Electronic Supplementary Material (ESI) for Physical Chemistry Chemical Physics.

Conformational and electronic effects on the formation of *anti* cyclobutane pyrimidine dimer in G-quadruplex structure

Wook Lee*[a], Spiridoula Matsika[a]

[a] Department of Chemistry, Temple University,

Philadelphia, Pennsylvania 19122, United States

Supplementary Information

Table S1. Vertical excitation energies and oscillator strengths of all frames. DP in bold indicates dimer precursor state and CT in bold indicates the charge transfer from guanine to adjacent thymine. The letters G, T1, T2 in front of excitation character represent the adjacent guanine base, Thy5, and Thy17, respectively. All energies are given in eV.

trans, anti

#1	**Energy (eV)**	**Character**	**#2**	**Energy (eV)**	**Character**	**#3**	**Energy (eV)**	**Character**
S1	4.815 (0.0003)	T1_nπ^*	S1	4.722 (0.0002)	T1_nπ^*	S1	4.915 (0.0005)	T1_nπ^*
S2	4.825 (0.0002)	T2_nπ^*	S2	4.921 (0.0003)	T2_nπ^*	S2	4.956 (0.0081)	T2_nπ^*
S3	5.043 (0.1376)	G_$\pi\pi^*_1$	S3	4.942 (0.1520)	G_$\pi\pi^*_1$	S3	4.963 (0.1358)	G_$\pi\pi^*_1$
S4	5.081 (0.0211)	**DP**	S4	5.217 (0.0753)	**DP**	S4	5.089 (0.0075)	**DP**
S5	5.436 (0.3009)	T1T2_$\pi\pi^*$	S5	5.516 (0.1960)	T1T2_$\pi\pi^*$	S5	5.359 (0.2707)	T1T2_$\pi\pi^*$
S6	5.615 (0.3715)	G_$\pi\pi^*_2$	S6	5.575 (0.1933)	G_nπ^*+G_$\pi\pi^*_2$	S6	5.580 (0.3905)	G_$\pi\pi^*_2$
S7	5.764 (0.0005)	G_nπ^*	S7	5.609 (0.1905)	G_nπ^*+G_$\pi\pi^*_2$	S7	5.720 (0.0004)	G_nπ^*
S8	6.151 (0.0048)	**CT**	S8	5.759 (0.0046)	**CT**	S8	6.046 (0.0260)	**CT**
#4	**Energy (eV)**	**Character**	**#5**	**Energy (eV)**	**Character**			
S1	4.823 (0.0001)	T1_nπ^*	S1	4.790 (0.0003)	T2_nπ^*			
S2	4.947 (0.0560)	T2_nπ^*	S2	4.977 (0.0002)	T1_nπ^*			
S3	4.949 (0.0821)	G_$\pi\pi^*_1$	S3	5.080 (0.1495)	G_$\pi\pi^*_1$			
S4	5.107 (0.0245)	**DP**	S4	5.183 (0.0018)	**DP**			
S5	5.443 (0.2286)	T1T2_$\pi\pi^*$	S5	5.463 (0.2905)	T1T2_$\pi\pi^*$			
S6	5.594 (0.4038)	G_$\pi\pi^*_2$	S6	5.601 (0.3741)	G_$\pi\pi^*_2$			
S7	5.639 (0.0136)	**CT**	S7	5.749 (0.0020)	G_nπ^*			
S8	5.769 (0.0005)	G_nπ^*	S8	6.109 (0.0214)	T1T2_CT			

cis, syn

#1	**Energy (eV)**	**Character**	**#2**	**Energy (eV)**	**Character**	**#3**	**Energy (eV)**	**Character**
S1	4.746 (0.0002)	T1_nπ^*	S1	4.692 (0.0002)	T1_nπ^*	S1	4.724 (0.0008)	T1_nπ^*
S2	4.820 (0.0003)	T2_nπ^*	S2	2.779 (0.0004)	T2_nπ^*	S2	4.800 (0.0014)	T2_nπ^*
S3	4.986 (0.1083)	G_$\pi\pi^*_1$	S3	4.989 (0.0495)	G_$\pi\pi^*_1$	S3	5.007 (0.0631)	G_$\pi\pi^*_1$
S4	5.083 (0.0104)	**CT**	S4	5.103 (0.1140)	**DP**	S4	5.120 (0.1028)	**DP**
S5	5.293 (0.0703)	**DP**	S5	5.213 (0.0618)	**CT**	S5	5.281 (0.1180)	**CT**+T1T2_$\pi\pi^*$
S6	5.394 (0.2526)	T1T2_$\pi\pi^*$	S6	5.343 (0.1936)	T1T2_$\pi\pi^*$	S6	5.366 (0.1199)	**CT**+T1T2_$\pi\pi^*$
S7	5.506 (0.0013)	G_nπ^*	S7	5.559 (0.0007)	G_nπ^*	S7	5.513 (0.0012)	G_nπ^*
S8	5.649 (0.2041)	G_$\pi\pi^*_2$	S8	5.639 (0.2364)	G_$\pi\pi^*_2$	S8	5.630 (0.2193)	G_$\pi\pi^*_2$
#4	**Energy (eV)**	**Character**	**#5**	**Energy (eV)**	**Character**			
S1	4.717 (0.0002)	T1_nπ^*	S1	4.844 (0.0009)	T1_nπ^*			
S2	4.856 (0.0037)	T2_nπ^*	S2	4.855 (0.0019)	T2_nπ^*			
S3	5.032 (0.0336)	**DP**+ G_$\pi\pi^*$	S3	5.029 (0.0796)	G_$\pi\pi^*_1$			
S4	5.113 (0.0603)	G_$\pi\pi^*_1$	S4	5.106 (0.0718)	**DP**			
S5	5.283 (0.3199)	T1T2_$\pi\pi^*$	S5	5.248 (0.2532)	T1T2_$\pi\pi^*$			
S6	5.510 (0.0601)	**CT**	S6	5.451 (0.0019)	G_nπ^*			
S7	5.532 (0.0002)	G_nπ^*	S7	5.468 (0.0299)	**CT**			
S8	5.646 (0.1903)	G_$\pi\pi^*_2$	S8	5.648 (0.2087)	G_$\pi\pi^*_2$			

Table S1. Continued

cis, anti **(1)**

#1	**Energy (eV)**	**Character**	**#2**	**Energy (eV)**	**Character**	**#3**	**Energy (eV)**	**Character**
S1	4.845 (0.0014)	T1_nπ^*	S1	4.794 (0.0002)	T2_nπ^*	S1	4.921 (0.0002)	T2_nπ^*
S2	4.849 (0.0001)	T2_nπ^*	S2	4.871 (0.0013)	T1_nπ^*	S2	4.922 (0.0021)	T1_nπ^*
S3	5.008 (0.0619)	G_$\pi\pi^*_1$	S3	4.959 (0.0719)	G_$\pi\pi^*_1$	S3	5.109 (0.0451)	G_$\pi\pi^*_1$
S4	5.119 (0.2157)	**DP**	S4	5.163 (0.1745)	**DP**	S4	5.224 (0.2635)	**DP**
S5	5.385 (0.0066)	**CT**	S5	5.298 (0.1486)	T1T2_$\pi\pi^*$	S5	5.331 (0.1320)	T1T2_$\pi\pi^*$
S6	5.400 (0.0007)	G_nπ^*	S6	5.508 (0.0105)	**CT**	S6	5.341 (0.0043)	G_nπ^*
S7	5.463 (0.1857)	T1T2_$\pi\pi^*$	S7	5.574 (0.2913)	G_$\pi\pi^*_2$	S7	5.621 (0.1295)	**CT**+ G_$\pi\pi^*_2$
S8	5.596 (0.2353)	G_$\pi\pi^*_2$	S8	5.627 (0.0007)	G_nπ^*	S8	5.751 (0.1161)	**CT**+ G_$\pi\pi^*_2$
#4	**Energy (eV)**	**Character**	**#5**	**Energy (eV)**	**Character**			
S1	4.728 (0.0003)	T1_nπ^*	S1	4.822 (0.0008)	T1_nπ^*			
S2	5.038 (0.0794)	G_$\pi\pi^*_1$	S2	4.886 (0.0045)	T2_nπ^*			
S3	5.063 (0.0084)	T2_nπ^*	S3	5.047 (0.0180)	**CT**+G_$\pi\pi^*_1$			
S4	5.240 (0.1588)	**DP**	S4	5.138 (0.1858)	**CT**+G_$\pi\pi^*_1$			
S5	5.352 (0.0021)	G_nπ^*	S5	5.212 (0.0682)	**DP**			
S6	5.390 (0.1519)	T1T2_$\pi\pi^*$	S6	5.360 (0.0034)	G_nπ^*			
S7	5.531 (0.0321)	**CT**	S7	5.374 (0.1610)	T1T2_$\pi\pi^*$			
S8	5.677 (0.2300)	G_$\pi\pi^*_2$	S8	5.636 (0.2602)	G_$\pi\pi^*_2$			

cis, anti **(2)**

#1	**Energy (eV)**	**Character**	**#2**	**Energy (eV)**	**Character**	**#3**	**Energy (eV)**	**Character**
S1	4.909 (0.0027)	T2_nπ^*	S1	4.957 (0.1091)	G_$\pi\pi^*_1$	S1	4.910 (0.0028)	T2_nπ^*
S2	5.005 (0.0979)	G_$\pi\pi^*_1$	S2	4.993 (0.0020)	T1_nπ^*	S2	5.053 (0.1183)	G_$\pi\pi^*_1$
S3	5.029 (0.0002)	T1_nπ^*	S3	4.996 (0.0049)	T2_nπ^*	S3	5.064 (0.0015)	T1_nπ^*
S4	5.124 (0.0509)	**DP**	S4	5.173 (0.1140)	**DP**	S4	5.100 (0.0603)	**DP**
S5	5.259 (0.2710)	T1T2_$\pi\pi^*$	S5	5.317 (0.2004)	T1T2_$\pi\pi^*$	S5	5.356 (0.2706)	T1T2_$\pi\pi^*$
S6	5.561 (0.2406)	G_$\pi\pi^*_2$	S6	5.534 (0.2380)	G_$\pi\pi^*_2$	S6	5.610 (0.0496)	G_nπ^*
S7	5.628 (0.0559)	G_nπ^*	S7	5.588 (0.0905)	G_nπ^*	S7	5.656 (0.2461)	G_$\pi\pi^*_2$
S8	5.772 (0.0004)	**CT**	S8	6.127 (0.0158)	T1T2_CT	S8	6.082 (0.0180)	T1T2_CT
#4	**Energy (eV)**	**Character**	**#5**	**Energy (eV)**	**Character**			
S1	4.860 (0.0067)	T2_nπ^*	S1	4.844 (0.0151)	T2_nπ^*			
S2	4.888 (0.1259)	G_$\pi\pi^*_1$	S2	4.904 (0.0011)	T1_nπ^*			
S3	5.013 (0.0511)	**DP**	S3	4.946 (0.1024)	G_$\pi\pi^*_1$			
S4	5.059 (0.0035)	T1_nπ^*	S4	4.964 (0.0467)	**DP**			
S5	5.360 (0.2567)	T1T2_$\pi\pi^*$	S5	5.243 (0.2435)	T1T2_$\pi\pi^*$			
S6	5.460 (0.3067)	G_$\pi\pi^*_2$	S6	5.471 (0.0233)	G_nπ^*			
S7	5.506 (0.0292)	**CT**	S7	5.514 (0.2986)	G_$\pi\pi^*_2$			
S8	5.586 (0.0027)	G_nπ^*	S8	5.631 (0.0009)	**CT**			

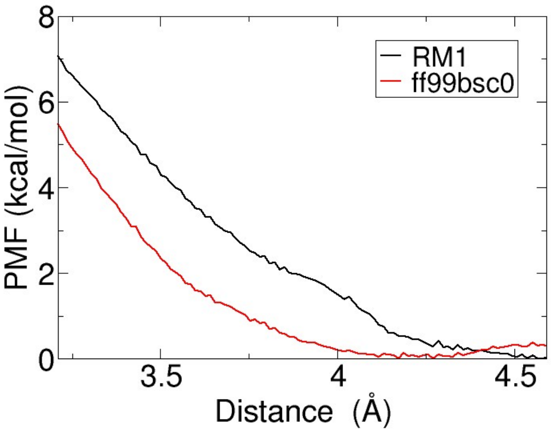

Figure S1. Comparison of PMF profiles between semi-empirical method (RM1) and ff99bsc0. The employed reaction coordinate is the distance between the midpoints of C5-C6 bonds from *cis, syn* isomer.

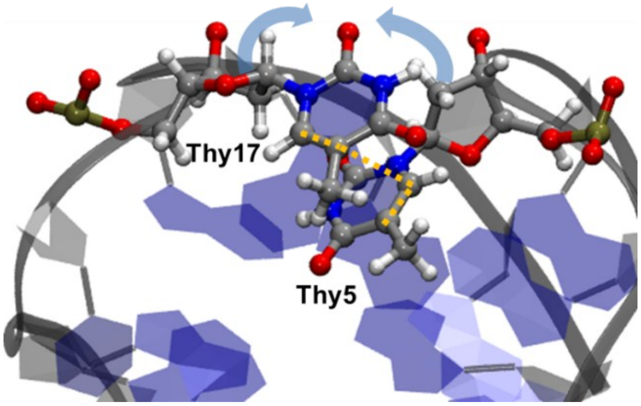

Figure S2. The conformation of Thy5 and Thy17 when the distance between the midpoints of two C5-C6 bonds is reduced to 3.4 Å. The dihedral angle C5-C6-C5'-C6' is shown in orange dotted line. The increase of this dihedral angle close to 180 degree would force the backbone of Thy5 and Thy17 closer as shown by blue arrows resulting in a huge structural strain.

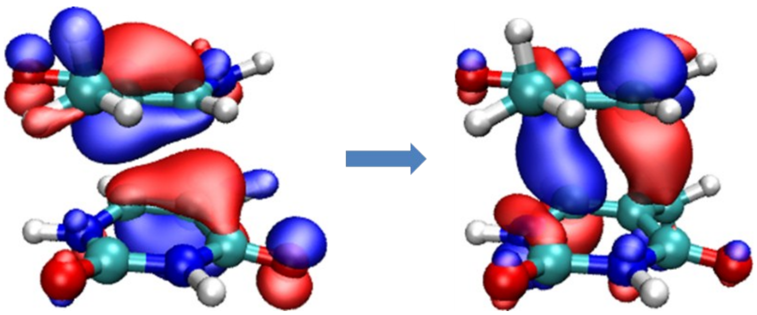

Figure S3. NTOs of DP state at longer distance. The transition occurs from the orbital with anti-bonding character to the one with bonding character.

Table S2. Distances between the centers of aromatic ring of Thy5 and Gua3 extracted from 5 QM/MM optimized geometries of each isomer. The averages from additional 100 geometries from MD simulations for each isomer are also shown at the bottom. The standard deviations are shown in the parenthesis, and all distances are given in angstrom.

	trans, anti	*cis, syn*	*cis, anti* (1)	*cis, anti* (2)
1	6.75	4.22	4.57	5.08
2	6.44	4.16	4.28	5.36
3	6.30	4.11	4.70	5.62
4	6.64	4.28	4.17	4.59
5	6.36	4.09	4.39	4.82
Average (STD)	6.42 (0.23)	4.32 (0.20)	4.38 (0.20)	8.08 (1.15)

Table S3. Angles formed by aromatic rings of Thy5 and Gua3 extracted from 5 QM/MM optimized geometries of each isomer. To obtain these angles, two vectors which are orthogonal to each aromatic ring were first computed, and then the angle formed by these two vectors was computed. The averages from additional 100 geometries from MD simulations for each isomer are also shown at the bottom. The standard deviations are shown in the parenthesis, and all angles are given in degree.

	trans, anti	*cis, syn*	*cis, anti* (1)	*cis, anti* (2)
1	97	22	19	131
2	100	18	28	130
3	91	16	30	132
4	100	20	22	128
5	84	18	22	122
Average (STD)	100 (8)	23 (6)	23 (5)	160 (13)

For all isomers, the distances and angles from 5 optimized structures are similar to the averages from 100 additional snapshots except *cis, anti* (2) isomer. In the case of *cis, anti* (2) isomer, conformational changes occur during the simulation where additional 100 snapshots were extracted. The angle becomes close to parallel rather than perpendicular during these conformational changes, but the two bases become much further away from each other (around 8 Å), making CT still very inefficient.

Computational details for the comparison between ff99bsc0 and Racife Model 1 (RM1)

In order to validate that the use of ff99bsc0 is legitimate for the extensive umbrella samplings in the present work, we compared ff99bsc0 to a semi-empirical method for one of the PMF profiles. Although the cost of semi-empirical methods is much cheaper than *ab initio* methods, it is still very expensive for extensive samplings like the one conducted in the present work, so we just compared one of the PMF profiles. The semi-empirical method of our choice is RM1[1] because it has been shown to give the most accurate results compared to other well-known semi-empirical methods.[2] The reaction coordinate we chose for this comparison is the distance between the midpoints of C5-C6 bonds from *cis, syn* isomer, which corresponds to a slice of the two dimensional PMF surface shown in Figure 4B. The QM region includes two thymine bases as well as sugar moieties of Thy5 and Thy17 because previous experiments showed that the conformation of the sugar moieties play an important role in the dimerization.[3] The simulations were conducted using sander module of AMBER 11,[4] and the simulation conditions are the same as the ones used in the original force field based umbrella sampling except the time step of 1fs. Therefore, the total sampling time becomes a half of the original force field based umbrella sampling which used the time step of 2fs. To make the same simulation conditions for both ff99bsc0 and RM1, we also redid the force field based umbrella sampling with a time step of 1fs and compared them. The resulting PMF profiles are shown in Figure S1, and the energy difference between two profiles along the reaction coordinate is around 1.5 kcal/mol. The small energy differences observed along the reaction indicate that our umbrella sampling results with ff99bsc0 are reliable.

References

1 G. B. Rocha, R. O. Freire, A. M. Simas and J. J. P. Stewart, *J. Comput. Chem.*, 2006, **27**, 1101–1111.
2 G. de M. Seabra, R. C. Walker and A. E. Roitberg, *J. Phys. Chem. A*, 2009, **113**, 11938–11948.
3 W. J. Schreier, J. Kubon, N. Regner, K. Haiser, T. E. Schrader, W. Zinth, P. Clivio and P. Gilch, *J. Am. Chem. Soc.*, 2009, **131**, 5038–5039.
4 D. Case, T. Darden, T. Cheatham, C. Simmerling, J. Wang, R. Duke, R. Luo, M. Crowley, R. Walker, W. Zhang, K. Merz, B. Wang, S. Hayik, A. Roitberg, G. Seabra, I. Kolossváry, K. Wong, F. Paesani, J. Vanicek, X. Wu, S. Brozell, T. Steinbrecher, H. Gohlke, L. Yang, C. Tan, J. Mongan, V. Hornak, G. Cui, D. Mathews, M. Seetin, C. Sagui, V. Babin and P. Kollman, *Amber 11, University of California, San Francisco*, 2010.