Supporting Information

for

DNA Preference of Indenoisoquinolines: A Computational Approach

Semiha Kevser Bali,[1,2]* Zeynep Pinar Haslak,[1,3] Gulsah Cifci,[1] Viktorya Aviyente[1]

[1] Bogazici University, Department of Chemistry, 34342, Istanbul, Turkey

[2] Michigan State University, Department of Chemistry, East Lansing, Michigan, 48824, USA

[3] Université de Reims Champagne-Ardenne, Reims, 51687, France

*balisem1@msu.edu, semihakevserbali@gmail.com

List of Tables

List of Figures

Table S1. The docking score of top 10 poses for L77 compound. E_{int} values are calculated using the single point QM/MM method after the minimization of the docked systems.

Pose	Docking Score	ΔE_{int} (kcal/mol)
1	-10.90	-120.95
2	-10.85	-105.06
3	-10.56	-96.17
4	-10.56	-100.73
5	-10.41	-119.00
6	-10.37	-121.28
7	-10.28	-108.28
8	-10.26	-98.54
9	-10.18	-114.04
10	-10.03	-84.95

Table S2. The atom types and partial charges of L77.

Atom Name	Atom Type	Seq.	Element	Partial Charge
C1	ca	1	6	-0.1597
H1	ha	2	1	0.1346
C2	ca	3	6	-0.1985
H2	ha	4	1	0.1486
C3	ca	5	6	-0.1533
C4	ca	6	6	0.2511
C5	ca	7	6	0.2263
C6	ca	8	6	-0.0866
O1	os	9	8	-0.3223
C7	c3	10	6	-0.0129
H3	h1	11	1	0.0784
H4	h1	12	1	0.0784
H5	h1	13	1	0.0784
O2	os	14	8	-0.3325
C8	c3	15	6	-0.0015
H6	h1	16	1	0.074
H7	h1	17	1	0.074
H8	h1	18	1	0.074
C9	cc	19	6	-0.0627
C10	cd	20	6	0.0187
C11	c	21	6	0.4937
O3	o	22	8	-0.5471
N1	n	23	7	-0.0294
C12	c	24	6	0.6005
O4	o	25	8	-0.5005
C13	ca	26	6	-0.2027
C14	ca	27	6	0.0221
C15	ca	28	6	-0.2429
H9	ha	29	1	0.2045
C16	ca	30	6	0.3394
C17	ca	31	6	0.2758
C18	ca	32	6	-0.3184
H10	ha	33	1	0.2385
O5	os	34	8	-0.4211
O	os	35	8	-0.4113
C19	c3	36	6	0.2853
H11	h2	37	1	0.0911
H12	h2	38	1	0.0911

C20	c3	39	6	-0.0403
H13	h1	40	1	0.0652
H14	h1	41	1	0.0652
C21	c3	42	6	0.0112
H15	hc	43	1	0.0409
H16	hc	44	1	0.0409
C22	c3	45	6	-0.0452
H17	h1	46	1	0.06
H18	h1	47	1	0.06
N2	na	48	7	0.0456
C23	cc	49	6	0.1777
H19	h5	50	1	0.1269
C24	cc	51	6	-0.3666
H20	h4	52	1	0.2009
C	cd	53	6	0.1117
H	h4	54	1	0.1252
N	nd	55	7	-0.5544

Table S3. The atom types and partial charges of CPT.

Atom Name	Atom Type	Seq.	Element	Partial Charge
C1	ca	1	6	-0.1524
H1	ha	2	1	0.1518
C2	ca	3	6	-0.1306
H2	ha	4	1	0.1534
C3	ca	5	6	-0.2332
H3	ha	6	1	0.1541
C4	ca	7	6	0.4162
C5	ca	8	6	0.0218
C6	ca	9	6	-0.1915
H4	ha	10	1	0.1488
N1	nb	11	7	-0.5578
C7	ca	12	6	0.3962
C8	ca	13	6	-0.0478
C9	ca	14	6	-0.2346
H5	ha	15	1	0.1927
C10	c3	16	6	-0.0274
H6	h1	17	1	0.1035
H7	h1	18	1	0.1035
N	n	19	7	-0.0338
C11	cc	20	6	-0.0534
C12	cd	21	6	-0.1801
H8	ha	22	1	0.1654
C13	cd	23	6	-0.0298
C14	cc	24	6	-0.0421
C15	c	25	6	0.4443
O1	o	26	8	-0.5972
C16	c3	27	6	0.0808
H9	h1	28	1	0.0982
H10	h1	29	1	0.0982
C17	c3	30	6	0.1748
O2	oh	31	8	-0.5883
H11	ho	32	1	0.4082
O3	os	33	8	-0.4269
C18	c	34	6	0.7264
O	o	35	8	-0.5679
C19	c3	36	6	0.0341
H12	hc	37	1	0.0085
H13	hc	38	1	0.0085

C	c3	39	6	-0.0822
H14	hc	40	1	0.0292
H15	hc	41	1	0.0292
H	hc	42	1	0.0292

Table S4. The atom types and partial charges of PTR residue.

Atom Name	Atom Type	Seq.	Element	Partial Charges
N	N	1	7	-0.42745
H	H	2	1	0.28402
CA	CX	3	6	-0.00332
HA	H1	4	1	0.111848
CB	CT	5	6	-0.06529
HB2	HC	6	1	0.045463
HB3	HC	7	1	0.045463
CG	CA	8	6	0.006319
CD1	CA	9	6	-0.21337
HD1	HA	10	1	0.136129
CE1	CA	11	6	-0.20822
HE1	HA	12	1	0.149042
CZ	CA	13	6	0.420808
OP	OS	14	8	-0.50491
P	P	15	15	1.208364
O1P	O2	16	8	-0.75422
O2P	O2	17	8	-0.75422
CE2	CA	18	6	-0.20822
HE2	HA	19	1	0.149042
CD2	CA	20	6	-0.21337
HD2	HA	21	1	0.136129
C	C	22	6	0.498833
O	O	23	8	-0.53098

Table S5. The average hydrogen bond percentages in each of the systems simulated. Res.1 and Res.2 correspond to the acceptor and donor atoms, respectively. W, M1, M2, and M4 are the systems simulated.

	Res.1(A)	Res.2(D)	W	M1	M2	M4
CPT	CPT@O	K532@NZ	11%	25%	26%	0%
	CPT@O2	K532@NZ	17.5%	0%	0%	0%
	575@O2	CPT@O2	28.5%	0%	0%	0%
	CPT@O1	576@O5'	36%	0%	49.5%	82.5%
	CPT@O3	576@O5'	0%	0%	0%	0%
	CPT@N1	R364@NH1	1%	0%	8%	1.5%
	CPT@O	R488@NH2	0%	0%	4%	0%
	CPT@O2	600@N2	0%	0%	0%	7.5%
	D533@OD2	CPT@O2	24%	0%	36.5%	79%
L77	L77@O4	R364@NH1	17.5%	16%	23%	65%
	L77@O1	576@O5'	1%	17%	26.5%	25%
	L77@O2	576@O5'	35%	40.5%	42.5%	20%

Table S6. The time-series plots for the distances calculated between the indicated residue-ligand atom pairs for CPT simulations. The running average values are calculated for every 100 frames.
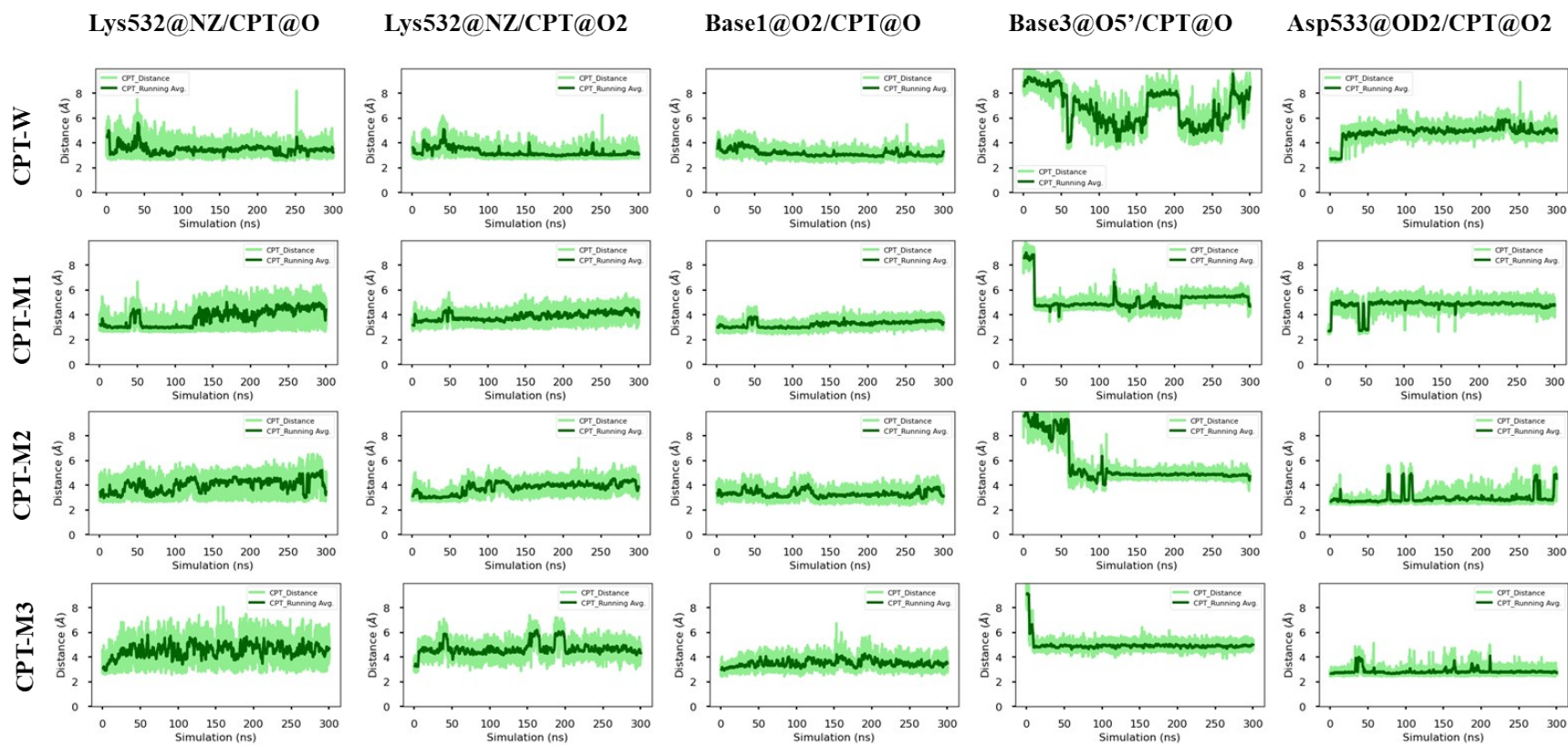
Table S7. Average values of distances reported in Figure 8 for the last 100 ns of the simulations of L77 compound. The naming of the columns follows the ResID@Atom nomenclature.

|	Arg364@NH1/L77@O4	Base3@O5'/L77@O1	Base3@O5'/L77@O2
L77-W	2.93 ± 0.16	3.51 ± 0.34	2.91 ± 0.24
L77-M1	2.99 ± 0.24	5.23 ± 1.77	4.94 ± 1.98
L77-M2	2.92 ± 0.17	3.43 ± 0.47	2.93 ± 0.37
L77-M3	2.91 ± 0.16	3.64 ± 0.56	3.04 ± 0.33

Table S8. Average values of distances reported in Table S6 for the last 100 ns of the simulations of CPT compound. The naming of the columns follows the ResID@Atom nomenclature.

	Lys532@NZ/CPT@O	Lys532@NZ/CPT@O2	Base1@O2/CPT@O	Base3@O5'/CPT@O	Asp533@OD2/CPT@O2
CPT-W	3.46 ± 0.53	3.25 ± 0.43	3.16 ± 0.36	6.79 ± 1.51	4.85 ± 0.63
CPT-M1	3.78 ± 0.85	3.86 ± 0.46	3.28 ± 0.31	5.17 ± 0.88	4.76 ± 0.55
CPT-M2	4.0 ± 0.72	3.78 ± 0.53	3.30 ± 0.38	5.65 ± 1.67	2.95 ± 0.56
CPT-M3	4.53 ± 0.80	4.66 ± 0.60	3.49 ± 0.40	4.95 ± 0.52	2.82 ± 0.30

Table S9. Average values of distances reported in Table S10-S12 for the last 100 ns of the simulations.

	Arg488	Lys532	Arg590	His632	Met428	Leu429	Asn430	Pro431	Ser432
CPT-W	13.89 ± 0.45	10.01 ± 0.37	16.54 ± 0.35	14.32 ± 1.27	12.54 ± 0.55	15.51 ± 0.46	14.58 ± 0.55	13.11 ± 0.57	17.43 ± 0.54
CPT-M1	13.76 ± 0.61	9.68 ± 0.51	16.76 ± 0.88	13.28 ± 1.11	12.17 ± 0.54	15.33 ± 0.58	14.55 ± 0.65	13.04 ± 0.56	17.34 ± 0.56
CPT-M2	12.82 ± 0.66	8.61 ± 0.51	15.89 ± 0.74	12.05 ± 1.06	13.58 ± 0.79	16.35 ± 0.83	14.68 ± 0.91	13.49 ± 0.85	17.74 ± 0.91
CPT-M3	13.46 ± 0.43	8.70 ± 0.31	15.76 ± 0.37	11.67 ± 0.31	13.20 ± 0.62	16.82 ± 0.61	16.02 ± 0.71	14.31 ± 0.61	18.66 ± 0.63
L77-W	17.14 ± 0.44	12.25 ± 0.27	18.41 ± 0.43	16.81 ± 0.43	9.79 ± 0.60	12.87 ± 0.60	12.52 ± 0.68	11.60 ± 0.73	15.81 ± 0.71
L77-M1	16.52 ± 0.66	12.87 ± 0.80	19.16 ± 0.74	16.46 ± 0.75	9.33 ± 0.53	13.08 ± 0.45	11.95 ± 0.67	14.31 ± 0.53	15.30 ± 0.60
L77-M2	16.87 ± 0.50	12.28 ± 0.35	18.63 ± 0.58	15.82 ± 0.40	9.52 ± 0.36	13.02 ± 0.55	12.29 ± 0.62	11.48 ± 0.57	15.61 ± 0.56
L77-M3	16.12 ± 0.37	12.00 ± 0.26	18.72 ± 0.45	16.10 ± 0.47	9.28 ± 0.46	13.29 ± 0.43	12.65 ± 0.42	12.05 ± 0.44	16.11 ± 0.40

Table S10. The center of mass distances between the hinge residues and the inhibitor compounds. All heavy atoms were considered for the residues and the ligand molecules.
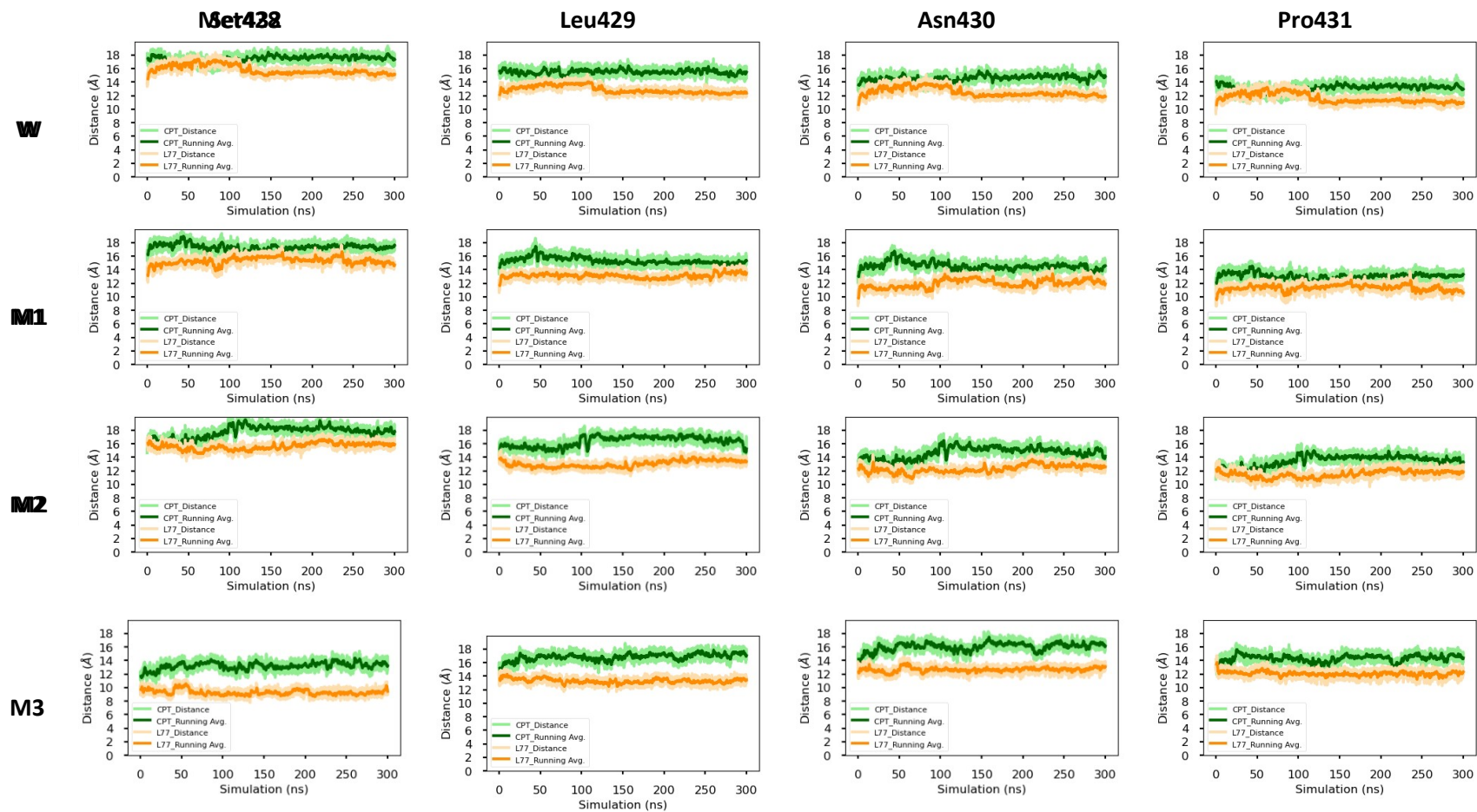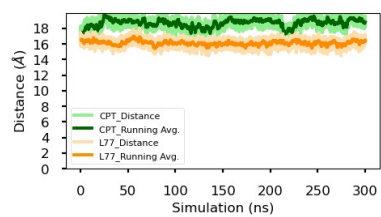
M3

Table S11. The center of mass distances between the hinge residues and the inhibitor compounds. All heavy atoms were considered for the residues and the ligand molecules.

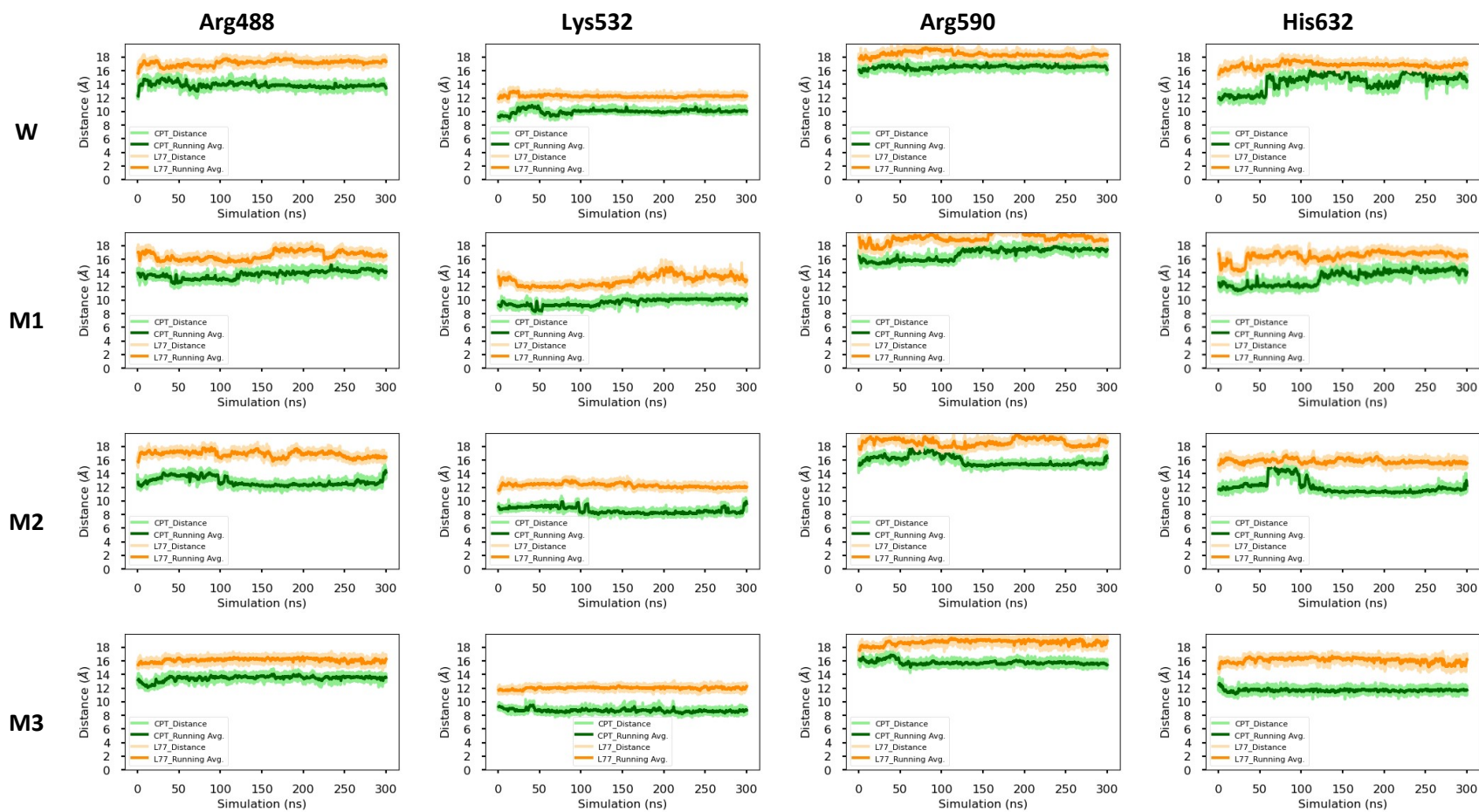**Table S12.** The center of mass distances between the catalytic tetrad residues and the inhibitor compounds. All heavy atoms were considered for the residues and the ligand molecules.
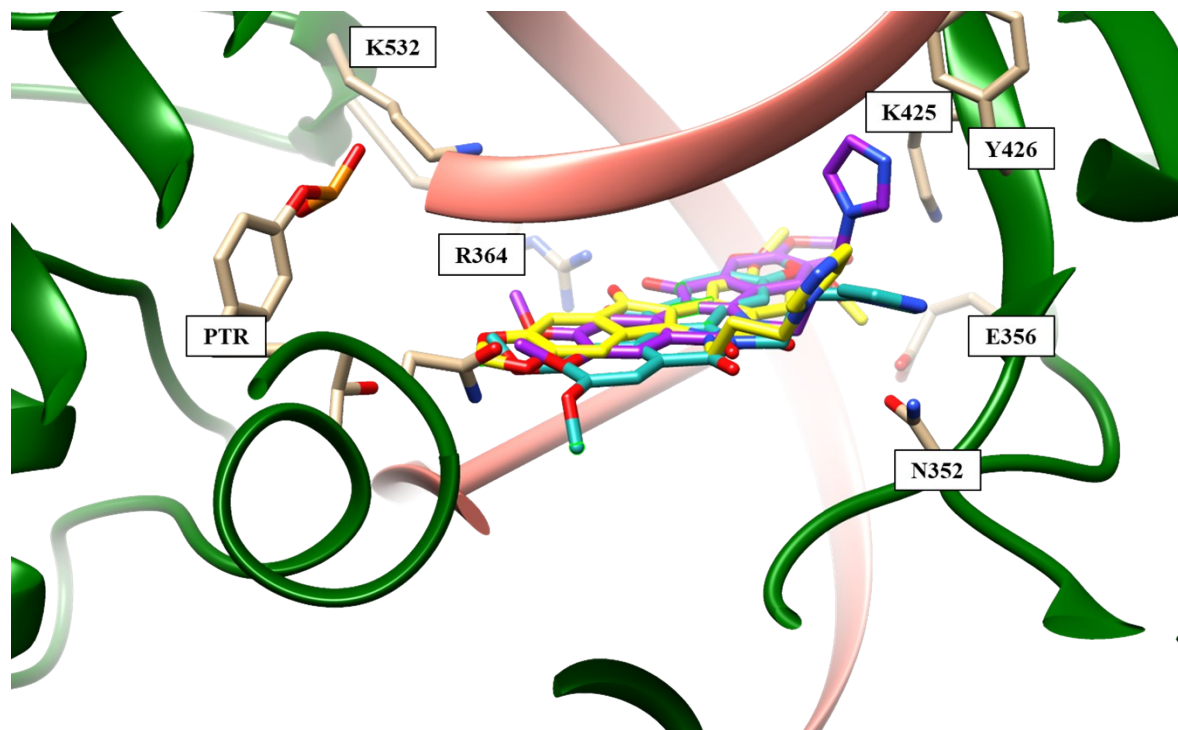
Figure S1. The docking poses of L77: Pose 1 (purple), Pose 5 (yellow), and pose 6 (cyan). The nearby residues (<5Å) are shown as well. The DNA backbone is colored in pink, hTopoIB backbone is colored in green. PTR: phosphorylated tyrosine residue. The figure is obtained using Chimera.
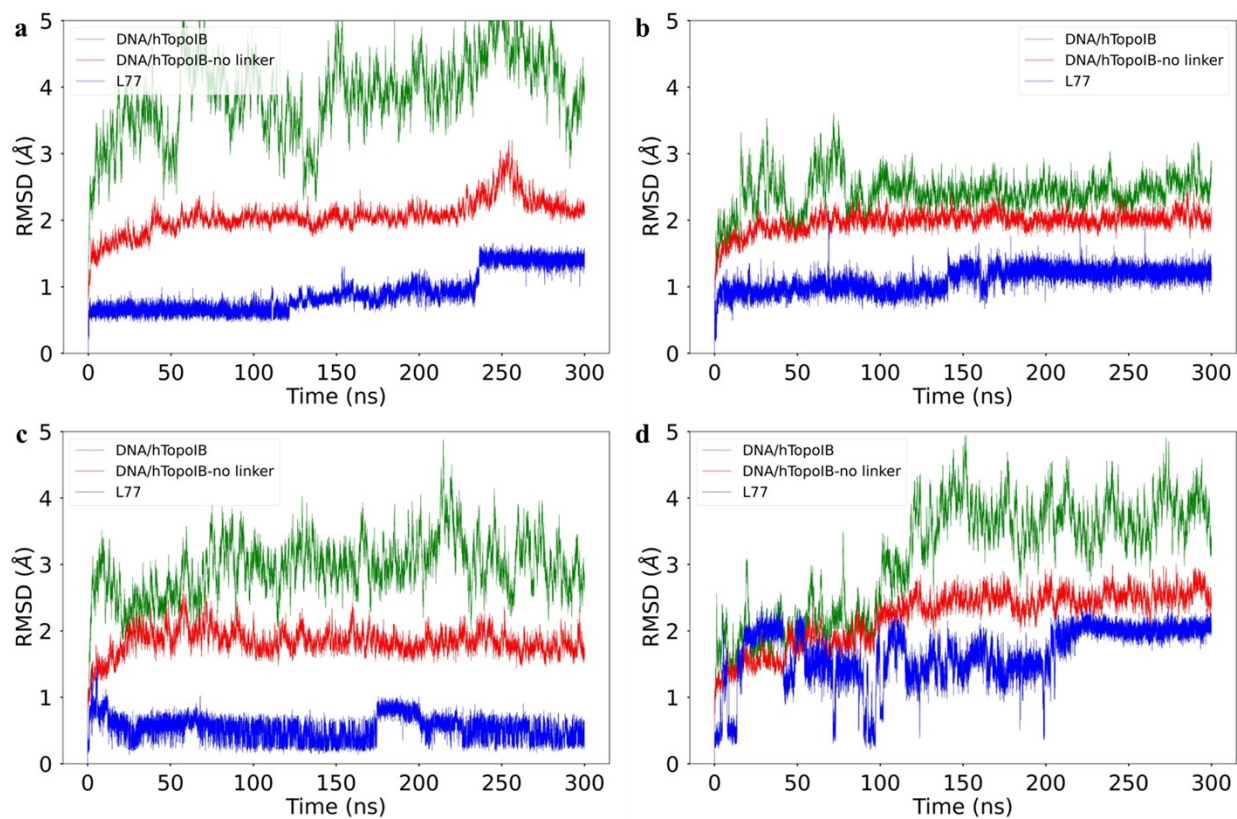
Figure S2. The RMSD plots of duplicate L77 simulations. **a**. M1; **b**. M2; **c**. M3; **d**. W. RMSD Fit plots were calculated by fitting the trajectories to the heated complex structures. DNA/TopoIB binary complex (green), DNA/TopoIB without the linker domain (red), and the L77 (blue) are shown.
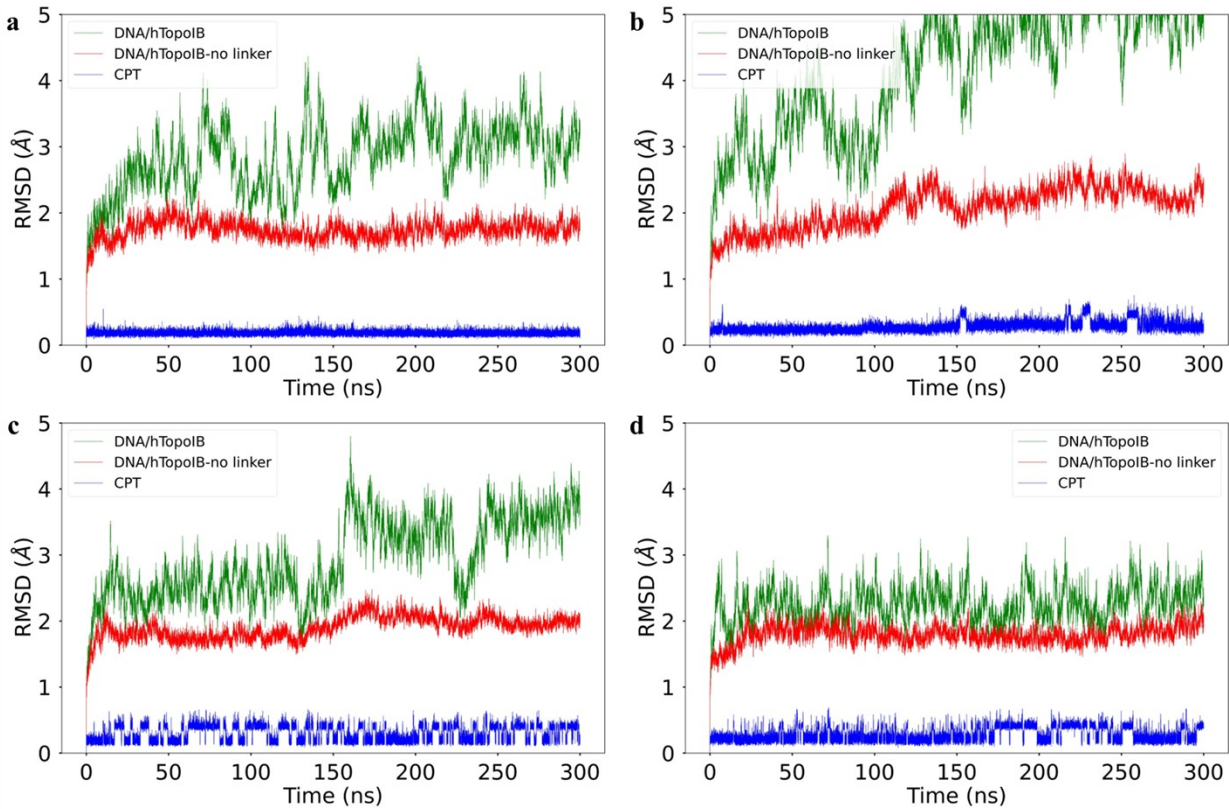
Figure S3. The RMSD plots of duplicate CPT simulations. **a.** M1; **b**. M2; **c**. M3; **d**. W. RMSD Fit plots were calculated by fitting the trajectories to the heated complex structures. DNA/TopoIB binary complex (green), DNA/TopoIB without the linker domain (red), and the CPT (blue) are shown.
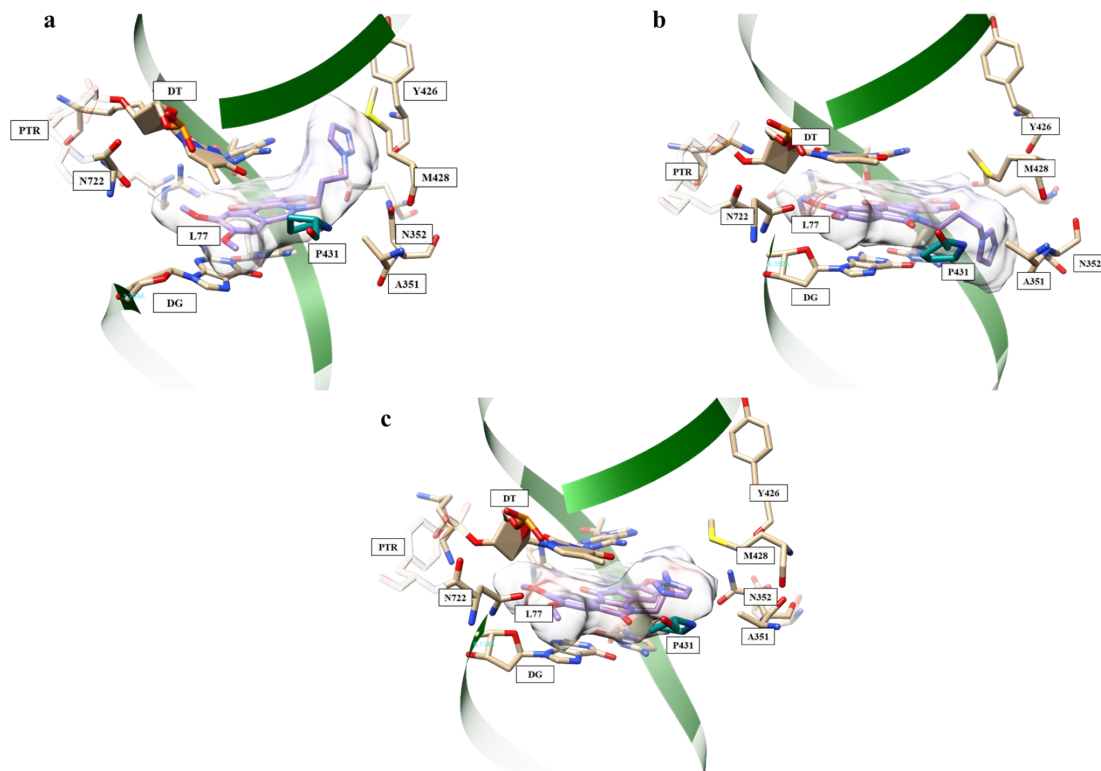
S18

Figure S4. Different orientations of L77 were observed in M1 simulations. **a**. The beginning of the simulation; **b**. at 170 ns; **c**. at 300 ns. The DNA backbone is shown in green ribbon. The surrounding residues as well as the stacking DNA bases are shown with stick representation. The L77 compound is shown in purple.
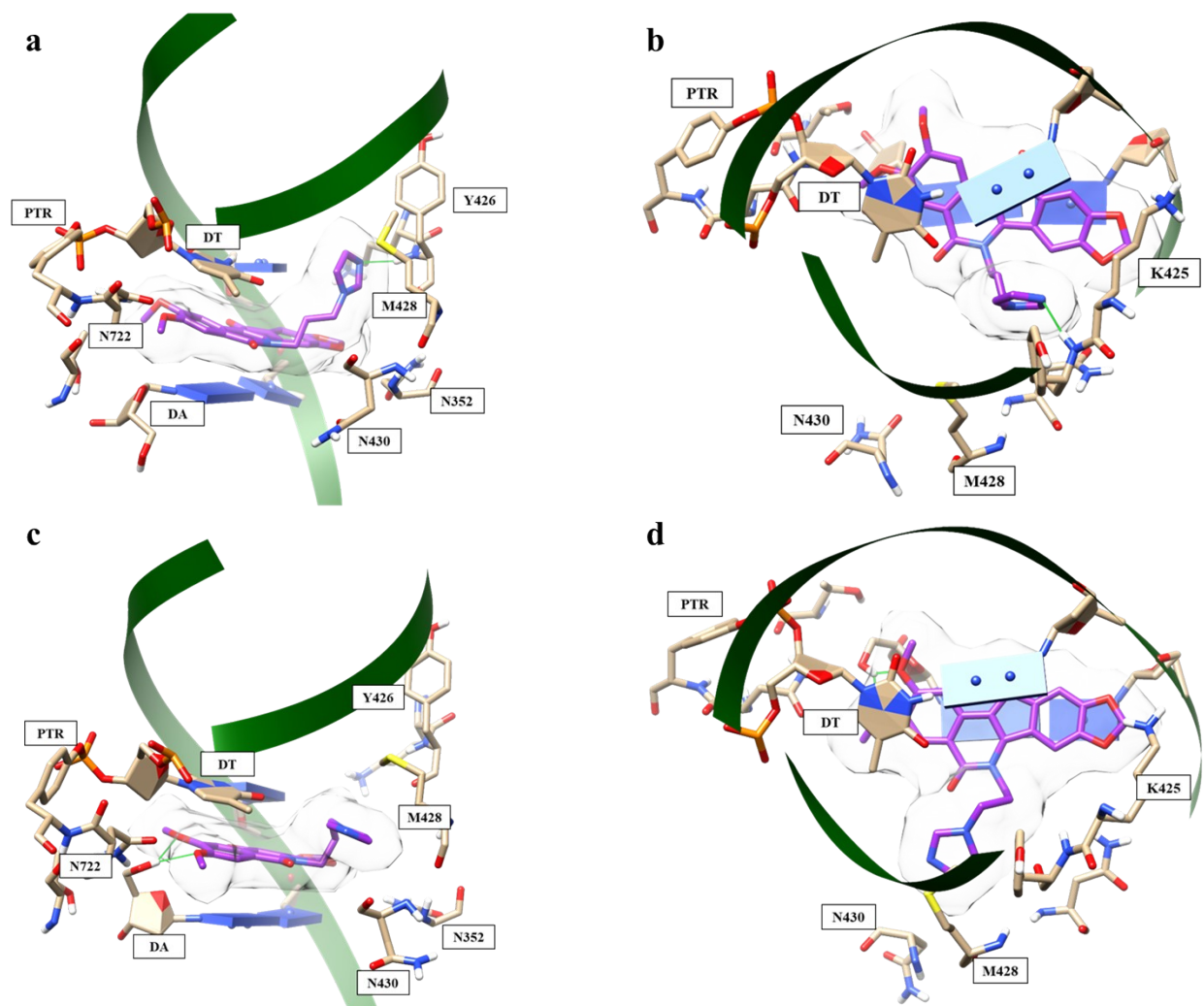
Figure S5. Different orientations of L77 were observed in M2 simulations. **a**. The beginning of the simulation side view; **b**. top view; **c**. side view at 100 ns; **d**. top view at 100ns. The DNA backbone is shown in green ribbon. The surrounding residues as well as the stacking DNA bases are shown with stick representation. The L77 compound is shown in purple.
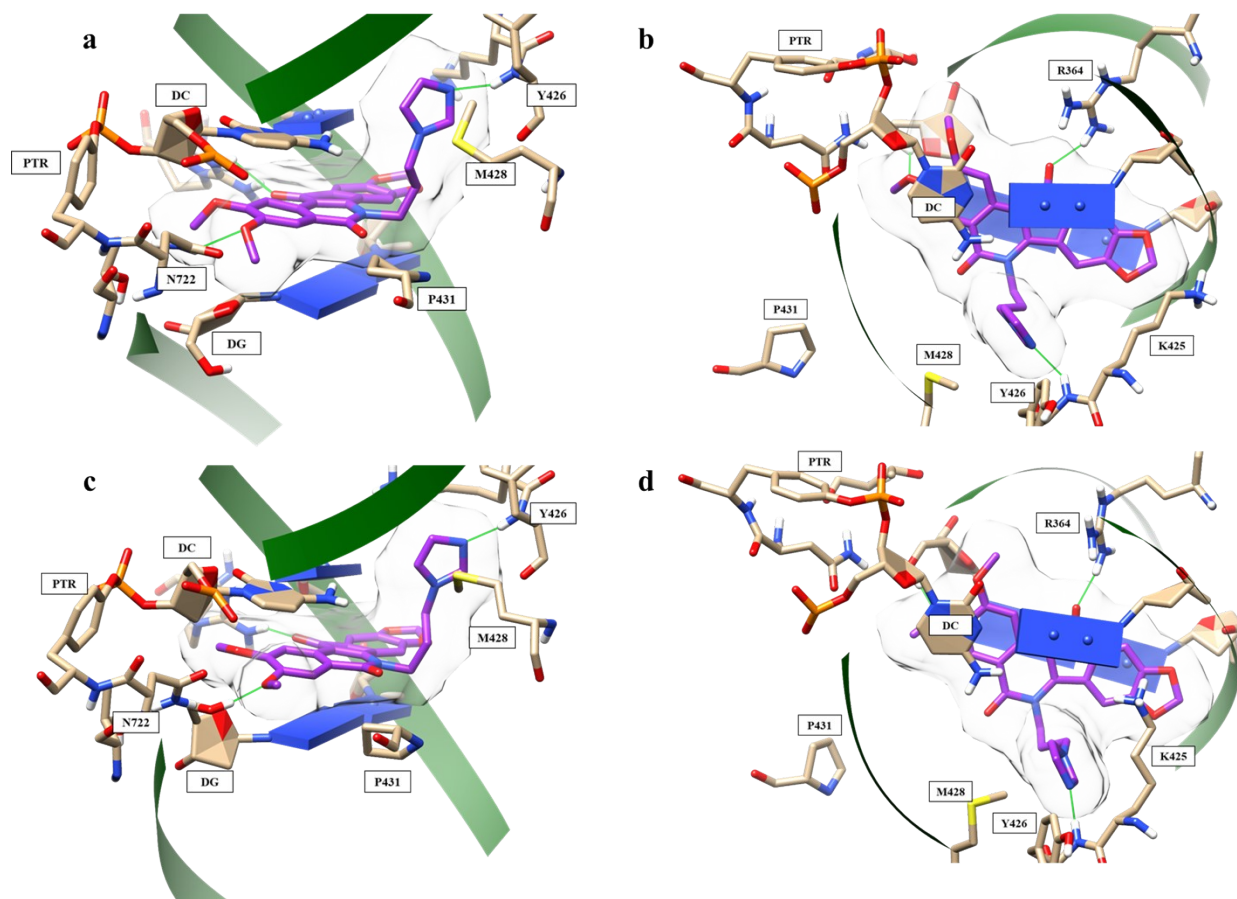
Figure S6. Different orientations of L77 were observed in M3 simulations. **a**. The beginning of the simulation side view; **b**. top view; **c**. side view at 300 ns; **d**. top view at 300ns. The DNA backbone is shown in green ribbon. The surrounding residues as well as the stacking DNA bases are shown with stick representation. The L77 compound is shown in purple.
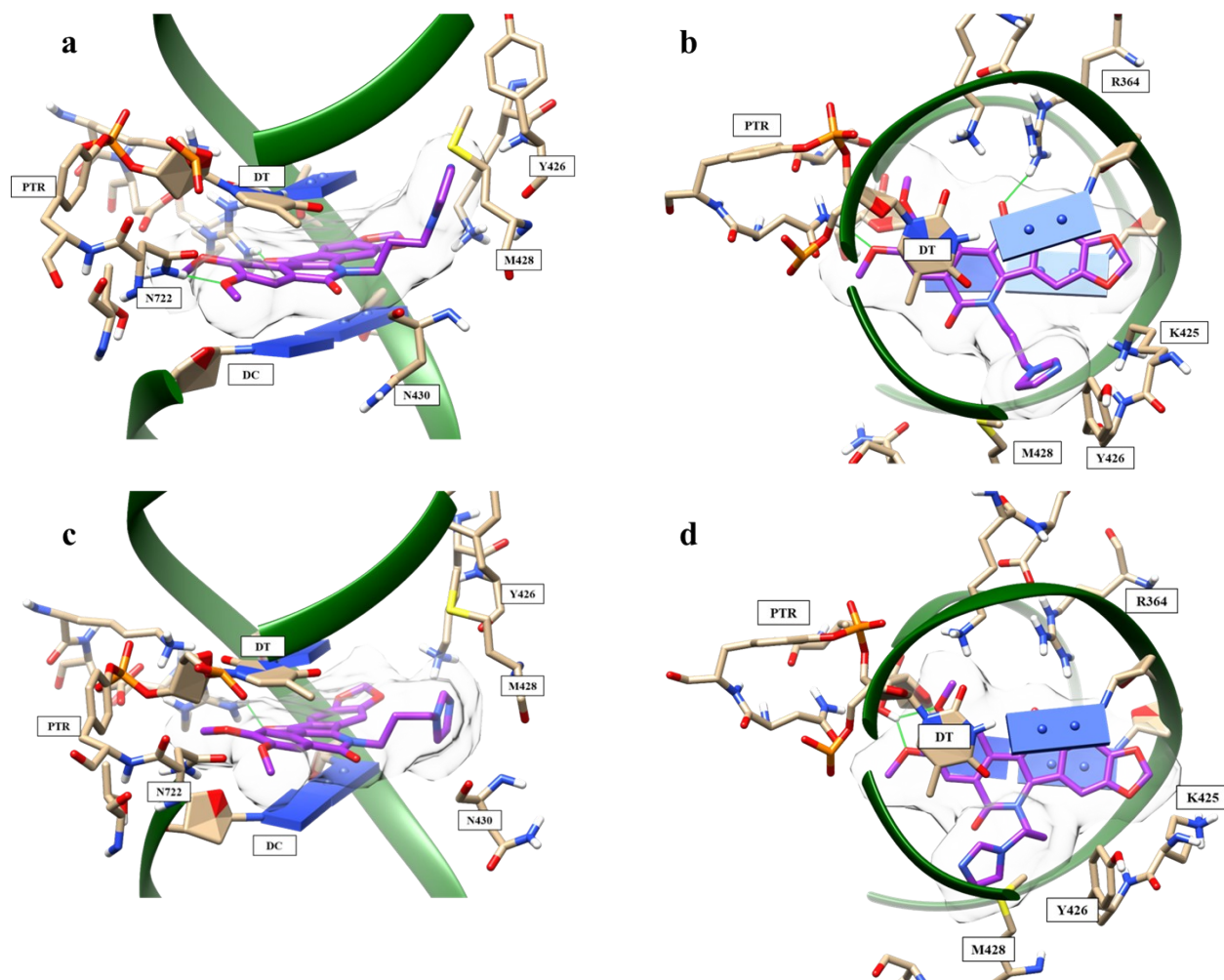
Figure S7. Different orientations of L77 were observed in W simulations. **a**. The beginning of the simulation side view; **b**. top view; **c**. side view at 100 ns; **d**. top view at 100ns. The DNA backbone is shown in green ribbon. The surrounding residues as well as the stacking DNA bases are shown with stick representation. The L77 compound is shown in purple.
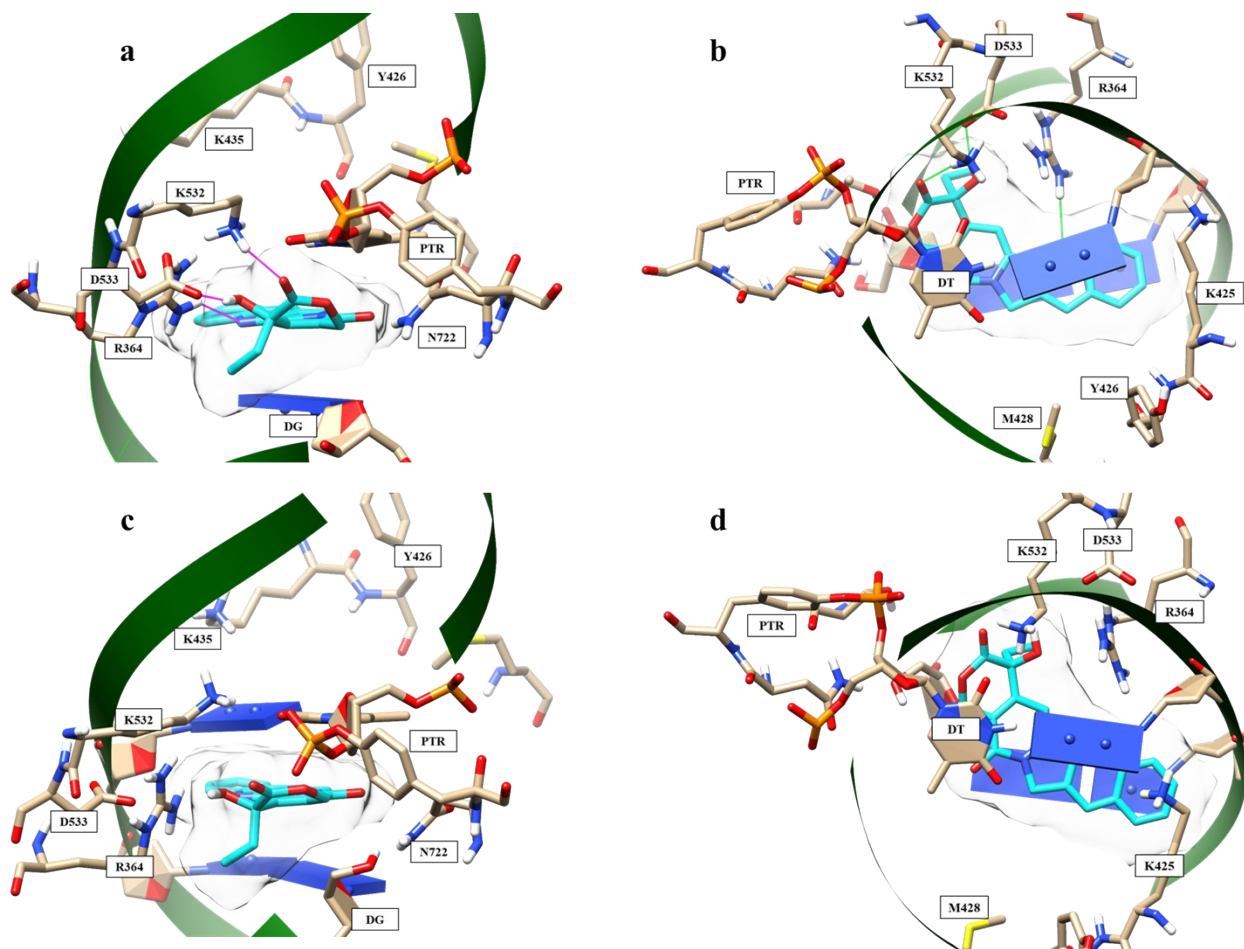
Figure S8. Different orientations of CPT were observed in M1 simulations. **a**. The beginning of the simulation side view; **b**. top view; **c**. side view at 300 ns; **d**. top view at 300ns. The DNA backbone is shown in green ribbon. The surrounding residues as well as the stacking DNA bases are shown with stick representation. The CPT compound is shown in cyan.
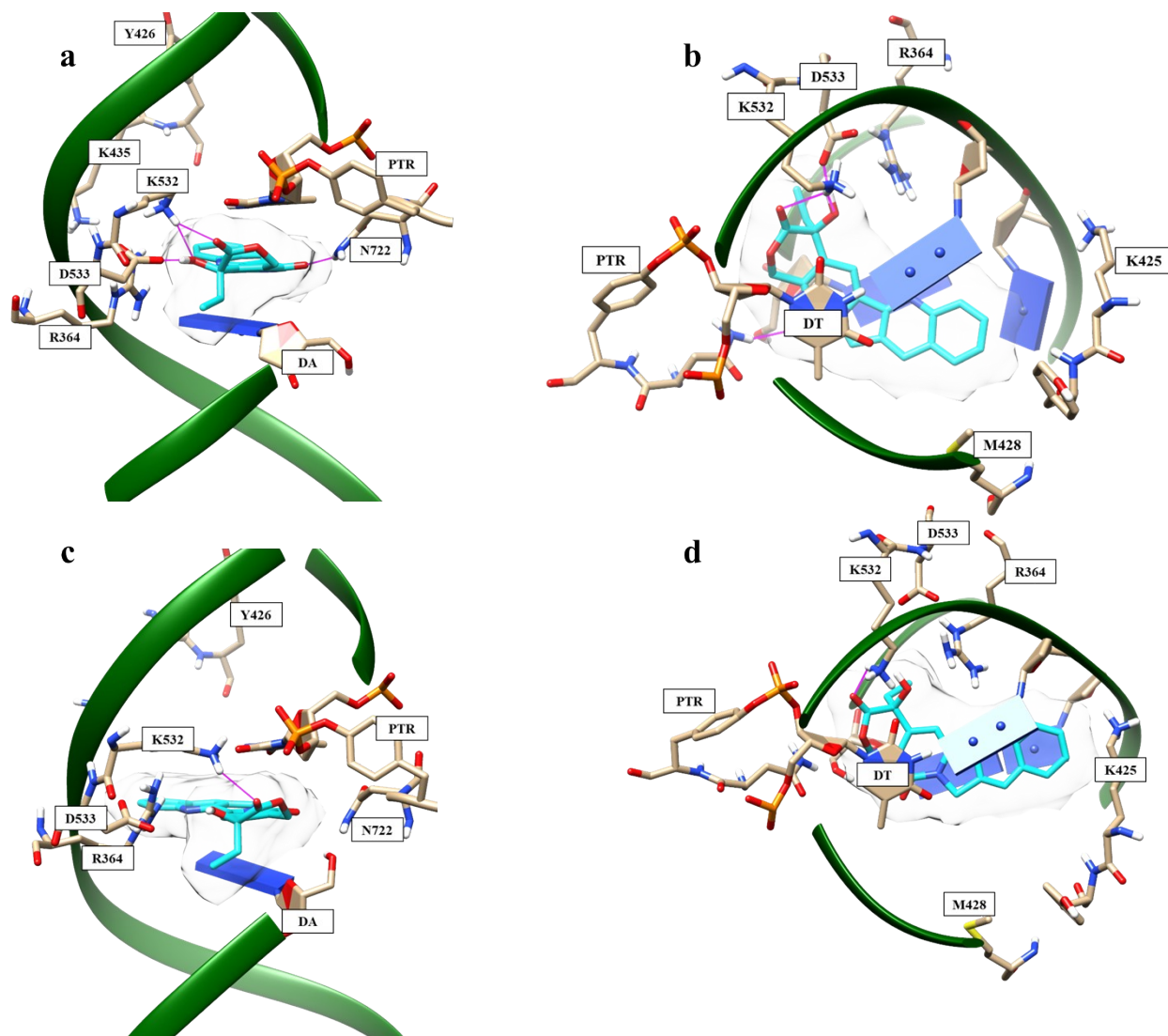
Figure S9. Different orientations of CPT were observed in M2 simulations. **a**. The beginning of the simulation side view; **b**. top view; **c**. side view at 300 ns; **d**. top view at 300ns. The DNA backbone is shown in green ribbon. The surrounding residues as well as the stacking DNA bases are shown with stick representation. The CPT compound is shown in cyan.
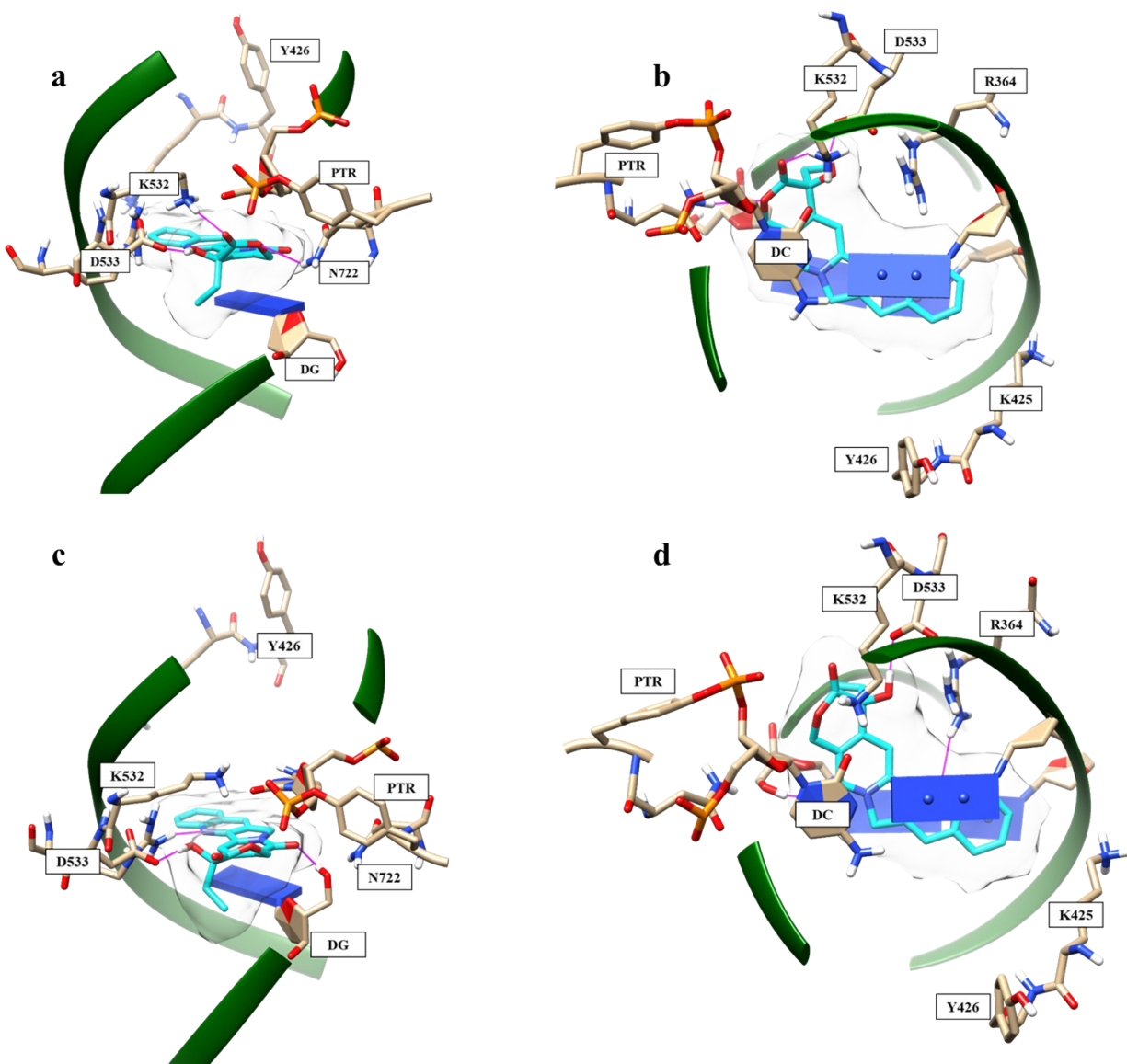
Figure S10. Different orientations of CPT were observed in M3 simulations. **a**. The beginning of the simulation side view; **b**. top view; **c**. side view at 300 ns; **d.** top view at 300ns. The DNA backbone is shown in green ribbon. The surrounding residues as well as the stacking DNA bases are shown with stick representation. The CPT compound is shown in cyan.
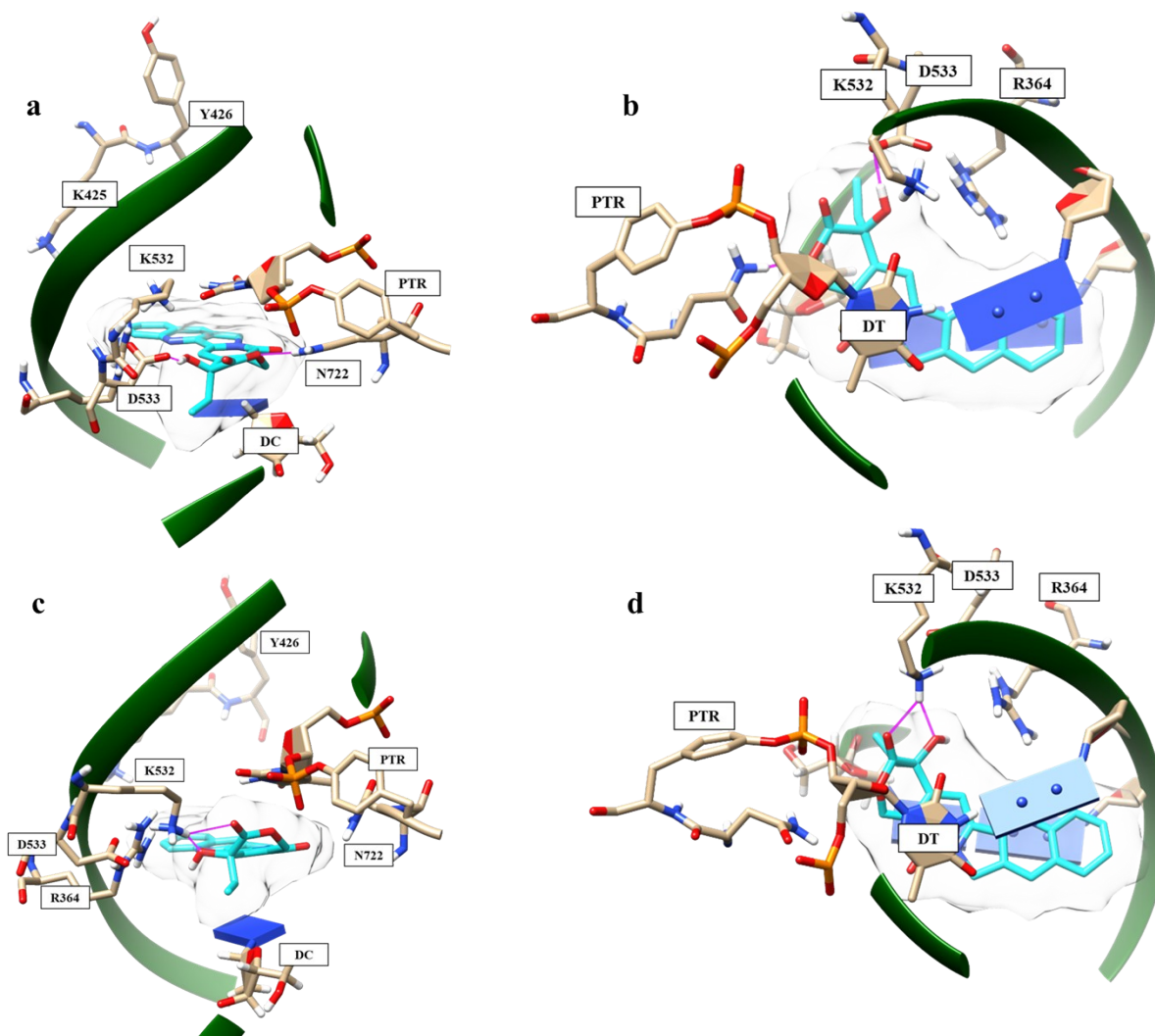
Figure S11. Different orientations of CPT were observed in W simulations. **a**. The beginning of the simulation side view; **b**. top view; **c**. side view at 300 ns; **d**. top view at 300ns. The DNA backbone is shown in green ribbon. The surrounding residues as well as the stacking DNA bases are shown with stick representation. The CPT compound is shown in cyan.
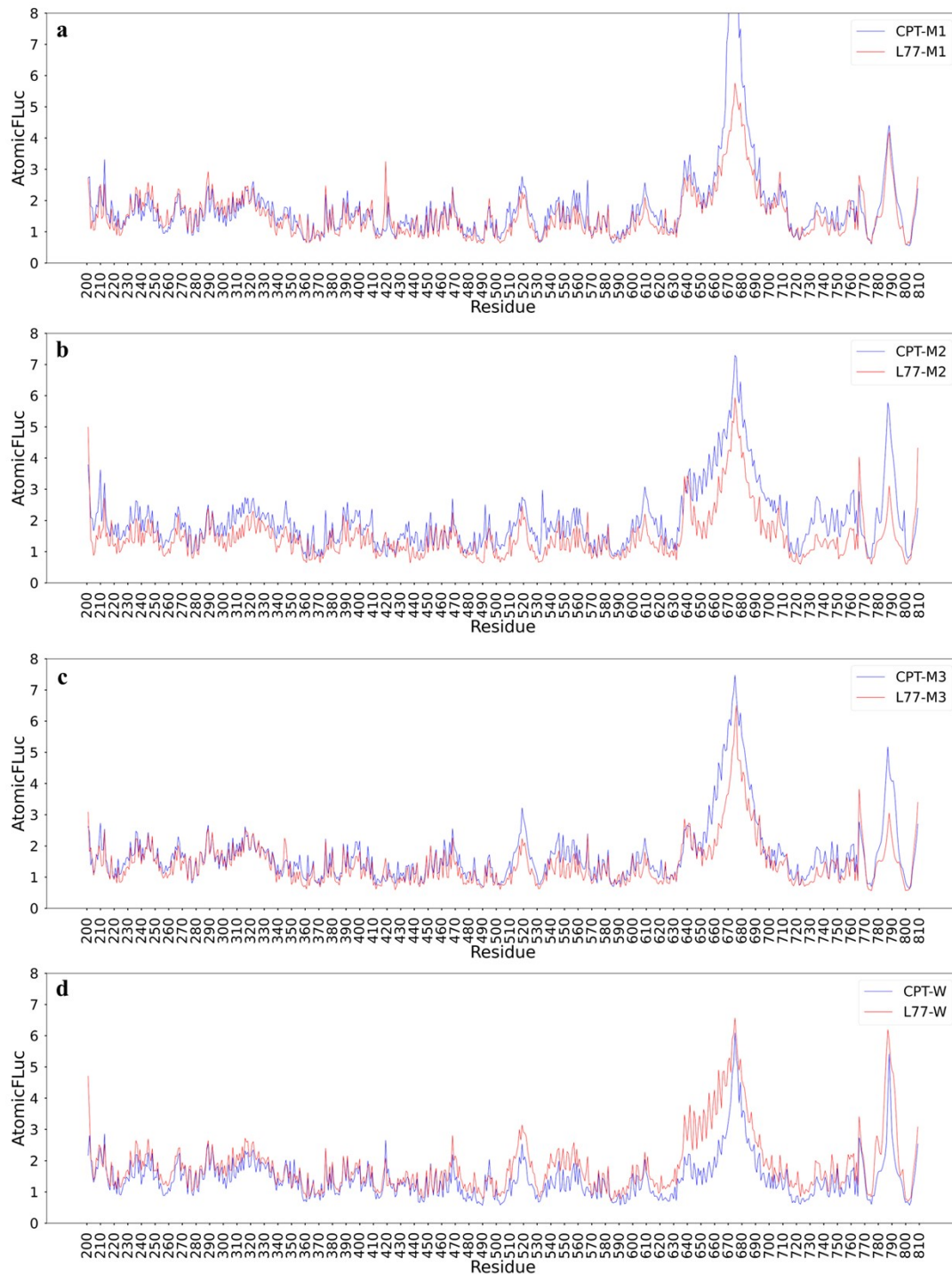
Figure S12. The RMSF plots of L77 (red) and CPT (blue) duplicate simulations for **a**. M1, **b**. M2, **c**. M3, and **d**. W simulations. N-terminal domain, core domain, linker domain, and C-terminal domain correspond to residues 1-214 (not shown here), 215-635, 636-712, and 713-765, respectively. The DNA bases start from 766-809, and 810 is the inhibitor compound.
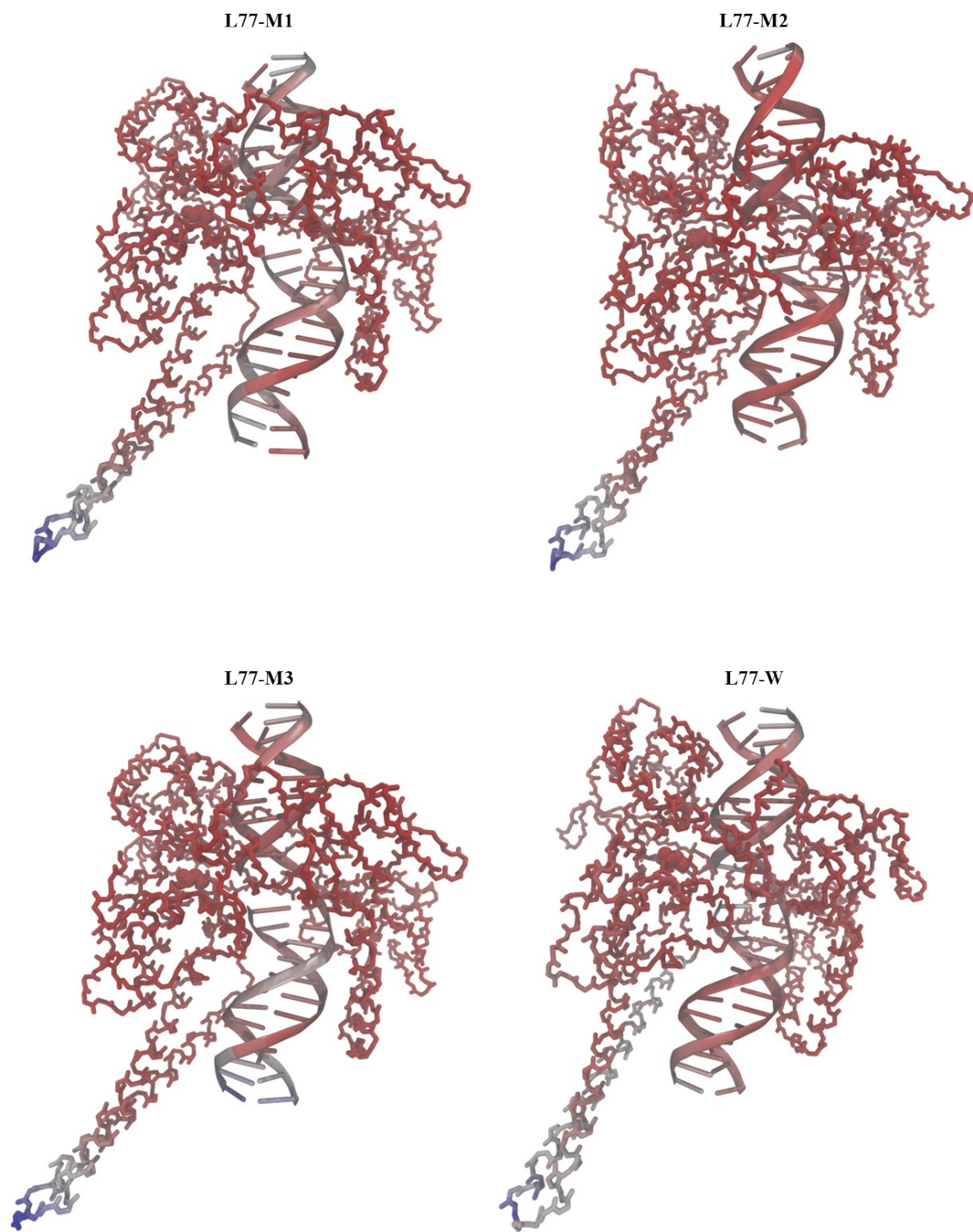
L77-M1

L77-M2

L77-M3

L77-W

Figure S13. The main dynamics obtained from Principal component analysis (PCA) of L77 simulations. The color scale from red to white to blue indicates the low mobility to high mobility.
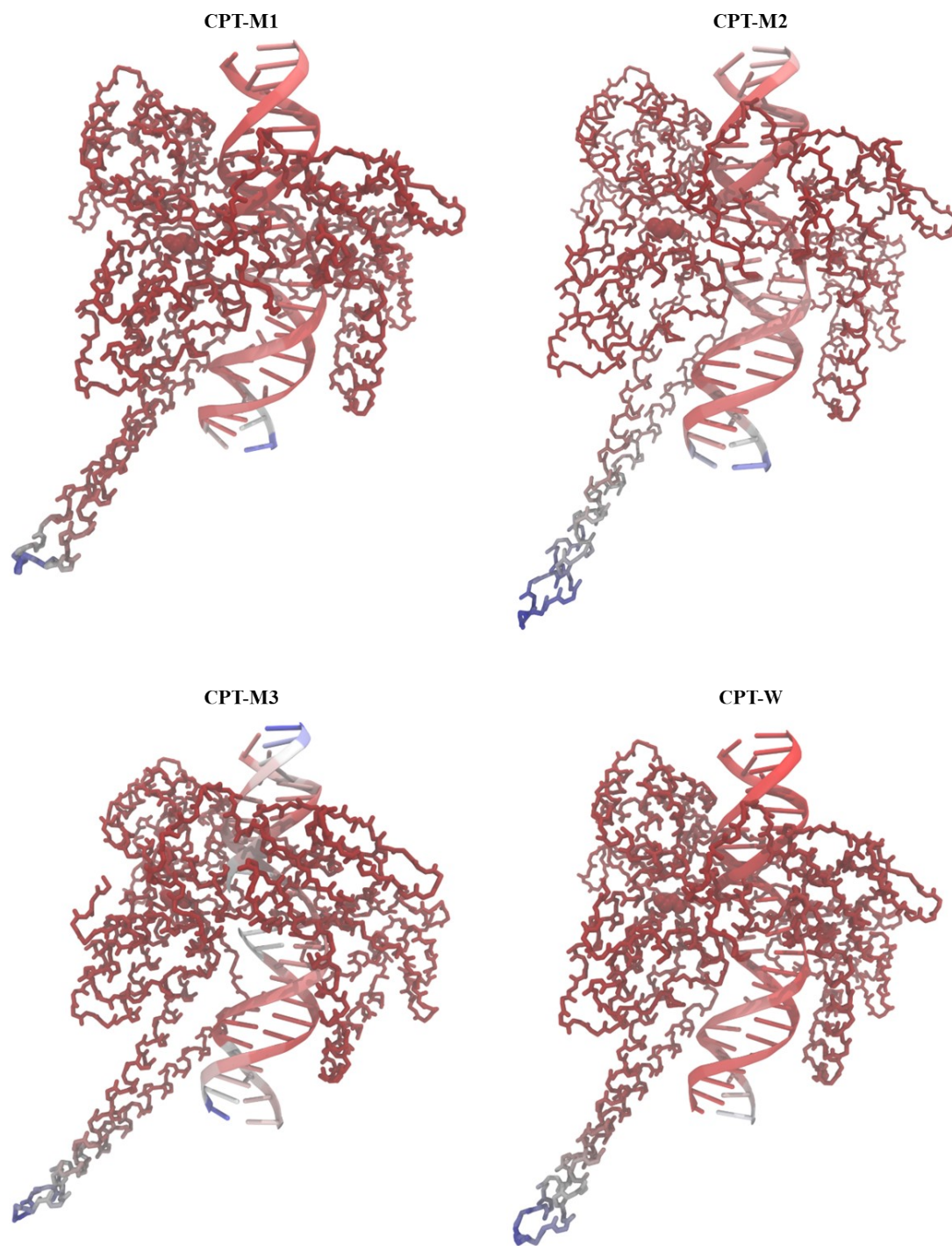
CPT-M1

CPT-M2

CPT-M3

CPT-W

Figure S14. The main dynamics obtained from Principal component analysis (PCA) of CPT simulations. The color scale from red to white to blue indicates the low mobility to high mobility.
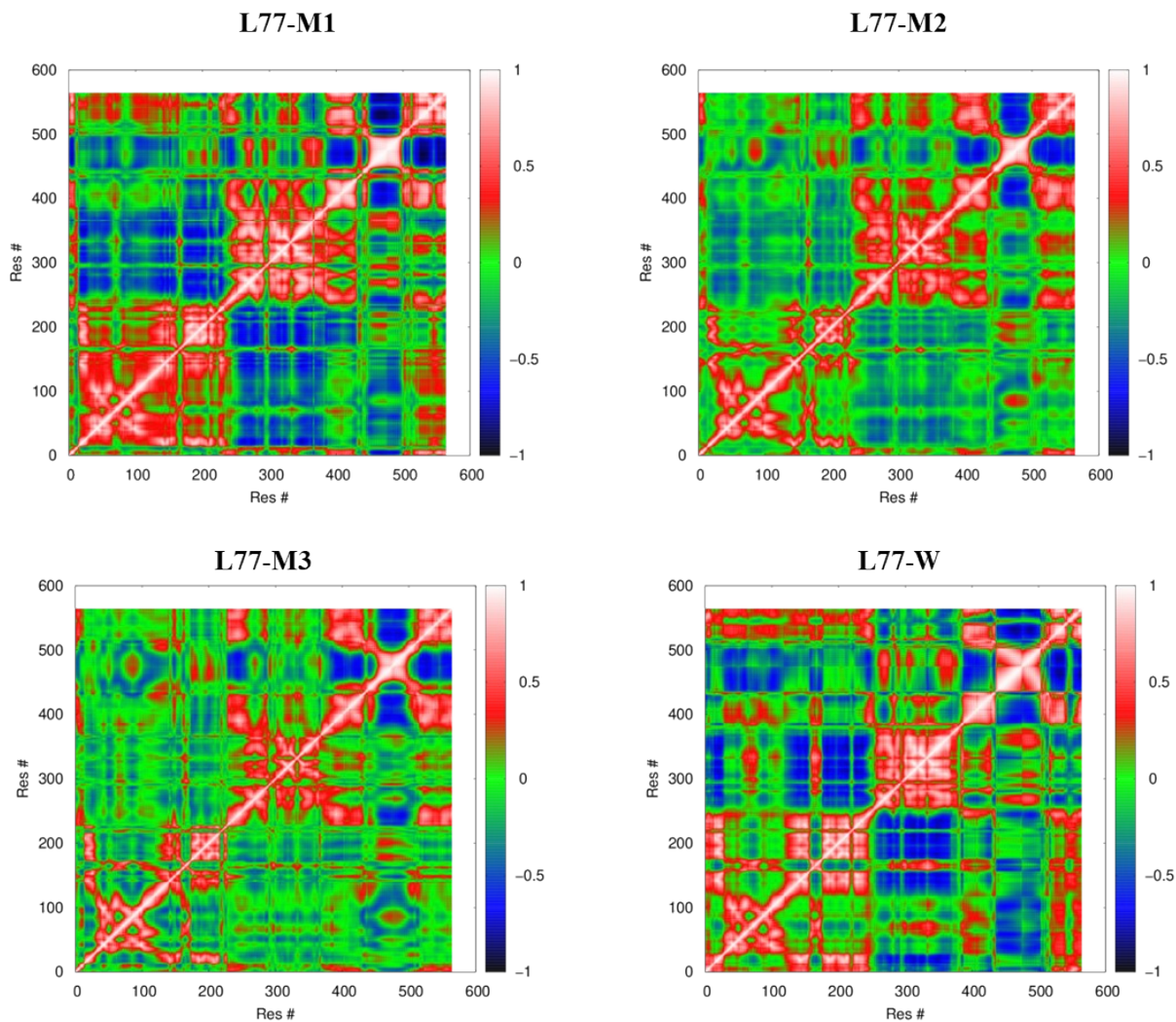
L77-M1

L77-M2

L77-M3

L77-W

Figure S15. The dynamic cross-correlation analysis of L77 simulations. From low correlation to high correlation, the residue pairs are colored from blue to red, respectively.
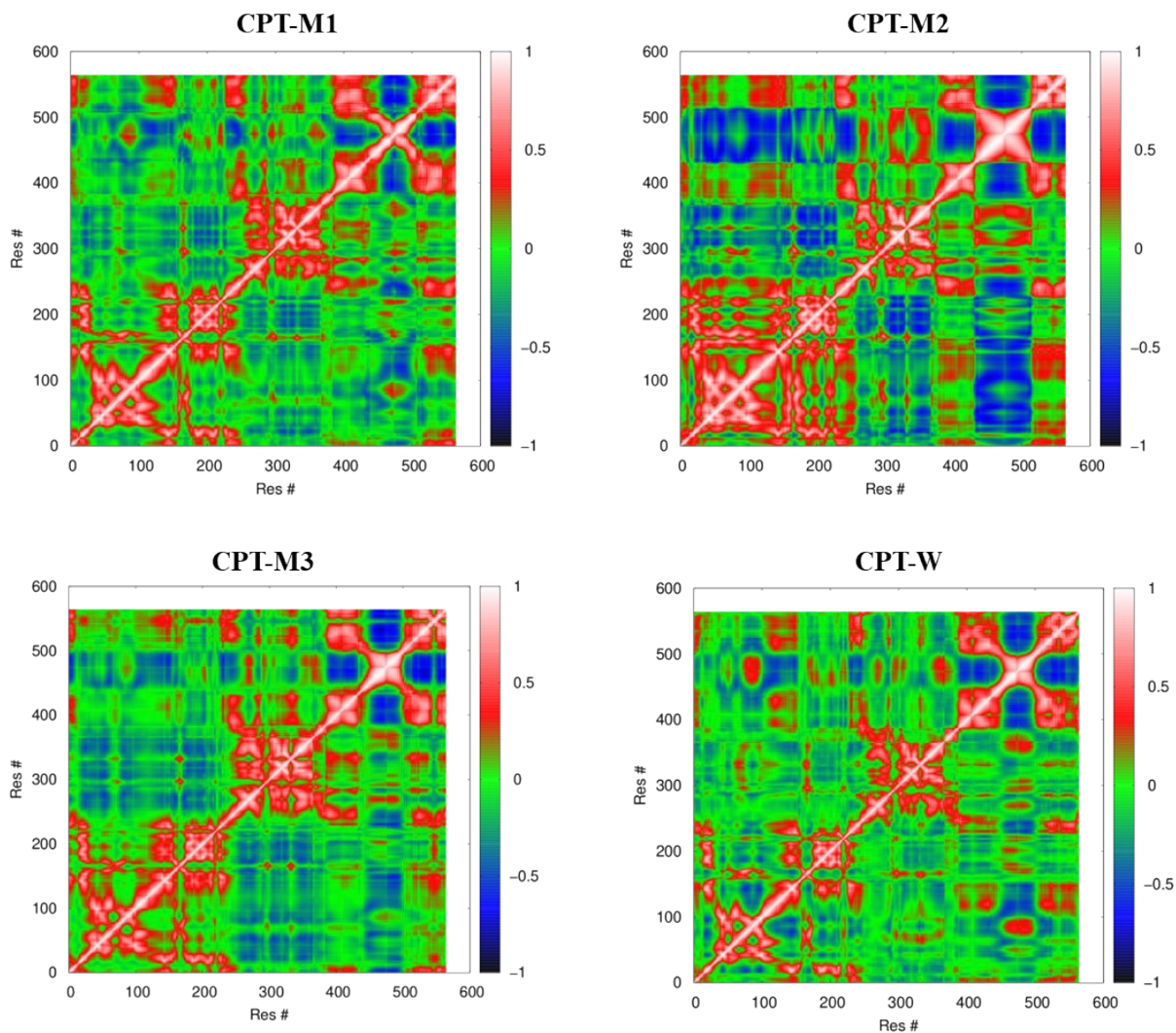
Figure S16. Dynamic cross-correlation analysis of CPT simulations. From low correlation to high correlation, the residue pairs are colored from blue to red, respectively.