Electronic Supplementary Material (ESI) for Journal of Materials Chemistry B.

Selectivity in the chiral self-assembly of arylazopyrazole photoswitches along DNA templates

Noemí Nogal, Santiago Guisán, David Dellemme, Mathieu Surin and Andrés de la Escosura.

Table of contents:

1. Experimental details 2

2. Synthesis and characterization 3

3. Optical spectroscopy 6

3.1. Samples preparation 6

3.2. UV-Vis absorption and Circular Dichroism spectroscopy 6

4. Molecular dynamics (MD) simulations 14

4.1 Reparameterization of the arylazopyrazole molecules 14

4.2 MD protocol 16

4.3 Analyses of the trajectories 18

5. NMR spectra 26

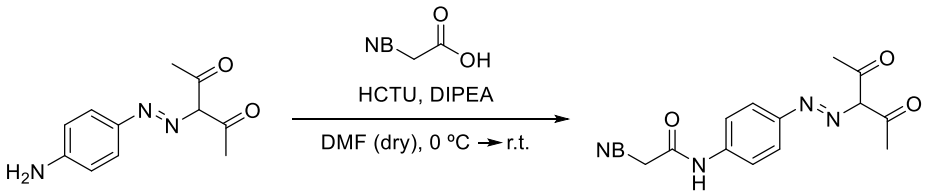
6. References 30

1. Experimental details

General information. Unless otherwise indicated, chemicals were purchased from Aldrich, Fluorochem, Scharlau, TCI or Bachem and were used without further purification. (E)-3-((4-aminophenyl)diazenyl)pentane-2,4-dione[S1,S2] and 2-[6-[Bis(tert-butoxycarbonyl)amino]-9H-purin-9-yl]acetic acid[S3] were prepared following the procedures described in the literature. Air sensitive reactions were performed using oven-dried glassware (at 150 °C) and performed under argon atmosphere using Schlenk techniques. Solvents were dried on a solvent purification system (PS-MD-5/7 Inert technology). Reactions were monitored by thin-layer chromatography (TLC) on silica-gel-coated aluminum foils (silica gel 60 F254, Merck). The TLC plates were visualized by UV light (λ = 254 nm). NMR spectra were recorded on a Bruker AV-300 or AV-500 spectrometer at 25 °C. Chemical shifts (δ) are reported in ppm relative to the residual solvent peak. Splitting patterns are indicated as (s) singlet, (d) doublet, (t) triplet, (q) quartet, (m) multiplet and (br) broad. Coupling constants (J) are reported in Hertz (Hz). High-resolution mass spectra (HRMS) were recorded on a Waters XEVO-G2 QTOF mass spectrometer, using electrospray ionization (ESI) in positive or negative mode, depending on the analyte. UV-Vis absorption and CD measurements were recorded using a ChirascanTM Plus CD Spectrometer from Applied Photophysics. The measurements were carried out using 10 mm quartz suprasil cells from Hellma Analytics. The spectra were recorded between 240 and 600 nm, with a bandwidth of 1nm, time per point 1 s, and one repetition. The aqueous solvent reference was used to get the baseline, which was automatically subtracted from the CD spectra of the samples.

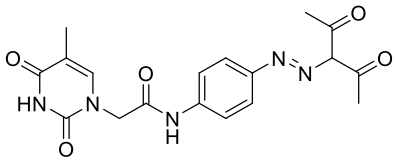
2. Synthesis and characterization

General procedure A for the synthesis of



The nucleobase carboxymethyl derivative (1.3 equiv) and HCTU (1.3 equiv) were dissolved in dry DMF [0.1 M] under Ar at 0 °C, and DIPEA (2.6 equiv) was added to the above solution with stirring. The mixture was stirred at the same temperature for 15 min. After this time (E)-3-((4-aminophenyl)diazenyl)pentane-2,4-dione[S1,S2] (1.0 equiv) was added to the reaction mixture. After stirring at room temperature for 18 h, water was added to the reaction mixture to precipitate the desired compound, which was then isolated by filtration and washed with more water. The product was obtained as a yellow-orange solid.

(E)-N-(4-((2,4-dioxopentan-3-yl)diazenyl)phenyl)-2-(5-methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl)acetamide (1).



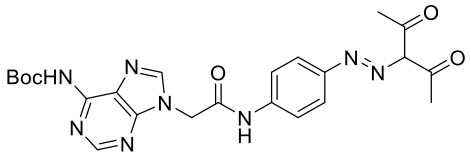
Following the **general procedure A**, using the carboxymethyl derivative of thymine (1.10 g, 5.93 mmol), HCTU (2.45 g, 5.93 mmol), DIPEA (1.53 g, 2.1 mL, 11.86 mmol), (E)-3-((4-aminophenyl)diazenyl)pentane-2,4-dione[S1,S2] (1.00 g, 4.56 mmol) and DMF (46 mL), **compound 1** was obtained as yellow solid **(1.27 g**, **72 %)**.

^{1}H-RMN (500 MHz, DMSO-d_6) δ: 14.23 (s, 1H), 11.34 (s, 1H), 10.38 (s, 1H), 7.65 (d, J = 8.7 Hz, 2H), 7.56 (d, J = 8.8 Hz, 2H), 7.51 (s, 1H), 4.50 (s, 2H), 2.47 (s, 3H), 2.41 (s, 3H), 1.78 (s, 3H).

^{13}C-RMN (126 MHz, DMSO-d_6) δ: 196.4, 196.2, 165.7, 164.4, 151.1, 142.4, 137.3, 136.3, 132.9, 120.1, 117.0, 108.0, 50.0, 31.1, 26.4, 11.9.

HRMS (ESI): Calculated for $C_{18}H_{20}N_5O_5$ $[M+H]^+$: 386.1459; Found: 386.1467.

(E)-(2-[6-[Bis(tert-butoxycarbonyl)amino]-9H-purin-9-yl)-N-(4-((2,4-dioxopentan-3-yl)diazenyl)phenyl)acetamide (2).



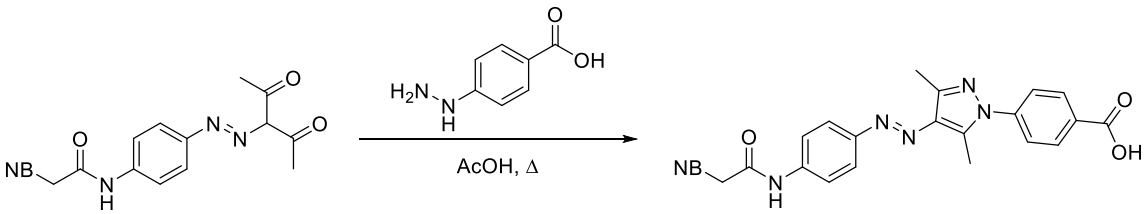
Following the **general procedure A**, using 2-[6-[bis(tert-butoxycarbonyl)amino]-9H-purin-9-yl]acetic acid[S3] (2.33 g, 5.93 mmol), HCTU (2.45 g, 5.93 mmol), DIPEA (1.53 g, 2.10 mL, 11.86 mmol), (E)-3-((4-aminophenyl)diazenyl)pentane-2,4-dione[S1,S2] (1.00 g, 4.56 mmol) and DMF (46 mL), **compound 2** was obtained as yellow solid **(1.51 g, 67 %)**.

^{1}H-RMN (500 MHz, DMSO-d_6) δ: 14.21 (s, 1H), 10.63 (s, 1H), 10.07 (s, 1H), 8.57 (s, 1H), 8.43 (s, 1H), 7.65 (d, J = 8.9 Hz, 2H), 7.56 (d, J = 9.0 Hz, 2H), 5.19 (s, 2H), 2.47 (s, 3H), 2.40 (s, 3H), 1.49 (s, 9H).

^{13}C-RMN (126 MHz, DMSO-d_6) δ: 196.5, 196.3, 164.9, 152.2, 151.6, 151.2, 149.8, 145.0, 137.5, 136.2, 133.1, 123.0, 120.3, 117.1, 80.2, 45.8, 31.2, 27.9, 26.5.

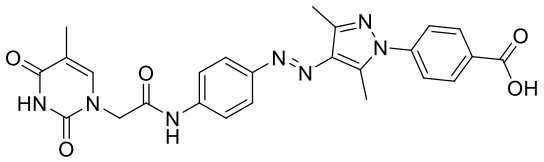
HRMS (ESI): Calculated for $C_{23}H_{27}N_8O_5$ $[M+H]^+$: 495.2099; Found: 494.2093.

General procedure B for the synthesis of arylazopyrazole derivatives (*pyrazole synthesis of Knorr*).



The corresponding acetylacetone derivative (1.0 equiv) and 4-hydrazinobenzoic acid (4.5 equiv) were placed in a sealed tube with a stirring bar. Acetic acid was added and the resulting heterogeneous solution was stirred at reflux temperature for 72 h. After this time, the reaction was cooled down to room temperature and placed in an Erlenmeyer. Water was then added over the reaction to precipitate the desired compound, which was isolated by filtration and washed with more water. The product was obtained as a yellow-orange solid.

(E)-4-(3,5-dimethyl-4-((4-(2-(5-methyl-2,4-dioxo-3,4-dihydropyrimidin-1(2H)-yl)acetamido)phenyl)diazenyl)-1H-pyrazol-1-yl)benzoic acid (Azo-T).



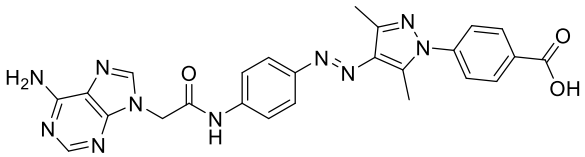
Following the **general procedure B**, using **compound 1** (385 mg, 1.0 mmol), 4-hydrazinobenzoic acid (685 mg, 4.5mmol) and acetic acid (40 mL), **Azo-T** was obtained as yellow solid (**261 mg, 52%**).

^{1}H-NMR (500 MHz, DMSO-d_6) δ: 11.34 (s, 1H), 10.55 (s, 1H), 8.10 (d, *J* = 8.3 Hz, 2H), 7.78 (m, 6H), 7.53 (s, 1H), 4.55 (s, 2H), 2.73 (s, 3H), 2.49 (s, 3H) 1.79 (s, 3H).

^{13}C-RMN (126 MHz, DMSO-d_6) δ: 166.6, 166.1, 164.4, 151.1, 148.7, 142.7, 142.4, 142.0, 140.3, 140.0, 135.8, 130.4, 123.8, 122.6, 119.3, 108.1, 50.1, 14.1, 11.9, 11.1.

HRMS (ESI): Calculated for $C_{25}H_{24}N_7O_5$ $[M+H]^+$: 502.1833; Found: 502.1837.

(E)-4-(4-((4-(2-(6-amino-9H-purin-9-yl)acetamido)phenyl)diazenyl)-3,5-dimethyl-1H-pyrazol-1-yl)benzoic acid (Azo-A).



Following the **general procedure B**, using **compound 2** (493 mg, 1.0 mmol), 4-hydrazinobenzoic acid (684.7mg, 4.5mmol) and acetic acid (40 mL), **Azo-A** was obtained as yellow solid (**240 mg, 47%**).

^{1}H-NMR (500 MHz, DMSO-d_6) δ: 10.75 (s, 1H), 8.16 (s, 1H), 8.15 (s, 1H), 8.11 (d, J = 8.6 Hz, 2H), 7.79 (m, 6H), 7.32 (bs, 2H), 5.12 (s, 2H), 2.72 (s, 3H), 2.49 (s, 3H).

^{13}C-RMN (126 MHz, DMSO-d_6) δ: 166.6, 165.4, 155.1, 151.4, 149.7, 148.8, 142.7, 142.3, 142.1, 140.2, 140.0, 135.9, 130.4, 129.7, 123.8, 122.6, 119.4, 45.8, 14.1, 11.1.

HRMS (ESI): Calculated for $C_{25}H_{23}N_{10}O_3$ $[M+H]^+$: 511.1949; Found: 511.1955.

3. Optical spectroscopy

3.1. Samples preparation

All samples were prepared from stock solutions of the corresponding oligonucleotide and nucleobase-arylazopyrazole derivative. For the stock solution of each azocompound, a concentration of 664 µM in DMSO was employed. For each measured sample, a final concentration of 20 µM was reached.The oligonucleotides were dissolved in Tris buffer at pH 7.5, according to the Eurogentec technical data for a stock concentration of 100 µM. In each experiment, the final oligonucleotide concentration was 2 µM. The final volume of each sample was 1.300 mL, and was completed by adding a solution of 1M or 5M NaCl, depending on the experiment.

To avoid false circular dichroism signals, the samples were subjected to a process of Vortex shaking for 2 min, and then to a cycle of irradiation with UV light for 5 min to dissolve possible aggregates and 15 min with blue light to populate the trans isomer. The samples were then stores for 24 h to let them self-organize.

3.2. UV-Vis absorption and Circular Dichroism spectroscopy

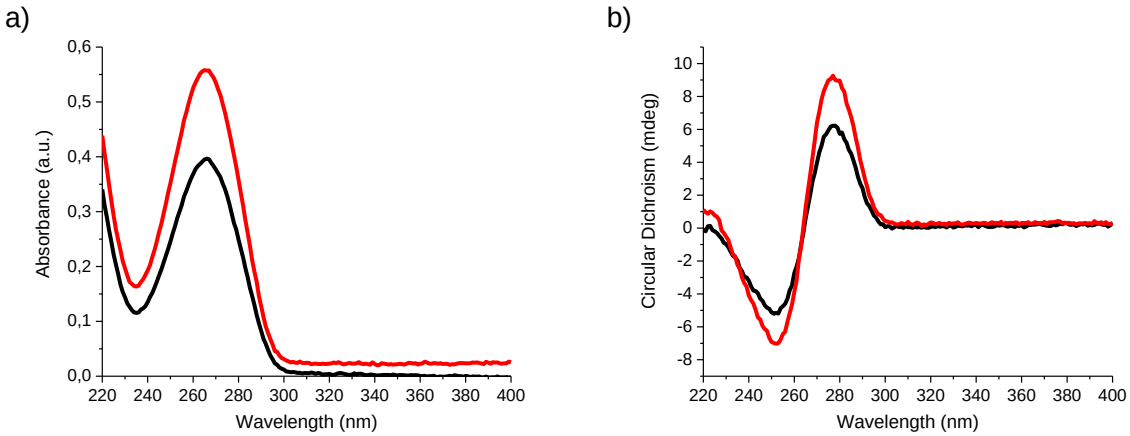


Figure S1. UV-Vis absorption (a) and CD spectra (b) of dT_{20} (black) and dT_{40} (red).

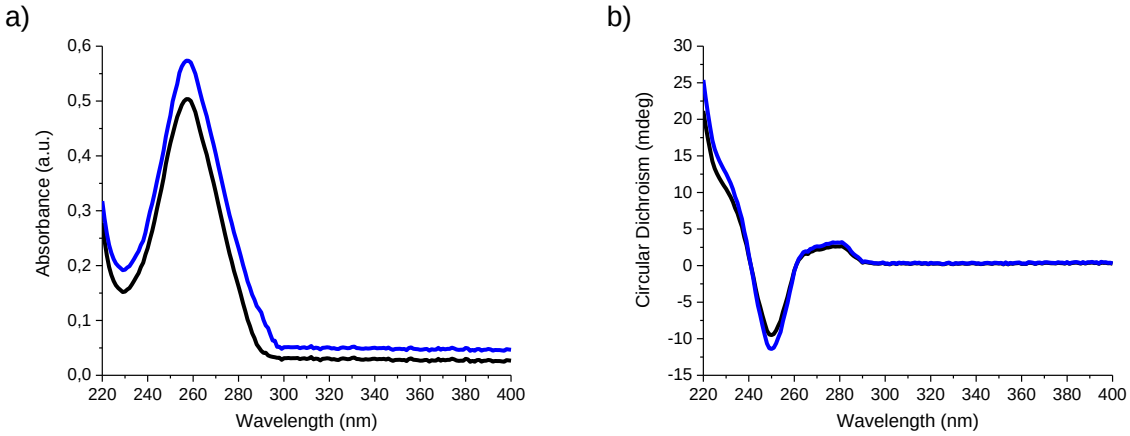


Figure S2. UV-Vis absorption (a) and CD spectra (b) of dA_{20} (black) and dA_{40} (blue).

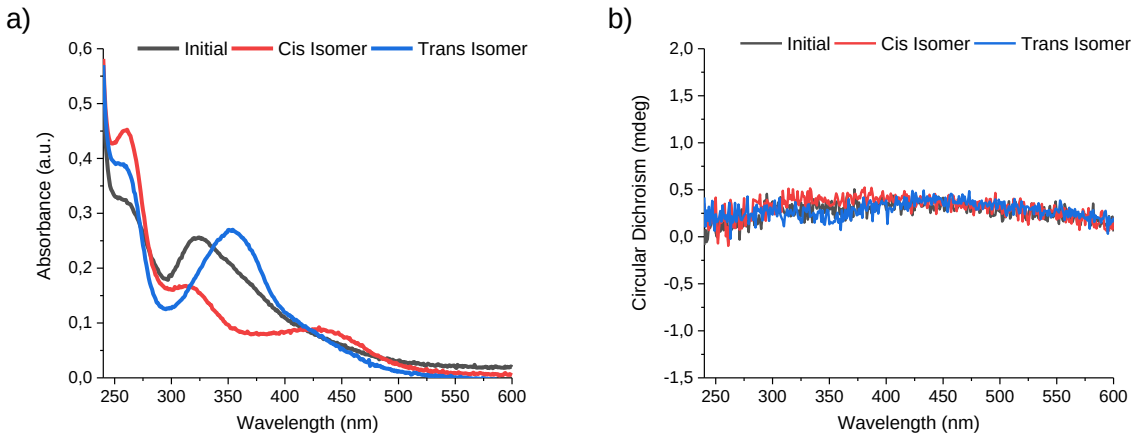


Figure S3. UV-Vis absorption (a) and CD (b) spectra of **Azo-A** irradiated at different wavelengths to enrich the photostationary state with either *cis* (365 nm) or trans (465 nm) isomer.

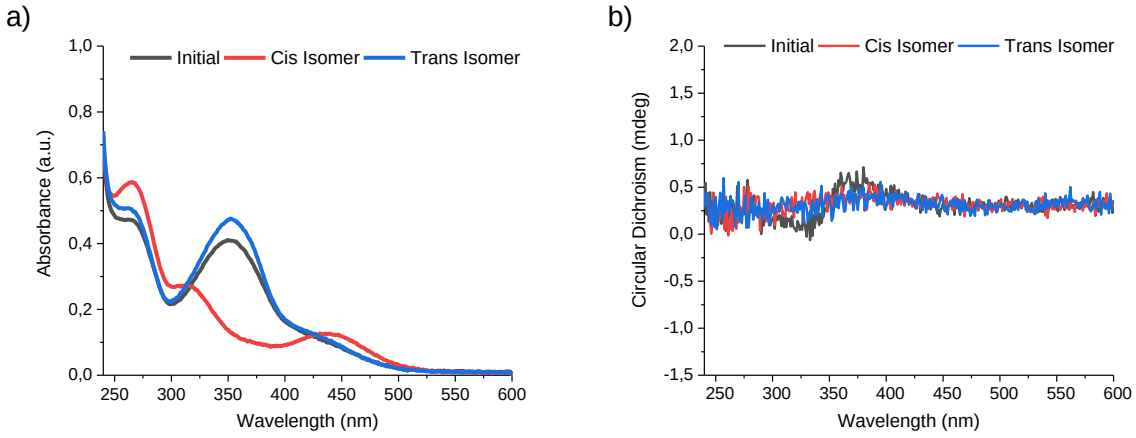


Figure S4. UV-Vis absorption (a) and CD (b) spectra of **Azo-T** irradiated at different wavelengths to enrich the photostationary state with either *cis* (365 nm) or trans (465 nm) isomer.

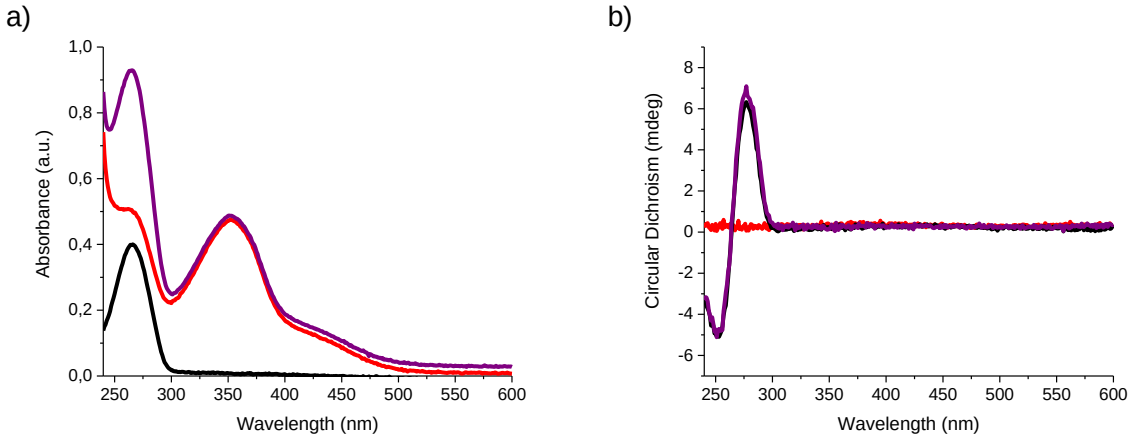


Figure S5. UV-Vis absorption (a) and CD spectra (b) of the **Azo-T** + **dT_{20}** systems and controls at 20 °C and 1M NaCl: **Azo-T** (red), **dT_{20}** (black) and **Azo-T/dT_{20}** (purple).

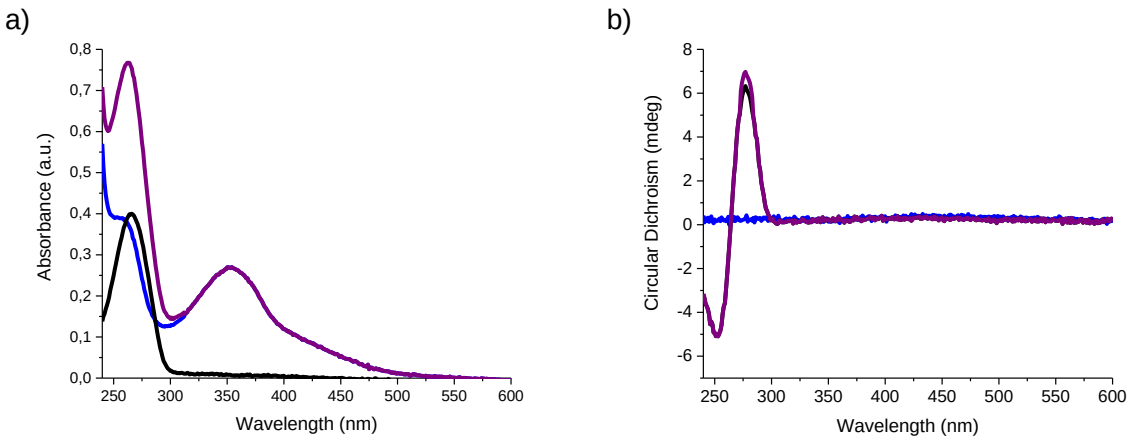


Figure S6. UV-Vis absorption (a) and CD spectra (b) of the **Azo-A** + **dA_{20}** systems and controls at 20 °C and 1 M NaCl: **Azo-A** (blue), **dA_{20}** (black) and **Azo-A/dA_{20}** (purple).

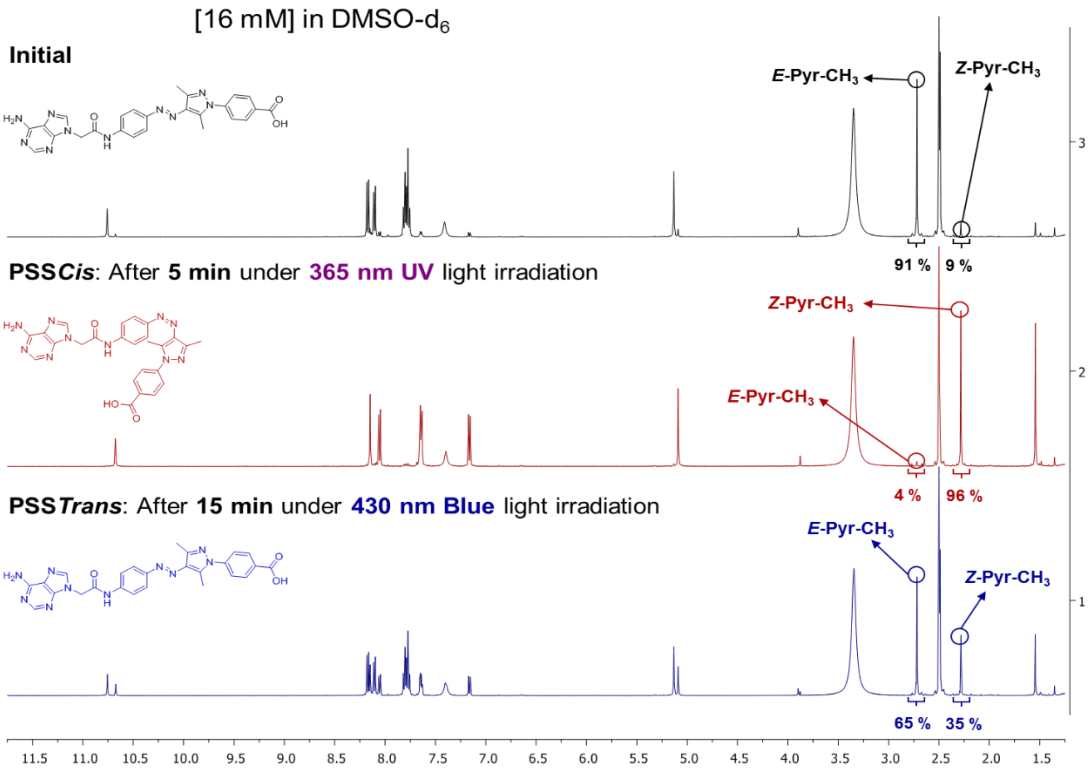


Figure S7. Photostationary states of **Azo-A** at 16 mM concentration**,** determined by ^{1}H-NMR using in DMSO-d6. **PSS *Cis*:** red spectrum. **PSS *Trans*:** blue spectrum.

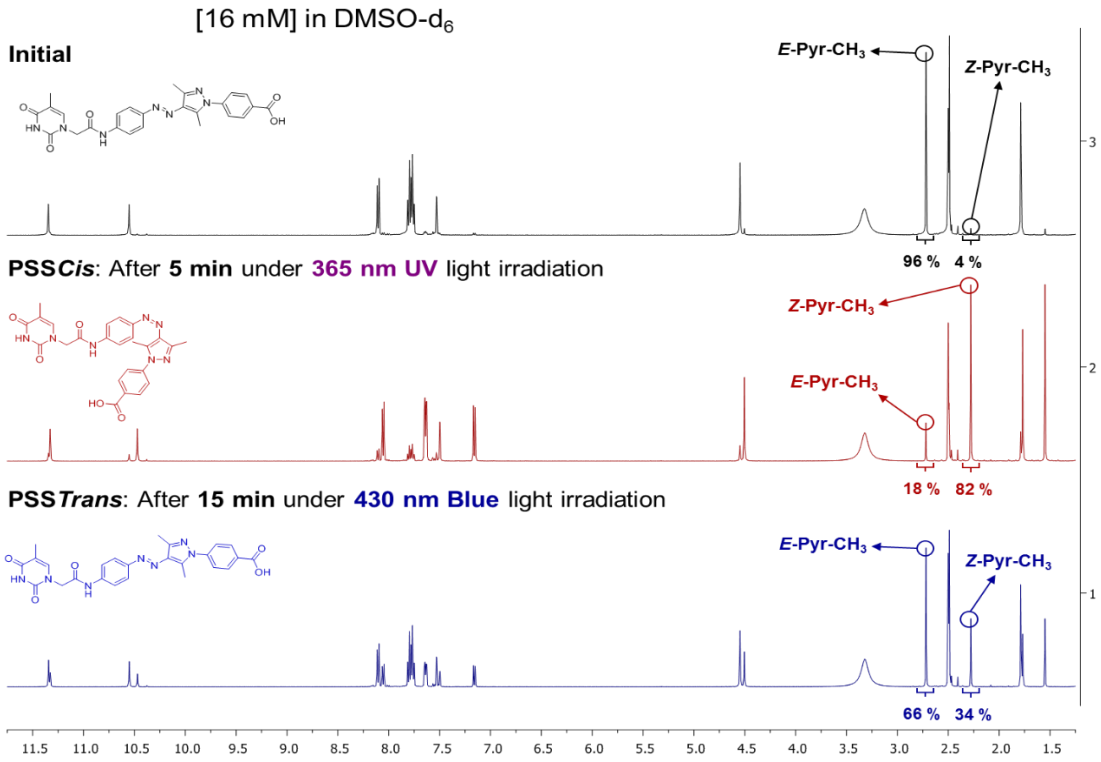


Figure S8. Photostationary states of **Azo-T** at 16 mM concentration**,** determined by ^{1}H-NMR using in DMSO-d6. **PSS *Cis*:** red spectrum. **PSS *Trans*:** blue spectrum.

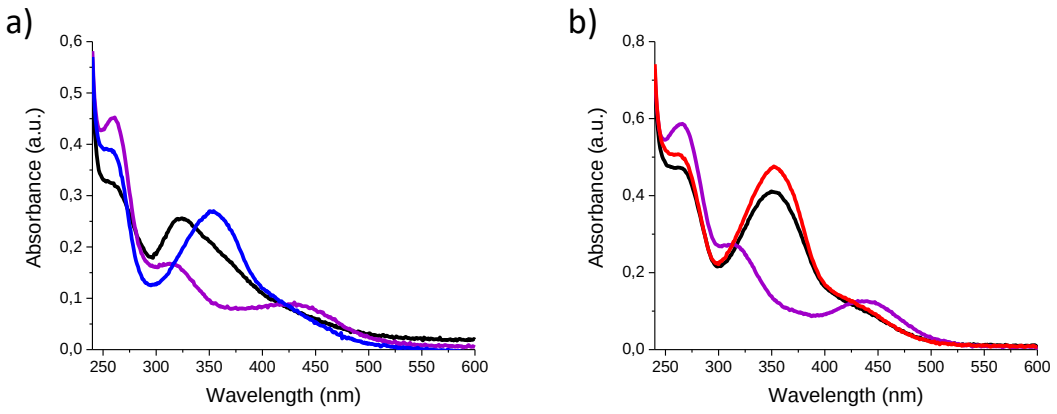


Figure S9. UV-Vis photoisomerization spectra of **Azo-A** (a) and **Azo-T** (b) at 20°C in aqueous solution, recorded after the initial solution preparation (black), irradiation with 365 nm UV light (purple) and 465 nm blue light (blue).

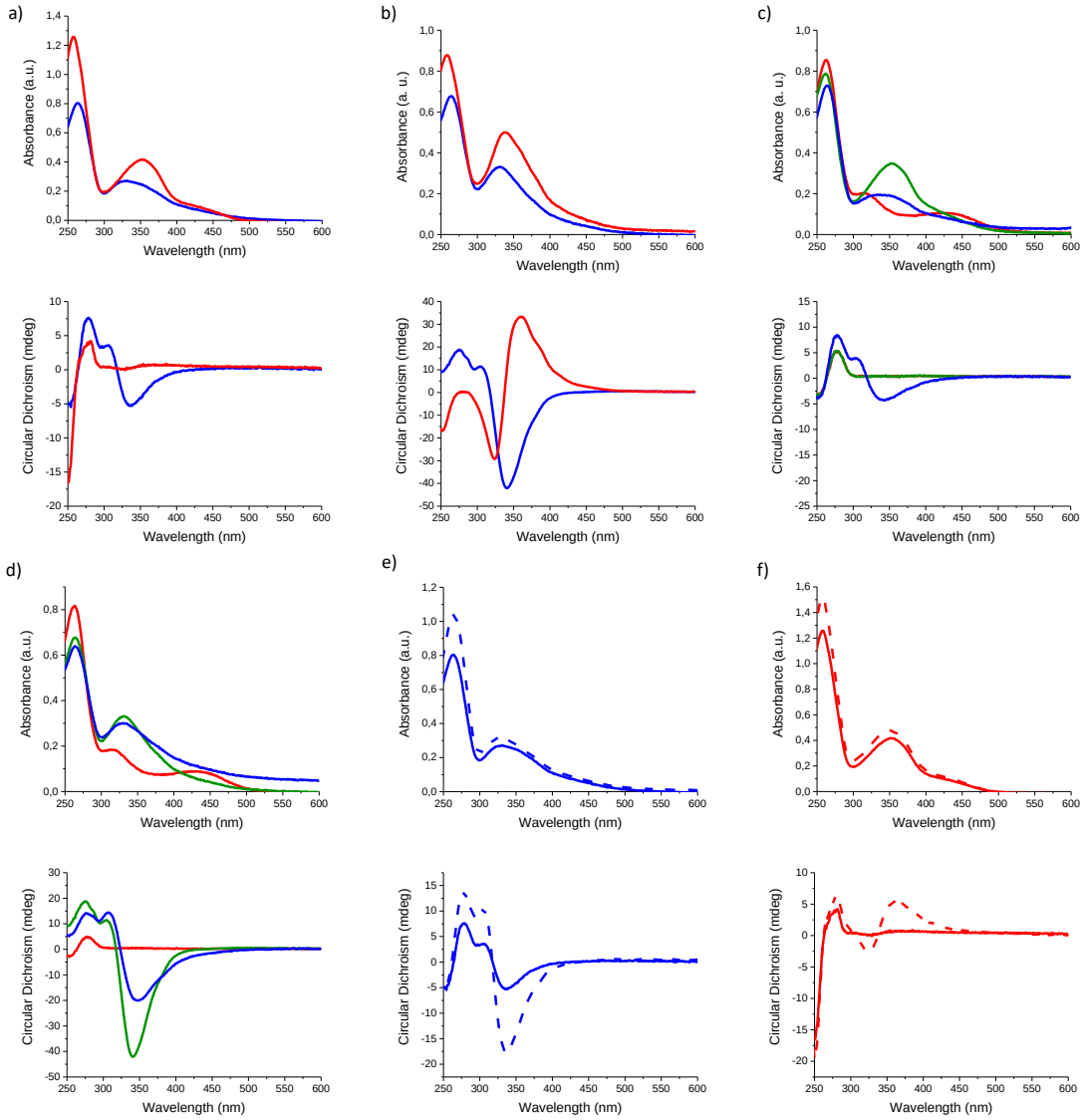


Figure S10. (a,b) UV-Vis and CD spectra of oligonucleotide/nucleobase-arylazopyrazole assemblies at 20 °C and (a) 1M or (b) 5M of NaCl. Blue and red curves correspond to **Azo-A/dT$_{20}$** and **Azo-T/dA$_{20}$**, respectively. (c,d) UV- Vis and CD spectra of **Azo-A/dT$_{20}$** assemblies at 20 °C and (c) 1M or (d) 5M of NaCl, recorded after the initial solution preparation, irradiation with 365 nm UV light (red) and 465 nm blue light (green), and after one week (blue). At 1M NaCl, an additional time of one week was required to get a partial recovery of the ICD signal. (e,f) UV-Vis and CD spectra showing the effect of oligonucleotide templates with a different number of bases at 20 °C and NaCl 1M, for (e) **Azo-A/dT$_{20}$** (solid line) and **Azo-A/dT$_{40}$** (dashed line); and for (f) **Azo-T/dA$_{20}$** (solid line) and **Azo-T/dA$_{40}$** (dashed line).

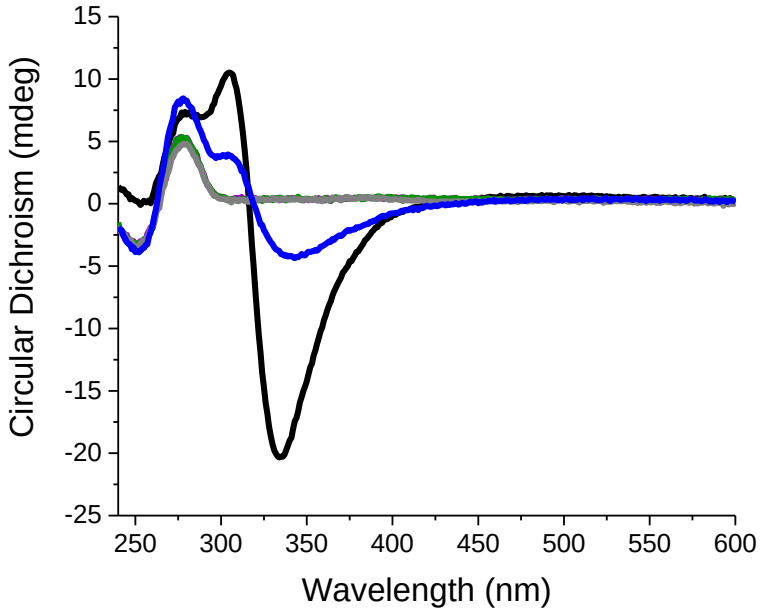


Figure S11. CD spectra of **Azo-A/dT$_{20}$** at 20°C and 1 M NaCl, recorded after the initial solution preparation (black), irradiation with 365 nm UV light (purple), 465 nm blue light (green), after 24 hours (grey) and after one week (blue). CD spectra recorded after irradiation with UV light, blue light and after 24 hours overlap.

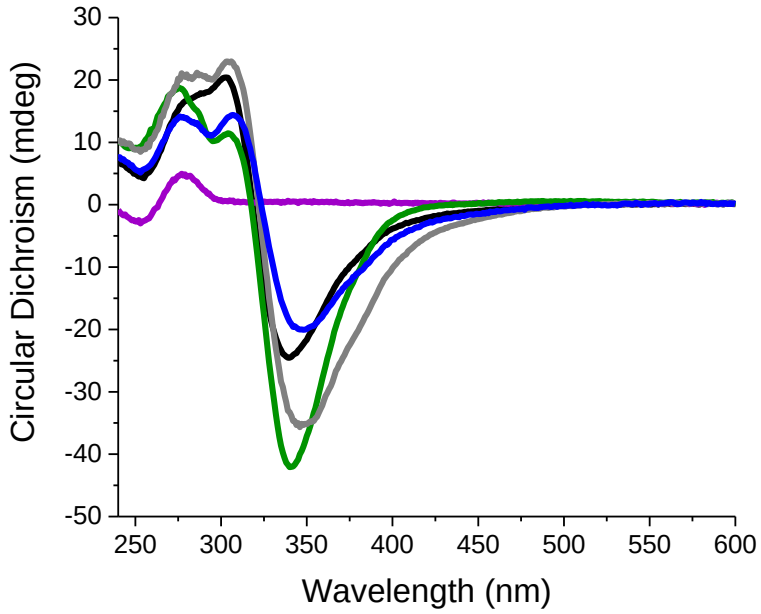


Figure S12. CD spectra of **Azo-A/dT$_{20}$** at 20°C and 5 M NaCl, recorded after the initial solution preparation (black), irradiation with 365 nm UV light (purple), 465 nm blue light (green), after 24 hours (grey) and after one week (blue).

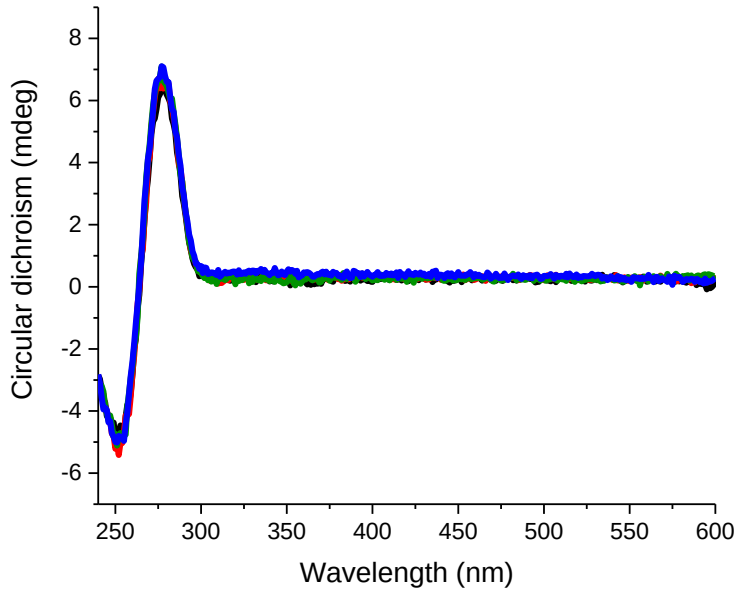


Figure S13. CD spectra of **Azo-T/dA$_{20}$** at 20°C and 1 M NaCl, recorded after the initial solution preparation (black), irradiation with 365 nm UV light (red), 465 nm blue light (green) and after one week (blue).

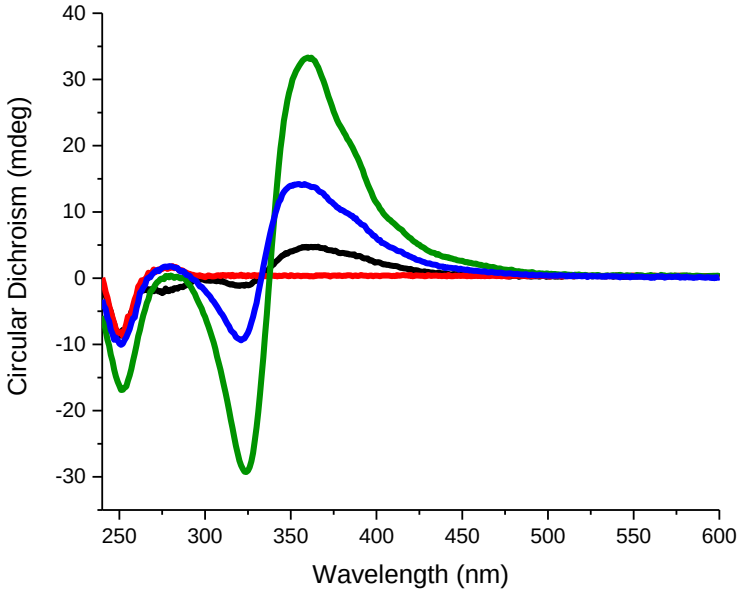


Figure S14. CD spectra of **Azo-T/dA$_{20}$** at 20°C and 5 M NaCl, recorded after the initial solution preparation (black), irradiation with 365 nm UV light (red), 465 nm blue light (green) and after one week (blue).

4. Molecular dynamics (MD) simulations

4.1 Reparameterization of the arylazopyrazole molecules

In order to sample correctly the conformational space of the arylazopyrazole molecules, a reparameterization of 4 dihedral angles was carried out. It was done with the *mdgx* module implemented in the AMBER suite of programs.[1] This tool allows one to generate a lot of conformers for a molecule by varying parameters (bonds, angles, dihedrals) selected by the user. The energy of these conformers is then calculated by molecular mechanics (MM), with the initial force-field parameters, which were in our case provided by the General Amber Force-Field (GAFF) 2.1, and compared to the energy of the same conformers calculated by quantum mechanics (QM).[2] The *Gaussian 16* software was used to perform the QM calculations, using the MP2 method and the cc-pvdz basis set.[3] A new set of parameters is generated by fitting the MM energies to the QM ones, using the linear least-squares method. This process can be repeated, the new set of parameters being used to generate a new set of conformers, which will lead to another, refined, set of parameters, and so on, until no significant improvement is achieved. The whole arylazopyrazole molecule was divided in 2 fragments (to lower the QM computational time and the number of degrees of freedom) containing the 4 reparameterized dihedral angles (Φ_1, Φ_2, Φ_3, Φ_4), as shown in the **Figure S15**. The Pearson correlation coefficients between MM and QM relative energies, before and after reparameterization, are shown in the **Table S1** and the torsional profiles are shown in the **Figure S16**. These results show clearly that the original parameters, issued from the GAFF 2.1 force-field, were doing a poor job in describing the torsional space of these molecules. This is particularly striking for the Φ_1 angle, which had completely inverted maxima and minima. It must be noted that, for this particular angle, the periodicity of the dihedral angle had to be modified manually from 3 to 2 to allow *mdgx* to find a correct agreement with QM data. It is also clear that GAFF 2.1 was not detecting the local minimum for the *cis* stereoisomer (Φ_2), which is correctly described after reparameterization. The fit for the angle Φ_4 is less satisfactory, but overall, the maxima and minima are better reproduced after the reparameterization, as seen on the torsional profile, even though the agreement is not perfect.

Figure S15. Chemical structure of the fragments used for the reparameterization.

Dihedral angle	Correlation (before reparameterization)	Correlation (after reparameterization)
Φ_1	-0.94	0.99
Φ_2	0.27	0.96
Φ_3	0.97	0.99
Φ_4	0.4	0.47

Table S1. Pearson correlation coefficients between the MM and QM relative energies, before and after reparameterization, for each dihedral angle.

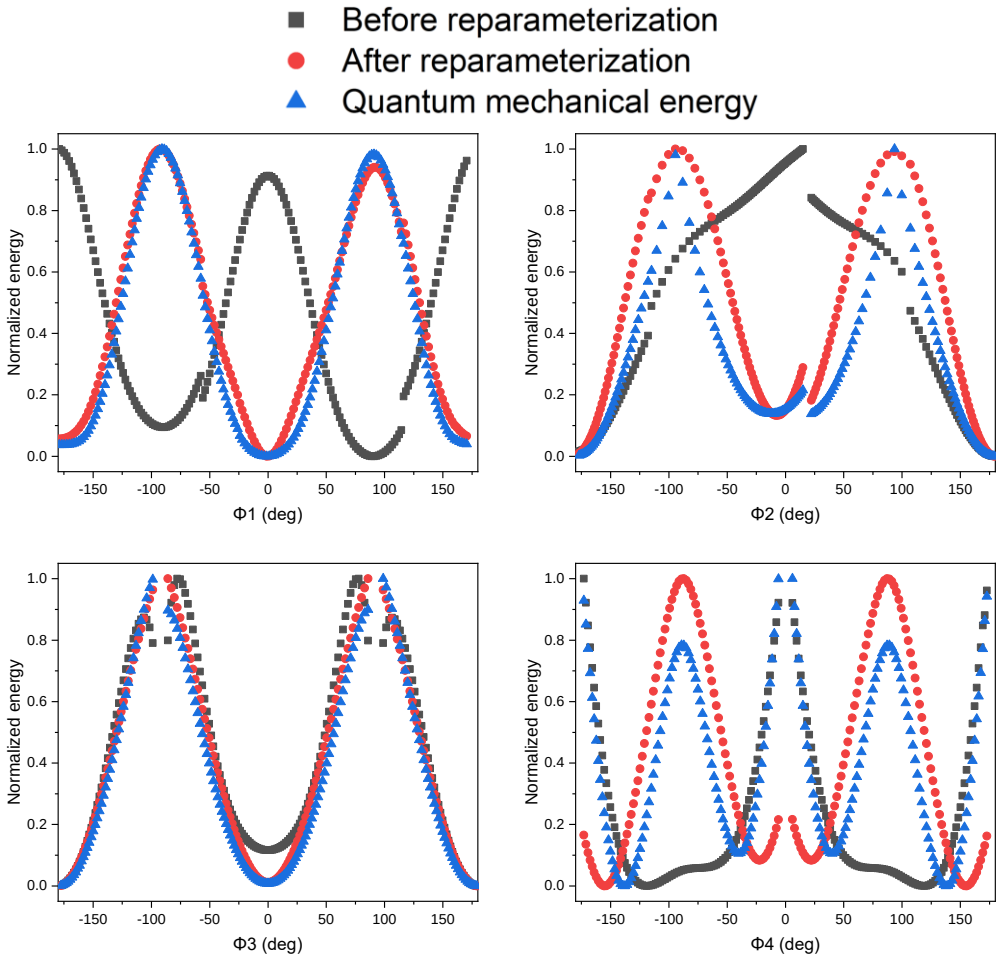


Figure S16. Normalized energy profile of the 4 investigated dihedral angles, showing the MM energies in black and red (before and after reparameterization, respectively) and the QM energy in blue.

4.2 MD protocol

To build the **Azo** compounds, the structure of the whole molecule was divided in 3 fragments, that were made with the *Avogadro 1.2.0* software **(Figure S17)**.[4] The molecule is built in its deprotonated form, as benzoic acid has a pKa value of about 4.[5] The assembly of the fragments and all subsequent operations were carried out with the AMBER simulation package and the associated programs included in AmberTools.[1] The division in fragments has the advantage that the arylazopyrazole part (in black in **Figure S17**) is built only once and can be kept for all molecules. Only the nucleobase subgroup (in blue in **Figure S17**) has to be built for each iteration. The calculations of the atomic partial charges were performed with the *antechamber* module of AMBER, using the semi-empirical AM1-BCC method.[6,7] All the force-field parameters for the **Azo** compounds were given by GAFF 2.1, except the reparameterized ones. The fragments were assembled using the *LEaP* module of AMBER. The oligonucleotides were built with the *Nucleic Acid Builder* (NAB) tool implemented in AMBER and the DNA force-field parameters were given by Parmbsc1.[8] The arylazopyrazole-oligonucleotide supramolecular complexes were built within LEaP, 10 **Azo** molecules being in interaction with an oligonucleotide of 20 bases. The **Azo** units were preorganized along the template (pairing of the complementary nucleobases) with *PyMol 2.5.4*.[9] The **Azo-A/dT$_{20}$** and **Azo-T/dA$_{20}$** complexes were simulated independently, in 2 replicas for each isomer (*trans* and *cis*), giving 8 independent simulations. They are subsequently referred to as **Azo-X** TRANS 1, **Azo-X** TRANS 2, **Azo-X** CIS 1, **Azo-X** CIS 2 (with **X** being **A** or **T** for **Azo-Adenine** and **Azo-Thymine**, respectively) **(Figure S18)**. All systems were solvated in truncated octahedral water boxes, with at least 25.0 Å between any solute atom and the edge of the box, in order to let enough space for the arylazopyrazole units to have the possibility to dissociate from the template. The 4-point OPC water model was used to describe the solvent and a concentration of NaCl of 5M was used to reproduce the experimental conditions, following the "SPLIT" method.[10,11]

« Azo-T » **« Azo-A »**

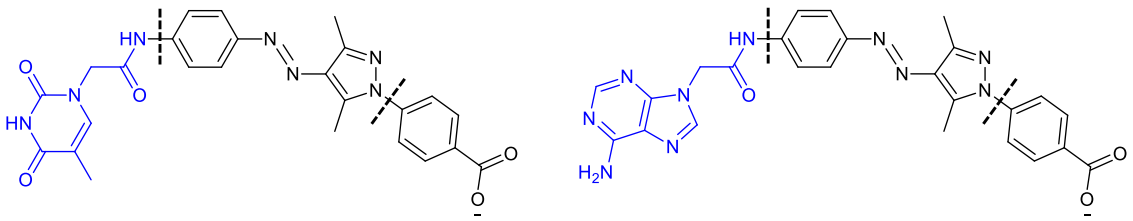


Figure S17. Chemical structure of the thymine-arylazopyrazole (**Azo-T**) and adenine-arylazopyrazole (**Azo-A**) molecules, divided in three fragments, as indicated by the black dots.

All MD simulations were performed with the GPU version of Amber.[12] They started with a geometry optimization performed by MM to get a stable starting structure. 1,000 steps of steepest descent were followed by 9,000 steps of conjugated gradient on the solvent and salt residues. A second geometry optimization was done with the same protocol, on the whole system. Then a heating step of 2 ns was performed in the NVT ensemble to bring the system to a temperature of 300 K. Positional restraints on the solute atoms were applied with a force-constant of 10 kcal.mol^{-1}.Å^{-2}. The temperature was maintained at 300 K with a Langevin thermostat, using a collision frequency of 1 ps^{-1}. The system was equilibrated during 10 ns in the NPT ensemble with a Monte-Carlo barostat, with restraints to maintain the hydrogen bonds between the complementary nucleobases of the arylazopyrazole units and the DNA template: a

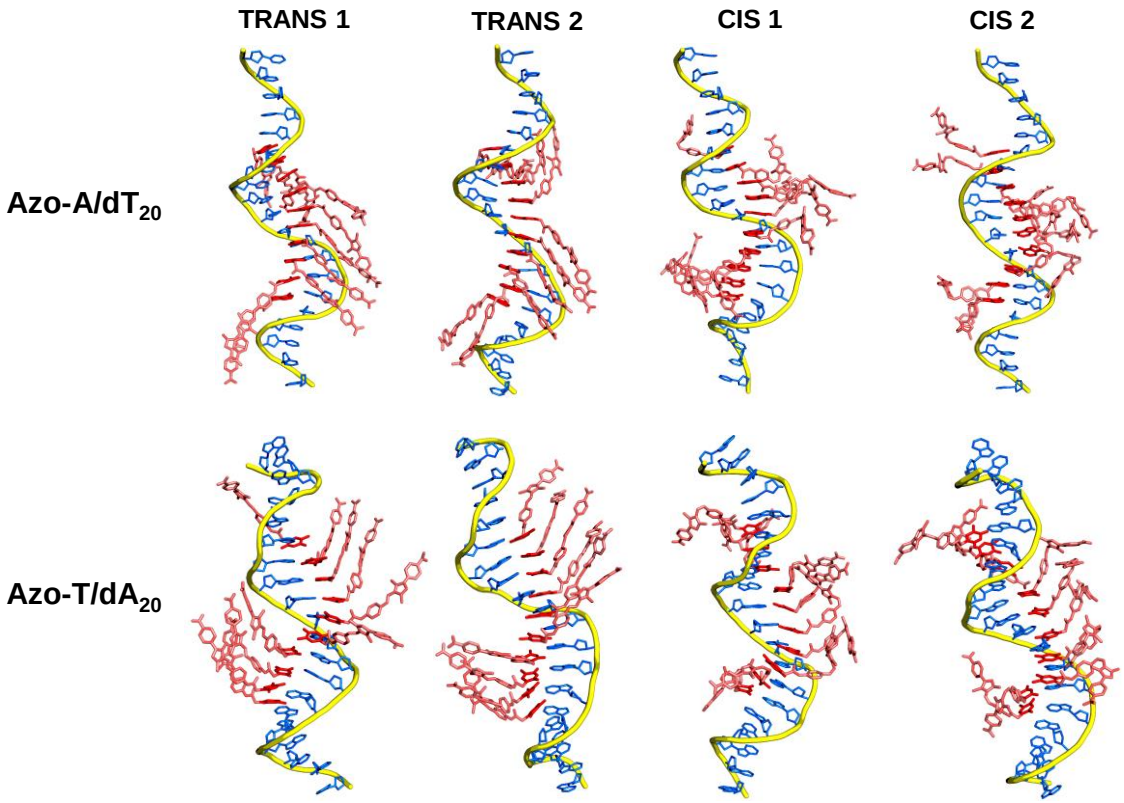


Figure S18. Snapshots of the first frame of the simulation for the 8 systems (2 replicas with the *trans* and *cis* isomers for the **Azo-A**/dT_{20} and **Azo-T**/dA_{20} complexes). The DNA backbone is shown as a yellow tube, its nucleobases are colored in blue, the **Azo** compounds and their nucleobase are colored in pink and red, respectively.

force constant of 40 kcal.mol^{-1}.Å^{-2} was applied as soon as the distance between the donor and the acceptor of the H-bond exceeded 2.2 Å. These restraints were maintained for 250 ns longer, to let the system stabilize while keeping the supramolecular assembly. Finally, the simulations were extended without any constraints. The simulations ran for 1.5 µs for all systems in *cis*, 2 and 1.25 µs for the 2 systems in *trans* for **Azo-A** and 1.75 and 1.25 µs for the 2 systems in *trans* for **Azo-T**. A time-step of 2 fs was used and the SHAKE algorithm was applied to constrain bonds involving hydrogen atoms. To switch the azoderivatives into their *cis* form, a constraint on the Φ_2 dihedral angle was imposed: a force constant of 100 kcal.mol^{-1}.Å^{-2} was applied as soon as the dihedral angle was going out of the range [-30.0 ; 30.0] degrees. In practice, this force constant helped the arylazopyrazole units to bypass the torsional barrier leading from the *trans* to *cis* configuration. A cut-off of 12.0 Å was selected for non-bonded interactions and the Particle Mesh Ewald (PME) scheme was used to treat electrostatic interactions. A snapshot was saved each ns and extracted for further analyses. *PyMol 2.5.4* was used to visualize the snapshots and to create images.[9]

4.3 Analyses of the trajectories

To analyze the trajectories, the *cpptraj* module implemented in AMBER was used.[13] Hydrogen bonds were detected with geometric criteria: the distance between the acceptor and the donor heavy atoms must be ≤ 3.0 Å and the angle between the donor, the hydrogen atom and the acceptor must be ≥ 135°. H-bonds were measured between the atoms of the nucleobases of the azoderivatives and of the oligonucleotide **(Figure S19)**. A fast decrease occurs when the restraints are removed. Heatmaps of hydrogen bonds were built over the last 500 ns of the simulations. This tool helps to visualize were the interactions are located, between which components of the system **(Figure S20)**. They work as follows: the arylazopyrazole molecules are numbered from 1 to 10 and are represented on the x axis. The bases of the oligonucleotide are numbered from T1 to T20 and A1 to A20, for oligothymine and oligoadenine, respectively, on the y axis. At the crossing of one Azo unit and one DNA nucleobase is found a colored rectangle: the color of this rectangle represents the number of H-bonds detected between these 2 residues, on average, by conformation. At the beginning of the simulation, the **Azo** 1 faces the base n°6, the **Azo** 2 faces the base n°7, and so on. These maps show mainly "diagonal patterns" for the *trans* isomers, meaning that consecutive **Azo** molecules interact with consecutive bases along the template. Even though all the interactions are not maintained, the systems in *trans* stay quite organized at the microsecond timescale. Conversely, the maps for the systems in *cis* (except **Azo-A** CIS 2, which is discussed in the text of the article) display far fewer H-bonds, indicating more difficulty for these systems to organize along the DNA template.

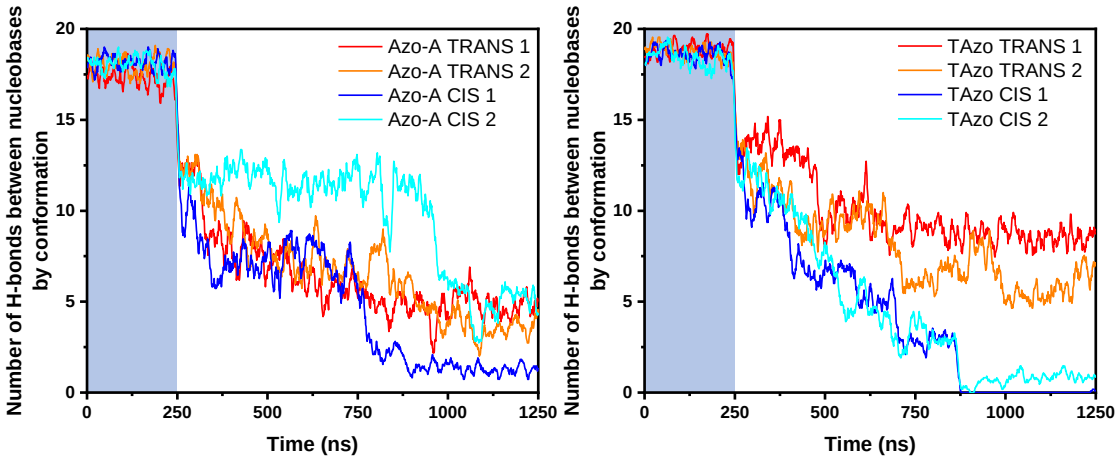


Figure S19. Evolution of the number of intermolecular H-bonds between the nucleobases of the **Azo** compounds and the complementary oligonucleotide (**Azo-A** on the left, **Azo-T** on the right) with respect to time. The running average including the 5 previous and 5 subsequent conformations is displayed, for ease of visualization.

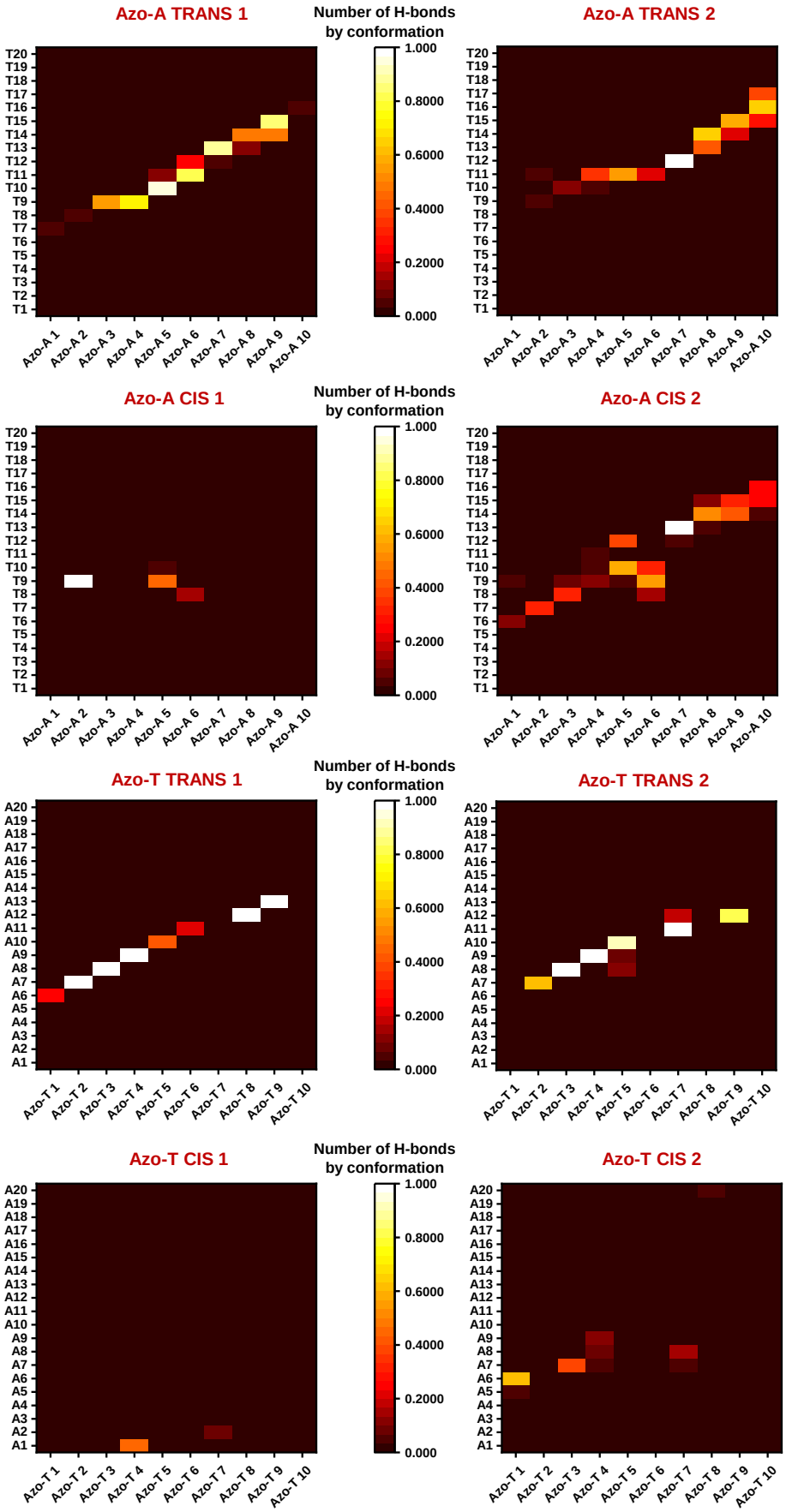


Figure S20. Heatmaps of the intermolecular H-bonds for the 8 simulated systems. The explanation to read them is given in the text.

Stacking interactions between the aromatic cycles were detected with geometric criteria: two cycles are considered stacked if the distance between their centres of mass is ≤ 5 Å and if the angle between them is < 45° or > 135°. The aromatic interactions were calculated for all pairs of cycles of the **Azo** molecules. The running average including the 5 previous and 5 subsequent conformations is displayed, for ease of visualization **(Figure S21)**. Heatmaps of stacking interactions were built over the last 500 ns of the simulations **(Figure S22)**. They work as follows: the aromatic cycles of the Azo compounds are numbered from 1 (first cycle of **Azo** 1) to 40 (last cycle of **Azo** 10), as illustrated for **Azo-T** in the **Figure S23**. In these maps, the same residues are represented on the x and y axes. At the crossing of two residues, a colored square appears: the color of this square indicates the number of aromatic interactions between these 2 residues, on average, by conformation. The heatmaps for the *trans* isomers show a high number of interactions following a "diagonal pattern", meaning that the **Azo** compounds are stacked on top of each other, consecutively. However, in most cases, the diagonal is not fully complete. This means that instead of one stack incorporating all the molecules, 2 or more smaller stacks are organized along the template. For example, for **Azo-T** TRANS 1, no interactions are detected between **Azo-T** 6 and **Azo-T** 7, indicating a discontinuity in the stacking. The maps for the molecules in *cis* configurations display more dispersed interactions, confirming the lack of organization in the stacking of these isomers.

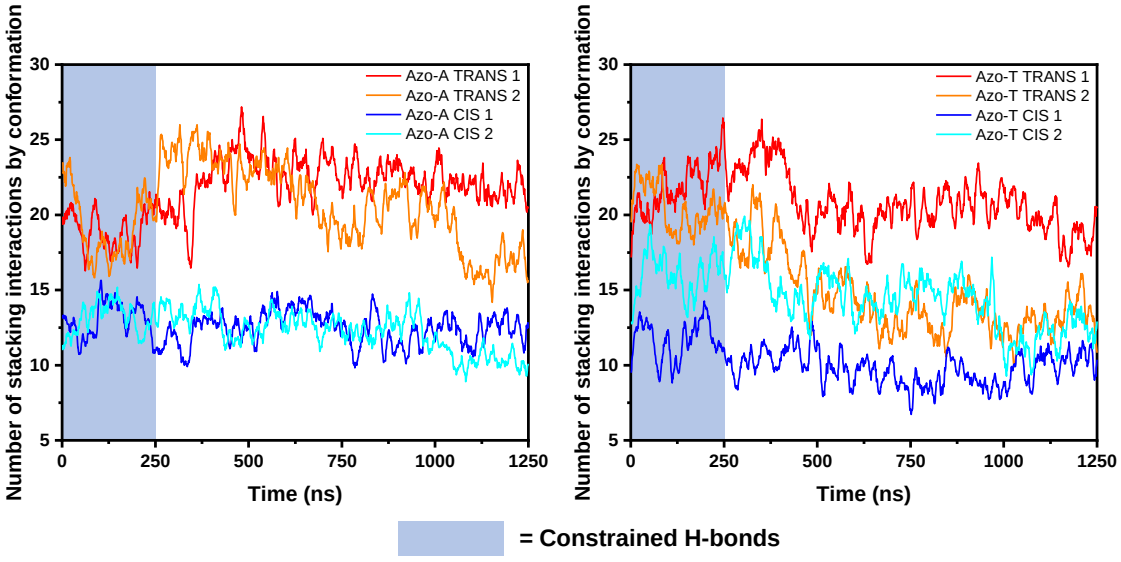


Figure S21. Evolution of the number of stacking interactions between the aromatic cycles of the **Azo** compounds (**Azo-A** on the left, **Azo-T** on the right) with respect to time. The running average including the 5 previous and 5 subsequent conformations is displayed, for ease of visualization.

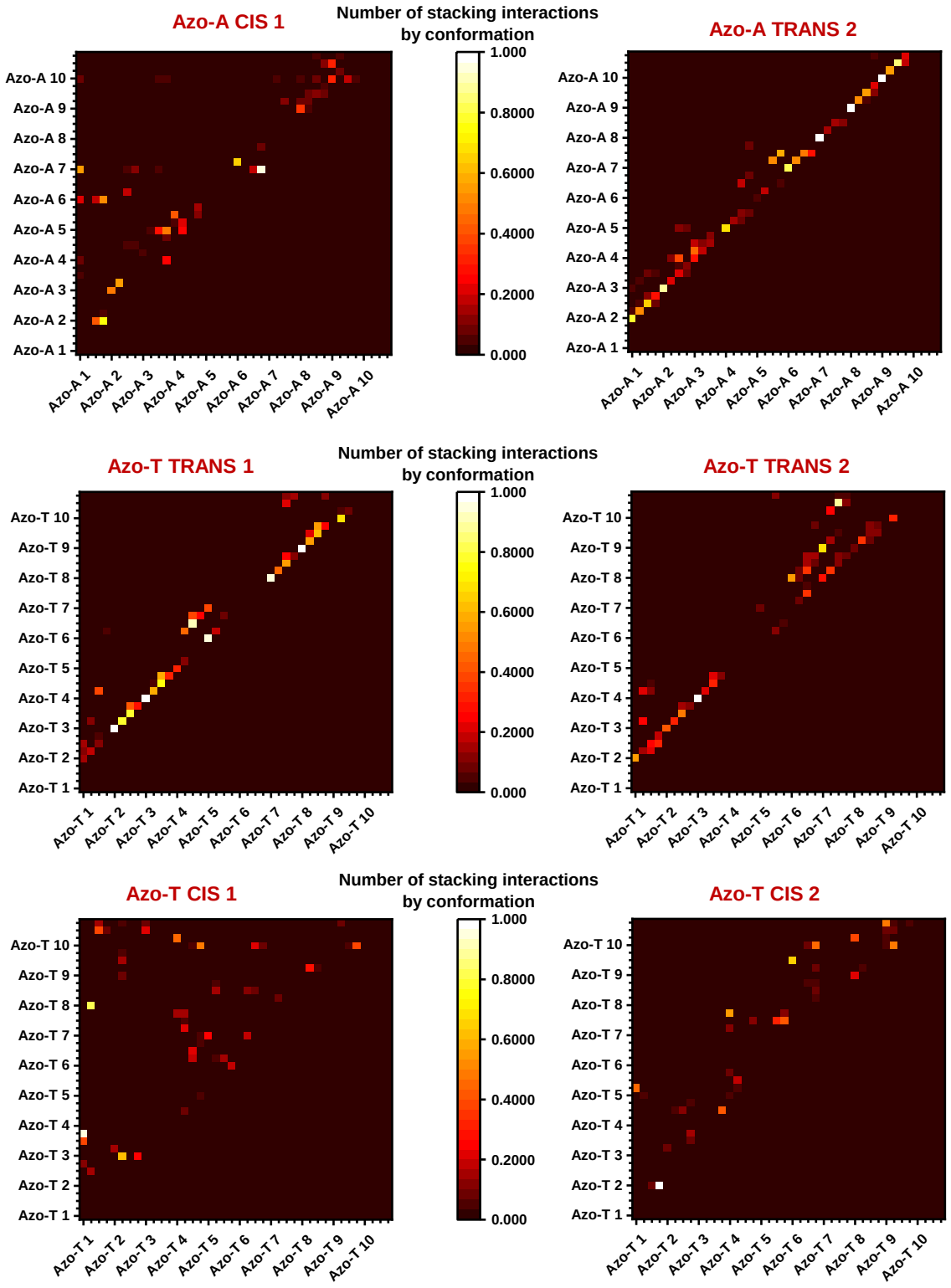


Figure S22. Heatmaps of the stacking interactions between the Azo compounds for 6 simulated systems (the maps for **Azo-A** TRANS 1 and **Azo-A** CIS 2 are shown in the article). The explanation to read them is given in the text.

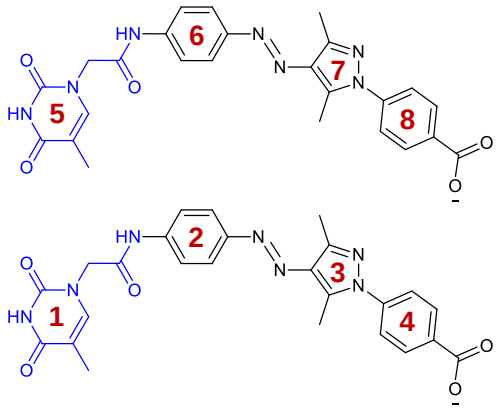


Figure S23. Representation of the numerotation pattern followed to build the heatmaps of stacking interactions. The same logic is followed for the **Azo-A** molecules.

Root-mean-square deviations (RMSD) were calculated, for each **Azo** unit, over the last 500 ns of the simulation. The first conformation of each molecule was used as a reference to suppress translational and rotational movements. RMSD average values increase from 1.7 Å (*trans*) to 2.3 Å (*cis*) for **Azo-A** and from 1.6 Å (*trans*) to 1.9 Å (*cis*), for **Azo-T**.

The dihedral angle Ψ_1 between the nucleobase of the azo compounds and their conjugated part was measured and compared for both systems, **Azo-A** and **Azo-T**, alone (in implicit solvation, 100 ns) and when stacked along the DNA template (explicit solvation, 250 ns) **(Figure S24)**. As shown, interconversion is possible at the single-molecule level, the two minima of Ψ_1 (around -80° and 80°) being occupied more or less equally for both Azo compounds. In contrast, most molecules undergo very limited conversion when stacked to the DNA, some of them being completely trapped in one minima (as shown for 2 **Azo-A** and 2 **Azo-T** compounds, distribution on the right of the figure).

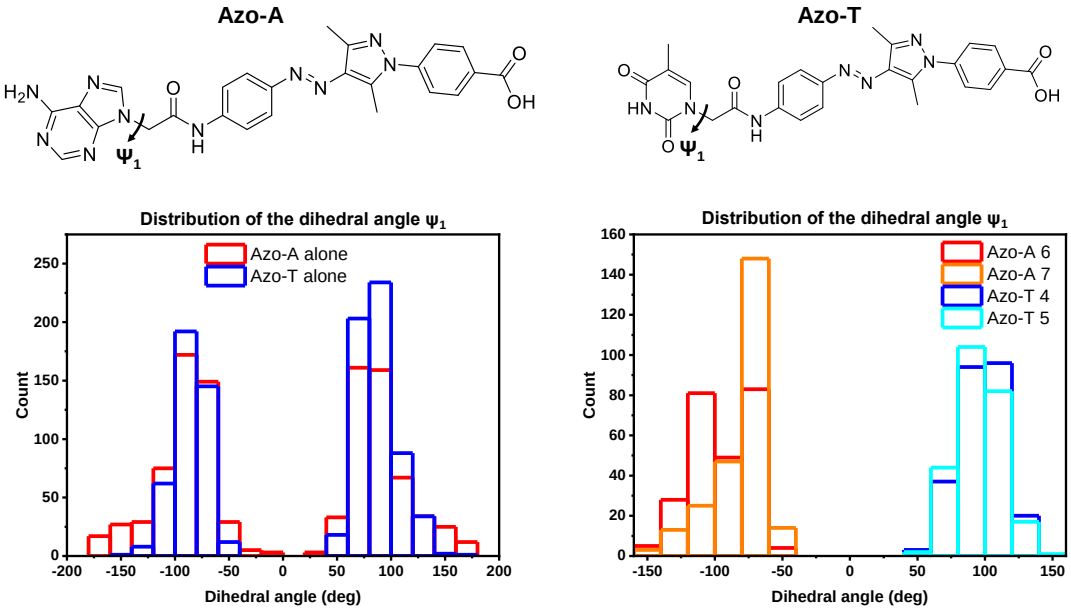
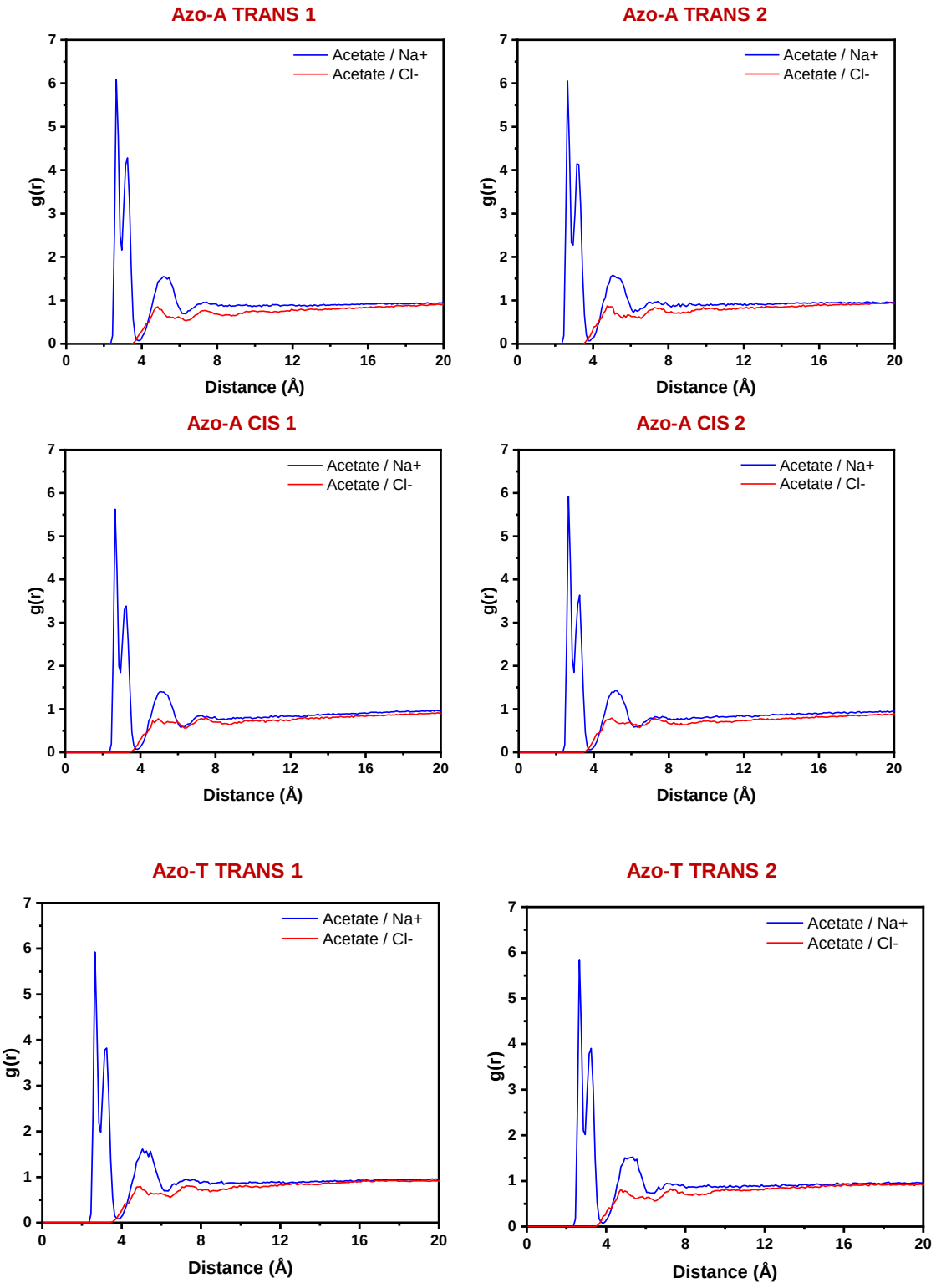


Figure S24. Distribution of the dihedral angle Ψ_1 (shown in the molecular structures above the graphs) for the azo compounds, when they are alone (left) and when they are stacked on the DNA template (right). As can be seen, two minima are occupied at the single-molecule level, while no interconversion occurs for some molecules in the stacks.

Radial distribution functions were measured between the carbon atom of the acetate moiety of each azoderivative and the Na+ or Cl- ions, with a bin spacing of 0.1 Å **(Figure S25)**. The density value used for normalization was calculated as the ratio between the number of ions in the box and the average volume measured during the entire simulation.



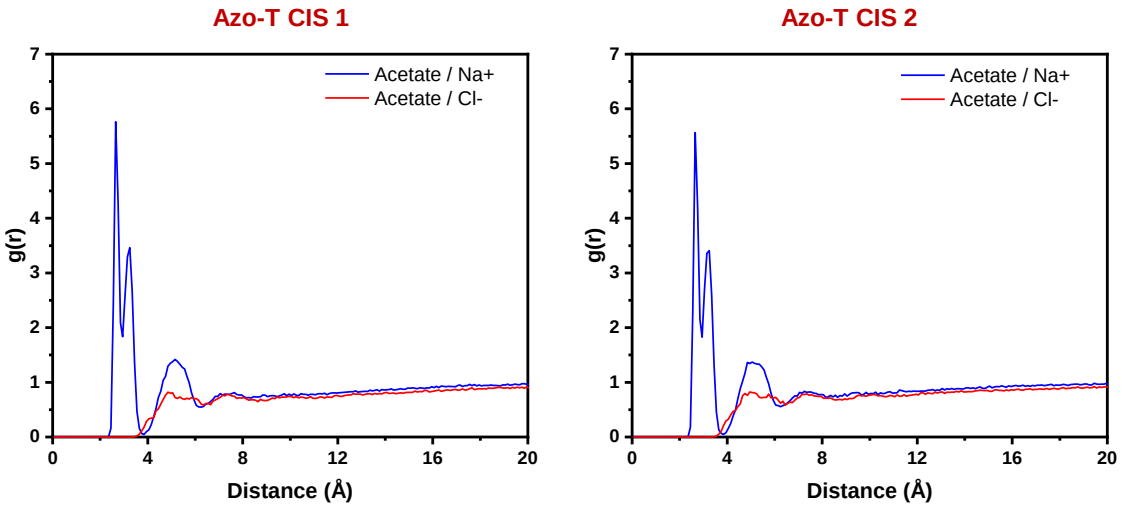


Figure S25. Radial distribution functions of the ions in solution (Na^+ in blue and Cl^- in red) as a function of their distance from the acetate moieties, located at the end of the **Azo** compounds, for the 8 simulated systems.

To measure the rotation between the conjugated parts of consecutive **Azo** derivatives, a vector was defined for each **Azo** unit, as represented by the blue arrows in the **Figure S26**. The angle of rotation between two stacked molecules is calculated from the dot-product of their vectors and was measured for each pair of consecutive **Azo** compounds (**Azo** 1 – **Azo** 2; **Azo** 2 – **Azo** 3; **Azo** 3 – **Azo** 4; and so on). Due to the flexibility of the system, reorganizations within the stacks are possible, and conjugated groups that are superimposed at the beginning of the simulation may dissociate after some time. These pairs were identified by visual inspection and were not taken into account in the calculation of the average angle, nor were those that were "inverted" (angle of ≈ 180°). Following this criterion, the angle was measured for all 9 pairs in **Azo-A** TRANS 1, 7 pairs in **Azo-A** TRANS 2, 6 pairs in **Azo-T** TRANS 1 and 1 pair in **Azo-T** TRANS 2. The distributions of angles for all these pairs are shown in the **Figure S27**. The first 4 bins (angle values comprised between 0 to 10°, 10 to 20°, 20 to 30° and 30 to 40°) account for 24, 29, 18 and 16 % of all angle values, respectively.

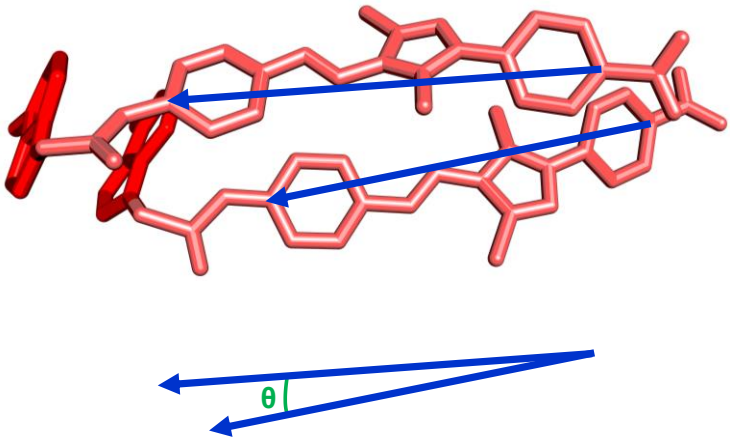


Figure S26. Representation of the vectors used to measure the rotation between two stacked **Azo** units. A vector is defined along the conjugated part of the molecule. The rotation is calculated as the angle, **θ**, between two vectors.

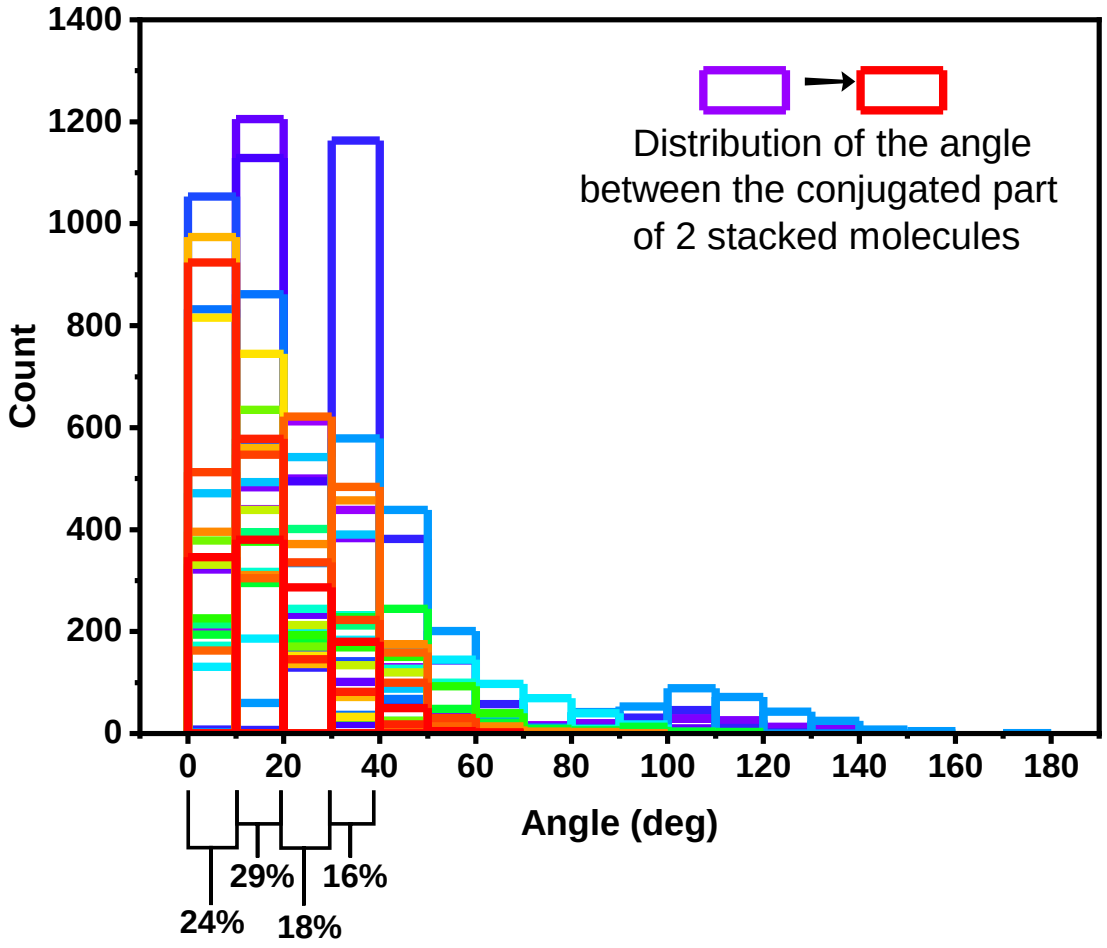
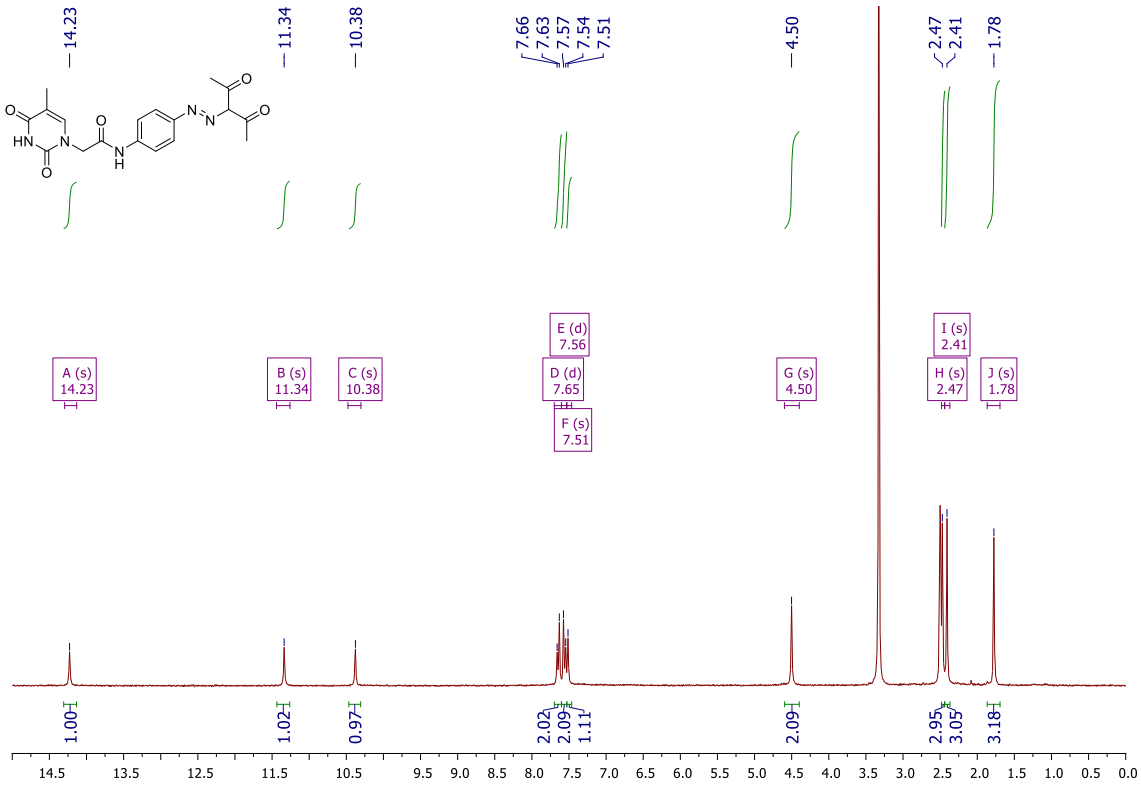


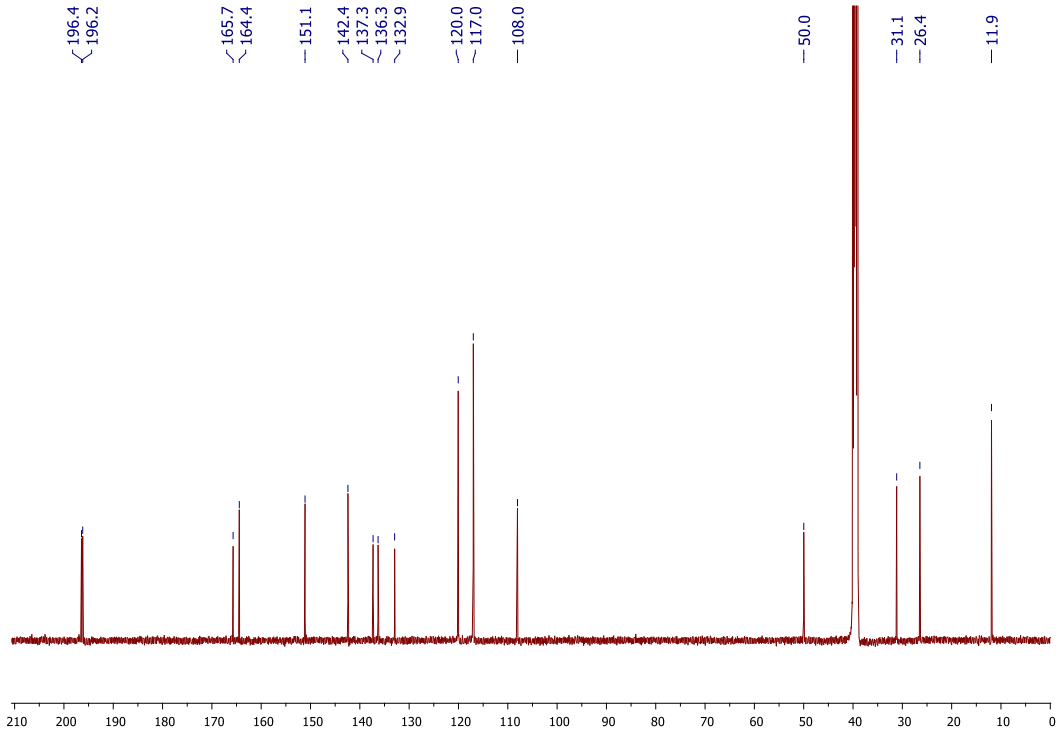
Figure S27. Distribution of the angle values measured for the 24 pairs of stacked molecules investigated, divided in bins of 10°. Each color represents the distribution of the angles for one pair of **Azo** compounds. The contribution of the first 4 bins to all angle values is given in percentage.

5. NMR spectra

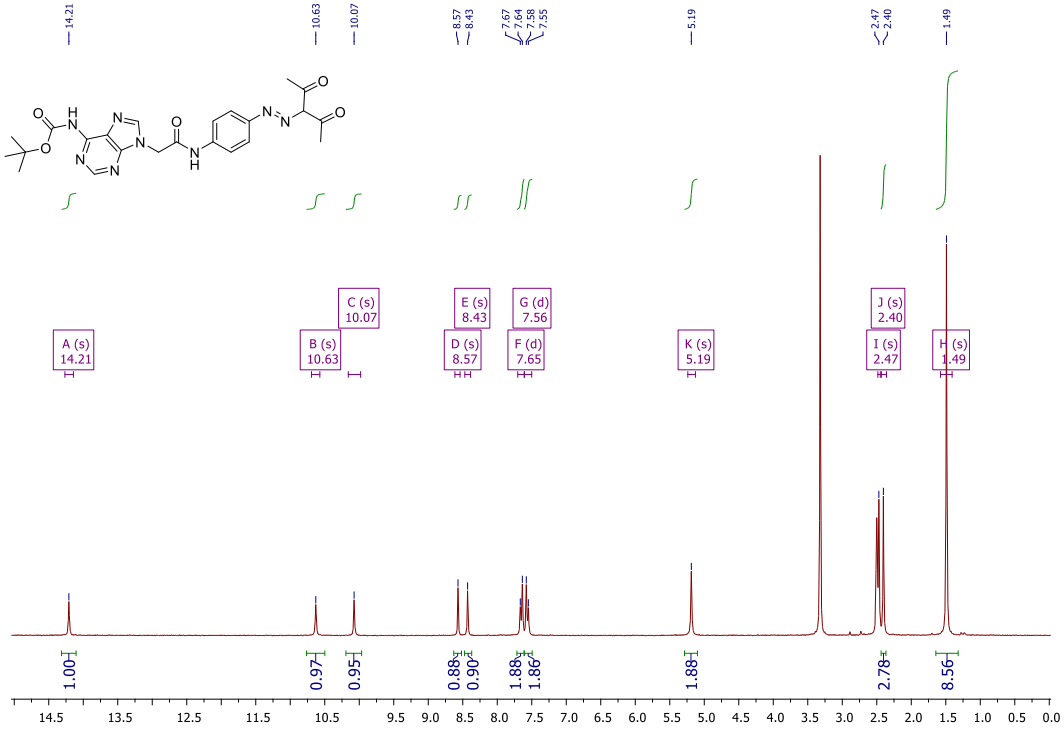
^{1}H-RMN (500 MHz, DMSO-d_6) **compound 1:**



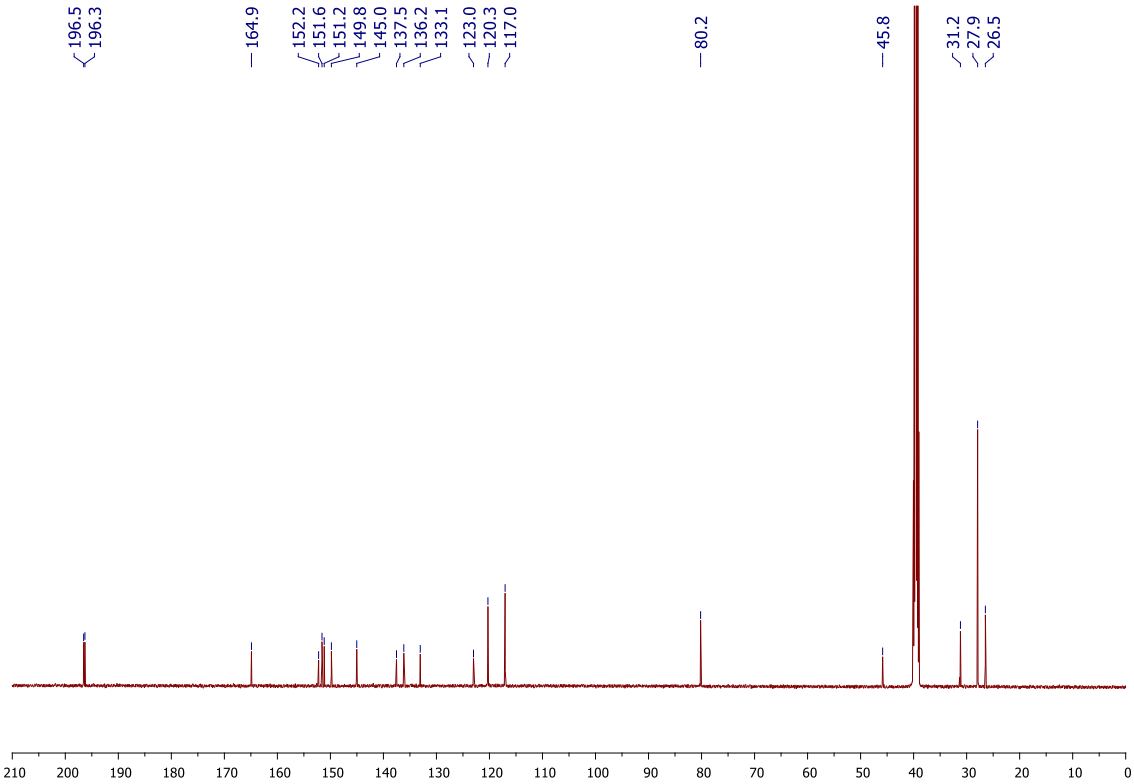
^{13}C-RMN (126 MHz, DMSO-d_6) **compound 1**



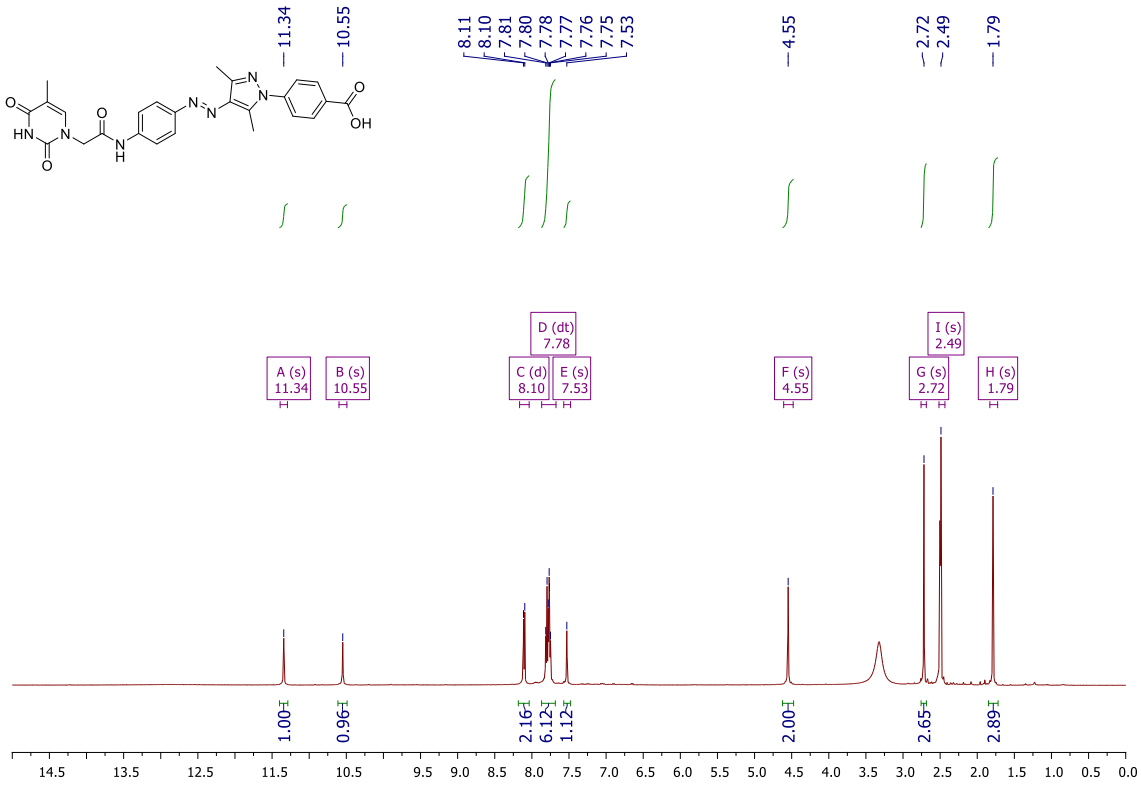
^{1}H-RMN (500 MHz, DMSO-d_6) **compound 2:**



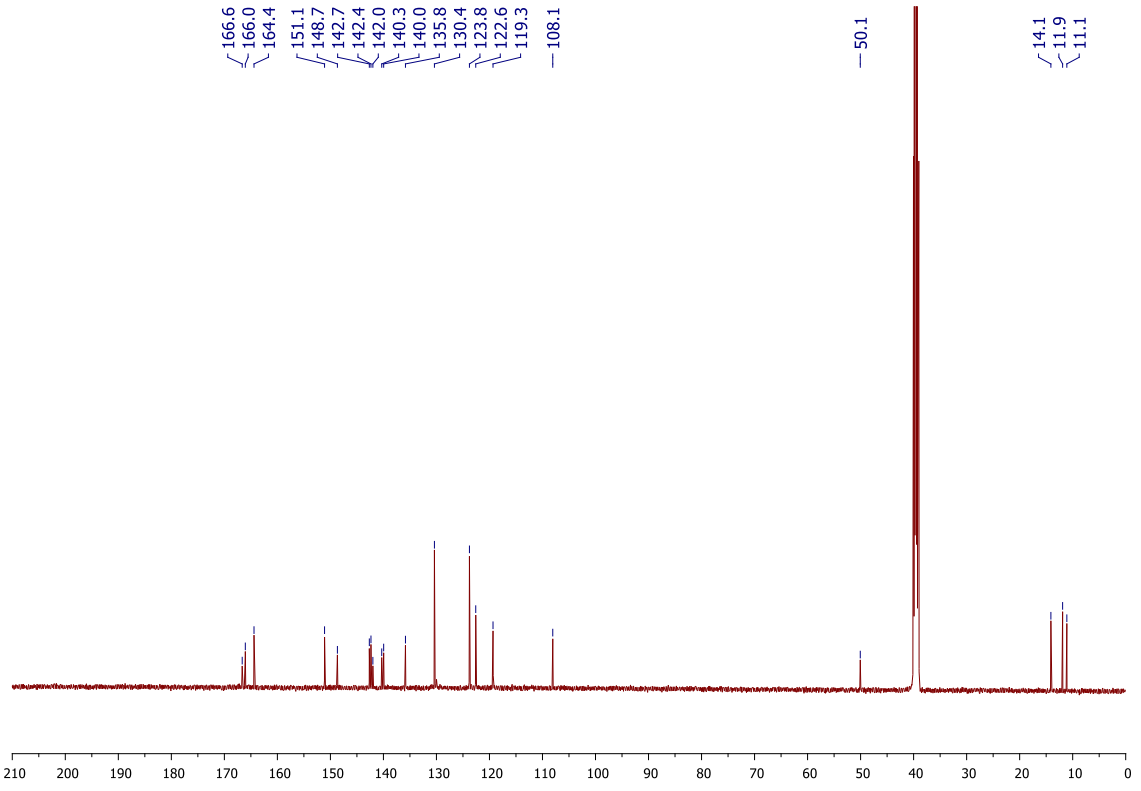
^{13}C-RMN (126 MHz, DMSO-d_6) **compound 2:**



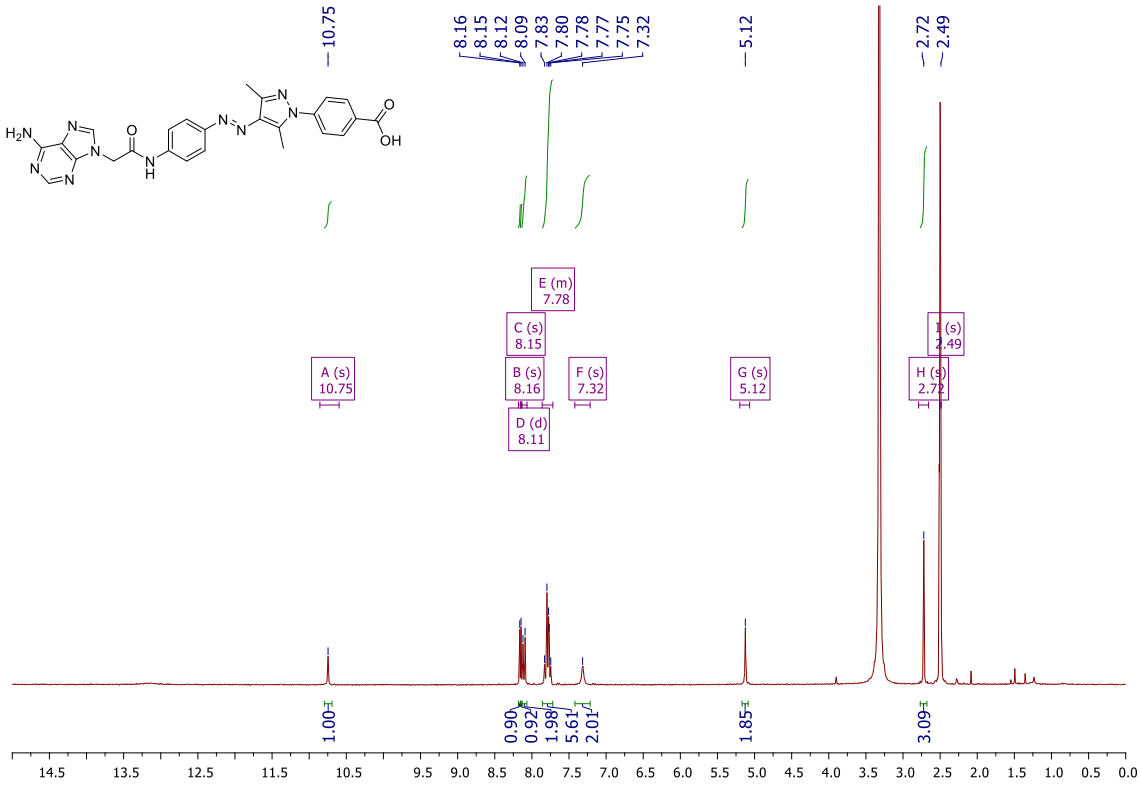
^{1}H-RMN (500 MHz, DMSO-d_6) **Azo-T:**



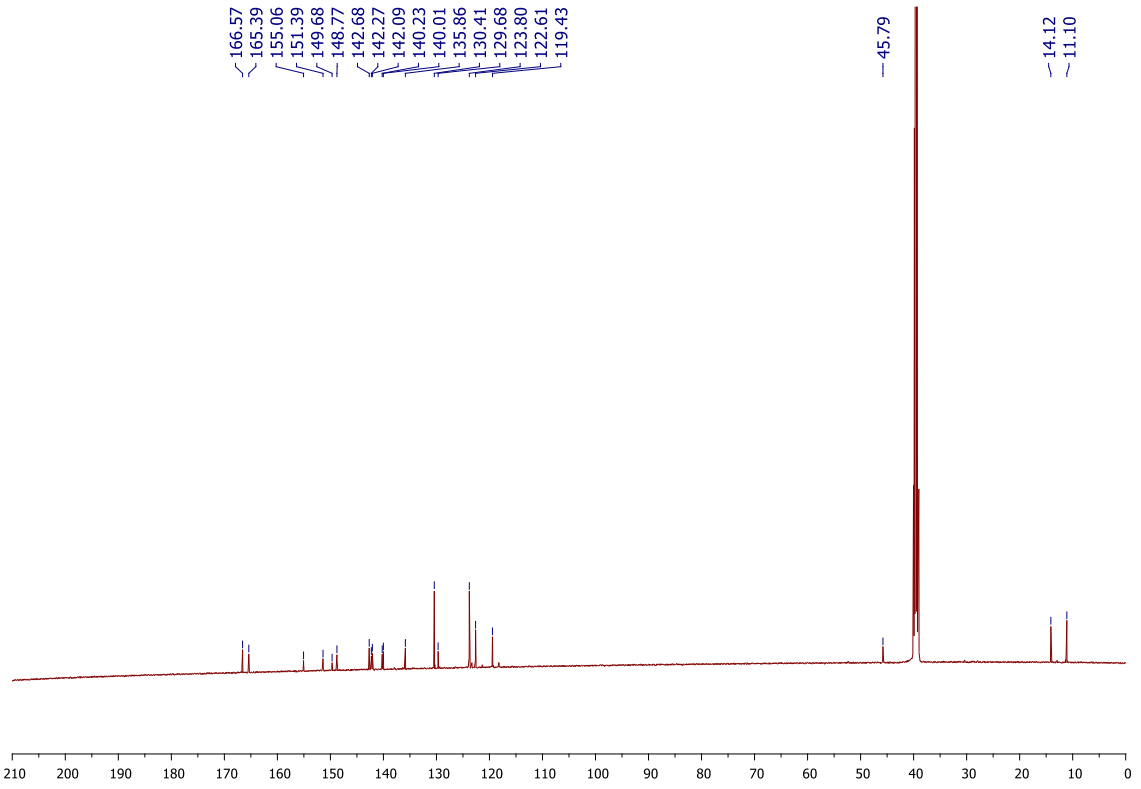
^{13}C-RMN (126 MHz, DMSO-d_6) **Azo-T:**



^{1}H-RMN (500 MHz, DMSO-d_6) **Azo-A:**



^{13}C-RMN (126 MHz, DMSO-d_6) **Azo-A:**



6. References

References for the synthetic procedures:

S1**.** S. Devi, M. Saraswat, S. Grewal and S. Venkataramani. *J. Org. Chem.* **2018***, 83*, 4307−4322.

S2. P. Kumar, A. Srivastava, C. Sah, Dr. S. Devi and Dr. S. Venkataramani. *Chem. Eur. J.* **2019**, *25*, 11924–11932.

S3. M. E. Mercurio, S. Tomassi, M. Gaglione, R. Russo, A. Chambery, S. Lama, P. Stiuso, S. Cosconati, E. Novellino, S. Di Maro and Anna Messere. *J. Org. Chem*. **2016**, *81*, 11612−11625.

References for the molecular dynamics simulations:

1 D. A. Case, T. E. Cheatham, T. Darden, H. Gohlke, R. Luo, K. M. Merz, A. Onufriev, C. Simmerling, B. Wang and R. J. Woods, *J Comput Chem*, 2005, **26**, 1668–1688.

2 J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman and D. A. Case, *J. Comput. Chem.*, 2004, **25**, 1157–1174.

3 Gaussian 16, Revision A.03, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, and D. J. Fox, Gaussian, Inc., Wallingford CT, 2016.

4 M. D. Hanwell, D. E. Curtis, D. C. Lonie, T. Vandermeersch, E. Zurek and G. R. Hutchison, *J. Cheminf.*, 2012, **4**, 1–17.

5 L. Settimo, K. Bellman and R. M. A. Knegtel, *Pharm. Res.*, 2014, **31**, 1082–1095.

6 J. Wang, W. Wang, P. A. Kollman and D. A. Case, *J. Chem. Inf. Comput. Sci.*, 2000, **222**.

7 A. Jakalian, B. L. Bush, D. B. Jack and C. I. Bayly, *J Comput Chem*, 2000, **21**, 132–146.

8 I. Ivani, P. D. Dans, A. Noy, A. Pérez, I. Faustino, A. Hospital, J. Walther, P. Andrio, R. Goñi, A. Balaceanu, G. Portella, F. Battistini, J. L. Gelpí, C. González, M. Vendruscolo, C. A. Laughton, S. A. Harris, D. A. Case and M. Orozco, *Nat Methods*, 2016, **13**, 55–58.

9 The PyMOL Molecular Graphics System, Version 2.5.4 Schrödinger, LLC.

10 S. Izadi, R. Anandakrishnan and A. V. Onufriev, *J Phys Chem Lett*, 2014, **5**, 3863–3871.

11 M. R. Machado and S. Pantano, *J Chem Theory Comput*, 2020, **16**, 1367–1372.

12 R. Salomon-Ferrer, A. W. Götz, D. Poole, S. Le Grand and R. C. Walker, *J Chem Theory Comput*, 2013, **9**, 3878–3888.

13 D. R. Roe and T. E. Cheatham, *J Chem Theory Comput*, 2013, **9**, 3084–3095.