

## **Supporting Information**

### **All-atomic Simulations on Human Telomeric G-quadruplex**

#### **DNA Binding with Thioflavin T**

Di Luo, Yuguang Mu\*

\*E-mail: [ygm@ntu.edu.sg](mailto:ygm@ntu.edu.sg) Telephone: (+65) 63162885

School of Biological Sciences,

Nanyang Technological University

Singapore 637551, Singapore

**Table S1 MD Simulation system sizes**

DNA	ThT	K <sup>+</sup>	Cl <sup>-</sup>	SOL	Cubic Box Volume(nm <sup>3</sup> )
1KF1	5	27	14	6837	236.81
143D	5	26	10	4781	158.73
2JSM	5	26	9	4438	149.58
2JSL	5	30	11	5562	184.49
2KF8	5	24	8	4191	140.91

**Table S2 Clustering analysis and MM-GBSA calculations for the determinations of binding modes and binding affinities**

(A) Clustering analysis and MM-GBSA calculations of 1KF1-ThTs

Structure (run-1)	No. of cluster	No. of structures in 1st clusters	Percentage (%)	RMSD	MM-GBSA (kcal/mol)	Binding mode
1KF1-ThT-1	20	14061	56.24	0.244	-24.14±1.90	3'-end
1KF1-ThT-2	25	15909	63.64	0.248	-22.06±2.60	5'-end
1KF1-ThT-3	205	1351	8.34	0.225	-9.29±1.45	
1KF1-ThT-4	330	2151	8.71	0.287	-11.12±1.21	
1KF1-ThT-5	59	11554	46.24	0.235	-24.53±3.71	5'-end

Structure (run-2)	No. of cluster	No. of structures in 1st clusters	Percentage (%)	RMSD	MM-GBSA (kcal/mol)	Binding mode
1KF1-ThT-1	31	10745	42.98	0.277	-23.91±1.93	5'-end
1KF1-ThT-2	96	8020	32.08	0.255	-18.81±3.06	3'-end
1KF1-ThT-3	12	20499	81.99	0.209	-22.73±2.15	3'-end
1KF1-ThT-4	349	4336	17.34	0.214	-17.35±2.93	DA19
1KF1-ThT-5	269	3559	14.24	0.232	-16.61±1.86	DT12, DA13

Structure (run-3)	No. of cluster	No. of structures in 1st clusters	Percentage (%)	RMSD	MM-GBSA (kcal/mol)	Binding mode
1KF1-ThT-1	194	3586	14.34	0.293	-12.38±1.62	
1KF1-ThT-2	166	3496	13.98	0.276	-12.05±2.32	
1KF1-ThT-3	161	5003	20.01	0.203	-22.88±3.58	3'-end
1KF1-ThT-4	11	17046*	68.18	0.215	-25.77±2.38	5'-end
1KF1-ThT-5	52	12804	51.21	0.262	-25.03±1.81	3'-end

(\*: the center structure of this cluster is the initial structure of the WT-MetaD simulation of 1KF1-ThT.)

(B) Clustering analysis and MM-GBSA calculations of 143D-ThTs

Structure (run-1)	No. of cluster	No. of structures in 1st clusters	Percentage (%)	RMSD	MM-GBSA (kcal/mol)	Binding mode
143D-ThT-1	73	12150	48.86	0.231	-14.85±1.90	
143D-ThT-2	124	10654	42.94	0.268	-12.77±1.60	
143D-ThT-3	60	10170	51.43	0.293	-19.04±3.63	DG10
143D-ThT-4	186	1162	13.12	0.298	-1.76±1.18	
143D-ThT-5	82	9176	36.76	0.238	-19.32±2.24	DT17

Structure (run-2)	No. of cluster	No. of structures in 1st clusters	Percentage (%)	RMSD	MM-GBSA (kcal/mol)	Binding mode
143D-ThT-1	66	17806	71.22	0.161	-16.35±1.50	DT11, DA13
143D-ThT-2	32	19038	76.15	0.235	-11.99±1.51	
143D-ThT-3	194	4028	16.11	0.290	-14.90±1.95	
143D-ThT-4	85	7319	29.27	0.257	-13.49±2.07	
143D-ThT-5	53	10976	43.90	0.241	-16.95±2.34	DA7, DT18

Structure (run-3)	No. of cluster	No. of structures in 1st clusters	Percentage (%)	RMSD	MM-GBSA (kcal/mol)	Binding mode
143D-ThT-1	206	1694	6.78	0.235	-7.69±1.27	
143D-ThT-2	140	12177	48.71	0.163	-14.97±1.92	
143D-ThT-3	64	11447	45.79	0.260	-13.54±1.59	
143D-ThT-4	86	7875	31.50	0.190	-14.66±1.79	
143D-ThT-5	123	6321	25.28	0.257	-3.36±1.97	

(C) Clustering analysis and MM-GBSA calculations of 2JSM-ThTs

Structure (run-1)	No. of cluster	No. of structures in 1st clusters	Percentage (%)	RMSD	MM-GBSA (kcal/mol)	Binding mode
2JSM-ThT-1	120	7909	42.78	0.216	-15.21±2.21	DA2
2JSM-ThT-2	194	9705*	40.08	0.239	-17.10±1.69	DT1, DT19
2JSM-ThT-3	20	20093	80.37	0.165	-35.19±6.22	sandwich
2JSM-ThT-4	182	6130	25.53	0.240	-22.31±2.61	DT7
2JSM-ThT-5	91	4346	26.59	0.208	-16.18±2.70	DT19

Structure (run-2)	No. of cluster	No. of structures in 1st clusters	Percentage (%)	RMSD	MM-GBSA (kcal/mol)	Binding mode
2JSM-ThT-1	237	7214	28.85	0.207	-15.14±2.13	DT18
2JSM-ThT-2	86	18429	73.71	0.182	-26.63±3.33	DG15 Groove
2JSM-ThT-3	97	11094	44.37	0.213	-15.20±1.81	DA2, DT19
2JSM-ThT-4	302	2063	8.25	0.304	-10.03±1.39	
2JSM-ThT-5	330	1370	5.48	0.297	-10.17±1.44	

Structure (run-3)	No. of cluster	No. of structures in 1st clusters	Percentage (%)	RMSD	MM-GBSA (kcal/mol)	Binding mode
2JSM-ThT-1	92	9796	39.18	0.231	-21.58±2.84	DT12, DA14
2JSM-ThT-2	62	12534	50.13	0.244	-15.02±1.99	DT19
2JSM-ThT-3	108	4343	17.37	0.177	-18.14±2.90	DG5 Groove
2JSM-ThT-4	230	4958	19.83	0.278	-12.70±1.78	
2JSM-ThT-5	183	1809	7.24	0.295	-3.14±1.12	

(\*: the center structure of this cluster is the initial structure of the WT-MetaD simulation of 2JSL-ThT.)

## (D) Clustering analysis and MM-GBSA calculations of 2JSL-ThTs

Structure (run-1)	No. of cluster	No. of structures in 1st clusters	Percentage (%)	RMSD	MM-GBSA (kcal/mol)	Binding mode
2JSL-ThT-1	33	18828	75.33	0.246	-17.02±2.43	DA8, DT25
2JSL-ThT-2	67	10563	42.38	0.151	-19.97±1.94	DT1, DA14
2JSL-ThT-3	254	2714	14.74	0.292	-10.12±1.20	
2JSL-ThT-4	109	15304	62.59	0.225	-15.85±1.78	DA20
2JSL-ThT-5	60	17282	69.16	0.184	-18.54±1.90	DT6

Structure (run-2)	No. of cluster	No. of structures in 1st clusters	Percentage (%)	RMSD	MM-GBSA (kcal/mol)	Binding mode
2JSL-ThT-1	52	15641	62.56	0.199	-20.12±1.84	DT13, DA14
2JSL-ThT-2	258	7406	29.62	0.265	-17.06±2.53	DG3-5, 21-23
2JSL-ThT-3	203	2022	8.09	0.242	-5.57±2.46	
2JSL-ThT-4	122	13331	53.32	0.224	-16.68±1.91	DA8, DT25
2JSL-ThT-5	245	3907	15.63	0.285	-12.50±1.62	DA2

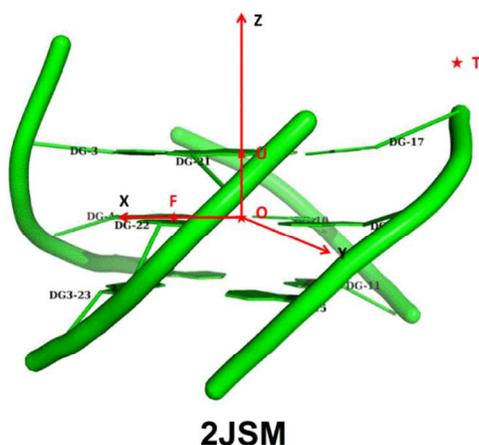
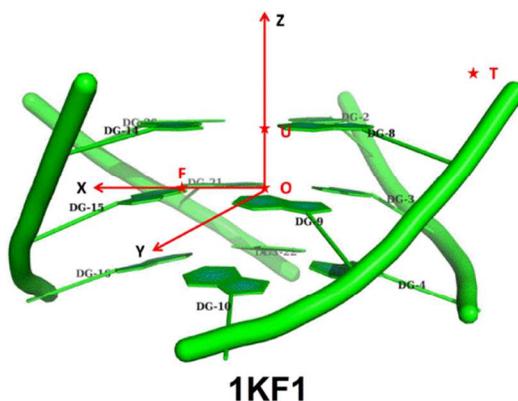
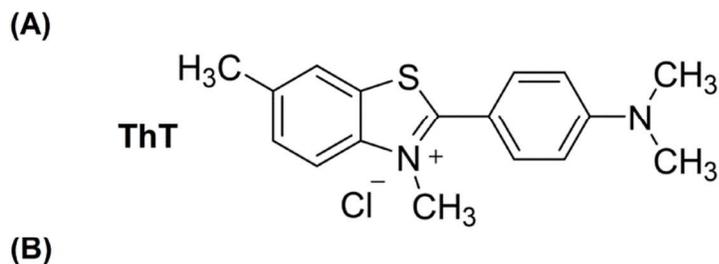
Structure (run-3)	No. of cluster	No. of structures in 1st clusters	Percentage (%)	RMSD	MM-GBSA (kcal/mol)	Binding mode
2JSL-ThT-1	112	2273	9.09	0.275	-3.38±1.27	
2JSL-ThT-2	30	15522	62.09	0.217	-16.86±1.67	DA2, DT13
2JSL-ThT-3	120	5428	21.71	0.239	-19.15±2.07	DT6
2JSL-ThT-4	19	16986	67.94	0.264	-19.40±2.42	DA8
2JSL-ThT-5	509	1083	4.33	0.307	-10.99±1.40	

## (E) Clustering analysis and MM-GBSA calculations of 2KF8-ThTs

Structure (run-1)	No. of cluster	No. of structures in 1st clusters	Percentage (%)	RMSD	MM-GBSA (kcal/mol)	Binding mode
2KF8-ThT-1	72	13162	53.67	0.257	-19.57±1.98	DT10
2KF8-ThT-2	212	7771	35.69	0.276	-11.58±1.30	
2KF8-ThT-3	87	305	21.49	0.254	-5.45±2.34	
2KF8-ThT-4	204	3449	20.22	0.262	-10.91±1.17	
2KF8-ThT-5	400	1522	9.75	0.267	-11.01±2.22	

Structure (run-1)	No. of cluster	No. of structures in 1st clusters	Percentage (%)	RMSD	MM-GBSA (kcal/mol)	Binding mode
2KF8-ThT-1	51	13241	52.96	0.225	-18.26±1.80	DT5, DT17
2KF8-ThT-2	222	5549	22.20	0.302	-10.63±1.41	
2KF8-ThT-3	154	3024	12.10	0.277	-12.30±1.96	
2KF8-ThT-4	165	5241	20.96	0.233	-17.67±2.23	DT10
2KF8-ThT-5	172	6098	24.39	0.257	-11.52±1.33	

Structure (run-1)	No. of cluster	No. of structures in 1st clusters	Percentage (%)	RMSD	MM-GBSA (kcal/mol)	Binding mode
2KF8-ThT-1	146	4896	19.58	0.273	-18.12±2.02	DT17
2KF8-ThT-2	233	3513	14.05	0.254	-11.57±1.22	
2KF8-ThT-3	143	5205	20.82	0.266	-11.72±1.22	
2KF8-ThT-4	196	5138	20.55	0.241	-11.98±1.29	
2KF8-ThT-5	149	8630	34.52	0.244	-14.79±1.78	



**Figure S1 (A) The chemical structure of ThT; (B) The CV definitions for the WT-MetaD simulations**

In (B), “O” represents the center of mass (COM) of the G-quadruplex (1KF1 or 2JSM). “U” stands for the COM of the 5’ end G-quartet of the G-quadruplex. “T” is the COM of ThT molecule. “F” is the COM of the flanking guanine bases (DG14-16 and DG20-22 for 1KF1; DG3-5 and DG21-23 for 2JSM). Therefore, CV1 “theta” is defined as the angle UOT; CV2 “dist” is defined as the distance OT; CV3 “phi” is defined as the dihedral angle between vectors  $\overline{UT}$  and  $\overline{OF}$ . (CV3 is not biased in the WT-MetaD simulations.)

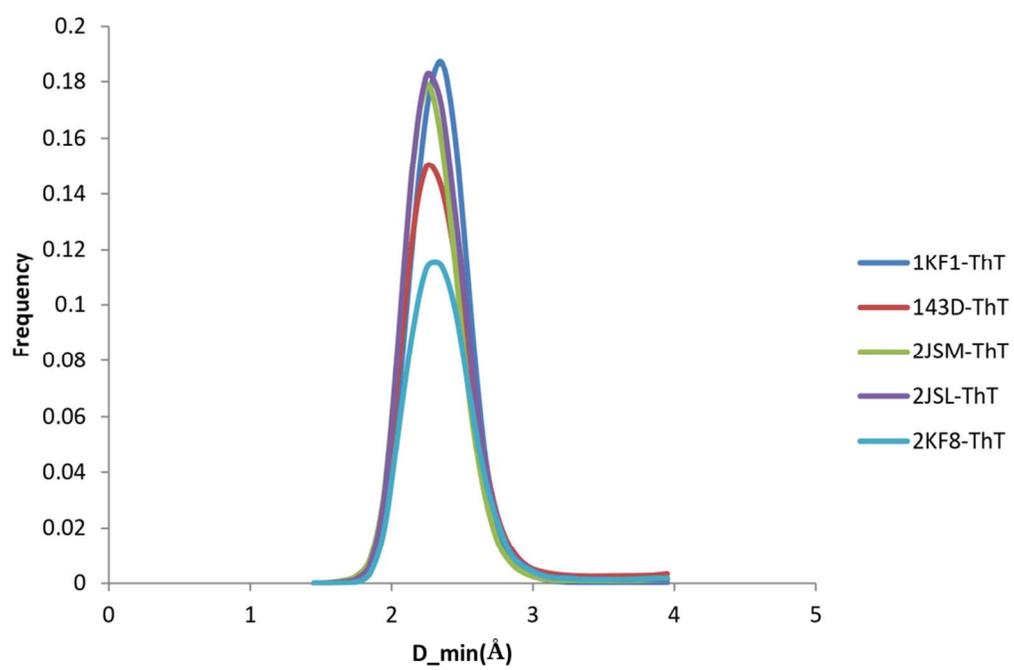
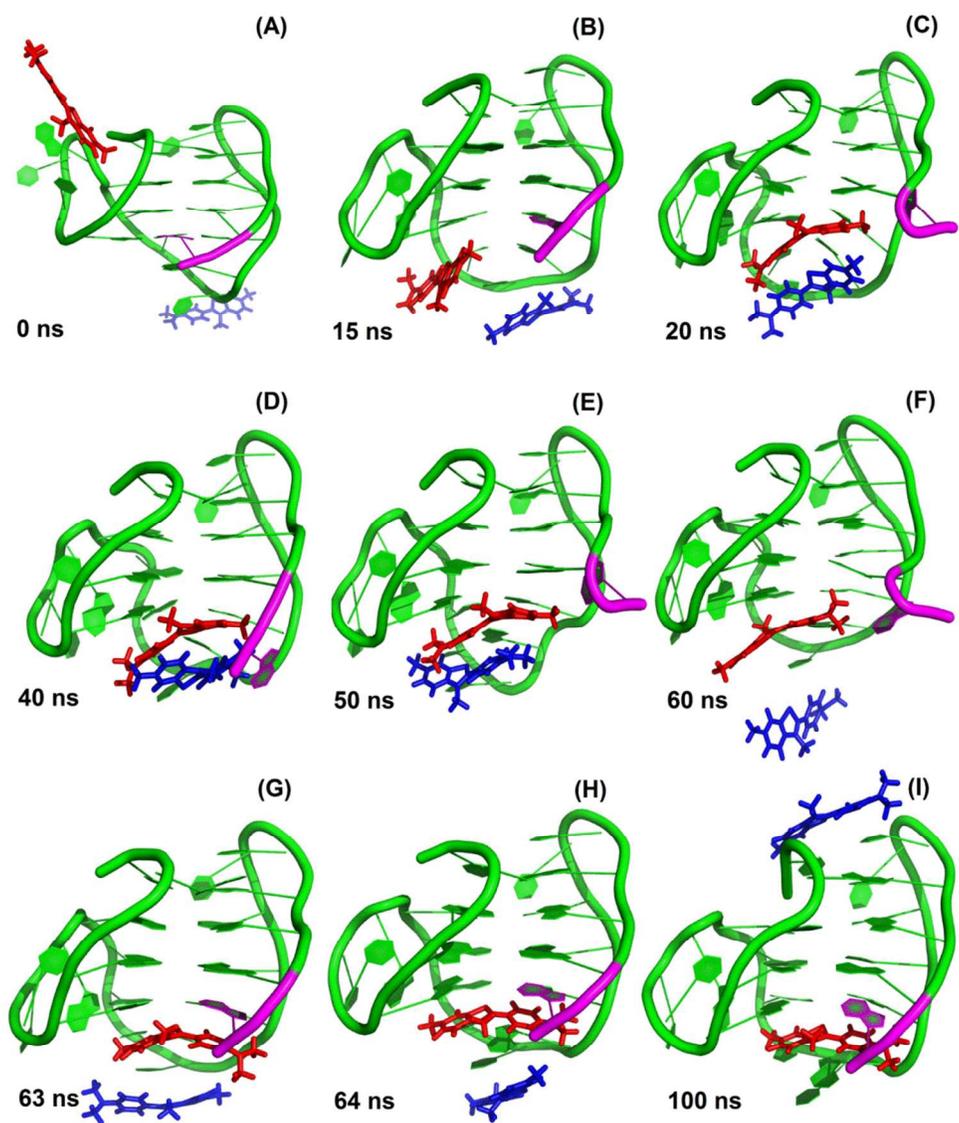
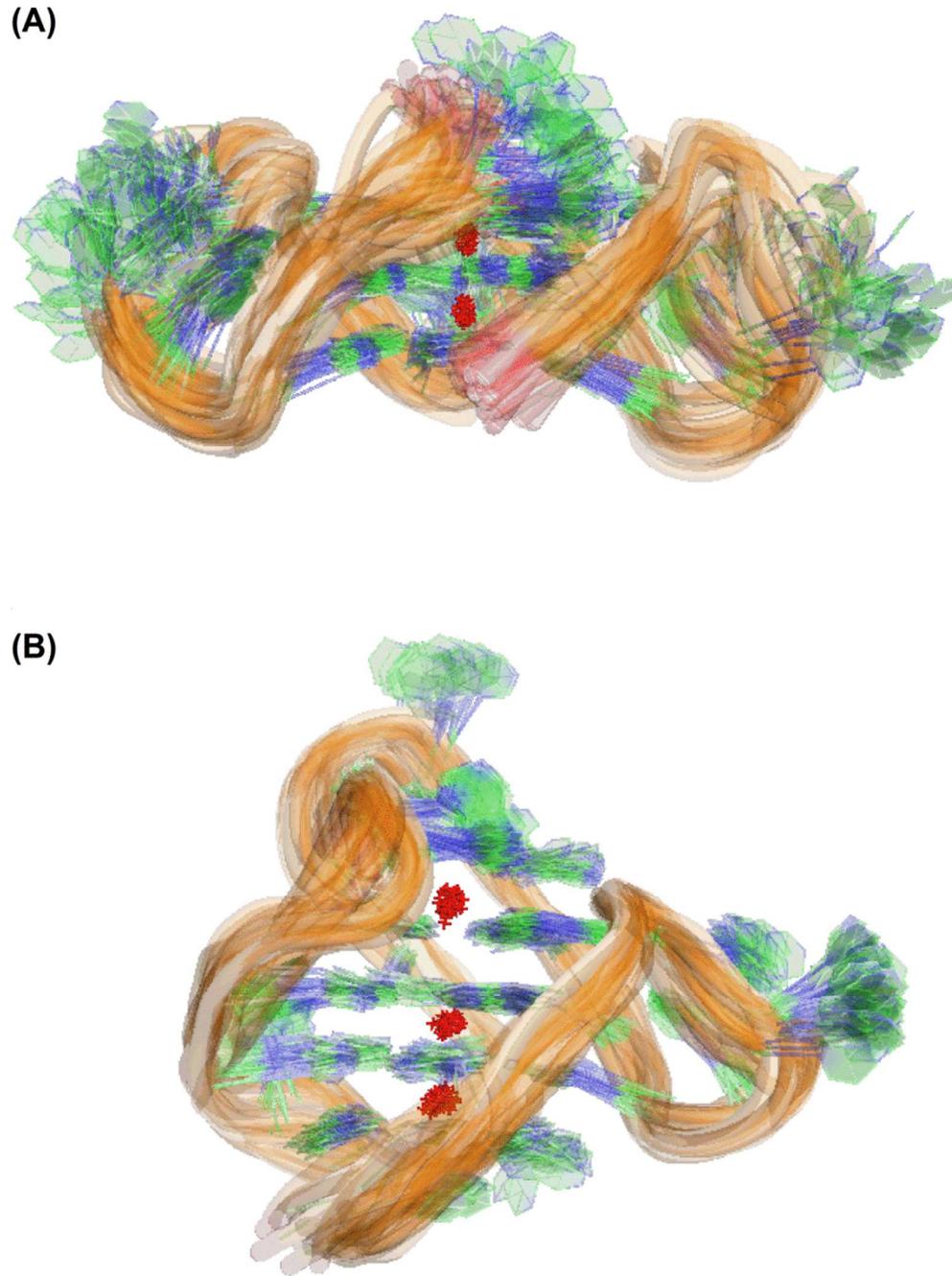


Figure S2 Minimum distance (d\_min) distributions



**Figure S3 Snapshots of the sandwich binding dynamics**

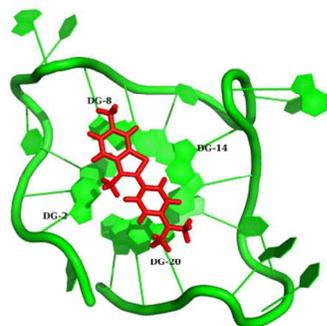
Selected snapshots of the sandwich binding process are shown. The initial conformation of 2JSM (green), ThT-2 (blue) and ThT-3 (red) is shown in (A). The DG23 (magenta) of 2JSM remains stable on the 3' end G-quartet until 15 ns as shown in (B). At 20 ns, DG23 tilts out of the 3' end G-quartet and ThT-3 bends itself to replace DG23 for stacking at the 3' end G-quartet as shown in (C). The stacking of ThT at 3' end G-quartet is facilitated by the underneath stacking with ThT-2 as shown in (D) and (E). At 60 ns, ThT-2 disassociates the stacking with ThT-3 and DG23 was about to resume its position at the 3' end G-quartet (shown in (F)). At 63 ns, DG23 returns back to the 3' end G-quartet and ThT-3 stacks underneath DG23 and DG5 (shown in (G)). At 64 ns, DA14 from the edgewise loop of 2JSM protruding inwards to act as the third layer of the sandwich binding structure. The sandwich binding structure remains stable until the end of the simulation (shown in (I)).



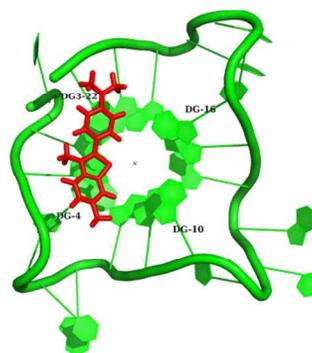
**Figure S4 Distributions of  $K^+$  (in red) on 1KF1 (A) and 2JSL (B)**

(A) exhibits the distribution patterns of  $K^+$  coordinating between the three G-quartets of 1KF1; (B) shows the distribution patterns of  $K^+$  coordinating above the 5' or below 3' end G-quartets of 2JSL as well as being sandwiched between the 2<sup>nd</sup> and 3<sup>rd</sup> G-quartets of 2JSL.

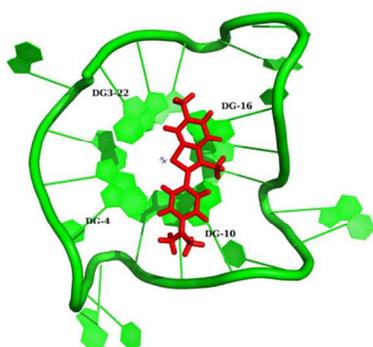
(A)



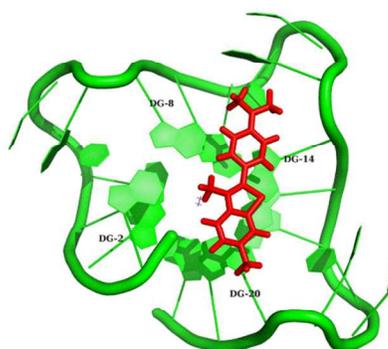
(B)



(C)

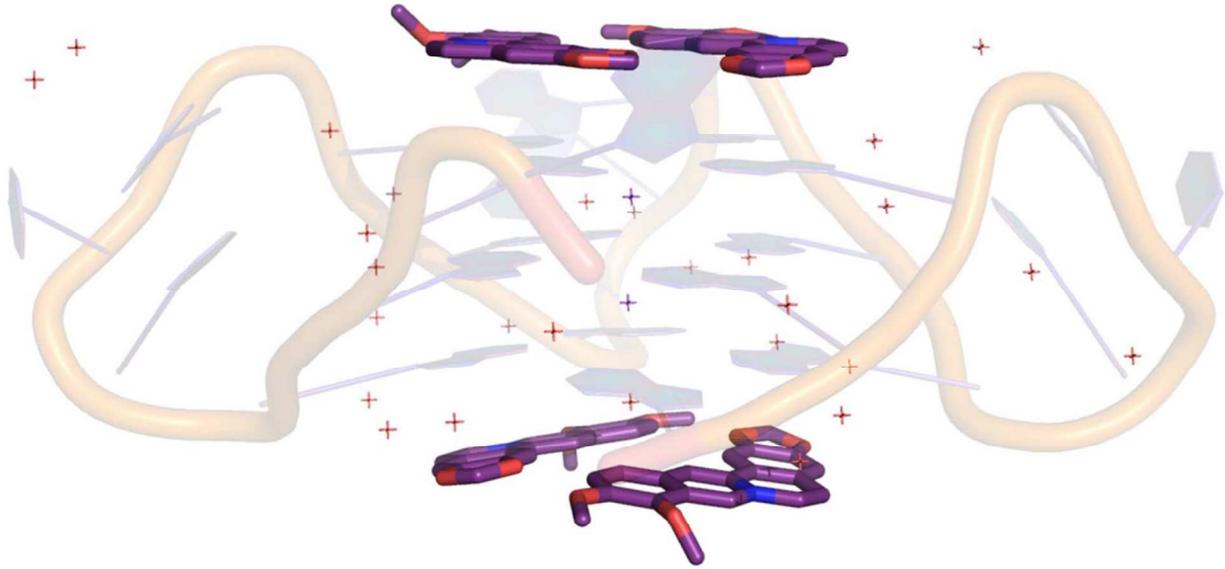


(D)



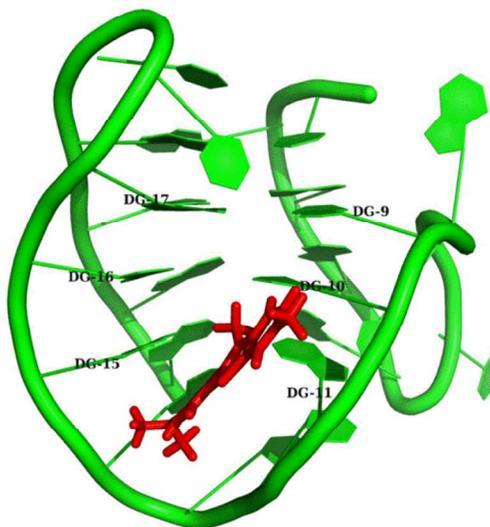
**Figure S5 Additional end-stacking conformations**

The end-stacking conformations observed in the other two parallel simulations are shown. The top views of end-stacking of ThT with DG8 & DG20 and DG14 & DG20 of 1KF1 are shown in (A) and (D), respectively; the bottom views of end-stacking of ThT with DG22 & DG4 and DG16 & DG10 of 1KF1 are shown in (B)-(C).

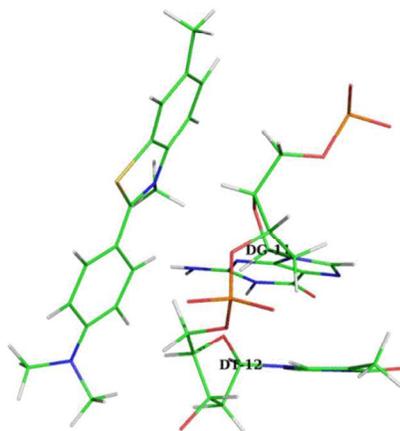


**Figure S6 Binding structure of 1KF1 with four berberine molecules (PDB: 3R6R)**

(A)

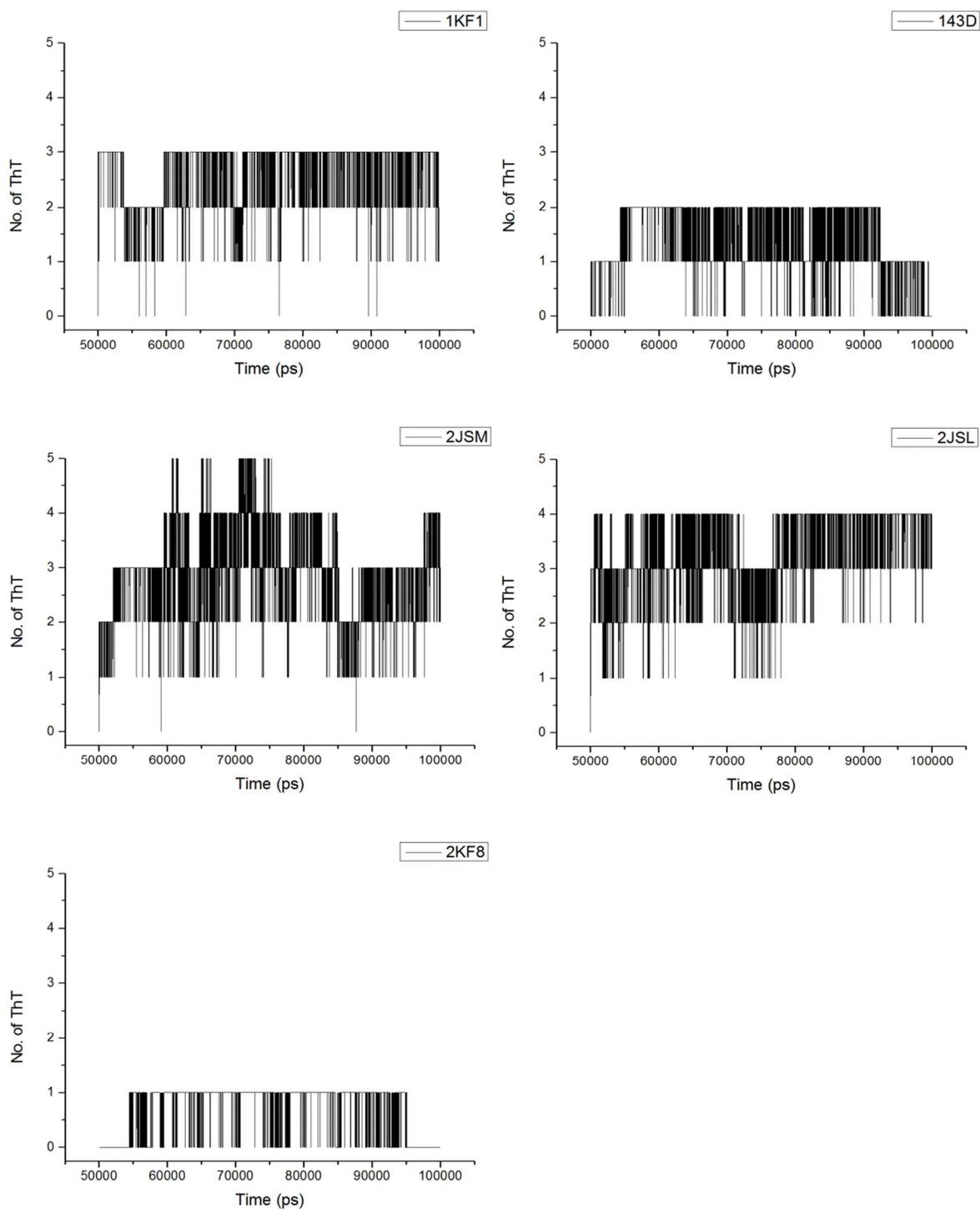


(B)



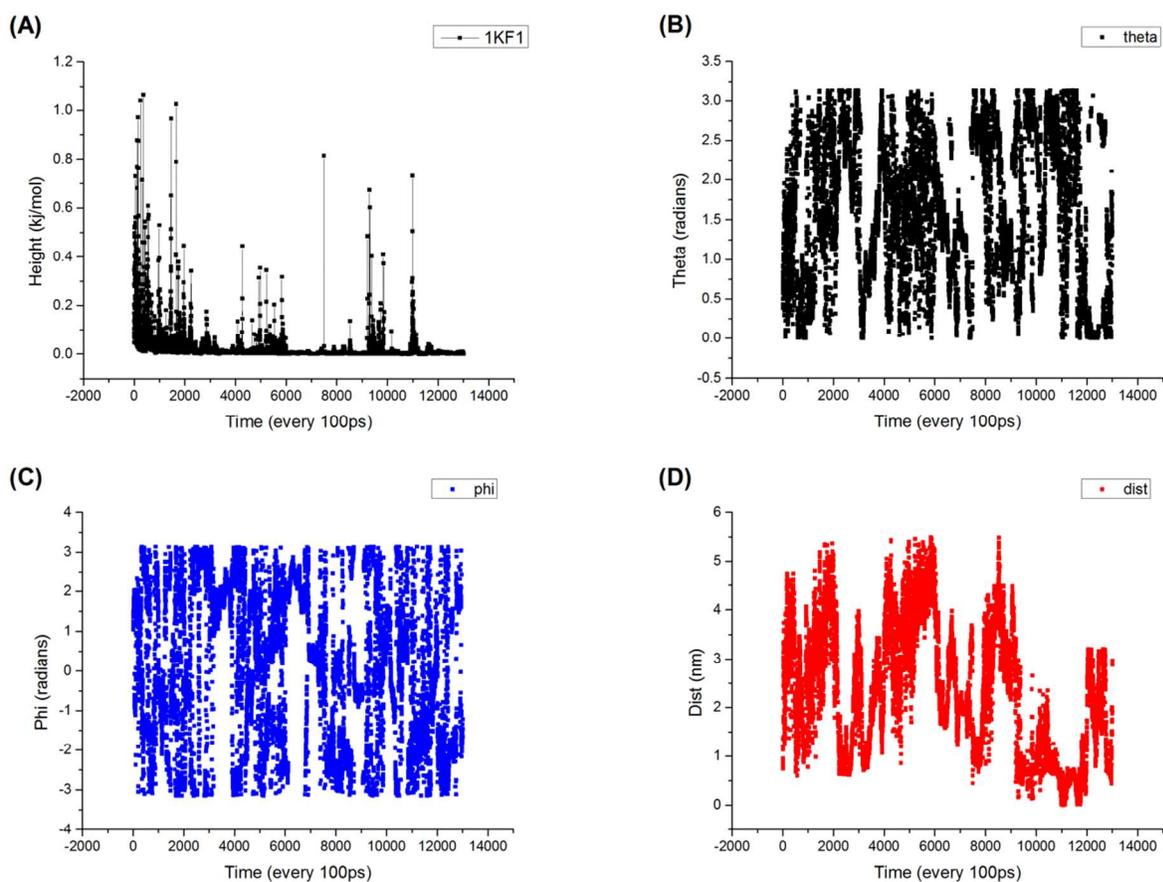
**Figure S7 Groove insertion conformation**

The groove insertion mode of ThT binding to the wide groove of 2JSM formed by DG9-DG17 is shown in (A). The groove insertion was driven by CH- $\pi$  interactions between the aromatic rings of ThT and the C5'-H5' groups of sugar rings of the nucleotides involving DG11 and DT12. The geometry of the CH-  $\pi$  interactions is shown in (B).



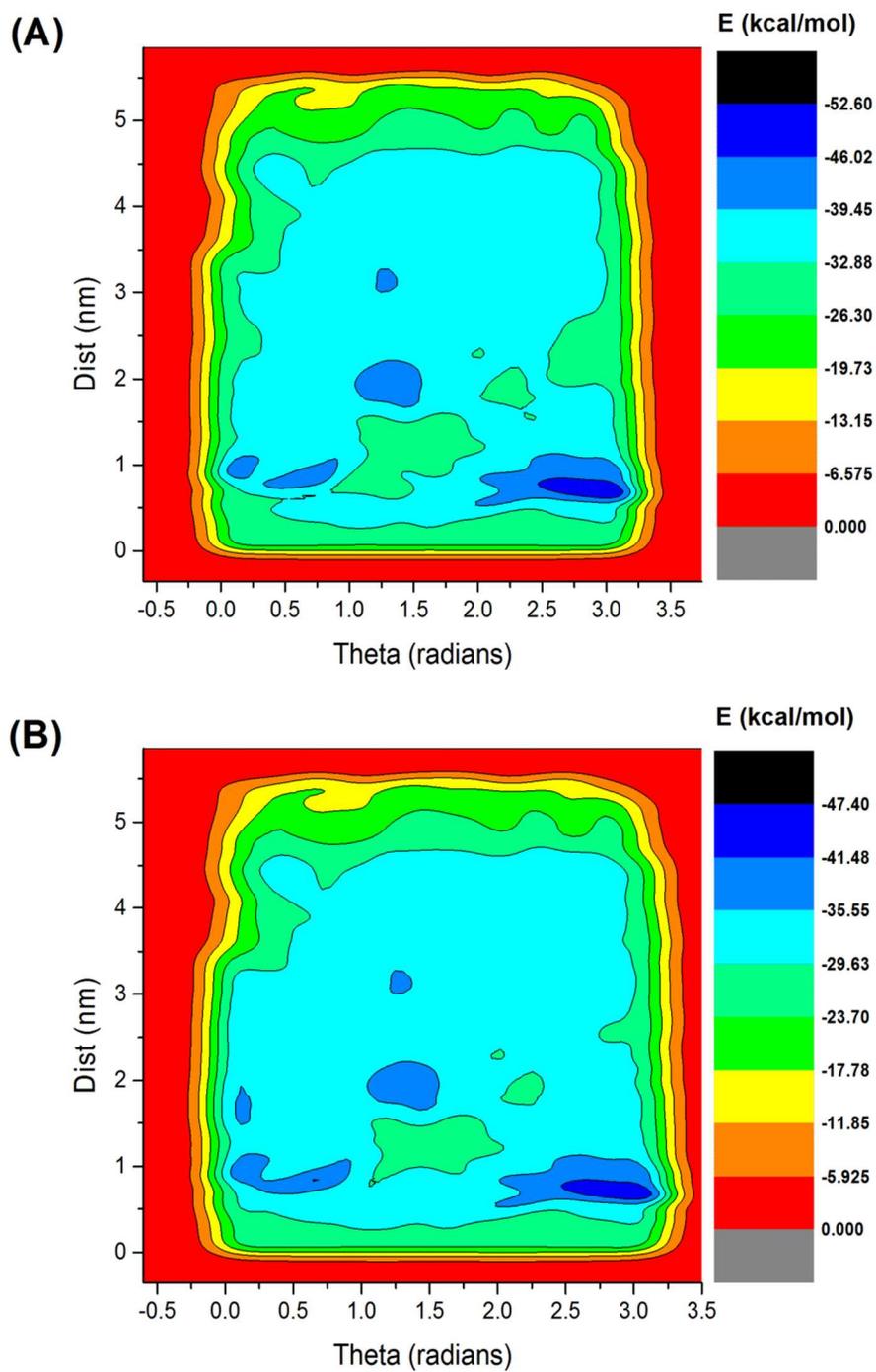
**Figure S8 Binding stoichiometry between G-quadruplexes and ThTs**

The number of ThTs showing effective binding to G-quadruplexes (indicated at the upper right corner) at each time step is shown.



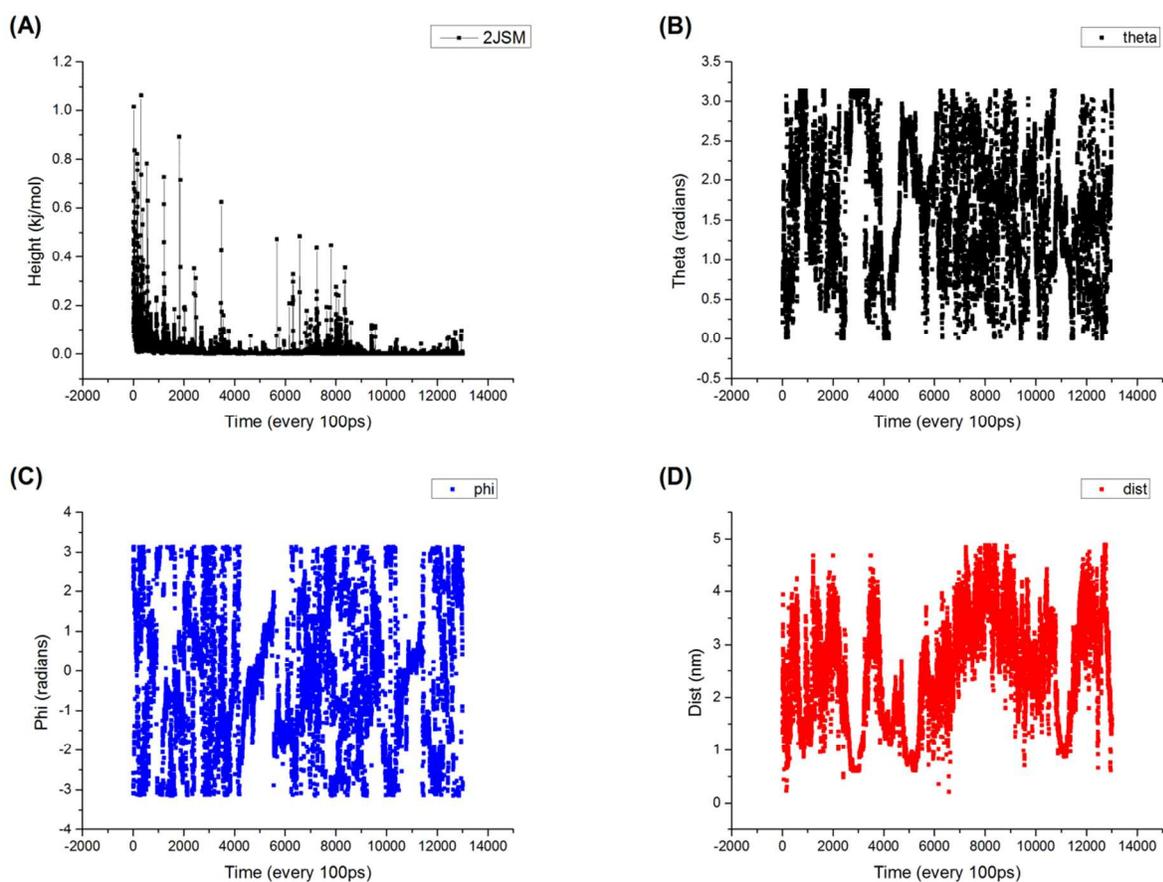
**Figure S9 Variations of Gaussian height (A) and CVs (B-D) as a function of time of the WT-MetaD simulation of 1KF1-ThT**

The rescaled Gaussian height as a function of simulation time is shown in (A). It is observed that after 1100 ns, the Gaussian height is approaching to zero and remained unchanged afterwards. In additions, the scattering variations of the three CVs with respect to simulation time are shown in (B) – (D).



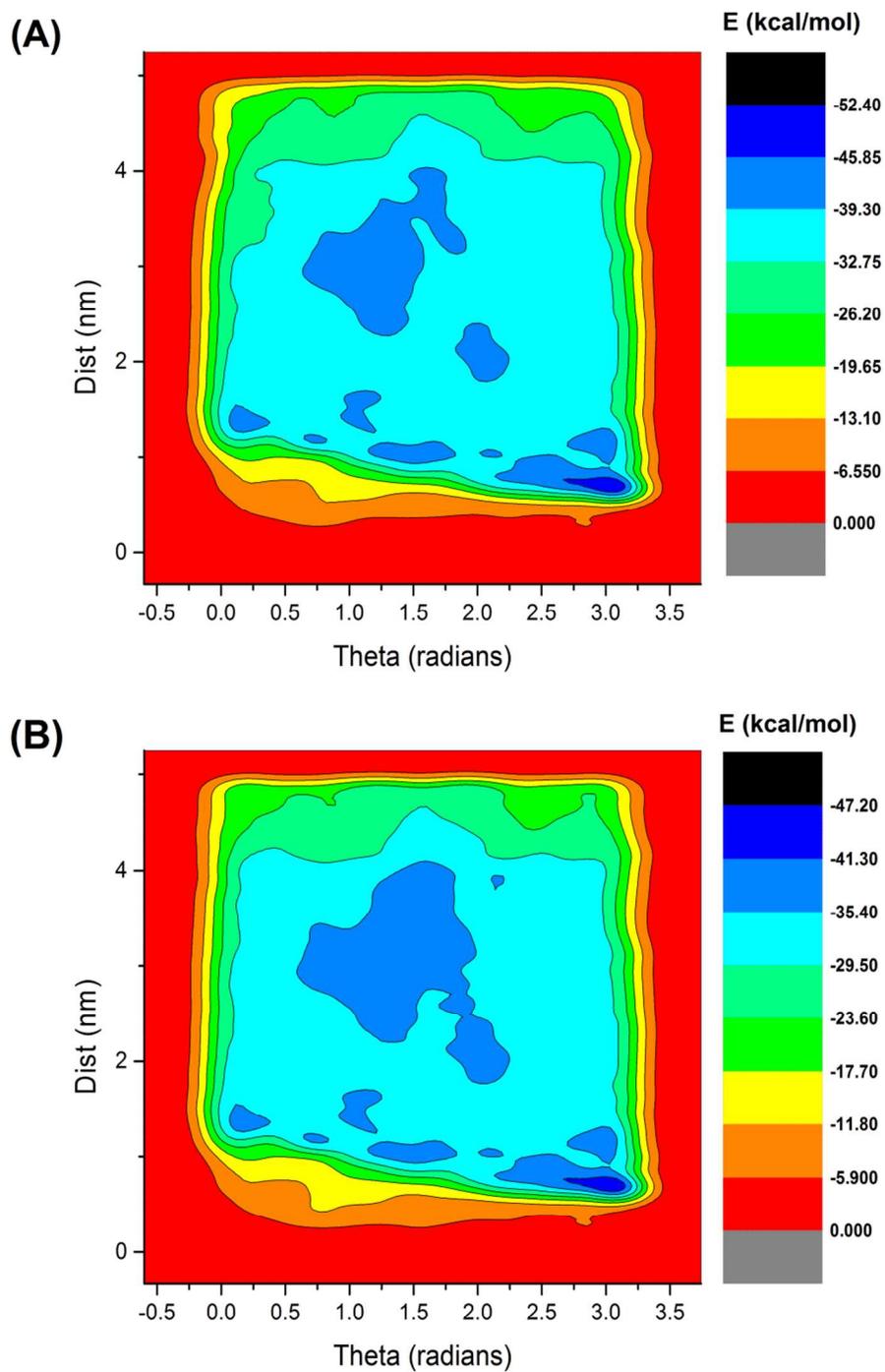
**Figure S10 Free energy landscapes of the WT-MetaD simulation of 1KF1-ThT at 1200 ns (A) and 1300 ns (B)**

The free energy landscape of 1KF1-ThT remained unchanged when comparing it at 1200 ns (A) and 1300 ns (B).



**Figure S11 Variations of Gaussian height (A) and CVs (B-D) as a function of time of the WT-MetaD simulation of 2JSM-ThT**

The rescaled Gaussian height as a function of simulation time is shown in (A). It is observed that after 900 ns, the Gaussian height undergoes minor fluctuations around zero. In addition, the scattering variations of the three CVs with respect to simulation time are shown in (B) – (D).



**Figure S12 Free energy landscapes of the WT-MetaD simulation of 2JSM-ThT at 1200 ns (A) and 1300 ns (B)**

The free energy landscape of 2JSM-ThT remained unchanged when comparing it at 1200 ns (A) and 1300 ns (B).