SUPPORTING INFORMATION FOR:

Hypoxia-selective dissociation mechanism of a nitroimidazole nucleoside in a DNA environment

Antonio Francés-Monerris,^{1,2}* Iñaki Tuñón² and Antonio Monari¹*

1. Université de Lorraine, CNRS, LPCT, Nancy F54000 (France)

2. Departamento de Química Física, Universitat de València, 46100 Burjassot, Spain

Corresponding Authors

*E-mail: A. F.-M. <u>Antonio.Frances@univ-lorraine.fr</u>, A. M.: <u>Antonio.Monari@univ-lorraine.fr</u>

Content

1. Computational Details	3
1.1. QM calculations	3
1.2. Classical molecular dynamics	4
1.3. QM/MM simulations	4
2. Neutral dissociations of the isolated modified nucleobase	5
3. CASSCF spin densities	6
4. Neutral dissociation of the model trimer	6
5. Intra base-pair parameters	7
6. Inter base-pair parameters	9
7. Calibration of the B3LYP functional and nature of the anion state 1	1
8. Histograms of O-H distances along the classical MD 1	2
9. O-H distances over selected windows of the umbrella sampling1	3
10. Radial distribution function of sodium ions around the nitro group	5

1. Computational Details

1.1. QM calculations

The neutral and anion dissociation mechanisms of **1** were studied through the DFT, TD-DFT and multiconfigurational CASPT2 methods. We employed both ω B97-XD and B3LYP functionals, as implemented in the GAUSSIAN 09 software package, in combination with the standard 6-31G* and 6-31++G(d,p) basis sets.

The dissociation process of **1** was studied in two different environments, in particular by considering the isolated system and by embedding **1** inside a trimer model **2** composed by two thymine molecules and the modified nucleobase **1**, being the latter at the central position (T**1**T). The model was built by using the AVOGADRO program, defining at first three π -stacked thymine molecules using the relative arrangements of canonical DNA, and subsequently modifying the central thymine moiety to include the non-natural nitroimidazole group (see blue atoms in Figure 1). Later, the system was subsequently optimized using DFT and freezing all atoms but the nitroimidazole chain.

To determine the dissociation profiles, a series of relaxed potential scans were performed by stretching the C7-N8 bond (see Figure 1 for atom numbering). For 2, only the C7 atom, their respective hydrogens, and the nitroimidazole chain were optimized. Three types of scans were performed: *i*) neutral system optimizing the lowest-lying singlet state (TD-DFT), *ii*) neutral system optimizing the triplet state (unrestricted DFT), and *iii*) anion system optimizing the lowest-energy doublet (unrestricted DFT).

To validate the dissociation profile of the isolated system **1**, the energies were also recomputed using the *ab initio* methods CASPT2 and DLPNO-CCSD(T) as implemented in the MOLCAS 8 and ORCA 4.0 software packages, respectively, on top of the structures scanned with the ω B97-XD/6-31G* method (see Figure S17). The CAS active space was composed of 11 electrons distributed into 10 molecular orbitals, only one root was demanded in the CASSCF procedure. Later, the CASPT2 method was used to compute the necessary dynamic electron correlation on top of the converged CASSCF wave functions. An imaginary level shift of 0.1 au was used to minimize the effect of weakly intruder states, whereas an IPEA shift of 0.25 au to the zeroth-order Hamiltonian has been used to improve the description of anionic systems, as described elsewhere. The ANO-S-VDZP basis set has been employed in all multiconfigurational calculations. Finally, the DLPNO-CCSD(T) computations were performed in combination with the aug-cc-pVTZ basis set and using the ORCA 4.0 default settings.

Results displayed in Figure S17 indicate that the inclusion of diffuse functions in the basis set does not significantly change the energy profile provided by the B3LYP/6-31G method, used in subsequent QM/MM calculations (see below). This fact can be explained in terms of the nature of the anion state of **1**, of valence-bound character, indicating that the extra electron is captured by a π -type valence molecular orbital. This is consistent with the small variations

(about 1%) found between the spatial extension of the electron density ($\langle r^2 \rangle$) of the neutral and the anionic solutions (see Table S1 below).

1.2. Classical molecular dynamics

The initial structures of the oligonucleotides **3** and **4** (the specific sequence is displayed in Figure 1) were built by using the NAB utility available in the AmberTools software package, taking the parameters from the parm99 force field and including bsc1 corrections for DNA. Both modified and canonical DNA strands **3** and **4** were solvated in a truncated octahedron containing 8544 TIP3P water molecules and 30 Na⁺ cations to have a neutral charge. All classical molecular dynamics were ran using the GPU CUDA version of Amber16. The initial structure was minimized applying 4000 steps of the steepest descent algorithm followed by additional 4000 steps of the conjugated gradient algorithm. Later, the system was slowly thermalized during 200 ps up to a final temperature of 300 K in the NVT ensemble and further equilibrated for 1 ns employing the NPT ensemble. Finally, a 200 ns production trajectory was run in the NPT ensemble. The pressure was set to 1 atm and maintained constant using the Monte Carlo barostat, while temperature conservation was ensured through Langevin dynamics. All MD simulations were conducted using periodic boundary conditions and using particle mesh-ewald (PME) with a cutoff of 9.0 Å and a time step of 1 fs. The results have been analyzed making use of the VMD and the Curves+ program packages.

In the case of **3**, the force field for the non-canonical residue **1** was reparametrized on the basis of the Amber force field using the restricted electrostatic potential (RESP) technique to obtain the atomic point charges. The nucleotide containing **1** was optimized at the B3LYP/6-311+G(d,p) level of theory, while the electrostatic potential was subsequently computed with the Hartree-Fock method. The point charges and the atom types of the nucleotide **1** can be found in the Supporting Information.

1.3. QM/MM simulations

The last structure of the classical MD simulation of the oligonucleotide **3** was used as a starting point to determine the potential of mean force (PMF) of the C7-N8 bond break *via* an umbrella sampling protocol. The Amber 16 program interfaced with the ORCA 4.0 program was used to run the QM/MM simulations. This means that the Amber 16 program performed all nuclear displacements using the forces of the QM part computed by ORCA 4.0. The QM-MM cutoff was set to 9 Å. The QM subspace comprised the modified nucleobase **1**, whereas the rest of DNA strand, water molecules and Na⁺ atoms were treated with the MM force field as described above (see Figure 2 of the main text). The QM level of theory was B3LYP/6-31G. We used the link atom procedure to saturate the valence of the system by adding hydrogens in the N1 and C1' atoms. An initial pre-equilibration run of 500 fs using neutral QM/MM settings was conducted prior the preparation of the simulation windows. Later, and to simulate the anionic dissociation, the charge and spin of the QM region was set to -1 and 2, respectively.

The simulation windows ranged from 1.45 to 3.55 Å and were separated by 0.05 Å, making a total of 43 windows. The C7-N8 bond distances were kept by applying a force constant of 1200 kcal/mol. The windows were prepared as follows. First, 400 minimization steps were

run at 1.50, 2.00, 2.50, 3.00, and 3.50 Å starting from the last snapshot of the pre-equilibrated neutral QM/MM run. Later, the system was thermalized during 8 ps up to a temperature of 300 K in the NVT ensemble, followed by 2 ps of additional simulation. The last structures were subsequently used as initial structures (seeds) for the rest of windows repeating the aforementioned minimization/equilibration protocol. Finally, 2 ps of equilibration in the NPT ensemble followed by a production run of 20 ps were performed for each simulation window.

The PMF was constructed from the ensemble of C7-N8 bond distances using the weighted histogram analysis method (WHAM) developed by the Grossfield lab.

2. Neutral dissociations of the isolated modified nucleobase



Figure S1. In vacuo relaxed scan of the C7-N8 bond breaking of the neutral system **1** relaxing S_1 . Geometries and energies are computed at the TD- ω B97-XD/6-31G* level of theory.



Figure S2. In vacuo relaxed scan of the C7-N8 bond breaking of the neutral system 1 relaxing T₁. Geometries and energies are computed at the unrestricted ω B97-XD/6-31G* level of theory.

3. CASSCF spin densities



Figure S3. CASSCF spin densities upon dissociation of the anion 1 in vacuo.

4. Neutral dissociation of the model trimer



Figure S4. In vacuo relaxed scan of the C7-N8 bond breaking of the neutral trimer model **2** relaxing S₁. Geometries and energies are computed at the TD- ω B97-XD/6-31G* level of theory.

5. Intra base-pair parameters



Figure S5. Shear angle (in degrees) measured over the classical MD trajectories of the oligonucleotides 4 (left panel) and 3 (right panel).



Figure S6. Stretch angle (in degrees) measured over the classical MD trajectories of the oligonucleotides 4 (left panel) and 3 (right panel).



Figure S7. Stagger angle (in degrees) measured over the classical MD trajectories of the oligonucleotides 4 (left panel) and 3 (right panel).



Figure S8. Buckle angle (in degrees) measured over the classical MD trajectories of the oligonucleotides **4** (left panel) and **3** (right panel).



Figure S9. Propeller angle (in degrees) measured over the classical MD trajectories of the oligonucleotides **4** (left panel) and **3** (right panel).



Figure S10. Opening angle (in degrees) measured over the classical MD trajectories of the oligonucleotides 4 (left panel) and 3 (right panel).

6. Inter base-pair parameters



Figure S11. Shift angle (in degrees) measured over the classical MD trajectories of the oligonucleotides 4 (left panel) and 3 (right panel).



Figure S12. Tilt angle (in degrees) measured over the classical MD trajectories of the oligonucleotides **4** (left panel) and **3** (right panel).



Figure S13. Slide angle (in degrees) measured over the classical MD trajectories of the oligonucleotides 4 (left panel) and 3 (right panel).



Figure S14. Roll angle (in degrees) measured over the classical MD trajectories of the oligonucleotides 4 (left panel) and 3 (right panel).



Figure S15. Rise angle (in degrees) measured over the classical MD trajectories of the oligonucleotides 4 (left panel) and 3 (right panel).



Figure S16. Twist angle (in degrees) measured over the classical MD trajectories of the oligonucleotides 4 (left panel) and 3 (right panel).

7. Calibration of the B3LYP functional and nature of the anion state



Figure S17. Benchmark of different levels of theory for the C7-N8 bond breaking of the anion system **1** in vacuo. Spin-density representations for selected points, computed with the B3LYP/6-31++G(d,p) method, are also shown.

Table S1. Values of the spatial extensions of the electron density $\langle r^2 \rangle$, in a.u., of selected points along the C7-N8 dissociation curves computed with the B3LYP/6-31++G(,d,p) method in Figure S17. The anionic wave functions are only ~1% larger than the corresponding neutral solutions, indicating valence-bound anionic states.

C7-N8 distance (Å)	Neutral	Anion
1.48	5596.2697	5657.6963
2.03	5945.8451	6001.6562
2.53	4587.9177	4636.7842
3.58	5285.4985	5338.9133

8. Histograms of O-H distances along the classical MD



Figure S18. Histogram of the interatomic distances between an oxygen of the NO₂ group of the anionic residue **1** and the external hydrogen of the NH₂ group of adenine #24 over the 200 ns of the classical MD trajectory. The hydrogen bonding is illustrated in Figure 5 of the main text.

9. O-H distances over selected windows of the umbrella sampling



Figure S19. Interatomic distances between an oxygen of the NO₂ group of the anionic residue **1** and the external hydrogen of the NH₂ group of adenine #24 along selected windows of the umbrella sampling corresponding to the non-H bonded PMF (red curve in Figure 4 of the main text). 1.45 Å correspond to the reactants; 2.10 and 2.20 Å belong to the transition state area, and 3.40 and 3.55 Å correspond to the products zone. The hydrogen bonding is illustrated in Figure 5 of the main text.



Figure S20. Interatomic distances between an oxygen of the NO₂ group of the anionic residue **1** and the external hydrogen of the NH₂ group of adenine #24 along selected windows of the umbrella sampling corresponding to the H bonded PMF (blue curve in Figure 4 of the main text). Top panel corresponds to 1.45 Å (reactants), medium panel corresponds to 2.10 Å (transition state), and bottom panel corresponds to 3.55 Å (products). The hydrogen bonding is illustrated in Figure 5 of the main text.

10. Radial distribution function of sodium ions around the nitro group



Figure S21. Normalized average radial distribution function of all sodium ions around the oxygen atoms of the nitro group of **1** over the 200 ns classical MD simulation.