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

Supplementary Information

Electron transfer characteristics of 2'-deoxy-2'-fluoro-arabinonucleic acid, a nucleic acid with enhanced chemical stability

Ruijie D. Teo,[a] Kiriko Terai,[a,b] Agostino Migliore,[a] * and David N. Beratan[a,c,d] *

[a] Department of Chemistry, Duke University, Durham, North Carolina 27708, United States

[b] Department of Natural Science, College of Liberal Arts, International Christian University, Osawa, Mitaka-shi, Tokyo 181-8585, Japan

[c] Department of Physics, Duke University, Durham, North Carolina 27708, United States

[d] Department of Biochemistry, Duke University, Durham, North Carolina 27710, United States

DNA and 2'F-ANA Initial Structures for MD Simulations. The original structure of the 2'F-ANA/ANA duplex (PDB 2LSC[1]) contains U nucleobases in the ANA strand. These bases were modified to T bases (using the Schrodinger's Maestro molecular modeling software[2]) to match the analogous Dickerson-Drew DNA sequence (PDB 4C64[3]). Finally, the sugar moieties in the ANA strand were replaced with the corresponding ones in 2'F-ANA to obtain a 2'F-ANA/2'F-ANA duplex.

MD Simulations. The dodecamers were solvated in a box of TIP3P water[4] that extended 10.0 Å on each side of the nucleic acid. The negative charge of the system was neutralized by Na^+ ions. AMBER parameters for FANA were derived from ref. 5, while the AMBER force field ff99[6] was used for DNA. The solvated DNA and FANA systems consisted of 13311 and 12069 atoms, respectively. We set the scaling factor for 1-4 electrostatic interactions to 0.833333. The cutoff distance for truncating van der Waals interactions was 12 Å. The maximum non-bonded interaction distance for the periodical calculation of the interaction energy was set to 14 Å. The electrostatic interactions were computed using the Particle Mesh Ewald method[7] (grid spacing = 1 Å). Full electrostatic interaction energy was calculated each 2 time steps. The water molecules were maintained rigid using the SHAKE algorithm. The unit cell vectors were (in Å): (46.008, 49.610, 65.744) for DNA and (48.000, 65.000, 46.000) for 2'F-ANA.

All MD simulations were carried out using the NAMD 2.11 software.[8] Both nucleic acid structures were first optimized by means of 8×10^4 energy minimization steps. For DNA, the minimization was followed by solvent equilibration (using a Langevin thermostat, with a damping coefficient of 1.0 ps^{-1} for the Langevin dynamics), by gradually increasing the temperature from 293 K (crystallization temperature at which the system was equilibrated for 225 ps) to 295.5 K (for 50 ps), and finally to 298 K (for further 50 ps). For 2'F-ANA, we only carried out 225 ps of solvent equilibration at 209 K, namely, the temperature at which the 2'F-ANA solution structure was obtained. Next, both systems were equilibrated at constant temperature and pressure for 1.5 ns, using Langevin thermostat and piston[9, 10] (temperature = 298 K, pressure = 1 atm; barostat period = 100 fs, characteristic damping time = 50 fs, damping coefficient = 2.0 ps^{-1}). The final MD production runs covered 50 ns, with a time step of 0.5 fs. The DNA and 2'F-ANA snapshots used for electronic coupling calculations were taken each nanosecond from 10 to 50 ns. Fig. S1 shows the RMSD evolutions over this time range for the backbone and the base pairs.

Calculation of the Electronic Couplings. The hole transfer mainly proceeds through the sequence of stacked nucleobases, and the electronic couplings between adjacent nucleobase pairs are critical parameters to determine the efficiency of the hole transport, regardless of the kinetic model used to study the double-strand conductance. Thus, for each selected MD snapshot we pruned dimers of base pairs (from the set highlighted in Fig. 2) from the dodecamer structures. The resulting dangling bonds were saturated by H atoms, whose positions were optimized by B3LYP/6-311G** calculations in the presence of the electron hole (that is, with a total charge of +1). All DFT computations were carried out using the NWChem package[11] (versions 6.5 and 6.6), except for the two calculations with the ωB97M-V functional[12], which were carried out using QChem[13] (version 4.4). Wave functions and energies of the diabatic states (where the excess positive charge is constrained on either the donor or the acceptor base pair) were obtained using CDFT as specified in the main text. The overlap parameters in eqn (1) were obtained exploiting the ET module in the NWChem code or the overlap matrix provided in the QChem output.

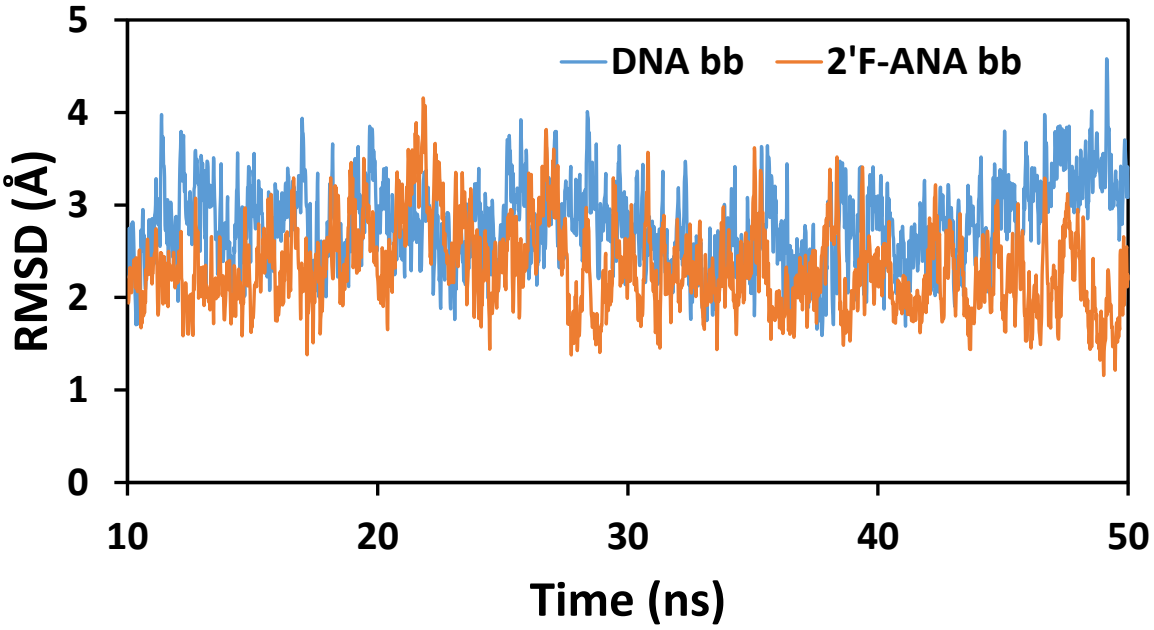
The electronic coupling values associated with the MD snapshots of DNA and 2'F-ANA are reported in Table S1 and diagrammed in Figs. S2-S5. The norms of the corresponding ground-state vectors (as expanded on the pertinent diabatic states) are given in Table S2. When the two-state approximation is exactly satisfied, the normalization condition $\sqrt{a^2 + b^2 + 2abS_{IF}} = 1$ holds. Table S2 shows that the norm of the ground-state vector is very close to 1 (namely, well above 0.99) in most cases, while a few values are in the range 0.97-0.99, thus supporting the validity of the two-state approximation.

Table S1. Effective electronic couplings (or generalize electron transfer integrals) for the CG-GC, GC-A_1T, A_1T-A_2T, and A_2T-TA dimers in the selected MD snapshots, at the simulation times indicated in the first column. All couplings values are in meV.

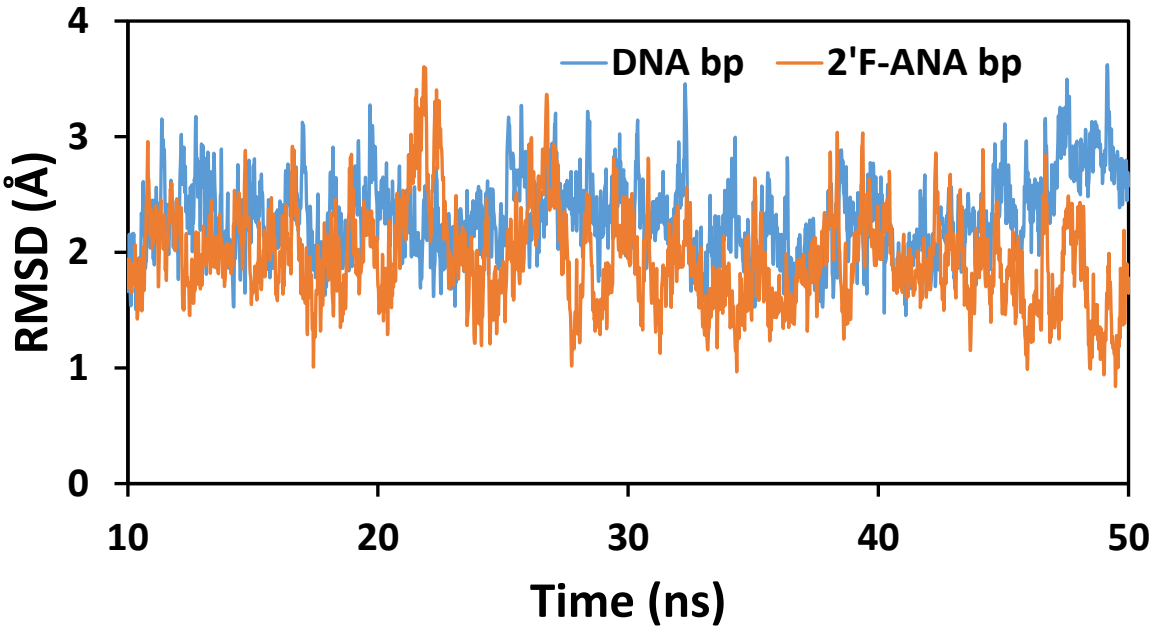
Time (ns)	V_{IF}^{DNA}				$V_{IF}^{2'F-ANA}$			
	CG-GC	**GC-A_1T**	**A_1T-A_2T**	**A_2T-TA**	**CG-GC**	**GC-A_1T**	**A_1T-A_2T**	**A_2T-TA**
10	8.5	78.8	4.1	2.4	27.6	35.5	47.3	68.1
11	9.0	100.1	55.8	140.1	4.8	46.1	18.9	10.1
12	11.0	89.2	45.6	23.1	5.2	27.8	85.5	260.5
13	15.7	22.4	46.4	42.3	6.0	49.2	22.3	38.4
14	2.0	22.4	79.0	29.7	22.5	27.4	16.9	16.3
15	2.9	6.6	45.1	19.6	13.1	35.5	64.6	3.0
16	21.6	11.2	0.1	10.8	10.5	54.5	19.2	3.2
17	18.2	70.7	43.3	0.6	9.1	28.0	0.4	4.9×10^{-2}
18	0.6	146.4	27.9	30.7	23.7	23.3	4.5	22.2
19	9.5	40.9	12.5	7.1	6.5	8.0	11.7	4.6
20	8.7	177.0	44.6	52.3	5.0	28.1	88.0	11.6
21	3.3	154.2	31.8	29.2	0.4	34.5	23.8	0.4
22	3.0×10^{-2}	151.7	57.5	52.5	6.3	2.3	212.9	0.1
23	5.9	141.0	4.2	7.6	0.6	60.8	26.5	20.5
24	2.2	49.7	29.6	14.1	7.4	9.5	22.5	3.6
25	1.3	13.6	39.0	7.2	31.0	48.4	20.5	13.0
26	6.6	6.7	50.9	4.8	16.3	33.4	15.8	16.6
27	4.2	8.9	8.0×10^{-2}	11.7	9.1	45.7	10.4	1.2
28	4.5	47.9	74.3	18.0	9.7	29.5	5.1	12.2
29	17.5	40.3	51.1	1.3	1.5	62.5	14.9	35.4
30	1.3	1.2	11.7	1.8	7.9	9.9	31.6	110.3
31	31.8	23.4	5.9	14.4	14.2	8.8	23.6	20.0
32	3.1	122.9	42.4	12.7	2.2×10^{-2}	14.1	20.2	13.1
33	9.3	64.8	17.6	44.3	5.4	9.5	51.0	5.4
34	44.1	175.6	32.5	51.6	6.7	8.7	71.8	2.9
35	12.1	10.6	48.8	8.6	30.7	41.7	3.0	62.9
36	24.6	40.9	75.7	24.6	10.2	30.2	4.4	8.2
37	0.4	41.4	9.0×10^{-3}	42.4	12.7	43.9	19.5	151.1
38	3.1	173.6	60.5	46.8	14.1	76.1	58.8	8.2
39	21.5	2.1	14.8	5.6	3.4	30.7	5.3	0.4
40	11.2	36.1	5.7	22.1	6.4	11.6	23.8	56.2
41	24.1	18.0	31.9	23.8	2.5	1.2	0.2	156.5
42	7.7	138.5	55.6	10.4	0.4	69.9	3.8	3.2
43	19.1	161.4	97.9	20.2	5.9	2.7	52.5	3.2
44	2.5	182.4	21.7	3.6	8.6	46.5	36.4	2.2
45	10.5	26.5	13.4	13.6	8.9	17.4	18.8	2.3
46	4.7	9.8	16.5	3.3	21.5	15.1	5.7	129.7
47	2.2	41.5	21.0	0.8	20.6	17.6	109.5	6.4
48	0.9	42.8	11.0	19.4	3.4	13.6	8.3	4.9
49	14.9	76.5	18.6	10.2	21.8	41.9	27.5	1.2
50	3.6	5.3	11.3	34.0	38.0	7.3	79.8	4.7

Table S2. Norm, *N*, of the ground state vector for the indicated base pair dimers across all selected MD snapshots.

Time (ns)	N^{DNA}				$N^{2'F\text{-}ANA}$			
	CG-GC	**GC-A_1T**	**A_1T-A_2T**	**A_2T-TA**	**CG-GC**	**GC-A_1T**	**A_1T-A_2T**	**A_2T-TA**
10	0.997	0.999	0.995	0.997	0.997	0.999	0.996	0.994
11	0.998	0.999	0.996	0.995	0.996	0.999	0.998	0.998
12	0.998	0.999	0.998	0.997	0.999	1.000	0.994	0.993
13	0.999	0.999	0.994	0.996	0.998	0.999	0.994	0.998
14	0.999	0.999	0.997	0.992	1.015	0.999	0.993	0.994
15	0.998	1.000	0.996	0.991	1.000	0.998	0.994	0.998
16	0.999	0.999	0.998	0.998	0.998	0.997	0.998	0.992
17	0.998	0.999	0.994	0.997	0.998	0.999	0.998	0.998
18	0.998	0.999	0.998	0.988	0.998	1.000	0.999	0.997
19	0.999	0.999	0.998	0.996	0.997	1.000	0.996	0.991
20	0.997	0.999	0.997	0.988	0.997	0.998	0.996	0.981
21	0.998	0.999	0.997	0.992	0.999	0.998	0.999	0.986
22	0.999	0.999	0.996	0.995	0.998	0.999	0.988	0.977
23	0.998	0.999	0.999	0.996	0.998	0.998	0.996	0.986
24	0.998	1.000	0.999	0.999	0.999	1.000	0.995	0.992
25	0.998	0.999	0.998	0.993	0.999	0.998	0.992	0.991
26	0.997	1.000	0.997	0.988	0.999	1.000	0.999	0.978
27	0.999	0.997	0.997	0.997	0.998	0.999	0.998	0.982
28	0.998	0.999	0.998	0.994	0.998	0.999	0.998	0.995
29	0.998	0.999	0.997	0.999	0.999	0.999	0.998	0.996
30	0.998	0.999	0.999	0.986	0.998	0.999	0.999	0.973
31	0.997	0.999	0.997	0.997	0.996	0.999	0.998	0.994
32	0.998	0.999	0.997	0.998	0.999	1.000	0.995	0.995
33	0.999	0.999	1.000	0.993	0.999	1.000	0.994	0.982
34	0.999	0.999	0.998	0.994	0.998	1.000	0.997	0.986
35	0.999	0.999	0.997	0.998	0.998	0.999	0.994	0.994
36	0.998	0.998	0.996	0.999	0.997	0.999	0.996	0.987
37	0.997	0.999	1.000	0.993	0.997	0.999	0.998	0.996
38	0.997	0.998	0.998	0.996	0.998	0.999	0.995	0.994
39	0.999	1.000	0.998	0.990	0.996	0.999	1.000	0.985
40	0.998	0.999	0.999	0.993	0.999	0.999	0.995	0.993
41	0.997	1.000	0.991	0.979	0.997	1.000	0.998	0.985
42	0.998	0.998	0.997	0.985	0.999	0.998	0.997	0.998
43	0.998	0.993	0.998	0.982	0.999	1.000	0.995	0.996
44	1.000	0.998	0.996	0.968	0.999	0.998	0.996	0.984
45	0.997	0.998	0.999	0.991	0.999	0.998	0.997	0.982
46	0.996	0.999	0.997	0.994	0.999	0.999	0.993	0.976
47	0.997	0.999	0.996	0.994	0.998	1.000	0.997	0.979
48	0.999	0.999	0.997	0.986	0.998	0.998	0.997	0.988
49	0.998	1.000	0.995	0.995	0.998	0.998	0.991	0.980
50	0.999	1.000	0.998	0.994	0.997	1.000	0.988	0.988



(a)



(b)

Fig. S1 RMSDs (excluding the H atoms) along the MD production simulations of DNA and 2'F-ANA. **(a)** RMSD of the nucleic acid backbone (bb). For DNA, the RMSD ranges from 1.6 Å to 4.6 Å (hence, with a maximum range of 3.0 Å), with an average value of 2.8 Å and a standard deviation of 0.45 Å. For 2'F-ANA, the RMSD ranges from 1.2 to 4.2 Å, with an average of 2.3 Å and a standard deviation of 0.46 Å. **(b)** RMSD of the base pair (bp) stack. For the natural DNA, the RMSD ranges from 1.45 Å to 3.62 Å, with an average of 2.3 Å and a standard deviation of 0.36 Å. For 2'F-ANA, the RMSD ranges from 0.84 Å to 3.61 Å, with an average of 1.9 Å and a standard deviation of 0.42 Å.

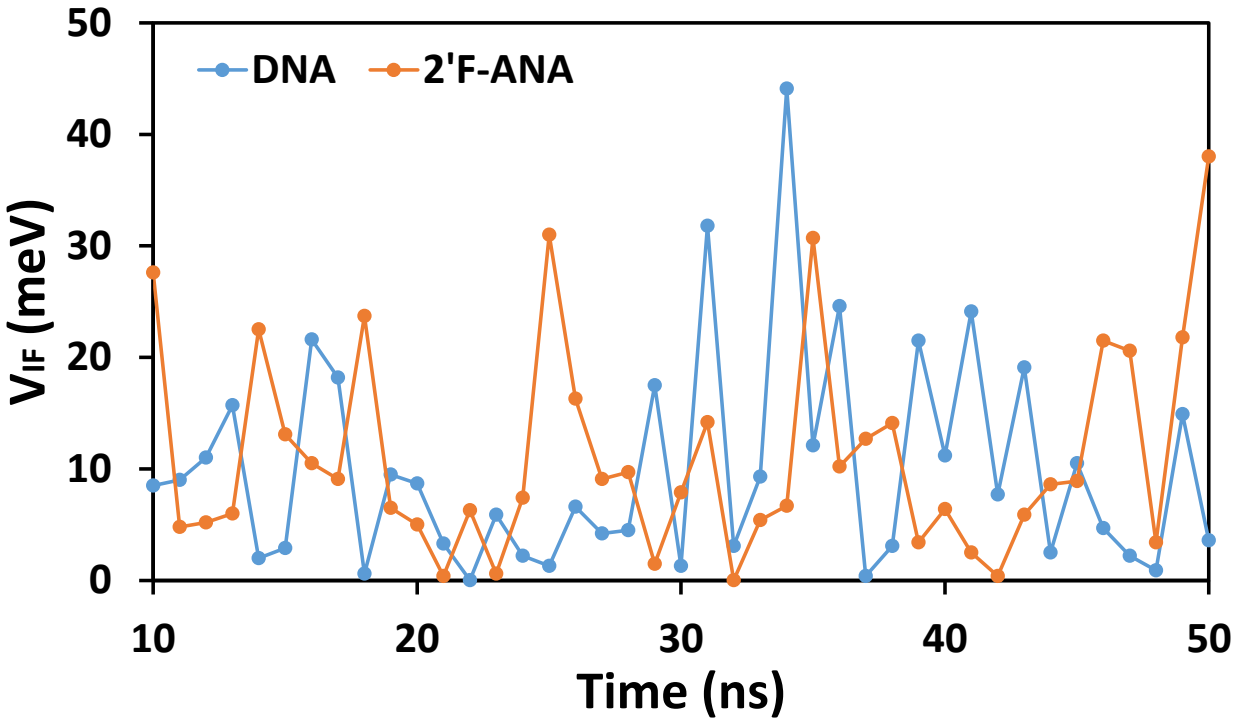


Fig. S2 Plot of V_{CG-GC} versus MD simulation time (ns) for DNA and 2'-FANA.

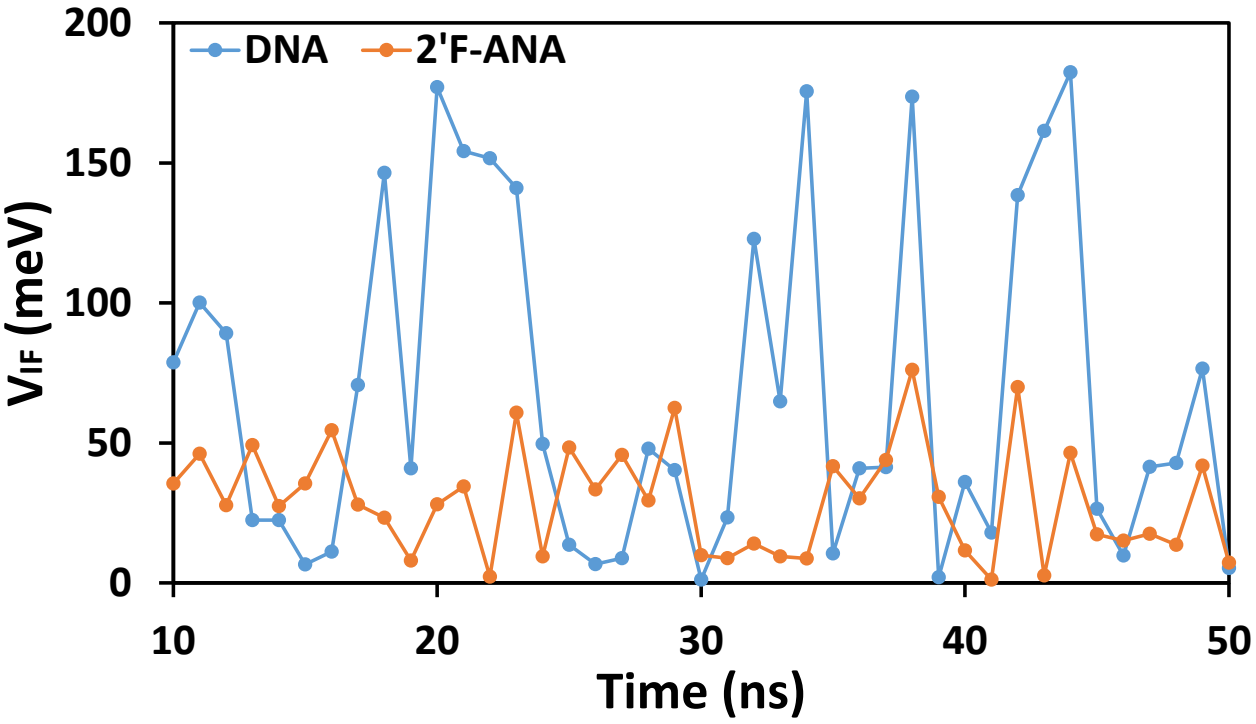


Fig. S3 Plot of V_{GC-A_1T} versus MD simulation time (ns) for DNA and 2'-FANA.

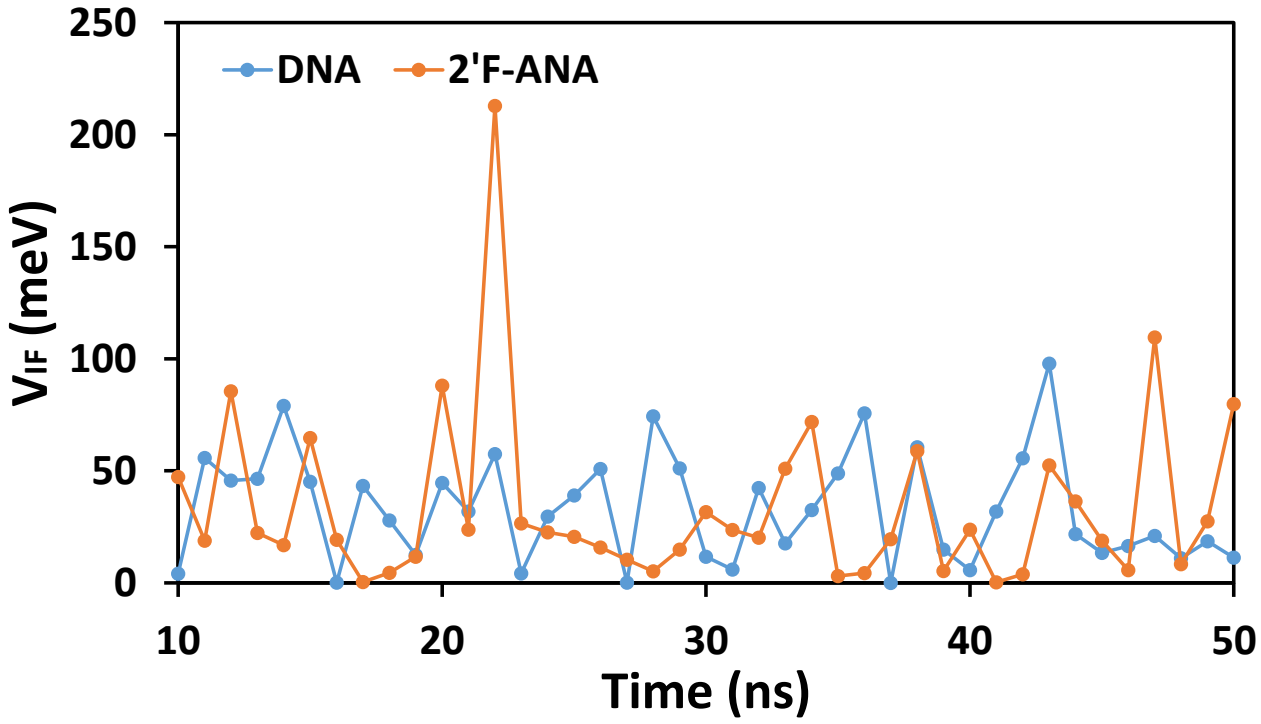


Fig. S4 Plot of $V_{A_1T-A_2T}$ versus MD simulation time (ns) for DNA and 2'-FANA.

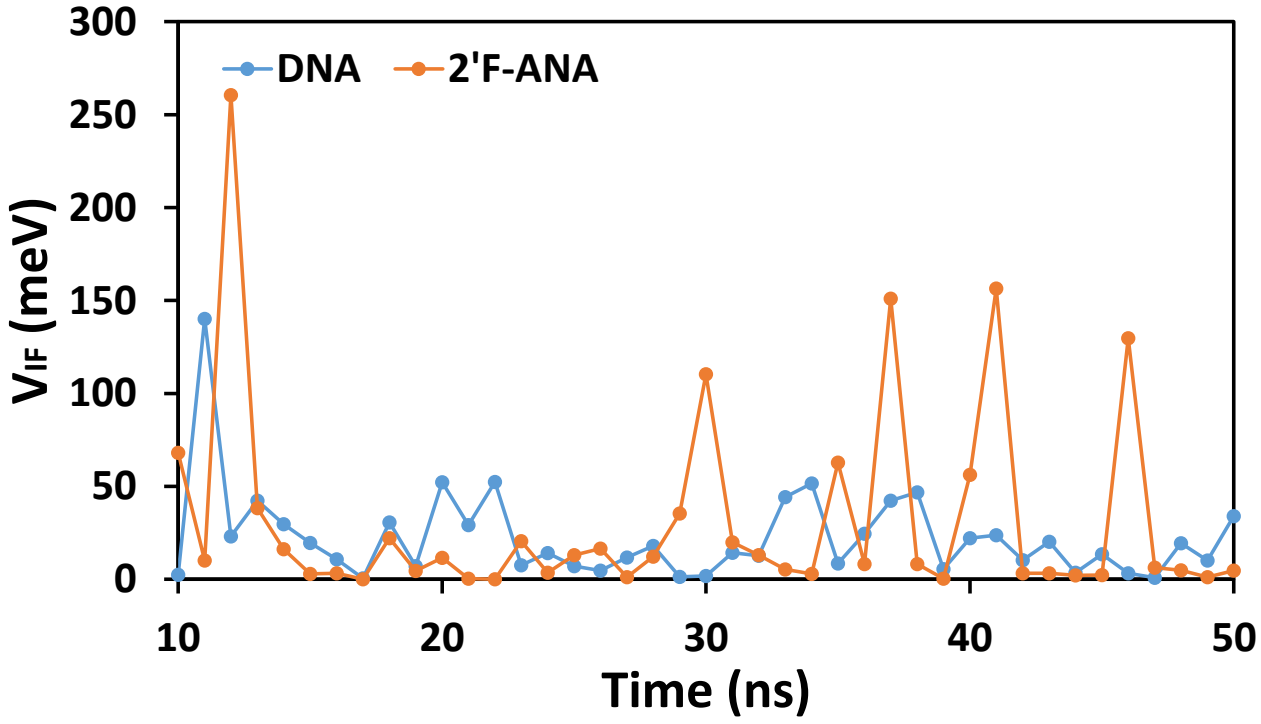


Fig. S5 Plot of V_{A_2T-TA} versus MD simulation time (ns) for DNA and 2'-FANA.

References

1. N. Martin-Pintado, M. Yahyaee-Anzahaee, R. Campos-Olivas, A. M. Noronha, C. J. Wilds, M. J. Damha and C. Gonzalez, *Nucleic Acids Res.*, 2012, **40**, 9329-9339.
2. S. Schrödinger Release 2018-1: Maestro, LLC, New York, NY, 2018.
3. L. Lercher, M. A. McDonough, A. H. El-Sagheer, A. Thalhammer, S. Kriaucionis, T. Brown and C. J. Schofield, *Chem. Commun.*, 2014, **50**, 1794-1796.
4. W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, *J. Chem. Phys.*, 1983, **79**, 926.
5. A. Noy, F. J. Luque and M. Orozco, *J. Am. Chem. Soc.*, 2008, **130**, 3486-3496.
6. J. Wang, P. Cieplak and P. A. Kollman, *J. Comput. Chem.*, 2000, **21**, 1049-1074.
7. T. Darden, D. York and L. Pedersen, *J. Chem. Phys.*, 1993, **98**, 10089-10092.
8. J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kalé and K. Schulten, *J. Comput. Chem.*, 2005, **26**, 1781-1802.
9. G. J. Martyna, D. J. Tobias and M. L. Klein, *J. Chem. Phys.*, 1994, **101**, 4177-4189.
10. S. E. Feller, Y. H. Zhang, R. W. Pastor and B. R. Brooks, *J. Chem. Phys.*, 1995, **103**, 4613-4621.
11. M. Valiev, E. J. Bylaska, N. Govind, K. Kowalski, T. P. Straatsma, H. J. J. Van Dam, D. Wang, J. Nieplocha, E. Apra, T. L. Windus and W. A. de Jong, *Comput. Phys. Commun.*, 2010, **181**, 1477-1489.
12. N. Mardirossian and M. Head-Gordon, *J. Chem. Phys.*, 2016, **144**, 214110.
13. Y. Shao, Z. Gan, E. Epifanovsky, A. T. B. Gilbert, M. Wormit, J. Kussmann, A. W. Lange, A. Behn, J. Deng, X. Feng, D. Ghosh, M. Goldey, P. R. Horn, L. D. Jacobson, I. Kaliman, R. Z. Khaliullin, T. Kuś, A. Landau, J. Liu, E. I. Proynov, Y. M. Rhee, R. M. Richard, M. A. Rohrdanz, R. P. Steele, E. J. Sundstrom, H. L. Woodcock, P. M. Zimmerman, D. Zuev, B. Albrecht, E. Alguire, B. Austin, G. J. O. Beran, Y. A. Bernard, E. Berquist, K. Brandhorst, K. B. Bravaya, S. T. Brown, D. Casanova, C.-M. Chang, Y. Chen, S. H. Chien, K. D. Closser, D. L. Crittenden, M. Diedenhofen, R. A. DiStasio, H. Do, A. D. Dutoi, R. G. Edgar, S. Fatehi, L. Fusti-Molnar, A. Ghysels, A. Golubeva-Zadorozhnaya, J. Gomes, M. W. D. Hanson-Heine, P. H. P. Harbach, A. W. Hauser, E. G. Hohenstein, Z. C. Holden, T.-C. Jagau, H. Ji, B. Kaduk, K. Khistyaev, J. Kim, J. Kim, R. A. King, P. Klunzinger, D. Kosenkov, T. Kowalczyk, C. M. Krauter, K. U. Lao, A. D. Laurent, K. V. Lawler, S. V. Levchenko, C. Y. Lin, F. Liu, E. Livshits, R. C. Lochan, A. Luenser, P. Manohar, S. F. Manzer, S.-P. Mao, N. Mardirossian, A. V. Marenich, S. A. Maurer, N. J. Mayhall, E. Neuscamman, C. M. Oana, R. Olivares-Amaya, D. P. O'Neill, J. A. Parkhill, T. M. Perrine, R. Peverati, A. Prociuk, D. R. Rehn, E. Rosta, N. J. Russ, S. M. Sharada, S. Sharma, D. W. Small, A. Sodt, T. Stein, D. Stück, Y.-C. Su, A. J. W. Thom, T. Tsuchimochi, V. Vanovschi, L. Vogt, O. Vydrov, T. Wang, M. A. Watson, J. Wenzel, A. White, C. F. Williams, J. Yang, S. Yeganeh, S. R. Yost, Z.-Q. You, I. Y. Zhang, X. Zhang, Y. Zhao, B. R. Brooks, G. K. L. Chan, D. M. Chipman, C. J. Cramer, W. A. Goddard, M. S. Gordon, W. J. Hehre, A. Klamt, H. F. Schaefer, M. W. Schmidt, C. D. Sherrill, D. G. Truhlar, A. Warshel, X. Xu, A. Aspuru-Guzik, R. Baer, A. T. Bell, N. A. Besley, J.-D. Chai, A. Dreuw, B. D. Dunietz, T. R. Furlani, S. R. Gwaltney, C.-P. Hsu, Y. Jung, J. Kong, D. S. Lambrecht, W. Liang, C. Ochsenfeld, V. A. Rassolov, L. V. Slipchenko, J. E. Subotnik, T. Van Voorhis, J. M. Herbert, A. I. Krylov, P. M. W. Gill and M. Head-Gordon, *Mol. Phys.*, 2015, **113**, 184-215.