

Secondary Structure Analysis

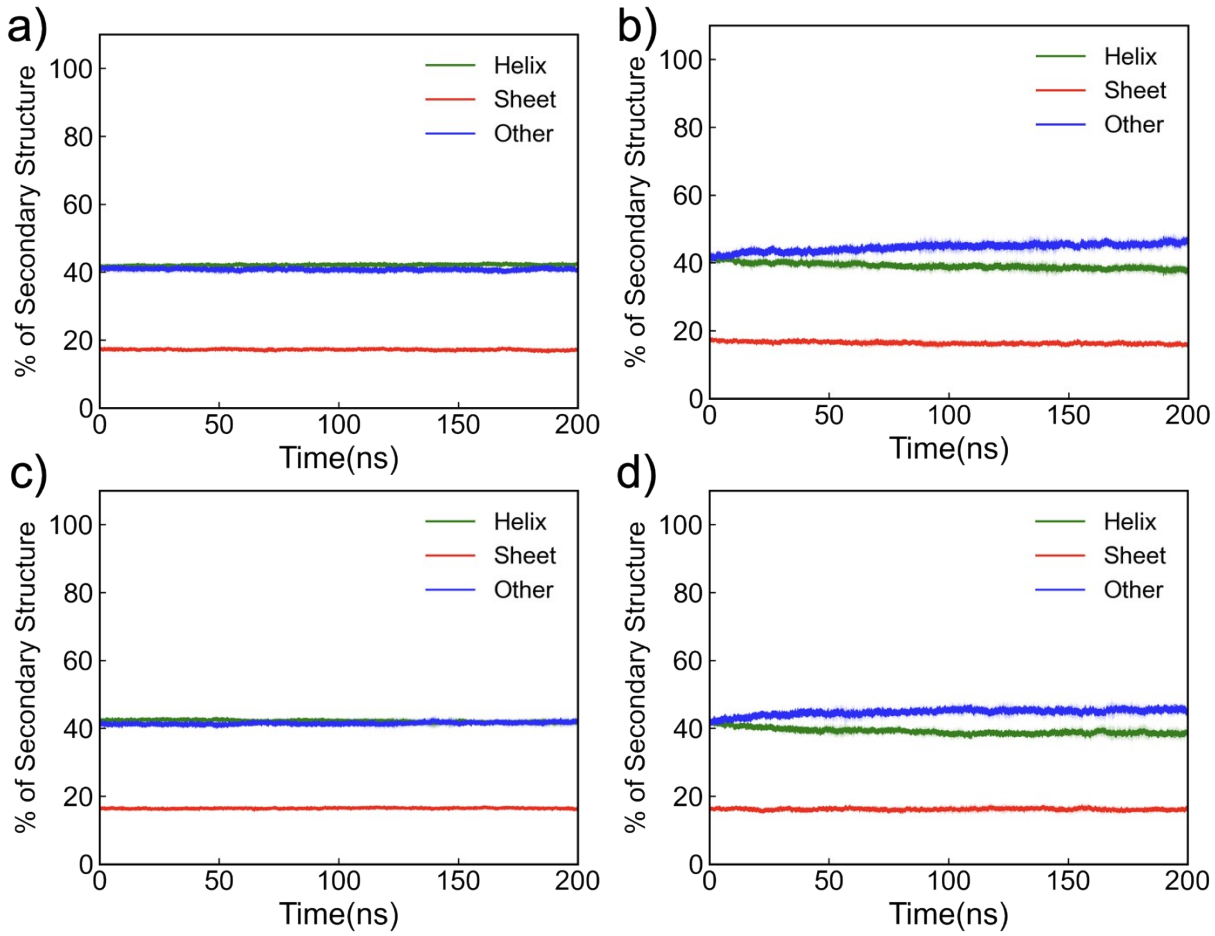

Figure S4. Secondary structure fraction of wild-type BrCas12b at (a) 300 K, (b) 400 K, and mutated-type BrCas12b at (c) 300 K, (d) 400 K. The solid lines indicate values averaged over three independent replicas, while the shaded region indicates the error bar calculated from the three replicas.

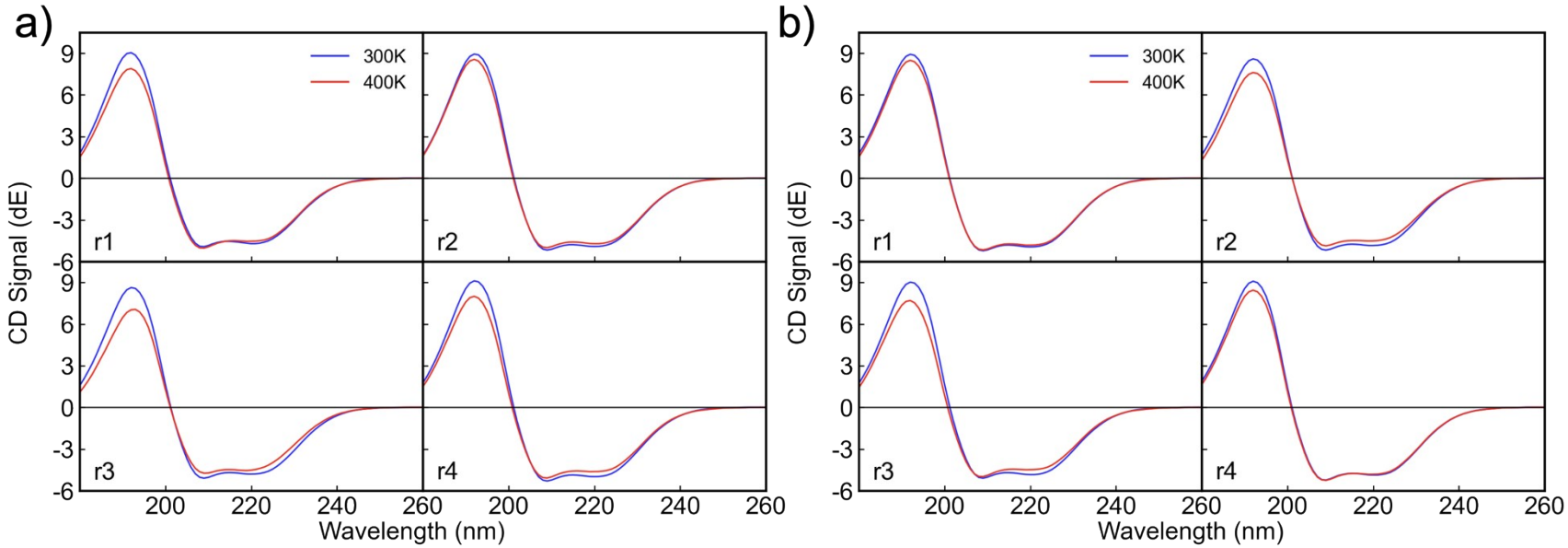

Figure S5. Predicted circular dichroism (CD) spectra of (a) wild-type BrCas12b and (b) mutated-type BrCas12b at 300 and 400 K after 200 ns simulation for four replicates (r1 – r4).

Table S1. Comparison of α-helix and β-sheet content in BrCas12b. The secondary structures of wild-type (WT) and mutant (MT) BrCas12b were analyzed from the last 100 ns of replicate simulations. These are compared with the secondary structures of existing Cas12b family crystal structures retrieved from the RCSB Protein Data Bank. All percentages were quantified using the STRIDE algorithm. Values for WT and MT BrCas12b represent the mean across replicates, with the standard deviation provided in parentheses.

PDB	sequence identity %	total residues modeled	helix %	sheet%
5WTI	41	982	44.6	16.29
5WQE	36	1011	42.63	18.49
5U30	36	1085	41.01	16.86
5U31	36	1085	41.01	16.86
5U33	36	1085	40.55	16.86
5U34	36	960	43.02	17.18
9KLN	27	1385	46.5	11.33
9KLO	27	1396	45.7	12.1
9KLP	27	1268	44.32	12.22
9KLQ	27	1284	44.94	11.06
BrCas12b(WT)	/	1090	42.16(0.49)	17.14(0.45)
BrCas12b(MT)	/	1090	41.90(0.75)	16.49(0.40)

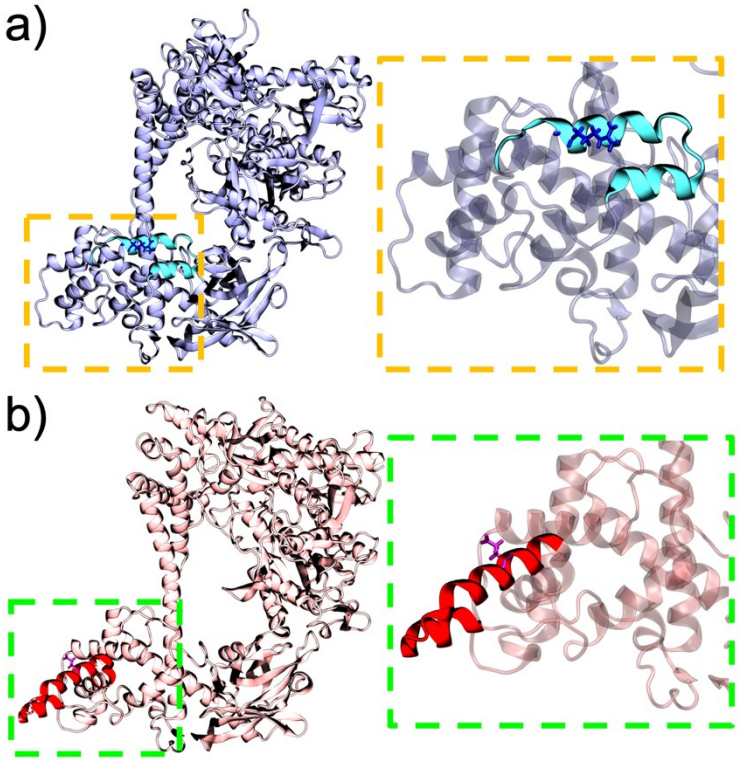

Figure S6. Snapshots of (a)WT and (b)MT BrCas12b at the end of the 200 ns simulation with the region near mutation R160E highlighted on the right.

Domain Level Analysis

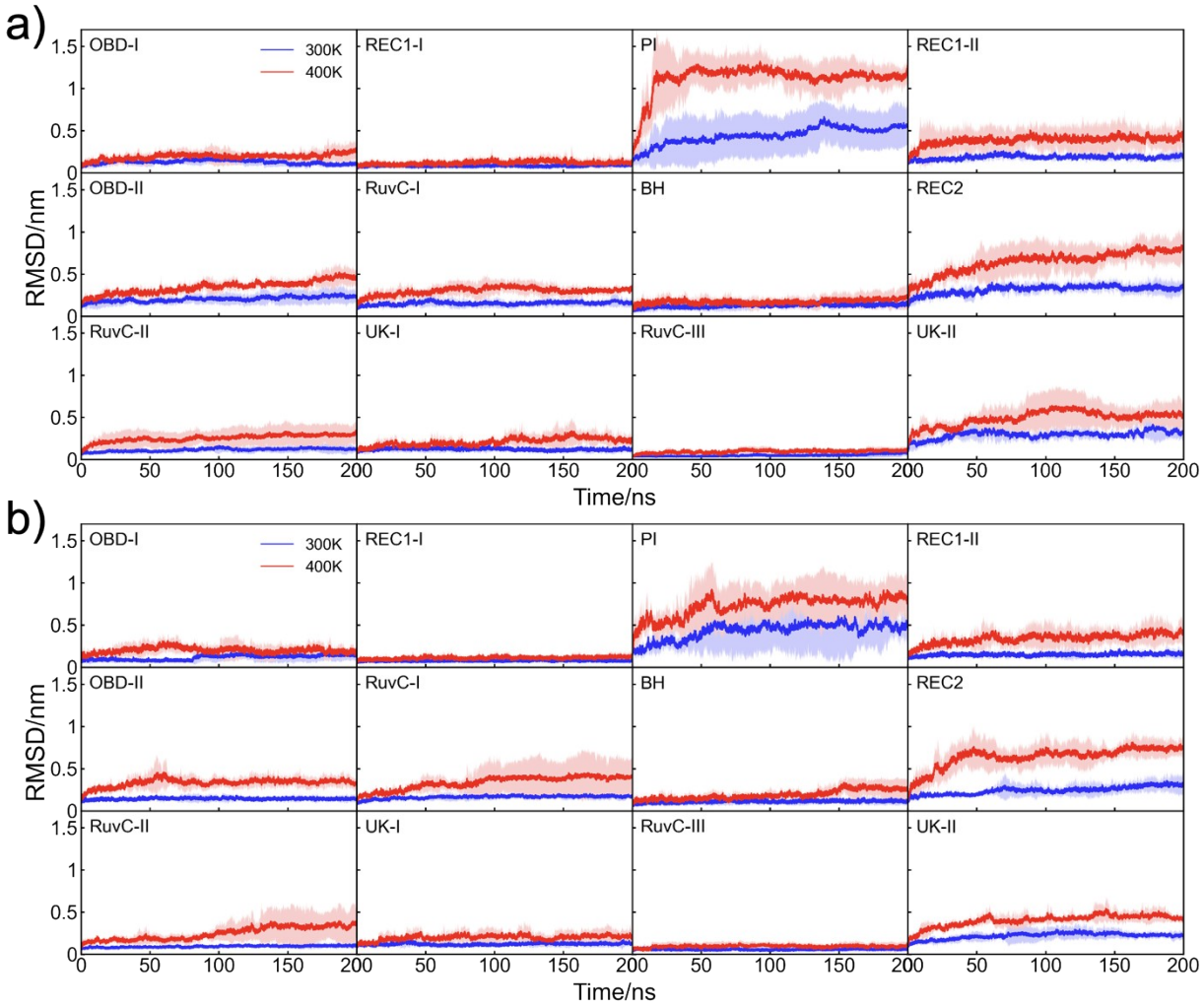

Figure S7. Backbone root mean squared deviation (RMSD) of each domain from (a) wild-type BrCas12b and (b) mutated-type BrCas12b, with the names of the domains shown at the top left corner. The solid lines indicate values averaged over three independent simulations, while the shaded region indicates the error bar calculated from the three replicas.

Table S2. Mean root-mean-square deviation (RMSD) over the final 100 ns of 4 replicas. The table reports the average value and standard deviation (in parentheses) across the four replicates.

Domain	Temperature	WT RMSD/nm	MT RMSD/nm	P Value
OBD1	300K	0.1216(0.0103)	0.1376(0.0458)	0.5226
OBD1	400K	0.2052(0.0491)	0.1959(0.0575)	0.8143
REC1-I	300K	0.0848(0.0041)	0.0777(0.0054)	0.0809
REC1-I	400K	0.1261(0.0343)	0.1162(0.0161)	0.6194
PI	300K	0.5125(0.2298)	0.4891(0.2733)	0.9002
PI	400K	1.1488(0.1503)	0.8022(0.2230)	0.0419
REC1-II	300K	0.1881(0.0190)	0.1549(0.0216)	0.0605
REC1-II	400K	0.4091(0.1312)	0.3735(0.1001)	0.6814
OBD-II	300K	0.2210(0.0703)	0.1422(0.0164)	0.0720
OBD-II	400K	0.4015(0.0423)	0.3458(0.0490)	0.1361
RuvC-I	300K	0.1560(0.0306)	0.1709(0.0203)	0.4497
RuvC-I	400K	0.3165(0.0354)	0.3931(0.2098)	0.4983
BH	300K	0.1355(0.0223)	0.1136(0.0192)	0.1871
BH	400K	0.1815(0.0627)	0.2311(0.0605)	0.2982
REC2	300K	0.3404(0.0480)	0.2707(0.0603)	0.1205
REC2	400K	0.7320(0.1117)	0.7061(0.0911)	0.7313
RuvC-II	300K	0.1245(0.0258)	0.0986(0.0137)	0.1269
RuvC-II	400K	0.2833(0.1033)	0.3167(0.1535)	0.7306
UK-I	300K	0.1149(0.0164)	0.1244(0.0217)	0.5144
UK-I	400K	0.2434(0.0667)	0.2057(0.0564)	0.4210
RuvC-III	300K	0.0559(0.0218)	0.0561(0.0090)	0.9897
RuvC-III	400K	0.0999(0.0109)	0.0925(0.0304)	0.6617
UK-II	300K	0.3113(0.0456)	0.2391(0.0444)	0.0636
UK-II	400K	0.5450(0.1459)	0.4460(0.0442)	0.2416

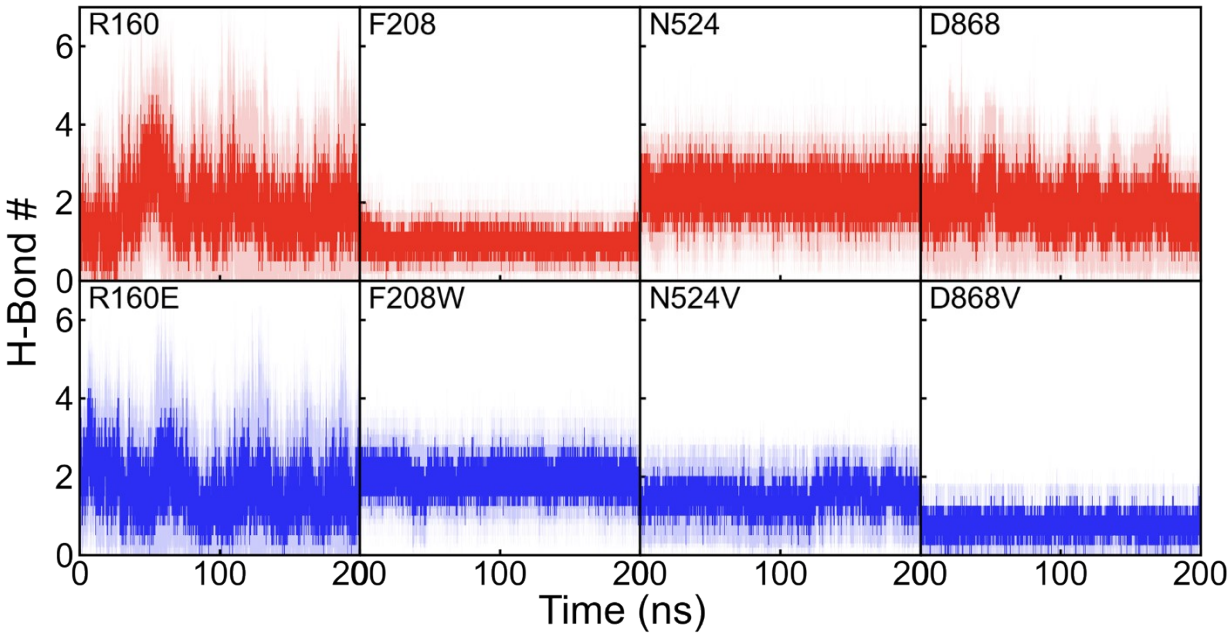

Figure S8. Number of hydrogen bond formed between the mutation point to other residues at 400 K, with each mutation point shown on the top left corner, as a function of simulation time. First row from wild-type, second row from mutated-type BrCas12b. The solid lines indicate values averaged over three independent replicas, while the shaded region indicates the error bar calculated from the four replicas.

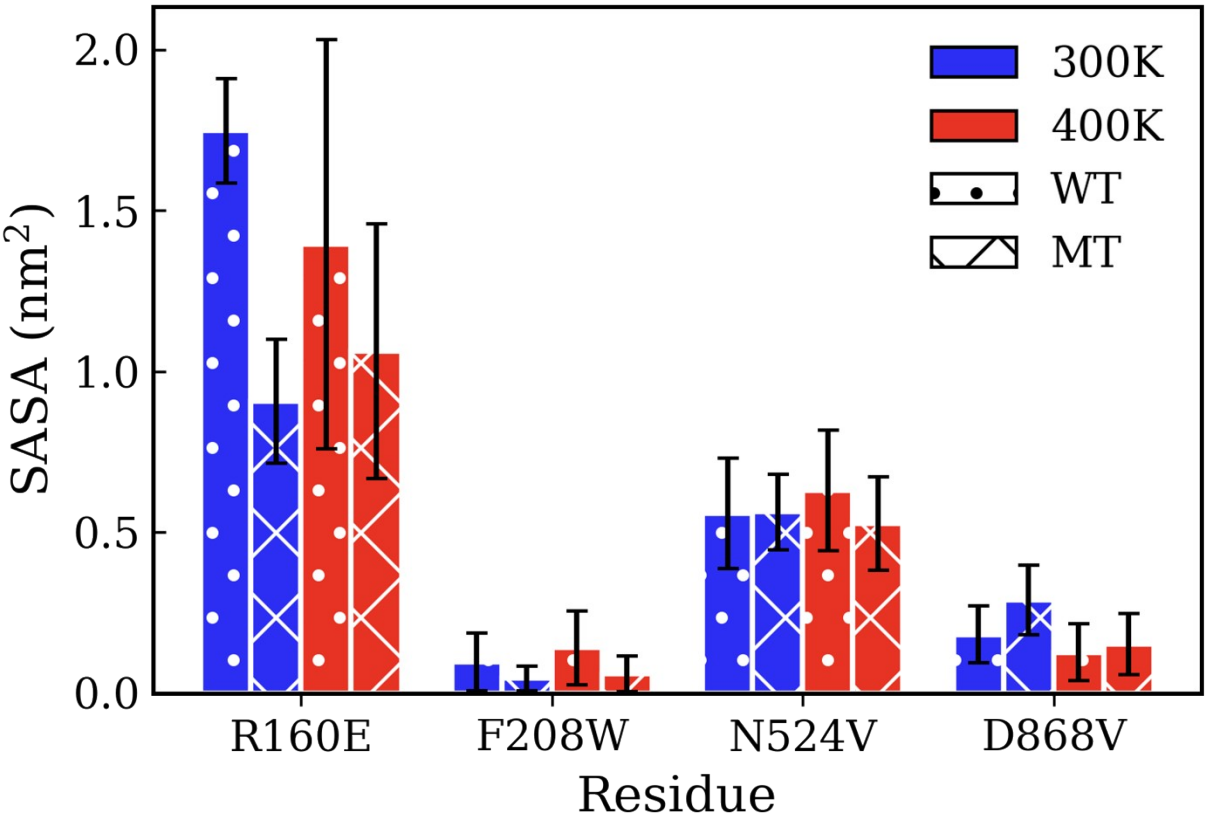

Figure S9. SASA of mutation point calculated from last 100ns of each simulation and averaged over 4 replicates. With error bar shown one standard deviation.

Principal Component Analysis

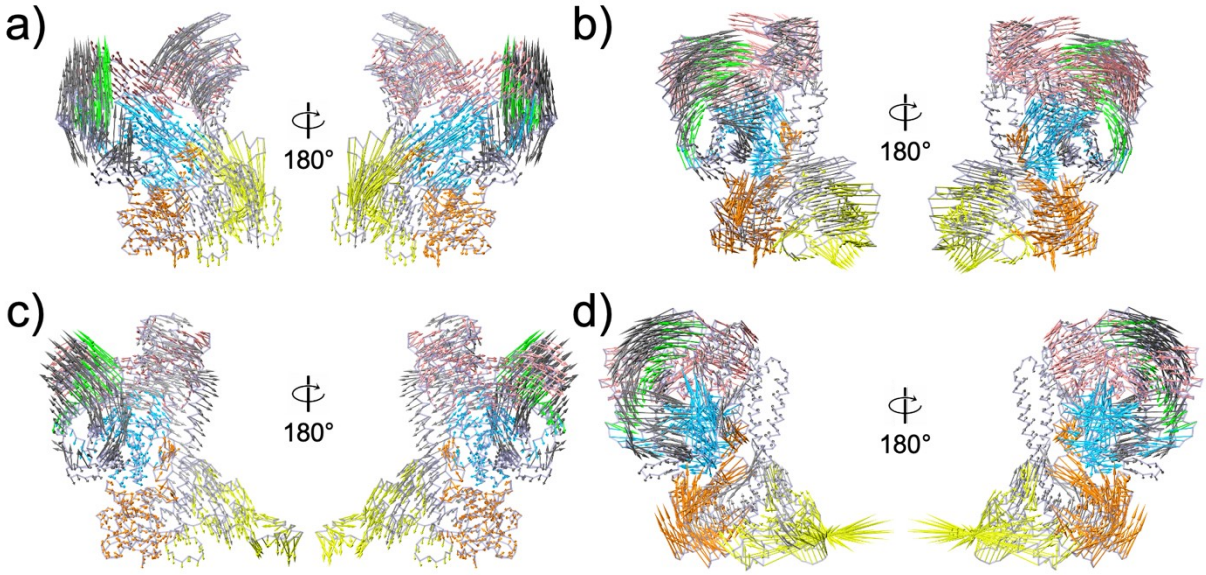

Figure S10. Motions obtained from principal component 2 (PC2) of wild-type BrCas12b at (a) 300 K, (b) 400 K, and mutated-type BrCas12b at (c) 300 K, (d) 400 K shown using arrows of sizes equivalent to the amplitude of motions, with colors adapted from Figure 1 to distinguish motions from different domains.

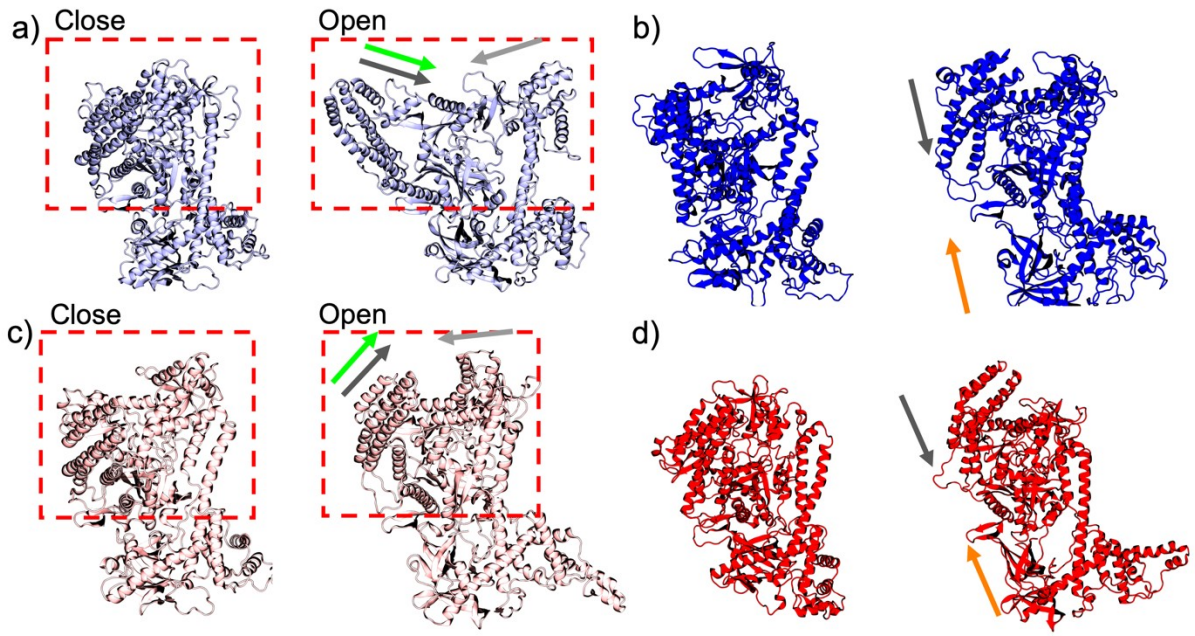

Figure S11 Snapshots representing the open and closed states derived from the first principal component (PC1) of motion for the (a) Wild-Type (WT) at 300 K, (b) WT at 400 K, (c) Mutant (MT) at 300 K, and (d) MT at 400 K. Arrows depict the direction and amplitude of the collective motion captured by PC1.

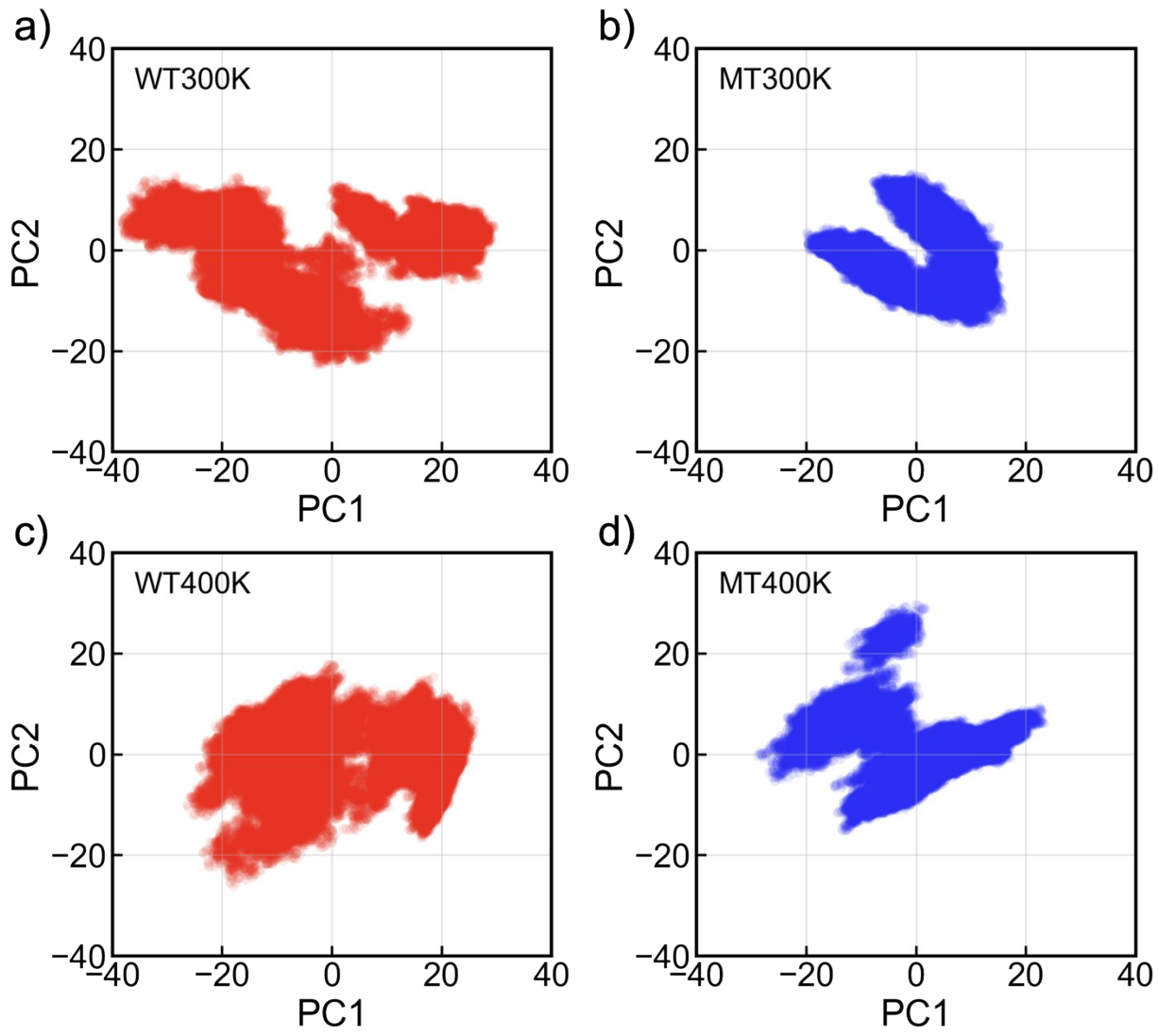

Figure S12. Projections of the simulation trajectory onto first and second principal component derived from (a) wild type BrCas12b at 300 K, (b) mutated type BrCas12b at 300 K, (c) wild type BrCas12b at 400 K, and (d) mutated type BrCas12b at 400 K.

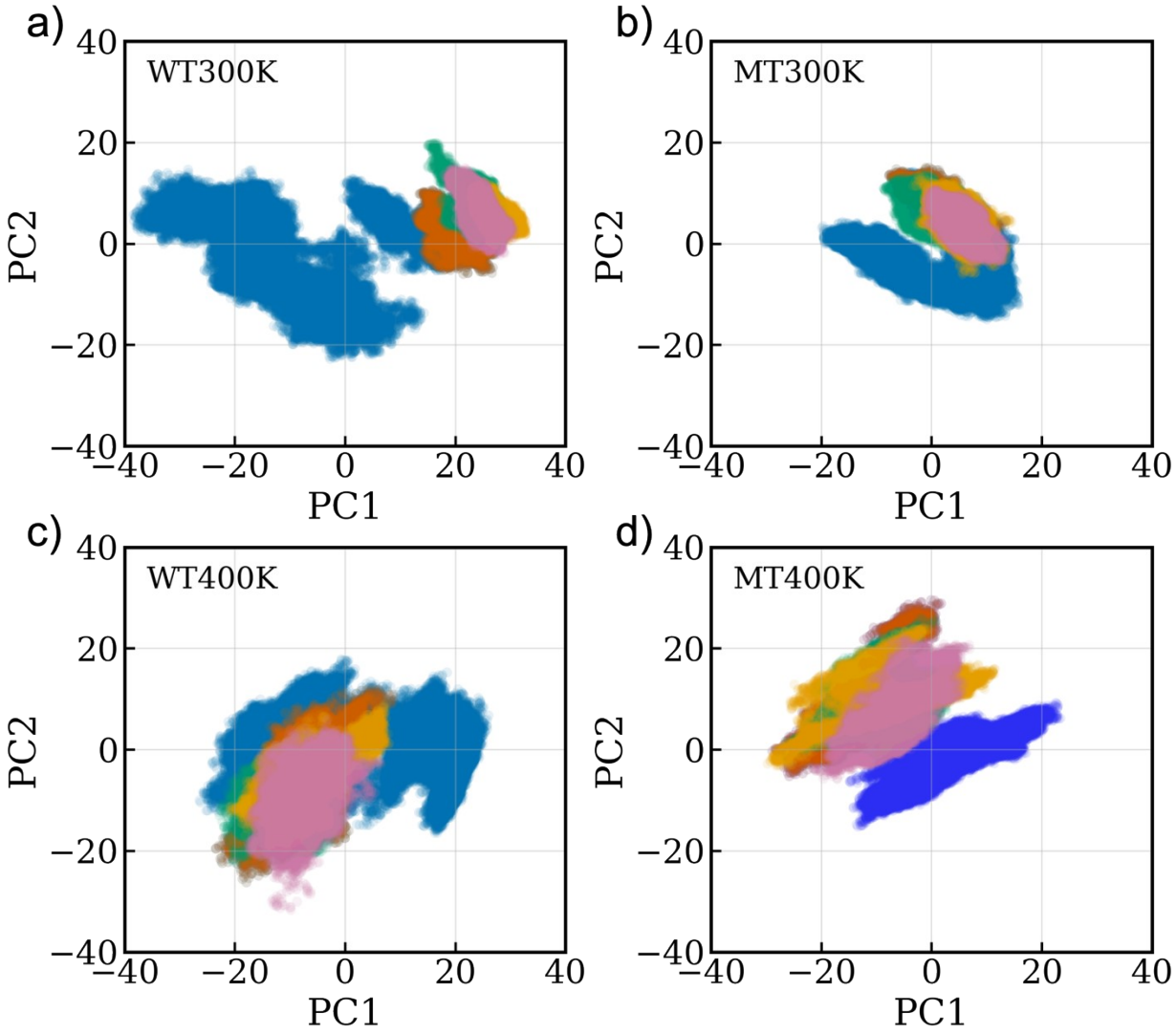

Figure S13. Projections of the replicas simulation trajectory onto first and second principal component derived from 1 μs trajectory of (a) wild type BrCas12b at 300K, (b) mutated type BrCas12b at 300 K, (c) wild type BrCas12b at 400 K, and (d) mutated type BrCas12b at 400 K, with dark blue presents the space explored by 1 μs simulation and other colors present the space explored by replicas.

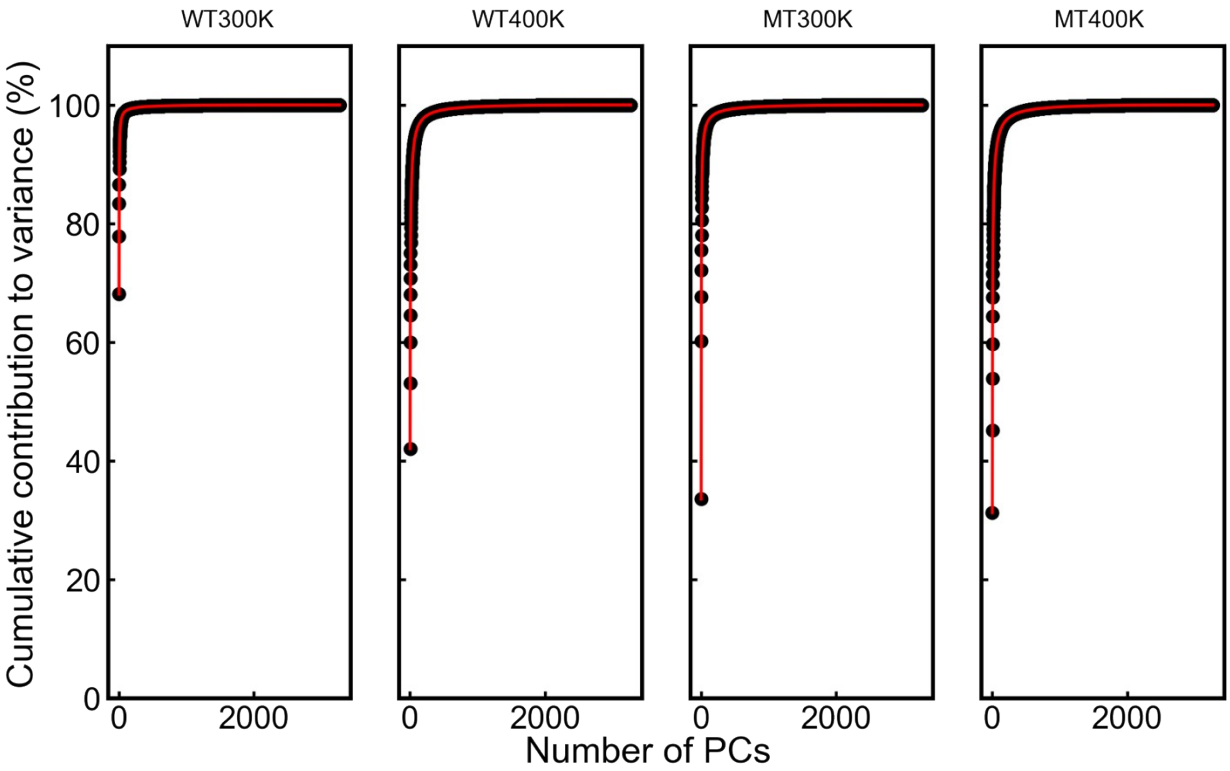

Figure S14. Cumulative contribution (%, y-axis) of all the principal components (PCs, x-axis) to the variance of the overall Cas12b motions calculated upon Principal Component Analysis (PCA) of WT and MT BrCas12b at 300 K and 400 K.