Structural and Dynamical Impact of a Universal Fluorescent Nucleoside Analogue Inserted Into a DNA Duplex

Loussiné Zargarian[1], Akli Ben Imeddourene[1], Krishna Gavvala[2], Nicolas P.F. Barthes[3], Benoit Y.Michel[3], Cyril A. Kenfack [4], Nelly Morellet[5], Brigitte René[1], Philippe Fossé[1], Alain Burger[3], Yves Mély[2] & Olivier Mauffret[1*].

Supporting Information

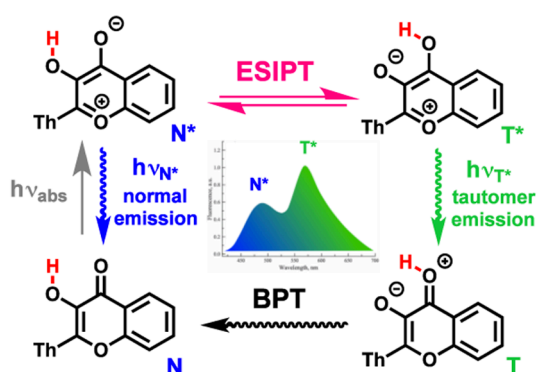**Table S1**: Spectroscopic data of the labelled duplex.

Labelled duplex[a] ODN	T_m (°C)[b]	λ_{abs} (nm)[d]	λ_{N*} (nm)[e]	λ_{T*} (nm)[f]	I_{N*}/I_{T*}[g]	Φ (%)[h]
5'-CCGCTT**S**AACGC-3'/ 5'-GCGTTTAAGCGG-3'	48.9 (51.3)[c]	382	437	541	0.037	40

a) 2 μM of ODN in 10 mM cacodylate pH 7.0, 100 mM NaCl, 1 mM EDTA; b) Melting temperature; ±0.5°C of the labelled duplex; c) Melting temperature ±0.5°C of the unlabelled ODN; d) Position of the absorption maximum; ±1 nm; e) Position of the emission maximum of the normal N* band; ±1nm; f) Position of the emission maximum of the tautomer T* band; ±1nm; g) I_{N*}/I_{T*} corresponds to the ratio of the intensities of the two emission bands at their maxima; ±2.5% mean standard deviation; h) Quantum yield determined using quinine sulfate (QS) in 0.1 M HCl solution (λ_{ex} = 350 nm, Φ = 0.54) as a reference, ±10% mean standard deviation.

Ref: W. H. Melhuish, *J. Phys. Chem.*(1961), **65**, 229–235.

Table S2 : Table I RMSD (Angstroms) of the 10 lowest energy structures have been superimposed for the 10 central base pairs (5-9 and 16-20 residues) and for all atoms.

	878	879	880	7605	11092	16626	17281	19173	19174	19576
878	0									
879	0.54	0								
880	0.62	0.60	0							
7605	1.31	1.31	1.36	0						
11092	1.73	1.77	1.81	2.15	0					
16626	1.90	2.01	1.98	2.41	1.69	0				
17281	2.03	2.06	2.05	2.25	1.91	1.37	0			
19173	1.85	1.91	1.88	2.15	1.78	1.22	1.18	0		
19174	1.74	1.82	1.78	2.00	1.68	1.12	1.11	0.58	0	
19576	1.36	1.45	1.54	1.44	1.58	1.80	1.93	1.70	1.58	0

Notes : BPT= back proton transfer, Th = 2-thienyl

Scheme 1. Schematic representation of the mechanism of Excited State Intramolecular Proton Transfer (ESIPT) reaction and dual emission.
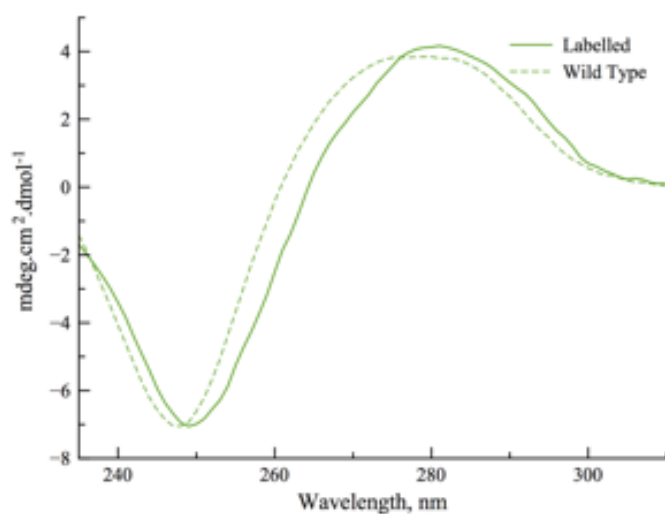

Figure S1. CD spectra of labelled (solid) and unlabelled (dashed) duplexes (2 mM) recorded in 10 mM cacodylate, pH 7.0, 100 mM NaCl, 1 mM EDTA.
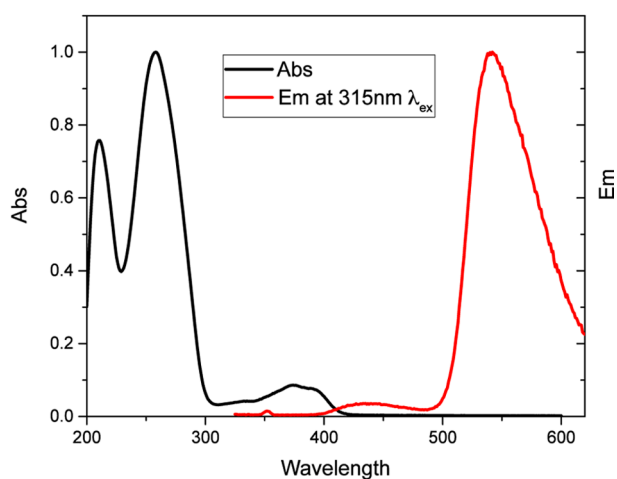
Figure S2. Absorption (black) and emission (red) spectra of the labelled duplex recorded in 10 mM cacodylate pH 7.0, 100 mM NaCl, 1 mM EDTA. The spectra were normalized at their maxima absorption and emission band.
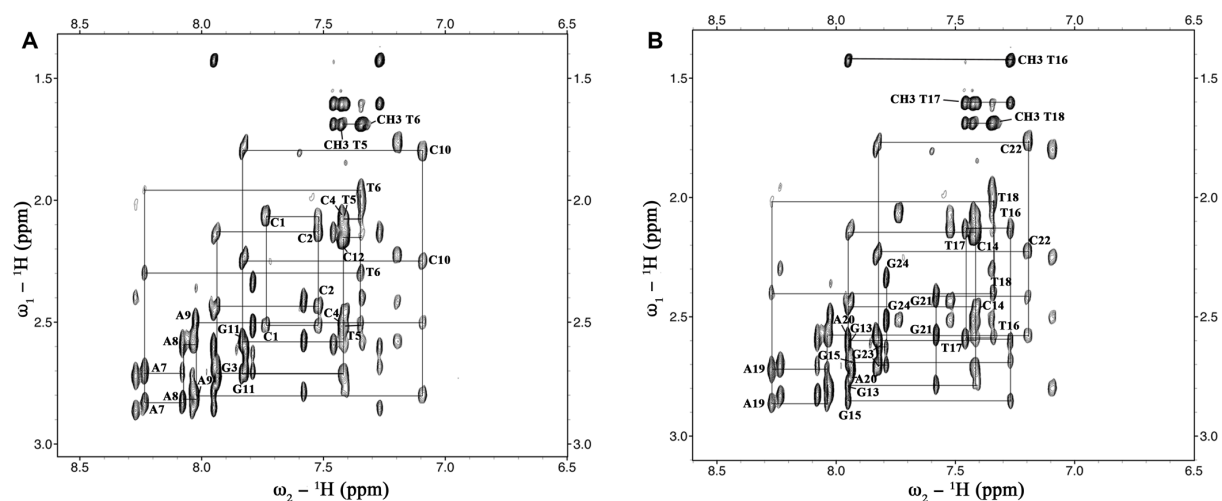

Figure S3. Expanded plot of a 500 MHz NOESY spectrum in D_2O buffer for the unmodified duplex showing the sequential nOes from the aromatic H6/H8 to H2'/H2"/CH3 protons. The spectrum was recorded at 20°C with a 300 ms mixing time. **(A)** Connectivities are traced for the strand 1 that is modified in the other duplex. The intra-nucleotide cross-peaks are indicated by the name of the residue. **(B)** Connectivities for the complementary strand.
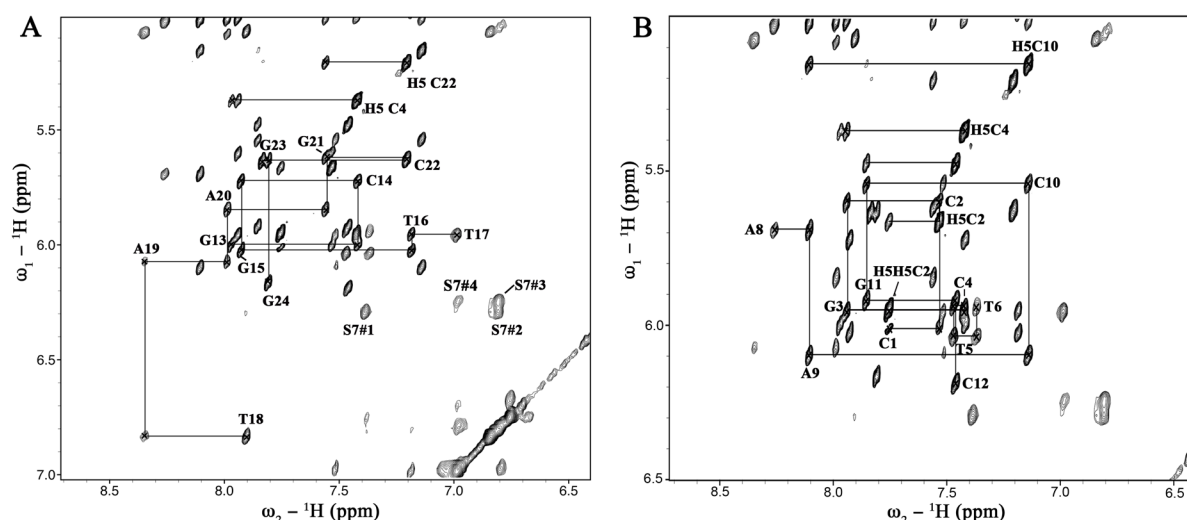
Figure S4. Expanded plot of a 950 MHz NOESY spectrum in D$_2$O buffer for the modified duplex showing the sequential nOes from the aromatic H6/H8 to H1' protons. The spectrum was recorded at 20°C with a 300 ms mixing time. **(A)** Connectivities are traced for the strand 1 that is unmodified relative to the other duplex. The intra-nucleotide cross-peaks are indicated by the name of the residue. Several cross-peaks involving resonances of the S7 probe are indicated by the symbols S7#; S7#1: H8S7-H7S7; S7#2: H7S7-H6S7; S7#3: H5S7-H6S7; S7#4: H5S7-H2A19. **(B)** Connectivities for the complementary strand.
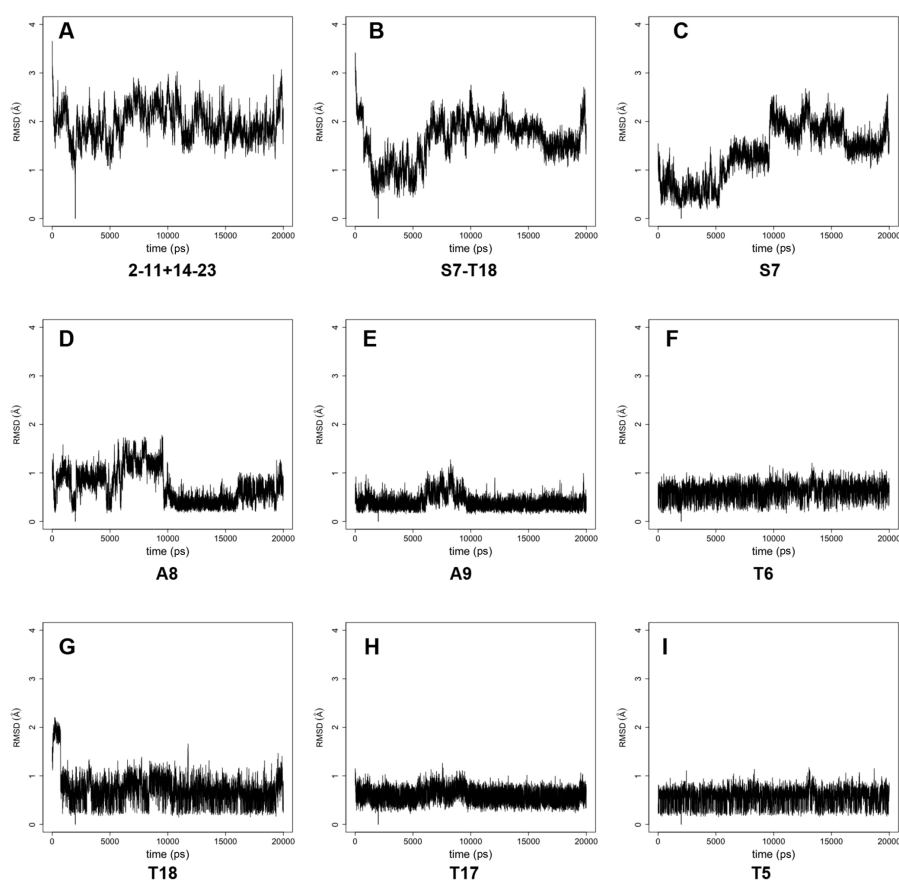
Figure S5. Local RMSD (Å) calculated for different subsets of the labelled duplex base pairs indicated under each graph during the 20 ns of the production phase performed with the NMR restraints. **(A)** The RMSD is calculated for residues 2-11 and 14-23 (i.e. the non terminal residues). **(B)** The RMSD is considered for the indicated residues that corresponds to 10 nucleotides distant from the probe. **(C)** The RMSD corresponds to the 10 central residues including the probe. From **(D)** to **(I)**, RMSD of various base pairs are shown. The reference molecule used for all the calculations is the molecule 2000 from the production phase (this corresponds to the molecule sampled at a time of 2000 ps). The results obtained with this latter snapshot display the best visualization of the differences existing during the course of the dynamic.
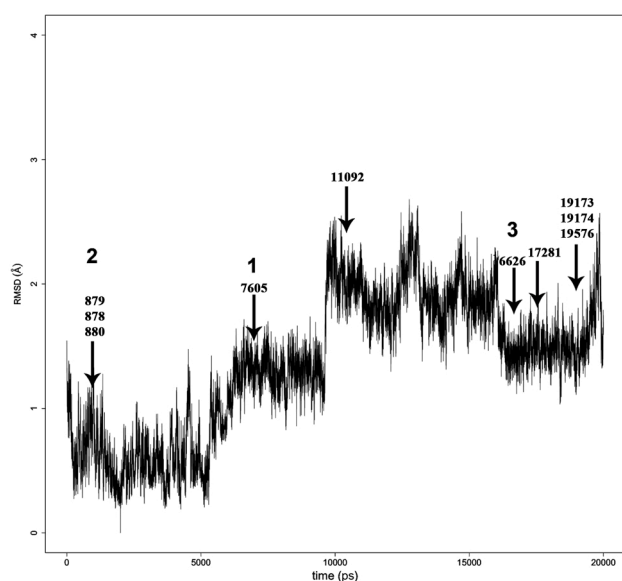

Figure S6. Local RMSD (Å) calculated for the probe S7 residue during the 20 ns of the dynamics performed with NMR restraints. Local RMSD (Å) is relative to a reference structure (similar to that used in Figure 5 and Supplementary Figure S6). This figure is the same as Figure 5C but here the positions of the 10 total energy minima (see text) are indicated along the dynamics. The 10 structures are indicated by their snapshot number (one snapshot is sampled every 1 ps and the dynamics contains therefore 20 000 structures). The numbers 1-3 indicate the ordering of the three best energy structures (snapshots 7605, 880 and 16626).
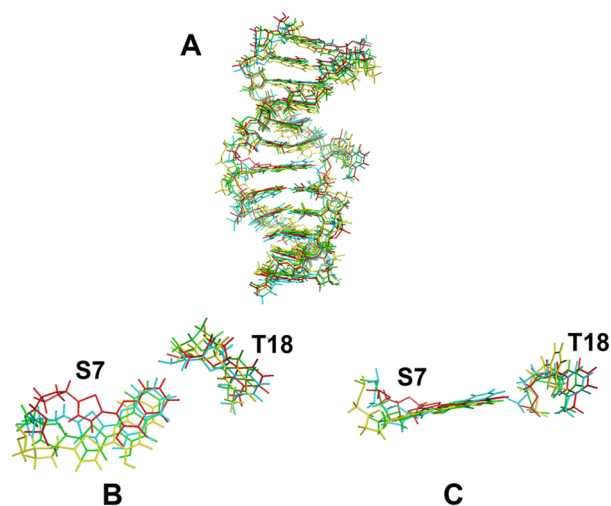
Figure S7. Superimposed views of four averaged structures. These structures were taken at minimum energy points along the dynamics (the different points are indicated in Figure S6). The four structures have been obtained following minimization with positional restraints of 1000 ps segments taken from the regions indicated on the plot of Figure S6. The colour code for the different regions indicated in Figure S6 is: yellow=7605; red=11902; blue=16626; green=880. The superimpositions have been made on the basis of the 10 central residues (including the probe). **(A)** View of the whole dodecamers. **(B)** View of the S7 and T18 residues illustrating the important shift of the S7 probe in a plane perpendicular to the helix axis of the dodecamer. **(C)** Perpendicular view from (B).