Chemical Science

EDGE ARTICLE

View Article Online
View Journal | View Issue

Cite this: *Chem. Sci.*, 2018, **9**, 4671

Azobenzene as a photoregulator covalently attached to RNA: a quantum mechanics/molecular mechanics-surface hopping dynamics study†

Padmabati Mondal,*[a] Giovanni Granucci,[b] Dominique Rastädter,[a] Maurizio Persico*[b] and Irene Burghardt*[a]

Received 7th January 2018
Accepted 25th April 2018

DOI: 10.1039/c8sc00072g

rsc.li/chemical-science

The photoregulation of nucleic acids by azobenzene photoswitches has recently attracted considerable interest in the context of emerging biotechnological applications. To understand the mechanism of photoinduced isomerisation and conformational control in these complex biological environments, we employ a Quantum Mechanics/Molecular Mechanics (QM/MM) approach in conjunction with nonadiabatic Surface Hopping (SH) dynamics. Two representative RNA–azobenzene complexes are investigated, both of which contain the azobenzene chromophore covalently attached to an RNA double strand *via* a β-deoxyribose linker. Due to the pronounced constraints of the local RNA environment, it is found that *trans*-to-*cis* isomerization is slowed down to a time scale of ∼10–15 picoseconds, in contrast to 500 femtoseconds *in vacuo*, with a quantum yield reduced by a factor of two. By contrast, *cis*-to-*trans* isomerization remains in a sub-picosecond regime. A volume-conserving isomerization mechanism is found, similarly to the pedal-like mechanism previously identified for azobenzene in solution phase. Strikingly, the chiral RNA environment induces opposite right-handed and left-handed helicities of the ground-state *cis*-azobenzene chromophore in the two RNA–azobenzene complexes, along with an almost completely chirality conserving photochemical pathway for these helical enantiomers.

1 Introduction

The light-induced control of conformational changes in nucleic acids by introducing covalently or non-covalently bound photoswitches has proven a promising route towards DNA and RNA functionalization.[1–6] Applications range from photoregulation of gene expression[7] and functional aptamers,[8] to the tailored design of photoresponsive oligonucleotide walkers,[9] tweezers,[10] and diverse nanoarchitectures.

In this context, azobenzene photoswitches play a prominent role,[1–6,11,12] besides other photo-switchable chromophores like spiropyrans which have also been employed in DNA control.[13] The azobenzene chromophore switches reversibly between the planar *trans* form and the non-planar *cis* form, upon absorption of UV light inducing *trans*-to-*cis* conversion, or visible light inducing *cis*-to-*trans* conversion.[14–16] While the planar *trans* azobenzene isomer intercalates readily between aromatic nucleobases, the non-planar *cis* form, which is shorter by ∼3 Å, tends to unstack from a DNA or RNA helix.[17–20] As a result, preferential destabilization of nucleic acid double helices has been observed after *trans*-to-*cis* conversion.[5,6,17,18,21] In a paradigm system involving azobenzene covalently linked to DNA or RNA oligomers *via* a D-threoninol linker,[5,17] differences in melting temperatures (T_m) between *trans vs. cis* azobenzene substituted oligomers of the order of $\Delta T_m = T_m^{trans} - T_m^{cis} \sim 10$–20 °C were observed, which can be augmented by inserting multiple azobenzenes into the duplex structure.[22]

Due to the widely different time scales of azobenzene isomerization (femtoseconds to picoseconds)[11] as compared with the much slower DNA or RNA conformational response (microseconds to milliseconds),[23] complementary techniques are required to simulate the relevant steps. In particular, Molecular Dynamics (MD) simulations including an effective switching potential[19,24] were employed to model the fast isomerization step, while enhanced sampling techniques like Replica Exchange Molecular Dynamics (REMD) were used to explore longer time scales where the effects of DNA and RNA unfolding become manifest.[20] Markov State Models (MSM) are the method of choice to provide an embedding of MD and REMD type information into a kinetic description adapted to long-time conformational changes.[25]

Here, we are concerned with the shortest, ultrafast time scale, in order to investigate whether azobenzene works as an efficient photoswitch in the highly specific nucleic acid

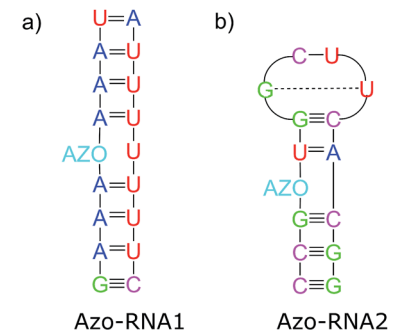
[a]Institute of Physical and Theoretical Chemistry, Goethe University Frankfurt, Max-von-Laue-Str. 7, 60438 Frankfurt, Germany. E-mail: padmabati.mondal@gmail.com; burghardt@chemie.uni-frankfurt.de

[b]Dipartimento di Chimica e Chimica Industriale, Università di Pisa, v. Moruzzi 13, I-56124 Pisa, Italy. E-mail: maurizio.persico@unipi.it

† Electronic supplementary information (ESI) available. See DOI: 10.1039/c8sc00072g

environment, and whether azobenzene isomerization proceeds similarly to the gas phase or solution phase. *A priori*, the DNA or RNA local environment could entail significant modifications of the mechanism and time scales. To this end, we use hybrid Quantum Mechanics/Molecular Mechanics (QM/MM) simulations[26–31] in conjunction with Surface Hopping (SH) dynamics[27,32,33] in order to obtain a detailed picture of the isomerization process in azobenzene–RNA complexes on a picosecond time scale. In a related approach, QM/MM–SH simulations were recently used to study peptide folding and unfolding driven by azobenzene photoisomerisation.[33] The simulations reported in the present work represent an important benchmark for more approximate Molecular Mechanics (MM) based treatments.[19,20,34,35]

Connecting to our recent work[20,34,36] in the context of RNA regulation, we focus on azobenzene-linker–RNA complexes – rather than complexes with DNA – in the present study. To reduce the bias associated with the study of a single system, we carried out parallel investigations of two azobenzene–RNA complexes as depicted in Fig. 1, both of which contain the azobenzene chromophore anchored to one of the RNA strands *via* a β-deoxyribose linker.[34,36] This linker-azobenzene combination has recently been shown to induce a destabilization effect comparable to the widely applied D-threoninol-linked azobenzene,[36,37] with excellent efficiencies and high thermal stability. In addition, the β-deoxyribose linker has been shown to impose minimal disturbance on the duplex stability.[34,36,37]



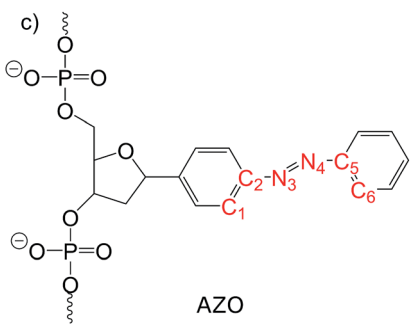


Fig. 1 Structure of the azobenzene-linker–RNA systems under study. The azobenzene chromophore is covalently bound to the RNA backbone *via* a β-deoxyribose linker. (a) Duplex structure denoted Azo–RNA1. (b) Hairpin structure, capped with a tetraloop, denoted Azo–RNA2. (c) Detailed structure of linker-azobenzene unit covalently bound to the RNA backbone.

The two RNAs that were selected correspond to (i) a double-helix structure representing a fragment of a longer double helix that was synthesized and studied both experimentally and theoretically,[34,36] and (ii) a 14-mer hairpin structure capped by a tetraloop motif – a frequently occurring RNA motif – which was studied in ref. 20 using REMD simulations. We refer to ref. 1 and 34 for a detailed discussion of the various factors that influence the efficiency of azobenzene photoswitches for nucleic acid functionalization. For the two types of experimentally established linkers mentioned above, *i.e.*, of deoxyribose and threoninol type, different anomeric and enantiomeric forms (*i.e.*, α-, β- and D-, L-, for the deoxyribose and threoninol linkers, respectively) are of critical importance, along with the presence or absence of an opposite base in the duplex structure.

2 Methods

Given that the electronic excitation remains localized on the azobenzene chromophore, a QM/MM treatment is a natural approach to the electronic structure of the azobenzene-linker–RNA complexes. As illustrated in Fig. 2, the QM part corresponds to the azobenzene chromophore, while the MM part combines the β-deoxyribose linker, the RNA duplex, and the surrounding water.

For the QM part, we employ a semi-empirical Configuration Interaction (CI) method based upon Floating Occupation Molecular Orbitals (FOMO–CI),[38] which has proven successful in the past in the study of azobenzene derivatives.[32] Calculations were performed with a development version of the MOPAC program, using a reparametrized AM1 Hamiltonian and additional features as detailed in ref. 38–40.

The linker and the RNA duplex are described by the Amber ff99SB[41] force field as implemented in the TINKER 6.1 package,[42] and the water molecules are of TIP3P type.[43] The connection atom scheme[44] is employed to describe the boundary between the covalently linked QM and MM parts. According to this scheme, the C1′ atom of the β-deoxyribose linker is the connection atom which behaves as a hydrogen atom in the QM part and as a normal carbon atom in the MM part (see Fig. 2). The total energy of the system is of additive type

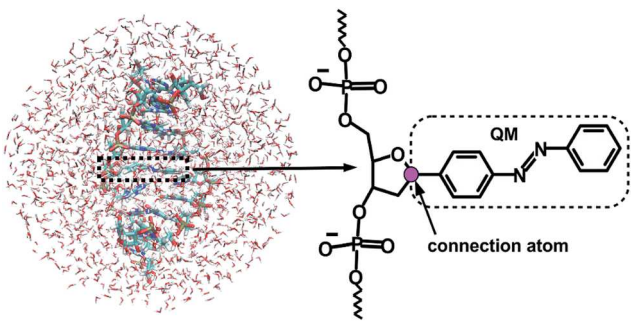


Fig. 2 Representation of the QM/MM set-up for the azobenzene-linker–RNA system in water (l.h.s.), with a detailed representation of the linker-azobenzene moiety, where the QM/MM connection atom is shown in magenta (r.h.s.).

and the electrostatic embedding scheme is considered for the calculation of the effective QM/MM Hamiltonian.[44]

The QM/MM scheme is combined with an approximate representation of the nonadiabatic dynamics using Tully's fewest switches surface hopping algorithm[45] including overlap based quantum decoherence corrections (SH–ODC).[32,46] Energy gradients and nonadiabatic couplings are calculated using the FOMO–CI method as a function of the time-evolving trajectories. Trajectories are initiated in the $S_2(\pi–\pi^*)$ state, whose squared transition dipole moment averaged over the ensemble is dominant (see ESI Section S1†). In fact, in the photostationary state a much higher fraction of the *cis* isomer is formed when exciting to the $\pi–\pi^*$ band than to the $n–\pi^*$ band, so this is the standard experimental situation we aim to simulate. An ensemble of 100 initial conditions is considered, which are sampled from equilibrated ground-state MD simulations for the relevant complexes. Following MD pre-equilibration, a QM/MM equilibration was carried out using Brownian dynamics in the electronic ground state, with an equilibration time of 50 ps at 300 K. To avoid nonequilibrium effects due to heating at the beginning of the QM/MM equilibration, the first 10 ps are excluded from the 50 ps equilibration data for generating initial conditions. During the simulation, the populations of the first six singlet states are monitored.

The QM/MM–FOMO–CI–SH approach has been previously used in various simulations of the photodynamics of azobenzene and its derivatives, often in more or less severely sterically hindered environments.[32,47–51] Excited state lifetimes obtained by transient spectroscopies were successfully reproduced.[49] Both experiments and simulations highlight the effect of caging, which can be extremely pronounced as in closely packed self-assembled monolayers[48] or much milder as in a viscous solvent (ethylene glycol) that hinders overall molecular rotation and, to a lesser extent, N=N double bond torsion.[49] Quantum yields, when experimentally available, are well reproduced for $n–\pi^*$ excitation[49] but tend to be overestimated for $\pi–\pi^*$ excitation.[32] The FOMO–CI method was calibrated specifically to reproduce energetic data for the ground and the first excited states of azobenzene, obtained experimentally and by calculations of the best affordable quality, performed by different authors.[39] As a result, the quality of the potential energy surfaces we are using is superior to what might be achieved by any *ab initio* approach of practical use in simulations of this kind. Furthermore, our surface hopping results compare well with quantum wavepacket calculations for azobenzene photoisomerization.[46]

In the present context, we further carried out selected benchmark calculations using the *ab initio* second-order Complete Active Space Perturbation Theory (CASPT2)[52] method for small fragments involving the azobenzene chromophore and the most important neighboring bases (see Section S3 of the ESI†). These calculations validate the QM/MM–FOMO–CI treatment of azobenzene embedded in the local RNA environment.

For completeness, we point out that triplet states were not included in the present calculations. Even though the azobenzene excited-state lifetimes will turn out to be considerably lengthened as compared with the gas-phase dynamics, intersystem crossing likely does not play an important role, given that the S_1 and T_1 potential energy surfaces (PESs) are roughly parallel and well separated in energy.[53] (Indeed, the latter reference addresses the role of the T_1 state in the thermally activated kinetics, not the photochemical pathways).

The ESI† provides further information regarding the simulation set-up, construction of initial conditions, and absorption spectra obtained within the QM/MM set-up (see ESI Section S1 and S2†).

3 Results and discussion

In the following, we first analyze the time-dependent electronic state populations and estimated isomerization quantum yields resulting from the QM/MM–SH simulations. Then, we turn to the isomerization mechanism, with emphasis on the role of the RNA environment, the subtle role of chirality, and the interplay of different dihedral angles whose combined dynamics results in a largely volume-preserving photoreactive path.

3.1 Photodynamical time scales and quantum yield

Fig. 3 illustrates one of the key results of this study: *trans*-to-*cis* isomerization of azobenzene is slowed down remarkably in the RNA environment, from 500 femtoseconds in the gas phase to ~10–15 picoseconds in RNA. By contrast, the time scale of *cis*-to-*trans* isomerization is almost unaffected and remains less than one picosecond. These findings are very similar for the two RNAs that we investigated. This suggests that the geometric changes accompanying *trans*-to-*cis* isomerization are strongly hindered in the RNA environment, for the present azobenzene-linker–RNA combination – pointing towards an important role of stacking interactions that stabilize the intercalated *trans* form of the chromophore.

In further detail, it is seen from Fig. 3a and b for the *trans*-to-*cis* dynamics that the decay from the initially excited $S_2(\pi–\pi^*)$ state to the $S_1(n–\pi^*)$ state happens within ~1 picosecond, while the chromophore subsequently remains in the S_1 state for ~10 picoseconds. From the build-up of S_0 population, we can infer that most trajectories return to the S_0/*trans* conformation, while a lesser portion isomerizes and yields the S_0/*cis* product.

For the trajectories starting in *cis* conformation (Fig. 3c and d), a very different scenario is observed: while the S_2 state again decays very rapidly, within ~1 picosecond, the S_1 state is only transiently populated – or barely populated at all – while the ground state (S_0) population rises within ~1 picosecond, concomitantly with the S_2 decay. This indicates that S_1 is at most a transient intermediate rather than a long-lived intermediate as in the *trans*-to-*cis* dynamics.

As also shown in Fig. 3, the population decay derived from the trajectory ensemble can be fitted to an approximate set of coupled first-order kinetic equations for the state populations P_n, $n = 0, 1, 2$, for a three-state system comprising the S_n, $n = 0, 1, 2$, electronic states (see also ref. 32),

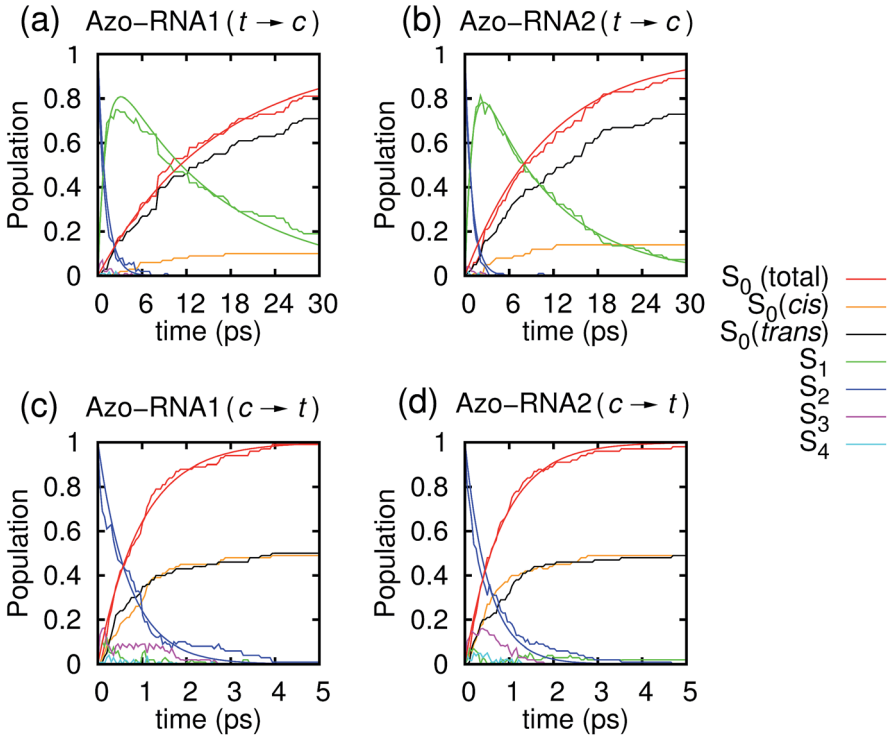


Fig. 3 Population dynamics for the first five singlet states S_n, $n = 0, \ldots, 4$, starting from (a and b) *trans* and (c and d) *cis* azobenzene for Azo–RNA1 (left panels) and Azo–RNA2 (right panels). Note the different time scales shown for *trans*-to-*cis* (30 ps) and *cis*-to-*trans* (5 ps) isomerisation.

$$P_2 = \exp(-t/\tau_2)$$
$$P_1 = \frac{\tau_1}{\tau_1 - \tau_2}(\exp(-t/\tau_1) - \exp(-t/\tau_2)) \quad (1)$$
$$P_0 = 1 - P_1 - P_2 \simeq 1 - \exp(-t/\tau_0)$$

where τ_2 and τ_1 are the lifetimes of the S_2 and S_1 states, and τ_0 describes the concomitant rise of the S_0 state (see Table 1). Here, τ_0 is treated as an independent decay parameter, and the single-exponential form is found to give good agreement with the relation $P_0 = 1 - P_1 - P_2$ for the data at hand, given that τ_1 and τ_2 typically differ by a factor of ten. In the case of *cis*-to-*trans* isomerisation, the fitting procedure for S_1 did not work well due to the small state population, such that only τ_2 and τ_0 were determined.

Table 1 Comparison of S_2 and S_1 lifetimes for *trans*-to-*cis* and *cis*-to-*trans* isomerisation *in vacuo* as compared with various solvents and the RNA environment. The values *in vacuo* and in solution are taken from ref. 32. In addition, the time scale of the growth of the S_0 population (τ_0) is given

Isomerisation	Environment	τ_1 (ps)	τ_2 (ps)	τ_0 (ps)
trans → *cis*	*In vacuo*	0.525	0.191	
	Methanol	0.643	0.145	
	Ethylene glycol	0.534	0.091	
	Azo–RNA1	14.590	1.120	16.10
	Azo–RNA2	10.233	0.952	11.30
cis → *trans*	*In vacuo*	0.085	0.446	
	Methanol	0.071	0.381	
	Ethylene glycol	0.082	0.352	
	Azo–RNA1	—	0.702	0.97
	Azo–RNA2	—	0.527	0.84

The nature of the population dynamics is very similar for the Azo–RNA1 and Azo–RNA2 complexes, even though the S_1 and S_2 lifetimes depend slightly on the environment. Since an analysis with the same FOMO–CI–SH set-up was previously carried out for the azobenzene chromophore in gas phase and in various solvents,[32] we compare in Table 1 the results for the relevant lifetimes τ_n, $n = 0, 1, 2$. This comparative analysis underscores that (i) S_1 lifetimes in the *trans*-to-*cis* isomerisation dynamics in an RNA environment differ by a factor of about 20 from previously observed lifetimes in the gas phase and solution phase, and (ii) in the case of *cis*-to-*trans* dynamics, S_2 lifetimes are slightly augmented, by a factor of two.

The quantum yield for *trans*-to-*cis* isomerisation ($\Phi_{t\to c} \sim 0.1$) in RNA drops to about half of typical values computed for non-viscous solvents such as methanol. For the *cis*-to-*trans* conversion the quantum yield remains high ($\Phi_{c\to t} \sim 0.5$) and its decrease with respect to methanol solution is modest. The quantum yields for azobenzene isomerisation in the gas phase, in solution, and in the RNA environment are given in Table 2 for comparison. As in Table 1, all results are obtained by the same simulation method.

These results are in qualitative agreement with recent experimental[55,56] and MD-based computational[35] studies on the photoisomerisation quantum yield of azobenzene in DNA where the quantum yield for the *trans*-to-*cis* conversion was found to depend largely on the local DNA sequence[35] as well as on temperature,[56] and strongly decreases with decreasing local

Table 2 Comparison of quantum yields for *trans*-to-*cis* ($t \rightarrow c$) and *cis*-to-*trans* ($c \rightarrow t$) isomerisation *in vacuo*, in solution, and in the RNA environment. The values *in vacuo* and in solution are taken from ref. 32; the experimental results are taken from ref. 54, referring to irradiation at 334 nm

Environment	$\Phi_{t\rightarrow c}$	$\Phi_{c\rightarrow t}$
In vacuo	0.21	0.54
Methanol	0.24	0.57
Ethylene glycol	0.24	0.67
Azo–RNA1	0.10	0.50
Azo–RNA2	0.14	0.52
Experiment (methanol)	0.155	0.39

volume in the nucleic acid environment. Specifically, the experimentally determined quantum yields reported in ref. 56 vary in the range $0.01 < \Phi_{t\rightarrow c} < 0.05$ for various azobenzene-substituted DNA double strands with a threoninol linker, below the melting temperature T_m. For temperatures above T_m, quantum yields for the different sequences converge towards a temperature-dependent value in the range $0.05 < \Phi_{t\rightarrow c} < 0.1$, without any sequence specificity. Our QM-based computational study of azobenzene photoisomerisation in an RNA environment points towards a "looser" local environment of azobenzene in RNA, leading to slightly higher quantum yields ($\Phi_{t\rightarrow c} \sim 0.1$) at room temperature.

The striking lengthening of the S_1 lifetime in the *trans*-to-*cis* photodynamics and the decrease in the $\Phi_{t\rightarrow c}$ quantum yield must clearly be due to steric constraints of the RNA (or DNA) environment that are much less effective in the *cis*-to-*trans* case. A slowing down of the azobenzene *trans*-to-*cis* photoisomerization and excited state decay in viscous solvents and constrained environments has been previously reported,[32,47,49–51,57] even though the effect was by far not as pronounced and typically limited to an increase of the isomerization time by a factor of two to three. Noticeably, though, a full supression of isomerisation,[48] mainly due to steric effects,[58] has been found for regularly packed self-assembled monolayers of *trans*-azobiphenyls, with S_1 lifetimes of the same order of magnitude as in the present work. In all cases the effect on the *cis*-to-*trans* photodynamics was much less important.

The reason why the interactions with the environment affect the *trans*-to-*cis* photodynamics much more than the *cis*-to-*trans* dynamics is twofold: first, the *trans* azobenzene isomer preferentially undergoes stacking interactions, which are especially pronounced in the nucleic acid environment. Second, the asymmetry of the PESs comes into play. In particular, the S_1 potential energy curve as a function of the CNNC torsional angle is much steeper on the *cis* side than on the *trans* side (see Fig. 4), according to various theoretical calculations.[39,57,59,60] This is in agreement with the S_1 lifetimes, that are much shorter for the *cis* than for the *trans* isomer, even in non-viscous solvents.[61,62] The S_1 Franck–Condon point for the *cis* isomer is about 0.7 eV higher than for the *trans* isomer[39,60] and the minimum energy pathway in S_1 from the *trans* conformation to the torsional conical intersection is practically flat, perhaps with a very small barrier.[39,60,63] Moreover, the S_2/S_1 conical intersection on the *trans* side is very close to the Franck–Condon point[39,63] whereas in the case of the *cis*-to-*trans* photochemical pathway the S_2/S_1 and S_1/S_0 conical intersection seams[64] are in immediate vicinity, permitting ultrafast $S_2 \rightarrow S_1 \rightarrow S_0$ conversions. Therefore, the driving force that allows the N=N double bond to twist even in the presence of steric hindrance is much stronger when starting from the *cis* conformation than from the *trans* conformation. Conversely, the shallow PES topology of the *trans* isomer enhances the susceptibility of the *trans* isomer to steric effects.

3.2 Photochemical pathways

To summarize our observations, Fig. 4 shows a schematic picture of the photodynamics of *trans* and *cis* azobenzene in RNA, as obtained from the ensemble of propagated QM/MM–SH trajectories. The effective reaction coordinate is a complex combination of local modes, especially involving a concerted motion of neighboring dihedrals, as further discussed below. In the following, we briefly recapitulate the photochemical pathways delineated in Fig. 4. Detailed information, *e.g.*, regarding the energetic distribution of the trajectory ensemble at the hopping geometries, is provided in the ESI (Section S4†).

After the initial excitation from the *trans* ground state to the S_2 state (with an excitation energy of 3.62 eV, averaged over the trajectory ensemble), the close-lying S_3, S_4 and S_5 states are temporarily populated to some extent by nonadiabatic transitions. Within less than a picosecond (*i.e.*, 0.7–1 ps), the population is transferred from S_2 and the partially populated upper states to the S_1 state, at geometries that remain close to the Franck–Condon region. A majority of trajectories are then found to stay on the S_1 surface for an extended time interval, presumably due to existence of a local minimum and shallow barrier preceding the S_1/S_0 conical intersection at CNNC = 90°. This PES topology, which is modified as compared with the gas phase, is most likely created by the stacking of *trans* azobenzene with the upper and lower bases, hindering *trans*-to-*cis* photoisomerisation (see the discussion below). At this point, a subset of reactive trajectories have high enough energy to proceed to the S_1/S_0 conical intersection. A trajectory is considered to reach the conical intersection if it meets a threshold criterion for the energy difference between two states, *i.e.*, $\Delta E < 0.1$ eV. The remaining trajectories are unreactive and return to S_0, mainly by hopping from the transoid region, where the S_1/S_0 energy difference is still large. In both reactive and unreactive cases, the dynamics tends to be slow, of the order of 10 picoseconds. Overall, almost 80% of the population is transferred back to the *trans* ground state *via* unreactive trajectories. The complementary subset of trajectories which reach the conical intersection are eventually equidistributed between the *trans* and *cis* ground-state geometries, rendering the overall quantum yield small, with $\Phi_{t\rightarrow c} \simeq 0.1$.

By contrast, starting in the S_2 state of the *cis* isomer (at an excitation energy of 4.1 eV), trajectories directly reach the three-state conical intersection region between S_2, S_1, and S_0, within <1 ps, promoted by the steep S_2 surface. No participation of the S_3, S_4 and S_5 states is noticed when starting from the *cis* isomer.

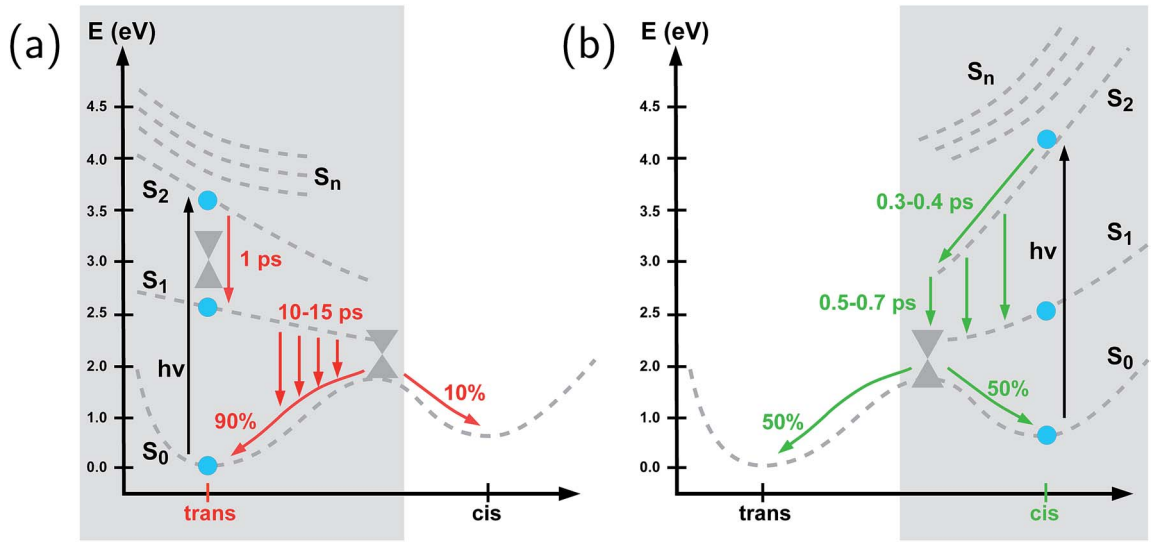


Fig. 4 Schematic representation of the photochemistry of the azobenzene-linker–RNA complexes as deduced from the present study, (a) for the *trans*-to-*cis* pathway (shown in red) which exhibits a remarkable slowing-down as compared to the chromophore *in vacuo* and (b) for the *cis*-to-*trans* pathway (shown in green) whose kinetics remains comparable to the situation *in vacuo*. The S_0 ground state, $S_1(n–\pi^*)$ state, and $S_2(\pi–\pi^*)$ state are shown explicitly (with calculated Franck–Condon energies indicated as solid blue circles), while the higher-lying singlet states (S_n, $n = 3, 4, 5$) are indicated schematically at the Franck–Condon point. The effective reaction coordinate is a complex combination of internal modes, especially involving a concerted motion of the dihedrals. Conical intersections which interconnect all relevant states are also indicated schematically.

At the conical intersection, population is transferred either directly to the S_0 state or *via* S_1 to S_0 almost immediately, such that the S_1 lifetime of the *cis* isomer could not be determined. Since the total population decays through the conical intersection which provides access to the *trans* and *cis* channels in a symmetric fashion, the final ground-state populations are found to be equally distributed, such that the *cis*-to-*trans* photoisomerisation quantum yield is 50%.

3.3 Isomerisation mechanism

To gain some insight into the effect of the local RNA environment on the isomerisation process, Fig. 5 shows selected snapshots from a reactive *trans*-to-*cis* trajectory in the Azo–RNA2 system. This trajectory is among the few realizations that isomerize rapidly, but similar scenarios have been observed for other reactive and unreactive trajectories. The figure illustrates the time-evolving center-of-mass distance between the azobenzene rings and the upper and lower base pairs. Despite the isomerisation event (at ~1.5 ps), the relative distances evolve smoothly and stacking interactions are preserved to a large extent. This underscores the important role of stacking, and also the flexibility of the chromophore's non-covalent binding interactions within its local RNA environment. As further discussed in the ESI,† the isomerization event indeed does not exert a major disrupting effect on the interaction with the nearest neighbouring RNA bases (see ESI Section S5†).

Fig. 5 further suggests that the isomerization event follows a volume-conserving mechanism, to exert as little perturbation on the nearest-neighbor environment as possible. This implies that a cooperative dynamics must be involved, similarly to the related pedal-like mechanism that was found in solution phase.[32,57] Fig. S9† shows that substantial changes in the CNN and NNC angles occur very soon after excitation, because in the $S_1(n–\pi^*)$ state, the equilibrium CNN/NNC value is larger than in the ground state.[39,49,63] Even in the gas phase, an inversion-assisted rotation mechanism of the photoisomerisation is observed.[63] To better understand the details of the mechanism, we now analyze the time evolution of several dihedrals – *i.e.*, the CNNC (C2–N3–N4–C5), CCNN (C1–C2–N3–N4), NNCC (N3–N4–C5–C6), and CCCC (C1–C2–C5–C6) dihedral angles – as illustrated in Fig. 6 for two selected trajectories (see Fig. 1c and ESI Section S6† for the atom labels and definition of the relevant dihedral angles).

Fig. 6a illustrates a reactive *trans*-to-*cis* trajectory in the Azo–RNA1 complex, where the sudden change in the CNNC dihedral (red trace) is seen to be accompanied by an almost equal change in the NNCC dihedral (black trace). Indeed, the absolute value of the NNCC dihedral changes more – from $\approx 180°$ to $\approx 0°$ – than the CCNN dihedral (cyan trace), which changes from $\approx 0°$ to $\approx 45°$ *via* $\approx 80°$ while going from structure 1 to 3 *via* 2. Note that structure 2 corresponds to a geometry near the conical intersection, preceding the hopping event. From the evolution of the dihedrals, we infer that the N3–N4 twist is immediately followed by a rotation around the N4–C5 single bond. The distant phenyl ring (as seen from the linker position) undergoes partial in-plane rotation around the N4–C5 bond. A very similar scenario is observed for the reactive *cis*-to-*trans* trajectory shown in Fig. 6b. The CNNC and NNCC evolution is again found to be closely correlated, even though the NNCC dihedral appears to be delayed until the hopping event occurs. In a complementary

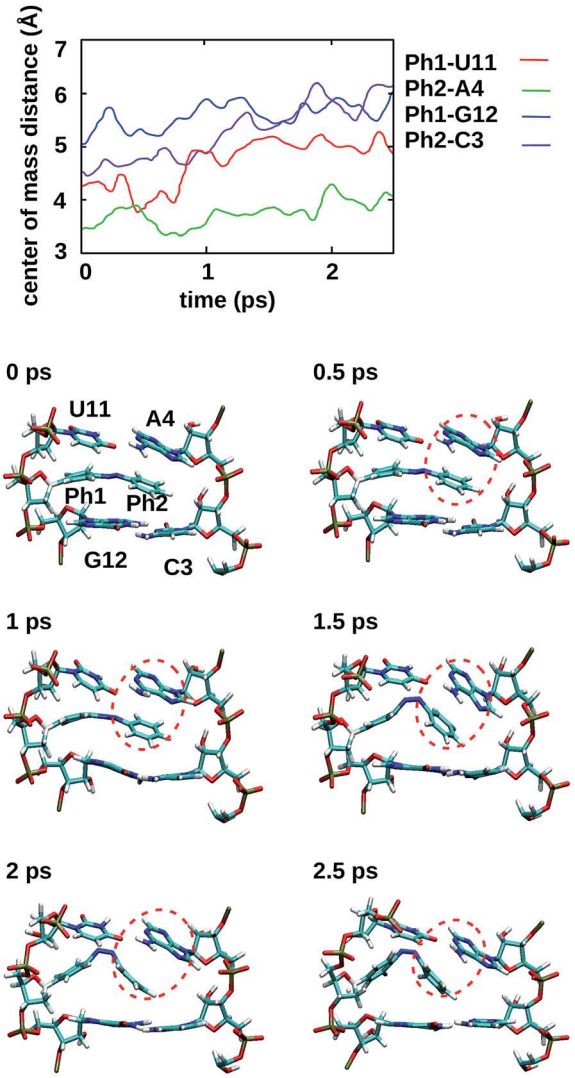


Fig. 5 For a representative trajectory taken from an Azo–RNA2 simulation, the center-of-mass distance between azobenzene and the upper and lower base pairs is shown, along with snapshots of the relevant structures for every 500 fs during the QM/MM–SH simulation. The azobenzene phenyl rings (denoted Ph1 and Ph2) are intercalated between the upper and lower bases. The snapshots show a slight decrease of stacking interactions after the switching event at ~1.5 ps, but stacking is preserved to a significant extent. In particular, Ph2-A4 stacking is highlighted (red circles), which persists even during and after the isomerisation event.

fashion to Fig. 6 and S8 of the ESI† shows the signed values of the relevant dihedrals.

Similar observations were made for the Azo–RNA2 complex, even though some details for the individual trajectories differ (see the ESI Section S7.2† for a detailed discussion). In all cases, a concerted mechanism involving several dihedrals is observed, which is consistent with the abovementioned pedal-like mechanism that was found for the azobenzene chromophore in solution.[32,57] As can be inferred from Fig. 6 and is further explained below, this mechanism is coupled to a collective rearrangement of the local RNA environment.

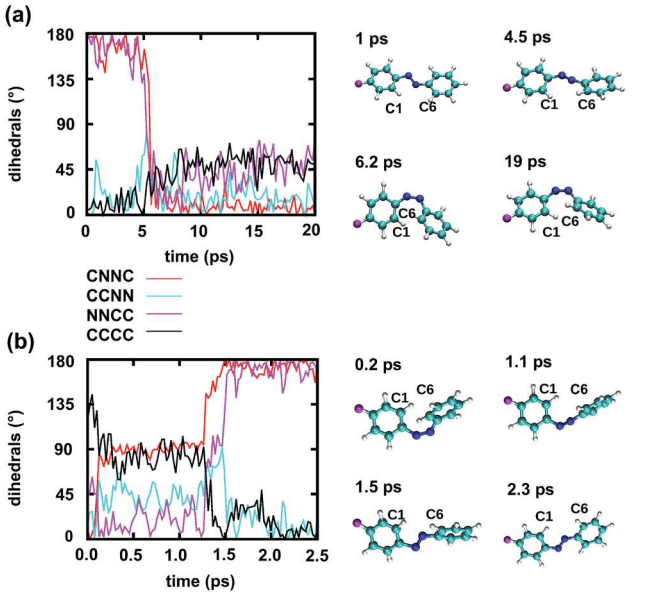


Fig. 6 Time dependence of different dihedrals for a reactive trajectory of the Azo–RNA1 complex, starting from (a) *trans* azobenzene and (b) *cis* azobenzene, along with snapshots of the relevant azobenzene structures. Note that the CNNC and NNCC dihedrals are strongly correlated during the isomerisation event.

3.4 Static *vs.* dynamic effects of RNA environment

In view of the pronounced effect of the RNA environment on the *trans*-to-*cis* isomerization kinetics, the question arises whether the local environment mainly exerts a static or dynamic influence. Therefore, we investigated the equilibrium (static) effects of the RNA environment on the shape of the effective torsional potential energy surface by Boltzmann inversion,[65] based on the CNNC dihedral angle distribution from an ensemble of ground state QM/MM equilibration trajectories for azobenzene in different environments (*in vacuo*, in ethylene glycol and in RNA). As detailed in the ESI (Section S5.3†), the analysis reveals that the shape of the ground state potential energy surface around the *cis* and *trans* equilibrium geometry in RNA is very similar to the corresponding potentials in solution and *in vacuo*. Small shifts in the equilibrium values of the CNNC dihedral are observed – about 5° for *trans*-azobenzene, and even less for *cis*-azobenzene – in opposite directions for Azo–RNA1 and Azo–RNA2, which can be related to the effect of the chiral environments (see the discussion below). Perhaps contrary to expectation, the effective torsional potentials in RNA are not steeper – but rather slightly broader – than in solution or *in vacuo*, showing no evidence of a constraining environment. Therefore, we conclude that the slowing down of the N=N double bond twisting and of the S_1 decay are due to dynamic caging, *i.e.*, to the relatively long time needed for the RNA bases to collectively relax geometrically in order to match the twisting of *trans*-azobenzene. In this sense, steric hindrance in the context of *trans*-to-*cis* isomerization should be thought of as a dynamic effect.

The comparatively slow time scale of collective reorientation can also be inferred from the time evolution of the distance between the azobenzene chromophore and the upper and lower

base pairs, see the ESI (Section S5.1†). Here, we analyse the fraction of time spent by the photochemical trajectories in configurations below the sum of van der Waals radii, using a model of interlocking spheres. In line with Fig. 5, the distal phenyl ring (Ph2) is found to be especially close to the upper bases, and spends about 20% of the simulation time at distances below the sum of van der Waals radii (and even up to 70% in the electronic ground state). On average, the distance steadily increases during the simulation time (see Fig. S4†), both for unreactive and reactive trajectories. As also suggested by Fig. 5, the reorientation of the local environment is gradual and leads to an increase in local RNA fluctuations (see Section S5.2†) due to the interaction with the vibronically hot chromophore.

3.5 Chiral selectivity

Finally, we address the role of chirality in the photoisomerization event, which is of particular interest in the present context since the RNA environment as such exhibits right-handed helical chirality. With regard to the chromophore, the *trans* azobenzene isomer is achiral, while the *cis* isomer exhibits helical chirality and exists in a P (plus, *i.e.*, right-handed) helical form and an M (minus, *i.e.*, left-handed) helical form (see ESI Section S6† for details). Likewise, the S_1/S_0 conical intersection exhibits symmetric P and M type precursor structures of the two *cis* isomers.[32,66,67] For the *trans*-to-*cis* photochemical pathway, it is therefore of interest to know whether the P or M type *cis* form is preferentially generated in the reactive trajectories. Conversely, for the *cis*-to-*trans* pathway, we will analyze whether a chirality preserving or a chirality inverting pathway is adopted. Even in an achiral environment, it was found that an asymmetry can exist between these pathways.[32,66,67] This asymmetry could be enhanced in the chiral RNA–linker environment.

Interestingly, a marked preference for M *vs.* P helicity appears in the *cis* species of the Azo–RNA1 *vs.* Azo–RNA2 complexes as a result of the initial equilibration procedure. That is, it turns out that the initial helicity of *cis*-azobenzene is of M type for Azo–RNA1 and of P type for Azo–RNA2. The reason for this intrinsic difference seems to lie in the preference of *cis* azobenzene to reside either in the major groove, as is the case for Azo–RNA1, or in the minor groove, as is the case for Azo–RNA2. In earlier MD studies of the Azo–RNA2 complex with a D-threoninol linker,[20] we also found a minor groove preference of *cis* azobenzene.

In the following analysis, we will make use of the fact that the algebraic sign of the CNNC dihedral during the isomerization process informs about the P and M helicity, *i.e.*, a positive sign correlates with P helicity while a negative sign correlates with M helicity.[66,68] A complementary analysis can be given in terms of the change of the NNCC and CCNN dihedrals,[32] which is especially useful when analyzing the initial chirality preserving or chirality inverting property of the isomerisation pathway (see ESI Section S8†): The pathway is chirality conserving if the CNNC twist is accompanied by a conrotatory torsion of the N–C bonds such that the CCNN and NNCC groups tend to become planar. By contrast, the pathway is chirality inverting if the initial conrotatory torsion of the N–C bonds brings the CCNN and NNCC dihedrals towards 90° and beyond. (Note that disrotatory phenyl torsions cannot be classified in this way).

Focusing on the distribution and sign of the CNNC dihedral, Fig. 7 shows the CNNC dihedral at the S_1/S_0 hopping point in the range [−180 : 180], for the trajectories starting from the *trans* (upper panels) and *cis* (lower panels) isomer, both for reactive (red) and unreactive (green) trajectories.

For the *trans* initial conditions (Fig. 7a and b), we focus upon the reactive trajectories (red bars) whose hopping geometries fall into a narrow range around ±90°, where the conical intersection is located. By contrast, for the unreactive trajectories (green bars), the CNNC dihedral at the S_1/S_0 hopping points is spread over a wide range, *i.e.*, −100° to −180° and 60° to 180°, consistent with the observation that most unreactive trajectories hop far before reaching the conical intersection, and consequently go back to the *trans* ground state, resulting in a small photoisomerisation quantum yield. As can be seen from Fig. 7a, there is no clear preference for the P-type or M-type *cis* photoproducts in Azo–RNA1, while a certain preference for the M-type *cis* photoproduct is found for the Azo–RNA2 complex. However, no definite conclusion can be drawn from the limited set of reactive trajectories. This is confirmed by a complementary analysis based upon the CCNN and NNCC dihedrals, as reported in the ESI (Section S8†).

Conversely, a pronounced chiral selectivity is observed when considering the *cis*-to-*trans* photoisomerisation dynamics, as can be inferred from Fig. 7c and d. Starting from the equilibrated M *vs.* P forms of the *cis* isomer of Azo–RNA1 and Azo–RNA2, it is seen that the M *vs.* P forms also prevail at the respective hopping geometries. Indeed, an almost perfectly

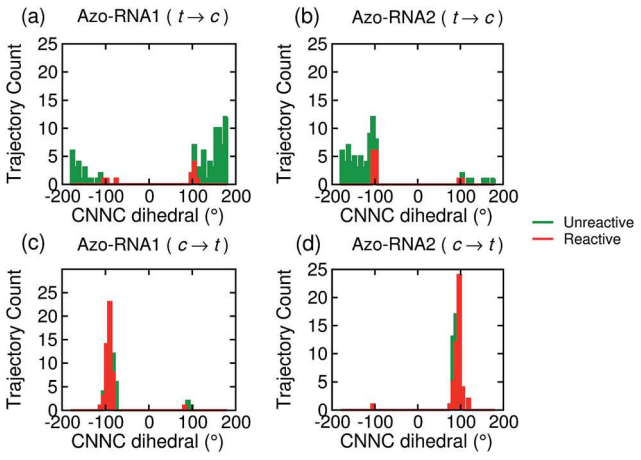


Fig. 7 Population distribution of the CNNC dihedral at the point of S_1/S_0 hopping for the trajectories starting from the *trans* (upper panels) and *cis* (lower panels) isomer for Azo–RNA1 (left panels) and Azo–RNA2 (right panels). The algebraic sign of the CNNC dihedral informs about the P-helical (CNNC > 0°) *vs.* M-helical (CNNC < 0°) form at the hopping geometry. The red and green traces correspond to reactive and unreactive trajectories, respectively. An almost perfectly chirality conserving pathway is observed for the *cis*-to-*trans* trajectory ensemble (see text for details).

chirality conserving pathway is found in both cases. This conclusion is confirmed by a detailed analysis of the NNCC and CCNN dihedral angles as described in the ESI (Section S8†): during the initial dynamics (~200 fs) of the Azo–RNA1 complex, the NNCC dihedral tends to 0° starting from negative values, and its counterpart in the Azo–RNA2 complex also tends to 0° starting from positive values. This confirms that the M-helical form of Azo–RNA1 and the P-helical form of Azo–RNA2 remain conserved on the path towards the S_1/S_0 conical intersection. Clearly, the chiral RNA environment is the determining factor inducing this pronounced selectivity. For comparison, a ~75% predominance of the chirality conserving pathway was found for the gas-phase *cis*-to-*trans* isomerization in ref. 66 and 67, while the study of ref. 32 did not provide evidence for such a preferential pathway in gas phase or solution. (However, a nearly complete stereospecificity was found for a related arylazopyrazole photoswitch[69]).

Notably, a narrow, almost normal-type distribution of CNNC dihedrals around ±90° is observed both for the reactive and unreactive *cis*-to-*trans* trajectories of Azo–RNA1 and Azo–RNA2 in Fig. 7. All trajectories reach the vicinity of the conical intersection and thereby the quantum yield of photoisomerisation is 50%.

4 Conclusions

To summarize, we have investigated the photoisomerization mechanism of an azobenzene chromophore covalently attached to an RNA double strand using state-of-the-art semiempirical QM/MM–SH methodology that has proven reliable in the treatment of azobenzene in various environments over recent years.[32,38–40,48,50] Concerted nonadiabatic dynamics in six singlet electronic states is followed after photoexcitation to the bright S_2 state. Our investigation highlights several key features: first, ultrafast internal conversion from the S_2 state to the S_1 state initiates the photochemical process (apart from minor participation of the S_3 and S_4 states), very similarly to the photochemistry of the bare azobenzene chromophore in gas phase[63,70] or in solution.[32,57] Second, the main effect of the extremely constrained RNA environment is felt in the S_1 dynamics: the local interactions with the RNA bases lead to an S_1 lifetime of ~15 picoseconds for the *trans* azobenzene chromophore preceding isomerization, as compared with 500 femtoseconds *in vacuo*. Further, the *trans*-to-*cis* photoisomerization quantum yield is reduced to ~10%. By contrast, the electronically excited *cis* isomer remains short-lived, on a time scale of ~500 femtoseconds, with a 50% *cis*-to-*trans* quantum yield. We attribute this pronounced difference to the fact that the upper and lower base pairs exert a highly stabilizing effect on the *trans* azobenzene isomer – but not on the *cis* isomer. This picture is confirmed by representative QM/MM trajectories. For the same reasons, the isomerization mechanism is found to be volume-preserving and proceeds by a concerted motion in several dihedral angles, resulting in a pedal-like motion that also prevails in the isomerisation in solution.[32,57] This motion is accompanied by a collective rearrangement of the local RNA environment that gives rise to steric constraints which are of dynamical rather than static nature.

Even though time-resolved spectroscopic observations for these and related systems are not yet available, our predictions are in line with similar trends observed for azobenzene photoisomerization in constrained environments and adsorbed on surfaces.[32,47–51]

Finally, we found that RNA chirality plays a subtle role in the process and leads to preferential P type (right-handed) or M type (left-handed) *cis*-to-*trans* isomerisation pathways. Strikingly, the equilibrated Azo–RNA1 *vs.* Azo–RNA2 *cis* species are of M type *vs.* P type, respectively, and conserve their helicity during their short-time dynamics such that the preferred M-type *vs.* P-type forms are also observed at the hopping geometry. While some degree of chiral selectivity has been previously found for *cis*-to-*trans* azobenzene isomerization in the gas phase or in less specific environments, the almost full conservation of M-type *vs.* P-type helicity during the excited-state dynamics of the Azo–RNA complexes is a unique feature of the chiral RNA environment.

Most of these conclusions are remarkably similar for the two types of chromophore-linker–RNA complexes that were investigated in this work, with the exception of the minor *vs.* major groove preference of the *cis* azobenzene chromophore, which is opposite for the two complexes, thus entailing opposite preferences for P *vs.* M helicity. Apart from this difference, the similarity of the mechanism and excited-state lifetimes is striking. Our results underline the crucial role of local stacking interactions in selectively stabilizing the *trans* form of the chromophore, which should be accounted for as one of the key guiding principles when designing new, optimally tailored RNA–linker-azobenzene combinations. Specific local modifications, using suitable linkers and substituents, could lead to reduced S_1 lifetimes and enhanced quantum yields, but possibly at the expense of reduced differences in melting temperatures associated with the *trans vs. cis* isomers. In view of connecting to longer time scales where RNA unfolding is observed,[23] the present benchmark study paves the way for multiscale simulations[31] to model light-responsive functional oligonucleotides and nanoassemblies.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

We gratefully acknowledge support by the DFG *via* the Research Center SFB 902 ("Molecular Principles of RNA-based Regulation") and the RTG 1986 "Complex Scenarios of Light Control", as well as by the Alfried Krupp von Bohlen und Halbach Foundation, associated with the German Scholar Organization (GSO) *via* grant II-98. M. P. acknowledges the financial support of Pisa University (grant PRA-2016-46). We thank K. Falahati for assistance with the preparation of the figures.

Notes and references

1 A. S. Lubbe, W. Szymanski and B. L. Feringa, *Chem. Soc. Rev.*, 2017, **46**, 1052–1079.
2 G. Mayer and A. Heckel, *Angew. Chem., Int. Ed.*, 2006, **45**, 4900–4921.
3 A. Jäschke, *FEBS Lett.*, 2012, **586**, 2106–2111.
4 W. Szymański, J. M. Beierle, H. A. V. Kistemaker, W. A. Velema and B. L. Feringa, *Chem. Rev.*, 2013, **113**, 6114–6178.
5 H. Ito, X. Liang, H. Nishioka and H. Asanuma, *Org. Biomol. Chem.*, 2010, **8**, 5519–5524.
6 J. Li, X. Wang and X. Liang, *Chem.–Asian J.*, 2014, **9**, 3344–3358.
7 Y. Kamiya, T. Takagi, H. Ooi, H. Ito, X. Liang and H. Asanuma, *ACS Synth. Biol.*, 2015, **4**, 365–370.
8 J. A. Phillips, H. Liu, M. B. O'Donoghue, X. Xiong, R. Wang, M. You, K. Sefah and W. Tan, *Bioconjugate Chem.*, 2011, **22**, 282–288.
9 T. Cha, J. Pan, H. Chen, H. N. Robinson, X. Li, C. Mao and J. H. Choi, *J. Am. Chem. Soc.*, 2015, **137**, 9429.
10 X. Liang, H. Nishioka, N. Takenaka and H. Asanuma, *ChemBioChem*, 2008, **9**, 702–705.
11 H. M. D. Bandara and S. C. Burdette, *Chem. Soc. Rev.*, 2012, **41**, 1809–1825.
12 A. A. Beharry and G. A. Woolley, *Chem. Soc. Rev.*, 2011, **40**, 4422–4437.
13 J. Andersson, S. Li, P. Lincoln and J. Andréasson, *J. Am. Chem. Soc.*, 2008, **130**, 11836–11837.
14 H. Rau, in *Photochemistry and Photophysics*, ed. RabekJ. F., CRC Press, Boca Raton, FL, USA, 1990, pp. 119–141.
15 E. Merino and M. Ribagorda, *Beilstein J. Org. Chem.*, 2012, **8**, 1071–1090.
16 K. Morgenstern, *Acc. Chem. Res.*, 2009, **42**, 213–223.
17 H. Kashida, X. Liang and H. Asanuma, *Curr. Org. Chem.*, 2009, **13**, 1065–1084.
18 H. Ito, H. Nishioka, X. Liang and H. Asanuma, *Nucleic Acids Symp. Ser.*, 2007, **51**, 171–172.
19 M. Biswas and I. Burghardt, *Biophys. J.*, 2014, **107**, 932–940.
20 D. Rastädter, M. Biswas and I. Burghardt, *J. Phys. Chem. B*, 2014, **118**, 8478–8488.
21 H. Asanuma, T. Ito, T. Yoshida, X. G. Liang and M. Komiyama, *Angew. Chem., Int. Ed.*, 1999, **38**, 2393–2395.
22 H. Asanuma, D. Matsunaga and M. Komiyama, *Nucleic Acids Symp. Ser.*, 2007, **49**, 35–36.
23 Y. Nakasone, H. Ooi, Y. Kamiya, H. Asanuma and M. Terazima, *J. Am. Chem. Soc.*, 2016, **138**, 9001–9004.
24 M. Böckmann, S. Braun, N. L. Doltsinis and D. Marx, *J. Chem. Phys.*, 2013, **139**, 084108.
25 B. G. Keller, A. Kobitski, G. Jäschke, U. Nienhaus and F. Noé, *J. Am. Chem. Soc.*, 2014, **136**, 4534–4543.
26 H. M. Senn and W. Thiel, *Angew. Chem., Int. Ed.*, 2009, **48**, 1198–1229.
27 O. Weingart, *Curr. Org. Chem.*, 2016, **21**, 586–601.
28 H. J. Kulik, J. Zhang, J. P. Klinman and T. J. Martínez, *J. Phys. Chem. B*, 2016, **120**, 11381–11394.
29 A. Strambi, B. Durbeej, N. Ferré and M. Olivucci, *Proc. Natl. Acad. Sci. U. S. A.*, 2010, **107**, 21322–21326.
30 V. A. Spata and S. Matsika, *J. Phys. Chem. A*, 2014, **118**, 12021–12030.
31 S. Osella and S. Knippenberg, *J. Am. Chem. Soc.*, 2017, **139**, 4418–4428.
32 V. Cantatore, G. Granucci and M. Persico, *Comput. Theor. Chem.*, 2014, **1040–1041**, 126–135.
33 S.-H. Xia, G. Cui, W.-H. Fang and W. Thiel, *Angew. Chem., Int. Ed.*, 2016, **55**, 2067–2072.
34 P. Mondal, M. Biswas, T. Goldau, A. Heckel and I. Burghardt, *J. Phys. Chem. B*, 2015, **119**, 11275–11286.
35 A. Kingsland, S. Samai, Y. Yan, D. S. Ginger and L. Maibaum, *J. Phys. Chem. Lett.*, 2016, **7**, 3027–3031.
36 T. Goldau, K. Murayama, C. Brieke, S. Steinwand, P. Mondal, M. Biswas, I. Burghardt, J. Wachtveitl, H. Asanuma and A. Heckel, *Chem.–Eur. J.*, 2015, **21**, 2845–2854.
37 T. Goldau, K. Murayama, C. Brieke, H. Asanuma and A. Heckel, *Chem.–Eur. J.*, 2015, **21**, 17870–17876.
38 G. Granucci, M. Persico and A. Toniolo, *J. Chem. Phys.*, 2001, **114**, 10608–10615.
39 T. Cusati, G. Granucci, E. Martínez-Núñez, F. Martini, M. Persico and S. Vázquez, *J. Phys. Chem. A*, 2012, **116**, 98–110.
40 C. Ciminelli, G. Granucci and M. Persico, *Chem.–Eur. J.*, 2004, **10**, 2327–2341.
41 V. Hornak, R. Abel, A. Okur, B. Strockbine, A. Roitberg and C. Simmerling, *Proteins*, 2006, **65**, 712–725.
42 J. W. Ponder, *TINKER 6.1*, Washington University School of Medicine, St. Louis, MO, 2012, http://dasher.wustl.edu/tinker.
43 W. L. Jorgensen and J. D. Madura, *J. Am. Chem. Soc.*, 1983, **105**, 1407–1413.
44 I. Antes and W. Thiel, *J. Phys. Chem. A*, 1999, **103**, 9290–9295.
45 J. C. Tully, *J. Chem. Phys.*, 1990, **93**, 1061–1071.
46 G. Granucci, M. Persico and A. Zoccante, *J. Chem. Phys.*, 2010, **133**, 134111.
47 G. Floss, G. Granucci and P. Saalfrank, *J. Chem. Phys.*, 2012, **137**, 234701.
48 V. Cantatore, G. Granucci, G. Rousseau, G. Padula and M. Persico, *J. Phys. Chem. Lett.*, 2016, **7**, 4027–4031.
49 T. Cusati, G. Granucci and M. Persico, *J. Am. Chem. Soc.*, 2011, **133**, 5109–5123.
50 E. Benassi, G. Granucci, M. Persico and S. Corni, *J. Phys. Chem. C*, 2015, **119**, 5962–5974.
51 L. Creatini, T. Cusati, G. Granucci and M. Persico, *Chem. Phys.*, 2008, **347**, 492–502.
52 K. Andersson, P.-A. Malmqvist, B. O. Roos, A. J. Sadlej and K. Wolinski, *J. Phys. Chem.*, 1990, **94**, 5483–5488.
53 A. Cembran, F. Bernardi, M. Garavelli, L. Gagliardi and G. Orlandi, *J. Am. Chem. Soc.*, 2004, **126**, 3234–3243.
54 V. Ladányi, P. Dvorák, J. Al Anshori, L. Vetráková, J. Wirz and D. Heger, *Photochem. Photobiol. Sci.*, 2017, **16**, 1757–1761.
55 Y. Yan, X. Wang, J. I. L. Chen and D. S. Ginger, *J. Am. Chem. Soc.*, 2013, **135**, 8382–8387.
56 S. Samal, D. J. Bradley, T. Choi, Y. Yan and D. S. Ginger, *J. Phys. Chem. C*, 2017, **121**, 6997–7004.

57 M. Böckmann, N. L. Doltsinis and D. Marx, *J. Phys. Chem. A*, 2010, **114**, 745–754.
58 E. Titov, G. Granucci, J. P. Gotze, M. Persico and P. Saalfrank, *J. Phys. Chem. Lett.*, 2016, **7**, 3591–3596.
59 M. Pederzoli, J. Pittner, M. Barbatti and H. Lischka, *J. Phys. Chem. A*, 2011, **115**, 11136–11143.
60 J. Casellas, M. J. Bearpark and M. Reguero, *ChemPhysChem*, 2016, **17**, 3068–3079.
61 T. Nägele, R. Hoche, W. Zinth and J. Wachtveitl, *Chem. Phys. Lett.*, 1997, **272**, 489.
62 H. Satzger, S. Spörlein, C. Root, J. Wachtveitl, W. Zinth and P. Gilch, *Chem. Phys. Lett.*, 2003, **372**, 216.
63 E. Tan, S. Amirjalayer, S. Smolarek, A. Vdovin, F. Zerbetto and W. J. Buma, *Nat. Commun.*, 2015, **6**, 5860.
64 *Conical Intersections: Theory, Computation, and Experiment*, ed. W. Domcke, H. Köppel and D. Yarkony, World Scientific Publishing Company, 2011.
65 D. Reith, M. Pütz and F. Müller-Plathe, *J. Comput. Chem.*, 2003, **24**, 1624.
66 O. Weingart, Z. Lan, A. Koslowski and W. Thiel, *J. Phys. Chem. Lett.*, 2011, **2**, 1506–1509.
67 Y. Ootani, S. Kiminori, A. Nakayama, T. Noro and T. Taketsugu, *J. Chem. Phys.*, 2009, **131**, 194306.
68 J. A. Gámez, O. Weingart, A. Koslowski and W. Thiel, *J. Chem. Theory Comput.*, 2012, **8**, 2352–2358.
69 Y. Wang, X. Liu, G. Cui, W. Fang and W. Thiel, *Angew. Chem., Int. Ed.*, 2016, **55**, 14009.
70 T. Schultz, J. Quenneville, B. Levine, A. Toniolo, T. J. Martínez, S. Lochbrunner, M. Schmitt, J. P. Shaffer, M. Z. Zgierski and A. Stolow, *J. Am. Chem. Soc.*, 2003, **125**, 8098–8099.

Supporting Information:

Azobenzene as a Photoregulator Covalently Attached to RNA: A Quantum Mechanics/Molecular Mechanics-Surface Hopping Dynamics Study

Padmabati Mondal,[*,†,¶] Giovanni Granucci,[‡] Dominique Rastädter,[†]

Maurizio Persico,[*,‡] and Irene Burghardt[*,†]

Institute of Physical and Theoretical Chemistry, Goethe University Frankfurt, Max-von-Laue-Str. 7, 60438 Frankfurt, Germany, and Dipartimento di Chimica e Chimica Industriale, Università di Pisa, v. Moruzzi 13, I-56124 Pisa, Italy

E-mail: padmabati.mondal@gmail.com; maurizio.persico@unipi.it;

burghardt@chemie.uni-frankfurt.de

[*]To whom correspondence should be addressed

[†]Institute of Physical and Theoretical Chemistry, Goethe University Frankfurt, Max-von-Laue-Str. 7, 60438 Frankfurt, Germany

[‡]Dipartimento di Chimica e Chimica Industriale, Università di Pisa, v. Moruzzi 13, I-56124 Pisa, Italy

[¶]Present address: Institut de Chimie Radicalaire, Université d'Aix-Marseille, 13013 Marseille, France

Supporting Information – Contents

S1. Generation of initial conditions for QM/MM-SH dynamics

The initial QM/MM-SH set-up for the Azo-RNA1 and Azo-RNA2 complexes was created as described in the following. In both cases, the azobenzene chromophore is attached to an RNA double helix *via* a β-deoxyribose linker.[1] For the linker, we used the parameters for ribose in the AMBER99 force field. The protocol was slightly different for the two complexes, as now detailed.

Azo-RNA1: A 9-mer azobenzene-RNA duplex (Azo-RNA1) is prepared from a snapshot of classical MD simulations of 200 ns duration for a 15-mer azobenzene-RNA complex,[2] from which three upper and three lower base-pairs were truncated. This azobenzene-linker-RNA system was then immersed into a TIP3P water sphere with a radius of 25 Å. The water sphere is further minimized energetically keeping the azobenzene-RNA system fixed. Finally, the complete system is minimized using TINKER,[3] and the minimized structure is used as the initial structure for the QM/MM equilibration.

Azo-RNA2: The UUCG tetraloop 14-mer structure of Azo-RNA2 is analogous to the system that was investigated by classical MD simulations in Ref. [4], except that a β-deoxyribose linker is now used. Following equilibration, a spherical cluster with a 25 Å radius was cut out around the RNA complex, including all ions. All water molecules with their center of mass within the sphere radius of the truncated cell were included.

Pre-equilibration using MD is followed by QM/MM equilibration as described in the main manuscript text. The ground-state distribution is obtained by taking snapshots from the QM/MM equilibration trajectory. Next, the ground state distribution is projected onto the electronically excited states, as further detailed below.

In the present simulations, the SH dynamics reported in the manuscript were started exclusively

from the $S_2(\pi-\pi^*)$ state, for 100 initial conditions. The rationale for this is that the $S_2(\pi-\pi^*)$ state exhibits the largest average squared transition dipole moment $\langle|\mathbf{d}|^2\rangle$, as can be inferred from the following data:

$\langle\|\mathbf{d}\|^2\rangle$ [a.u.]	S_1	S_2	S_3	S_4	S_5
trans-Azo-RNA1	0.15	8.75	0.78	0.64	0.22
cis-Azo-RNA1	0.39	3.15	1.51	1.36	0.91

Table S1. Average squared transition dipole moment (in [a.u.]) for all accepted geometries.

Trajectories are selected for QM/MM-SH propagation using the criterion that the vertical excitation energy falls into a predefined transition energy window,[5] which is here specified for excitation to the bright $S_2(\pi-\pi^*)$ state as $\Delta E = 3.5\pm0.5$ eV for the *trans*-azobenzene substituted system and $\Delta E = 4.0\pm0.5$ eV for the *cis*-azobenzene substituted system. Preliminary values for the transition energy windows were taken from previous studies of azobenzene in condensed phase,[6] which were subsequently refined using the available range of geometry-dependent vertical excitation energies. The average vertical excitation energies for the *trans* and *cis* $S_2(\pi-\pi^*)$ states are 3.62 eV (342 nm) and 4.10 eV (300 nm), respectively, in line with the spectra reported below (see Fig. S1(c) and (d)).

To reduce the computational cost, trajectories are stopped when one of the following two criteria is fulfilled: (i) the trajectory has been running on the ground state potential surface for a total time longer than 2.5 ps, in the transoid or cisoid geometry (i.e., CNNC dihedral close to 180° or else 0°, CNN angles smaller than 150°); (ii) the total simulation time is larger than 30 ps.

S2. Absorption spectra

To ensure reasonable agreement between theory and experiment, the absorption spectra of the azobenzene chromophore in the two azobenzene-RNA model systems under study and an experimentally measured reference system[1] were compared. Figure S1 presents a comparison between the experimental and computed spectra. The experimental spectrum[1] was measured for an azobenzene with a β-deoxyribose linker in *para* position (denoted "pAzo" in Fig. S1(a)), both in solution (Fig. S1(a)) and in an RNA double strand (Fig. S1(b)). The azobenzene-linker-RNA double strand is closely related but not identical to the Azo-RNA1 system under study (see Ref. [1] for details). As can be seen from the comparison of Fig. S1(a) and Fig. S1(b), the RNA environment induces a spectral redshift of about 10 nm in the $S_2(\pi-\pi^*)$ absorption of the *trans*-azobenzene species (from $\sim$325 nm to $\sim$340 nm). Also, note that Fig. S1(b) shows the RNA absorption at noticeably higher energies ($\sim$250 nm); this is not reproduced in our QM/MM setting where electronic excitations are restricted to the QM subsystem.

The theoretical spectra shown in Fig. S1(c) and Fig. S1(d) were computed as histograms of the dipole radiative transition probability versus the excitation energy, taken from the QM/MM 40 ps equilibration trajectory (40000 snapshots). Thereby, vibrational broadening effects were taken into account through the thermal distribution of nuclear configurations in the classical way (no zero point vibrational motion).

In the calculated spectra, the $S_2(\pi \rightarrow \pi^*)$ transitions of *trans* and *cis* azobenzene in RNA are found at 342 nm and 300 nm (see above), in qualitatively good agreement with the measured spectra. The *trans* transition at 340 nm is even in very good agreement with the experimental spectrum of Fig. S1(b). By contrast, the $S_1(n \rightarrow \pi^*)$ transition exhibits a redshift of around 100 nm as compared with the experimental spectrum.

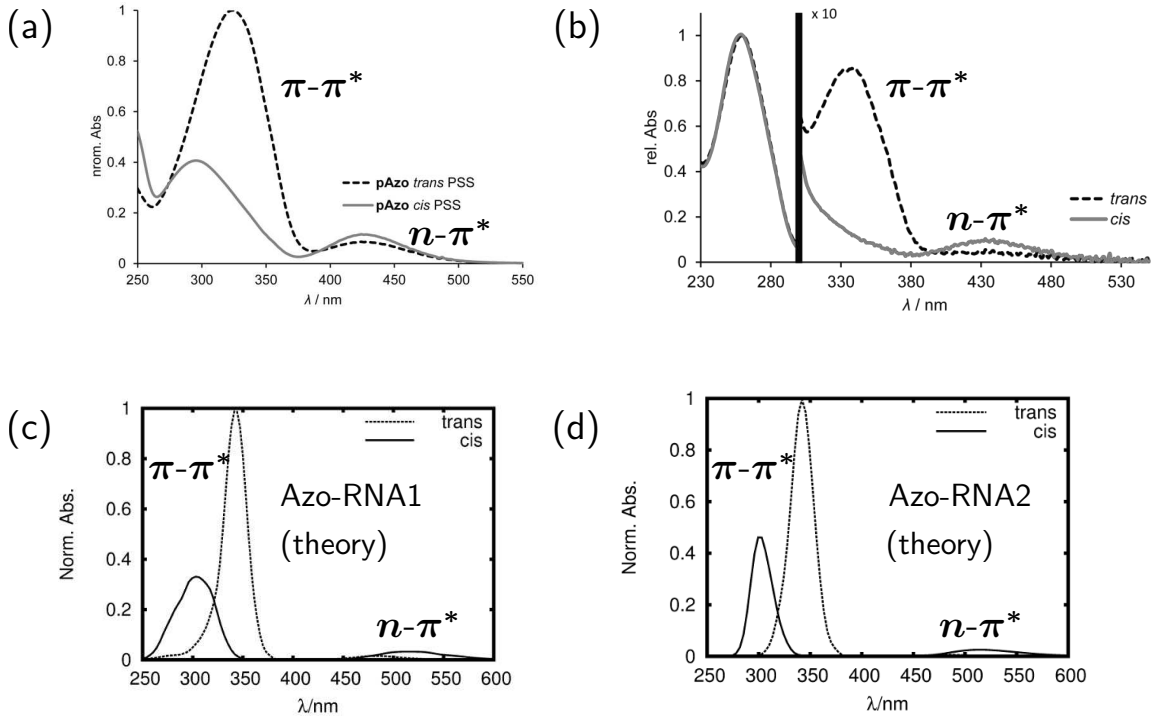


Figure S1: Experimental (upper panels) and calculated (lower panels) absorption spectra of *trans* and *cis* azobenzene with the β-deoxyribose linker under study. The $S_2(\pi - \pi^*)$ and $S_1(n - \pi^*)$ transitions are indicated explicitly. (a) Experimental spectrum of azobenzene with a β-deoxyribose linker in *para* position (pAzo) in solution (from Supp. Material of Ref. [1]), (b) experimental spectrum of pAzo in a double-stranded RNA (from Supp. Material of Ref. [1]), exhibiting a spectrally shifted azobenzene absorption and a high-energetic RNA absorption at $\sim$250 nm; (c) calculated spectrum of azobenzene in Azo-RNA1, (d) calculated spectrum in Azo-RNA2.

S3. FOMO-CI vs. CASPT2 energetics with/without RNA environment

In this section we compare FOMO-CI with CASPT2 computed energies. Moreover, we assess the effect of the environment as computed by the semiempirical QM/MM procedure or by including selected RNA bases in a full QM calculation. The FOMO-CI and CASPT2 calculations are run at the same geometries in all cases. The CASPT2 calculations are based on a SA-CASSCF(14,12) with the 6-31G* basis set, as in Conti et al.,[7] averaging over the first four singlet states.

In Table S2 we compare the vertical excitation energies at the equilibrium geometries of ground state *trans*- and *cis*-azobenzene, as a reference for further comparisons. In the last column we show the experimental energies inferred from the absorption spectra.[8] The S_0-S_1 transition energies are in good agreement, while larger discrepancies are observed for the S_0-S_2 transition: in fact, the CASPT2 results are 0.2-0.5 eV higher than the FOMO-CI energies, and the latter are much closer to the experimental ones, as a result of the *ad hoc* reparameterization specifically developed for azobenzene.[8]

		FOMO-CI	CASPT2	Exp.[8]
trans	S_0-S_1	2.83	2.60	2.82
	S_0-S_2	4.04	4.56	4.12
cis	S_0-S_1	2.89	2.91	2.92
	S_0-S_2	4.59	4.81	4.60

Table S2. Vertical transition energies (eV) obtained by gas phase calculations at the FOMO-CI equilibrium geometry for the ground state.

The reliability of the semiempirical QM/MM calculations used in the simulations is tested in two different situations along a photochemical trajectory: (i) at the beginning, when the only distortions with respect to the gas phase equilibrium geometry are due to the environment and to the thermal fluctuations (see Table S3), and (ii) at the time of the (first) hop from S_1 to S_0, which is the crucial internal conversion event which determines whether the trajectory is reactive or unreactive (see Table S4). We consider several fragments relating to the Azo-RNA1 system, notably involving the two bases (A13 and U6, see Figure 2) that most strongly interact with azobenzene.

Table S3 shows the S_0-S_1 and S_0-S_2 transition energies for *trans*- and *cis*-azobenzene with geometries corresponding to the beginning of two randomly selected trajectories of the Azo-RNA1 system. A set of calculations was run for the bare azobenzene molecule, where the connection atom was replaced by an hydrogen atom, the coordinates of which were optimized at FOMO-CI level. All the transition energies computed with both methods are lower than those of Table S2, because of the distorsion of the molecular geometry induced by the environment and by the thermal motions. The FOMO-CI and CASPT2 S_0-S_1 transition energies are in good agreement and the small difference for *trans*-azobenzene is close to the one observed at the equilibrium geometry (see Table S2). Again, for the S_0-S_2 transition the discrepancies are larger. In particular, for *trans*-azobenzene the difference is 0.95 eV, but about half of it can be ascribed to the overestimation of the S_0-S_2 transition energy by CASPT2 already seen in Table S2.

The effect of the environment is taken into account at the QM/MM level in the FOMO-CI calculations and at a full QM level in the CASPT2 calculations. In the former we include the closest RNA bases (A13 and U6 separately and together) as well as the whole Azo-RNA1 system, whereas in the latter we test the effect of the closest RNA bases (A13 and U6 separately and

together). For *cis*-azobenzene, the environmental shifts are found to be practically negligible. For *trans*-azobenzene, by contrast, positive increments of the S_0-S_1 and S_0-S_2 transition energies of the order of $\sim$0.2 eV (and somewhat larger for S_0-S_1 in the CASPT2 results) are found at both the *ab initio* and the semiempirical computational levels. This is in line with the role of stacking interactions that typically take values between 0.1-0.4 eV in nucleic acids.[9] As mentioned above, the environmental shift at the CASPT2 level as compared with FOMO-CI, is larger for the S_0-S_1 transition and smaller for the S_0-S_2 one. The discrepancy between the FOMO-CI and the CASPT2 results therefore tends to decrease when the environmental effects are taken into account.

		FOMO-CI				
		azo	azo+A13	azo+U6	azo+A13U6	azo-RNA1
trans	S_0-S_1	2.34	2.50	2.49	2.50	2.54
	S_0-S_2	3.36	3.73	3.68	3.71	3.65
cis	S_0-S_1	2.14	2.15	2.13	2.13	2.22
	S_0-S_2	4.09	4.14	4.07	4.12	4.11
		CASPT2				
		azo	azo+A13	azo+U6	azo+A13U6	
trans	S_0-S_1	2.03	2.58	2.55	2.60	
	S_0-S_2	4.31	4.40	4.37	4.51	
cis	S_0-S_1	2.17	2.19	2.30	2.31	
	S_0-S_2	4.61	4.85	4.80	4.78	

Table S3. Vertical transition energies (eV) obtained at the starting geometries of two typical trajectories of Azo-RNA1, for the *trans*→*cis* and the *cis*→ *trans* photoisomerization, respectively. The upper block shows FOMO-CI results, while the lower block shows CASPT2 results.

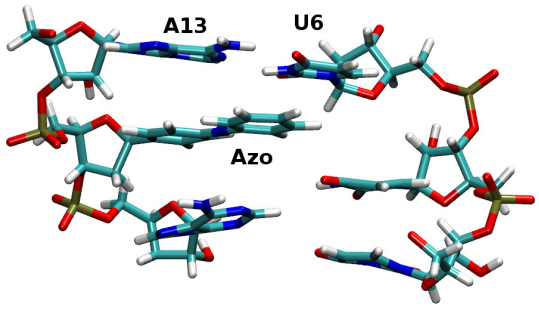


Figure S2: Snapshot of the truncated *trans* Azo-RNA1 complex, to illustrate the location of the closest bases A13 and U6 with respect to azobenzene (Azo).

In Table S4 we show the results obtained at the S_1-S_0 hopping geometries for a *trans*→*cis* and a *cis*→*trans* trajectory. We see that the S_1-S_0 energy differences for bare azobenzene, computed at the FOMO-CI vs. CASPT2 levels, agree very well. The environmental effects are almost negligible at both levels of theory.

	FOMO-CI				
	azo	azo+A13	azo+U6	azo+A13U6	azo-RNA1
trans→*cis*	0.15	0.22	0.25	0.16	0.16
cis→*trans*	0.21	0.21	0.24	0.22	0.20
	CASPT2				
	azo	azo+A13	azo+U6	azo+A13U6	
trans→*cis*	0.14	0.26	0.30	0.24	
cis→*trans*	0.29	0.30	0.27	0.23	

Table S4. S_1-S_0 energy gaps (eV) obtained at the hopping geometries of two typical trajectories of Azo-RNA1, for the *trans*→*cis* and the *cis*→*trans* photoisomerization, respectively.

Overall, we see that the differences between FOMO-CI and CASPT2 results are quite small for the S_0-S_1 transition energies and larger for the S_0-S_2 ones, especially at transoid geometries. The comparison with experimental data shows that at least in part the discrepancies are due to the overestimation of the S_0-S_2 transition energy by CASPT2. The distorsions of the molecular geometry due to the RNA cage and to thermal fluctuations tend to lower the transition energies, while the perturbations of the electronic structure caused by interactions with the RNA bases have the opposite effect. Both effects are correctly reproduced by the FOMO-CI QM/MM treatment, at least semiquantitatively. The small S_0-S_1 energy gaps computed at hopping points are very closely reproduced at the CASPT2 level.

S4. Surface Hopping dynamics: Energy gap distributions

Here, we analyze the energy gaps between the electronic states at the SH hopping points, both for the *trans*→*cis* transition (Figure S3a/b) and the *cis*→*trans* transition (Figure S3c/d).

Starting from *trans* azobenzene (Figure S2a/b), most of the trajectories hop from S_2 to S_1 (green circles) within the first picosecond, close to the Franck-Condon region. The S_1/S_0 hopping events (red circles) occur in a wide time range from 2 ps up to 30 ps. Only few direct S_2/S_0 hops (blue circles) take place, remaining as rare events. A small number of trajectories start to hop from S_1 to S_0 near $\Delta E = 0$, which is in the region of the conical intersection. Most of the trajectories reach the

ground state far before the CI and therefore end back in *trans* form. This explains the low quantum yield of the *trans*-to-*cis* photoisomerization.

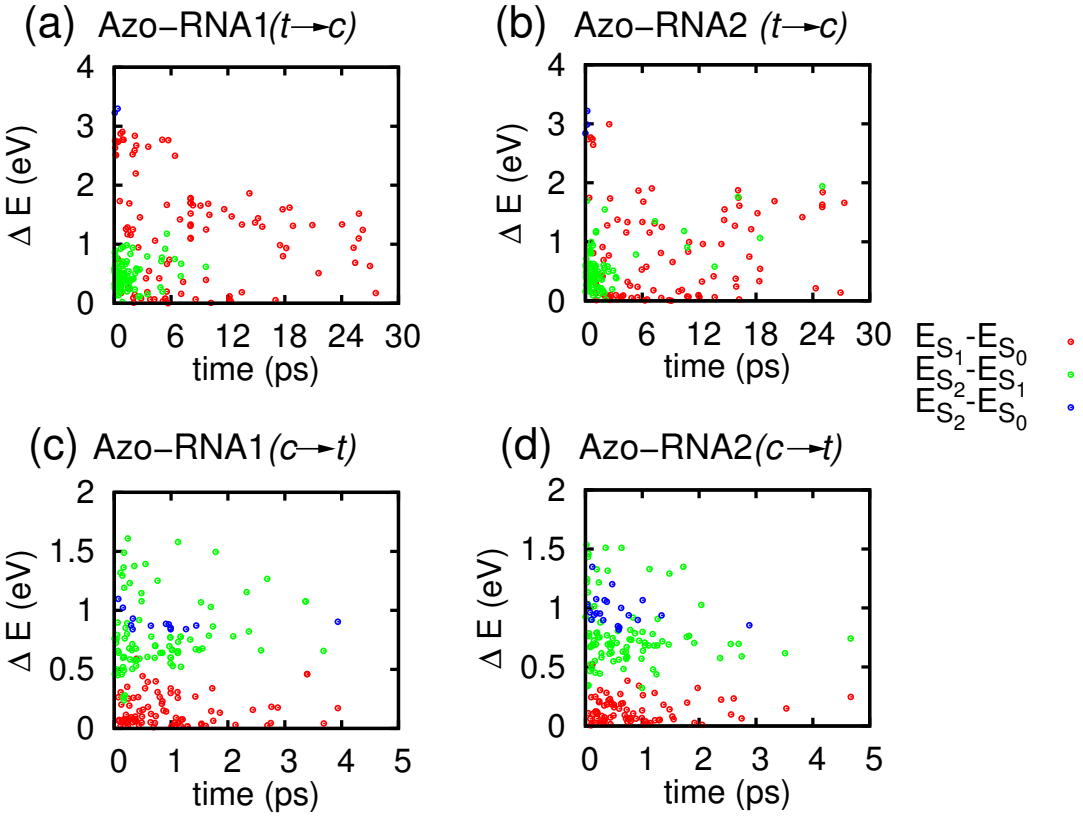


Figure S3: Energy differences between S_1/S_0 (red), S_2/S_1 (green) and S_2/S_0 (blue) states at the respective hopping points, starting from (a, b) *trans* and (c, d)*cis* azobenzene for Azo-RNA1 (left panels) and Azo-RNA2 (right panels).

In the dynamics starting from the *cis* isomer (Figure S3c/d), both the S_2/S_1 and S_1/S_0 hopping events are scattered in the time window from 0 ps to 5 ps. The figure indicates that while the S_2/S_1 conical intersection on the *trans* side is very close to the Franck-Condon region – but far from S_1/S_0 conical intersection – the situation is different for the *cis* side: Here, the corresponding S_2/S_1 and S_1/S_0 conical intersections lie in close vicinity, which explains the small population of the S_1 state in Figure 3 of the manuscript. Unlike the *trans* case, many trajectories reach very close to the S_1/S_0 conical intersection (red circles near $\Delta E = 0$), thereby explaining the higher *cis*-to-*trans* photoisomerisation quantum yield. Differently from the *trans* case, the S_1 state is very short-lived in the *cis* case due to the above-mentioned characteristics of potential energy surface. Either a direct S_2/S_0 population transfer (10%) or else S_2/S_1 hops immediately followed by a S_1/S_0 hop are observed.

The two azobenzene-RNA systems, i.e., Azo-RNA1 and Azo-RNA2, show a very similar dynamics overall. The only noticeable difference relates to the fact that while all S_2/S_1 hopping

events occur within 0-8 ps for Azo-RNA1 (Figure a), a certain percentage (5%) of the S_2/S_1 hopping events also occur in the range of 8-30 ps for Azo-RNA2.

S5. Interaction between azobenzene and local RNA environment

Much of the present study is concerned with understanding how the rigid RNA environment influences the lifetime and isomerisation pathways of azobenzene. Here, we analyse some dynamic and static aspects of the mutual influence between the azobenzene chromophore and its local RNA surroundings, notably by a time-dependent distance analysis (Sec. S5.1), by root mean square fluctuation (RMSF) analysis (Sec. S5.2), and by Boltzmann inversion analysis (Sec. S5.3).

S5.1 Time-dependent distance analysis of azobenzene and neighboring bases

In Table S5 we report the fraction of the simulation time spent below a given threshold by the distances between pairs of atoms, one belonging to an RNA base and the other to the distal phenyl ring of the *trans* azobenzene chromophore (i.e., the ring that is not covalently bound to β–deoxyribose and undergoes the largest displacements). The threshold is defined by the sum of the van der Waals radii of the two atoms. We chose the base atoms with the shortest distances from each of the phenyl C atoms, and these turned out to belong to the adenine and uracil lying above azobenzene in the Azo-RNA1 complex. The same distances, averaged over all the unreactive trajectories, are shown in Fig. S4. We see that in the ground state the phenyl ring (Ph2) and the two RNA bases lying above it are well stacked together, whereas in the excited states all the distances increase in the average and less time is spent at short distances. Quite clearly the torsion of the N=N double bond occurs in the more favorable direction, i.e. by moving the phenyl ring away from the closest base pair.

C_{phenyl} (Ph2)	base atom	$T_<$ [ground-state tr.]	$T_<$ [unreactive tr.]	$T_<$ [reactive tr.]
C5	N1 (A13)	0.052	0.017	0.020
C6 (*ortho*)	N10 (A13)	0.374	0.216	0.185
C7 (*meta*)	N10 (A13)	0.164	0.077	0.156
C8 (*para*)	O8 (U6)	0.137	0.057	0.046
C9 (*meta*)	C4 (U6)	0.706	0.211	0.263
C10 (*ortho*)	N3 (U6)	0.343	0.197	0.161

Table S5. Fraction of the simulation time spent by different pairs of atoms at a distance smaller than the sum of the respective van der Waals radii (denoted $T_<$), for the Azo-RNA1 for the thermal ground state dynamics and for the photochemical trajectories. In the columns labelled unreactive and reactive trajectories we display averages over all the trajectories of each kind. The numbering of the phenyl C atoms is sequential and starts as in Figs. 1 (main text) and S7. The base numbering

is done starting from the lower right base and ending at the lower left base. For the adenine and uracil atoms, the numbering is the standard one.

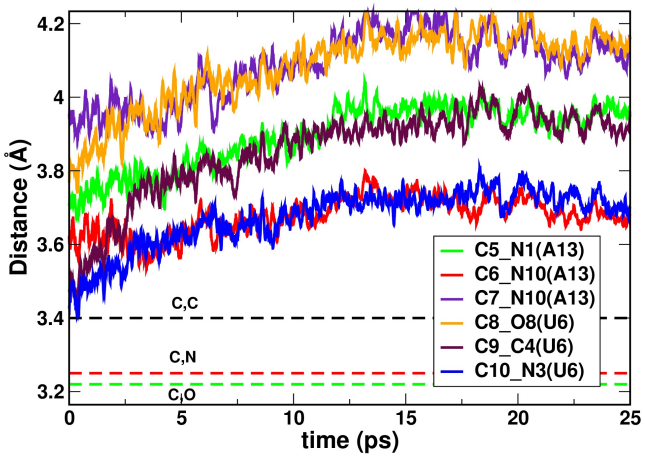


Figure S4: Distances between atoms of upper bases (A13, U6) and carbon atoms of the Ph2 ring of azobenzene in Azo-RNA1, averaged over all unreactive trajectories starting from the *trans* isomer. For reference, the sums of van der Waals radii for (C, C), (C, N), and (C, O) pairs are indicated.

S5.2 Root mean square fluctuation (RMSF) analysis

In a complementary fashion, Figure S5 shows the changes in the root mean square fluctuations (RMSF) of the four nearest-neighbour bases before and after the photoisomerisation for a typical reactive trajectory of Azo-RNA2. The figure indicates that especially the two upper bases (A4 and U11) exhibit more pronounced fluctuations after the photoisomerisation. This is most likely due to the fact that these bases show increased mobility to sustain stacking interactions during and even after the photoisomerisation, as demonstrated in Figure 5 of the main text.

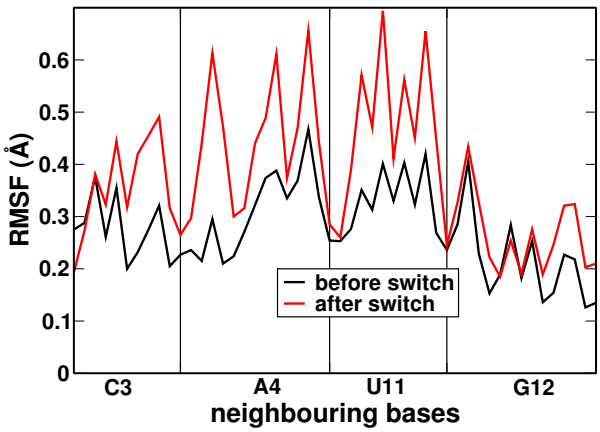


Figure S5: RMSF of the centres of mass of four nearest-neighbouring bases before (black) and after (red) photoisomerisation for a typical reactive trajectory of Azo-RNA2.

Overall, the above observations suggest that azobenzene isomerization does not exert a major disrupting effect on the local RNA environment. This is also confirmed by an analysis of the low-frequency backbone dihedrals, which do not exhibit any significant effect of the isomerization.

S5.2 Effective torsional potential by Boltzmann inversion

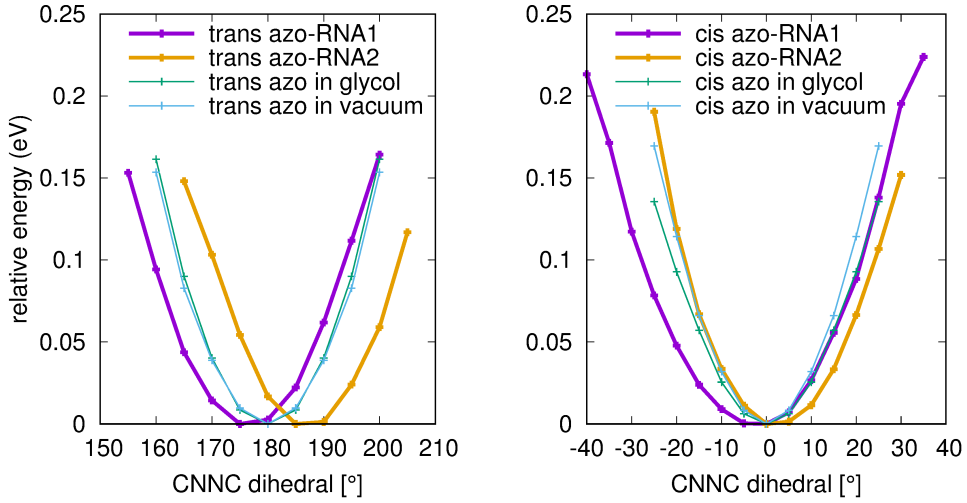


Figure S6: Relative effective torsional potentials around the *trans* (left) and *cis* (right) ground state equilibrium geometry of azobenzene in vacuum (cyan), ethylene glycol (green), azo-RNA1 (violet) and azo-RNA2 (orange).

.

In this analysis, the probability distribution $\rho(\theta)$ of the CNNC dihedral is used to obtain an effective torsional potential as a function of $\theta \equiv$ CNNC in the form $U(\theta) = -k_B T \ln\rho(\theta)$. In this way, the shape of the effective potential energy of the ground state potential around the *trans* and *cis* equilibrium geometry was obtained for azobenzene in different environments (in *vacuo*, in ethylene glycol and in RNA). Figure S6 shows the relative potential energy of the *trans* (left) and *cis* (right) azobenzene in vacuum (cyan), ethylene glycol (green), azo-RNA1 (violet) and in azo-RNA2 (orange) around their corresponding equilibrium geometry obtained by Boltzmann inversion. Interestingly, the curvature of the near-harmonic local potentials is almost unaffected by the environment, and do not reflect any pronounced effects of a constraining environment. A more tangible effect is a shift in the equilibrium positions which deviates from the *vacuo* value by $\sim 5^\circ$, in opposite directions for the Azo-RNA1 and Azo-RNA2 species. See the main text for a further discussion of these results.

S6. Definition of reaction coordinates and chirality

To describe azobenzene isomerization, several angular coordinates are important,[10,11] notably the CNNC, CCNN and NNCC dihedrals and the two CNN angles. In the following, we will refer to the labeling of Figure S7.

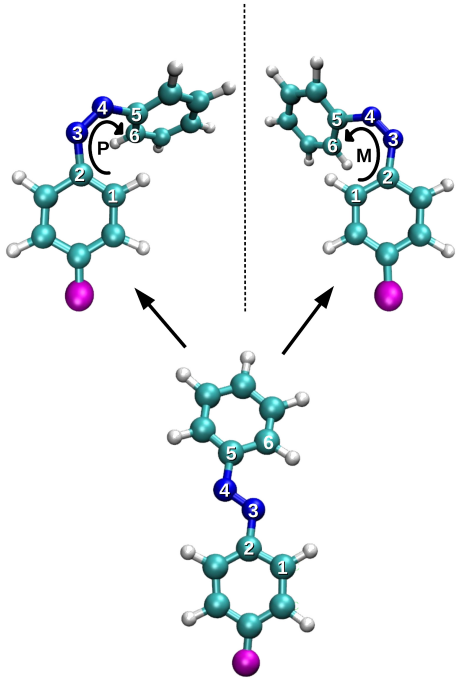


Figure S7: Molecular structure of the azobenzene chromophore, highlighting the atoms involved in the isomerization. Starting from the achiral *trans* form, azobenzene can isomerize to the P-type and M-type helical *cis* enantiomers.

The main reaction coordinate describing the *trans*-*cis* isomerization is the dihedral angle CNNC $\equiv$ *C*2-*N*3-*N*4-*C*5. Further, the isomerisation involves (i) the CNN angles, i.e., the *C*2-*N*3-*N*4 and *C*5-*N*4-*N*3 angles between the NN bond and the phenyl rings and (ii) the NNCC and CCNN dihedrals. Here, the *N*3-*N*4-*C*5-*C*6 (NNCC) dihedral angle describes the torsion of the free phenyl ring around the NC bond, and the *C*1-*C*2-*N*3-*N*4 (CCNN) dihedral describes the torsion of the phenyl ring connecting the deoxyribose linker with azobenzene.

While the *trans* isomer is achiral, the *cis* isomer exhibits helical chirality. (Note that chirality is not only restricted to stereogenic atoms: there are also planar and axial chiral elements.) The chirality of *cis* azobenzene results from the helicity along the CCNNCC fragment (atoms 1-6 in Figure S7), involving the torsions of the two CN and the NN bonds. The helix rotation can be detected by looking along the CCNNCC chain, with one phenyl ring in front and the other one in the back. When moving along the chain — from atom 1 to atom 6 — either a clockwise or an anti-clockwise rotational motion results. The clockwise helical rotation is denoted P (plus) helicity, while the anti-clockwise rotation is denoted M (minus) helicity. In one enantiomer, the CCNN and

NNCC dihedrals are in between -90° and 0° (for the M form), while in the other one they are in between 0° and 90° (for the P form).[6] Due to the mutual dependency of the three dihedrals along the CCNNCC chain, the algebraic sign of the CNNC dihedral during the isomerization process informs about the P-helical vs. M-helical form (with a positive or negative sign, respectively).

Both chiral *cis* enantiomers can undergo two pathways leading to the achiral *trans* form during CNNC torsion: either a chirality conserving pathway or a chirality inverting pathway.[6] To monitor these pathways, it is convenient to refer to the CCNN and NNCC dihedrals. For ease of notation, we employ the following definitions:[6] $\phi_1 = C1\text{-}C2\text{-}N3\text{-}N4$; $\phi_2 = 180^\circ - \phi_1$; $\phi_3 = N3\text{-}N4\text{-}C5\text{-}C6$; $\phi_4 = 180^\circ - \phi_3$. Since the CCNN dihedral does not change much due to its connectivity to the RNA strand, we will focus here only on the evolution of ϕ_3. The two enantiomers of *cis* azobenzene correspond to the structures where $-90^\circ < \phi_3 < 0^\circ$ (M) and $0^\circ < \phi_3 < 90^\circ$ (P), respectively. The *cis*-to-*trans* isomerisation starting from the M-helical enantiomer is chirality conserving if ϕ_3, starting from a negative value, decreases and approach to 0° or even a positive value. Otherwise, the pathway is chirality inverting. Similarly, the *cis*-to-*trans* isomerisation starting from the P-helical enantiomer is chirality conserving if ϕ_3, starting from a positive number, decreases and approach to 0° or even to a negative number.

As detailed in the manuscript and in Secs. S7 and S8, we found a chirality conserving pathway for both Azo-RNA1 and Azo-RNA2. Presumably this can be interpreted as an effect of the chiral RNA environment.

S7. Analysis of trajectory dynamics

S7.1 Azo-RNA1 complex

Complementary to Figure 6 of the main manuscript text, signed values of the relevant dihedrals are shown in Figure S8. These exhibit frequent sign changes between $\pm 180^\circ$ in the *trans* conformation, due to the fluctuating twist of the N3-N4 bond.

As can be seen from Figure S8b for a *cis*-to-*trans* reactive trajectory, the CNNC dihedral takes negative values, indicating conservation of the M-type helical form of the chromophore. Further, during the initial dynamics ($\sim$200 fs), the NNCC dihedral and the CCNN dihedral tend to zero, starting from negative initial values. As further explained in the main text and in Sec. S8, this confirms that a chirality conserving pathway is taken by the M-type *cis* azobenzene in the Azo-RNA1 complex.

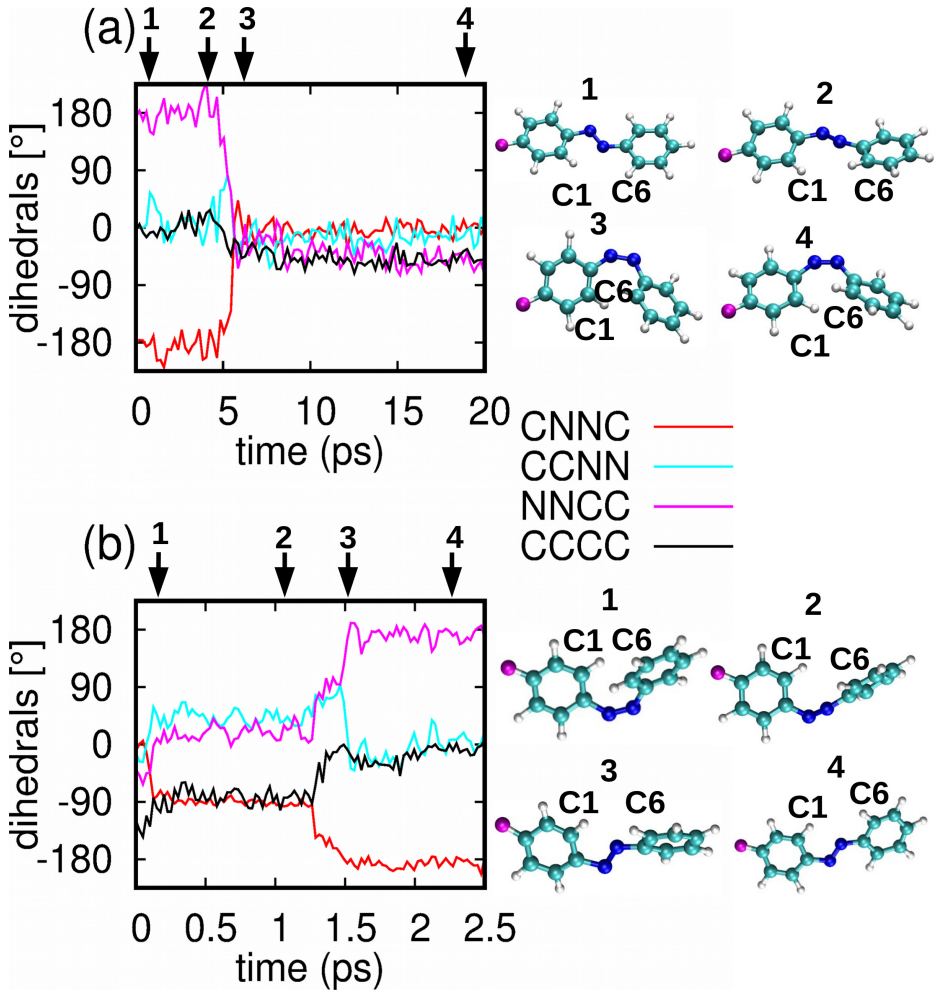


Figure S8: Changes of different dihedrals (with signs) and corresponding structure of azobenzene for a reactive trajectory starting from (a) *trans* and (b) *cis* azobenzene to explore the mechanism of photoisomerisation of Azo-RNA1.

S7.2 Azo-RNA2 complex

To compare with Figure 6 of the manuscript as well as Figure S8 addressed above, showing a representative photoisomerizing trajectory of the Azo-RNA1 complex, we now illustrate an analogous trajectory of Azo-RNA2 in Figure S9.

During *trans*-to-*cis* isomerisation, the absolute values of both the CCNN and NNCC dihedrals change by about $\pm 90^{\circ}$, which is slightly different from the case of Azo-RNA1 (see Figure 6 in the main text). The isomerisation in Azo-RNA2 thus involves a N3-N4 twist followed by rotation of both the $C2-N3$ and $N4-C5$ bonds. Both phenyl rings undergo partial in-plane rotation around the $N-C$ bonds.

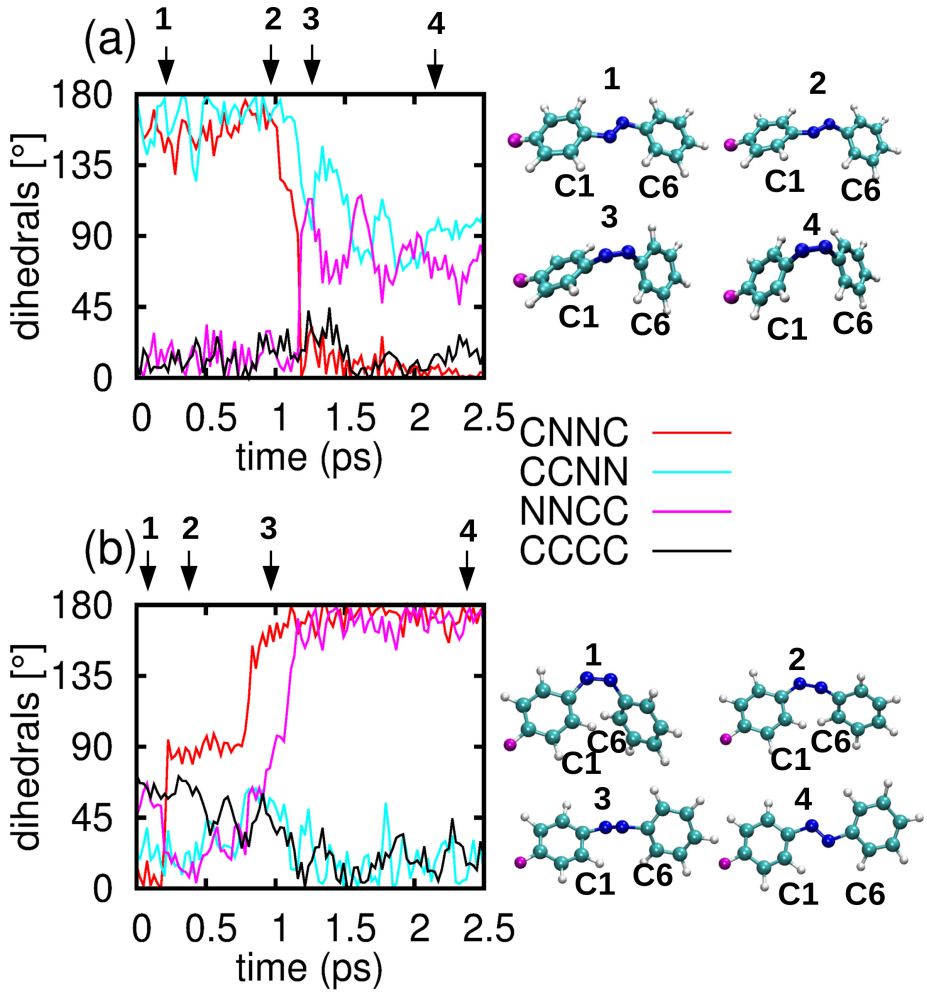


Figure S9: Changes of different dihedrals (absolute value) and corresponding structure of azobenzene for a reactive trajectory starting from (a) *trans* and (b) *cis* azobenzene to explore the mechanism of photoisomerisation of Azo-RNA2.

The *cis*-to-*trans* isomerisation (Figure S9) occurs via an N-N twist followed by a N4-C5 rotation. In this case, the change in the NNCC dihedral is larger than the change of the CCNN dihedral. The azobenzene chromophore reorients by rotating the distant phenyl ring, keeping the ring connected to the backbone almost fixed.

For a clearer picture, the dihedrals in Figure S9 are plotted in the range of [-180:180] with their corresponding signs in Figure S10. Here, we will mainly focus on *cis*-to-*trans* isomerisation and the selection of pathways in terms of chirality conservation and chirality inversion.

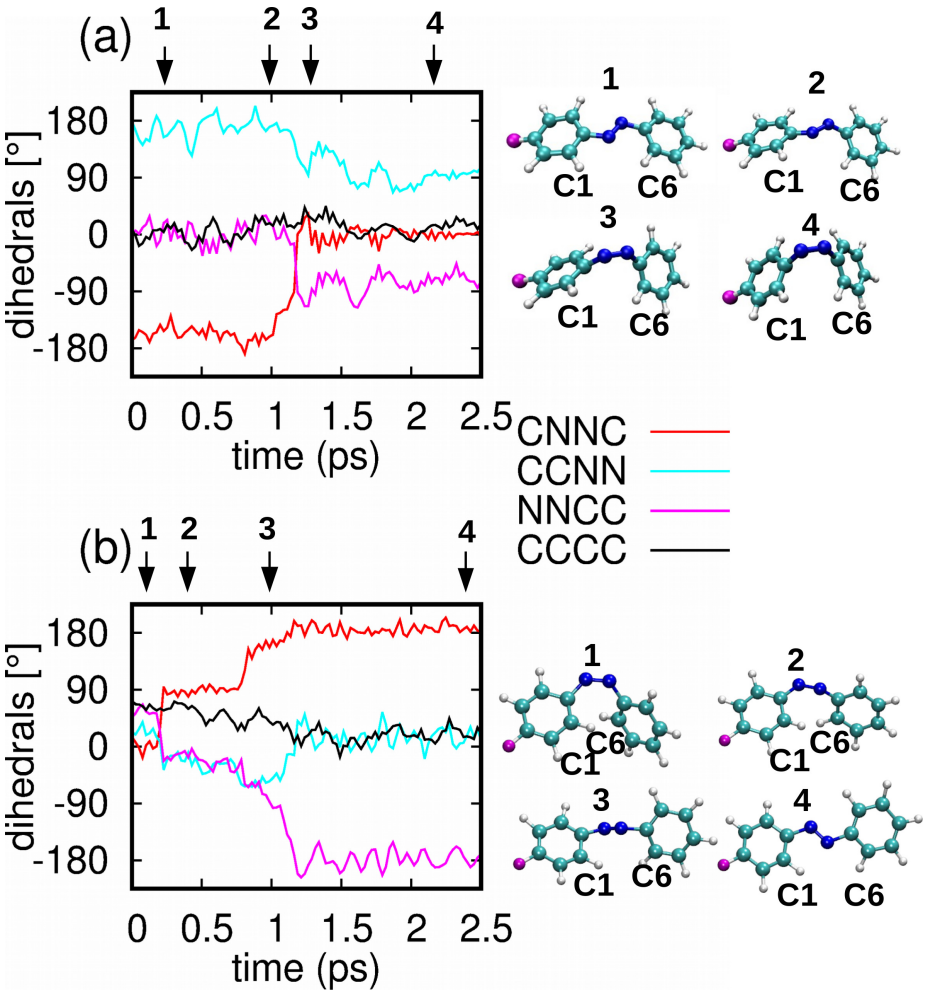


Figure S10: Changes of different dihedrals (with signs) and corresponding structure of azobenzene for a reactive trajectory starting from (a) *trans* and (b) *cis* azobenzene to explore the mechanism of photoisomerisation of Azo-RNA2.

The CNNC dihedral in Figure S10, starting from a small positive value, goes to $+180^\circ$ via $+90^\circ$ which indicates the selection of the P-type pathway during isomerisation of *cis* azobenzene in Azo-RNA2. Also, the CCNN and NNCC dihedrals, starting from a positive value approach to a negative value via 0°, again confirming the selection of the P-type pathway. Therefore, the *cis*-to-*trans* isomerisation of azobenzene in Azo-RNA2 proceeds via a chirality conserving pathway, starting from a P-type intial structure. Overall, we may conclude that the *cis* azobenzene in two types of RNAs under consideration selects a chirality conserving pathway during the photoisomerisation.

S8. Concerted dynamics of dihedrals

Here, we detail the QM/MM-SH time evolution of the CNNC dihedral and the two CNN angles. The evolution of these dihedral angles during the photoisomerisation is shown in the following figures, for averages over the total, reactive and unreactive trajectory sub-ensembles.

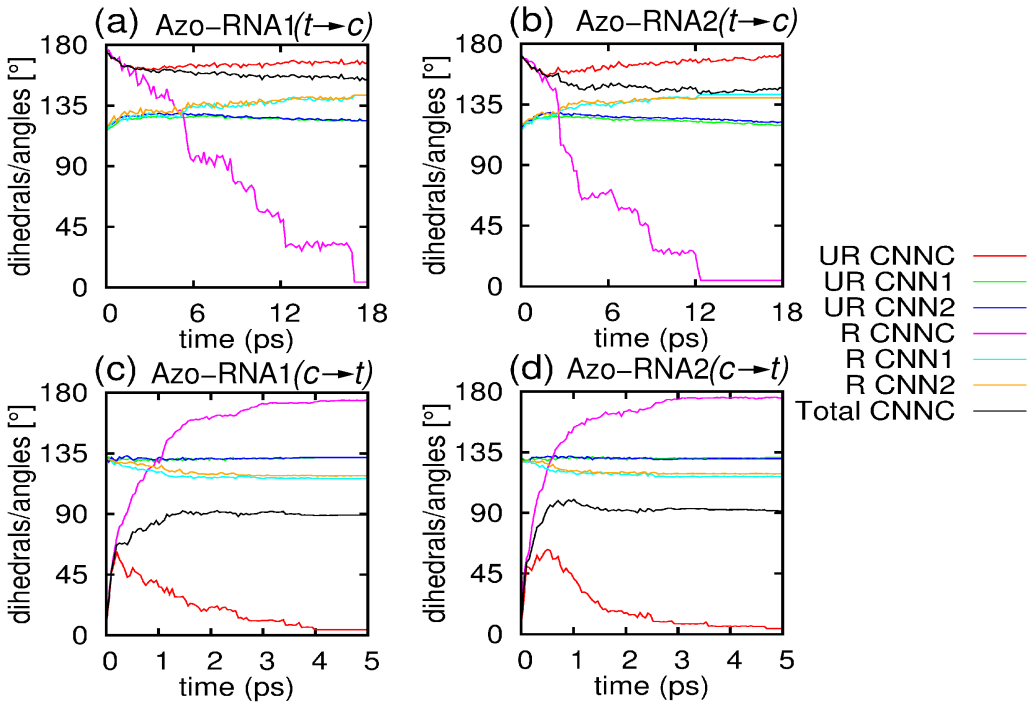


Figure S11: Changes of absolute CNNC and CNN angles during the dynamics starting from *trans* (a,b) and *cis* (c,d) azobenzene for Azo-RNA1 (left panels) and Azo-RNA2 (right panels). In the legends, UR and R refer to unreactive and reactive trajectories, respectively.

As observed in Figure S11a for *trans*-to-*cis* photoisomerization in Azo-RNA1, the absolute value of the CNNC dihedral angle is changing from about 180° to 0° for the reactive trajectories, while it remains within the 165°-180° range for the unreactive ones. In Figure S11c, concerning the *cis*-to-*trans* conversion in Azo-RNA1, the opposite occurs: the averaged CNNC angle goes from 0° to about 180° for the reactive trajectories, while it reaches a maximum of about 60° and then goes back to 0° for the unreactive ones. Since in this case the isomerisation quantum yield is around 50%, the average over all trajectories stabilizes around 90°. The corresponding picture of Figures S10b and S10d for Azo-RNA2 is very similar, indicating a similar mechanism of photoisomerisation despite the differences in RNA sequences.

Both CNN angles (labelled CNN1 and CNN2 in the figures) remain between 115° and 150°, which is consistent with a photoisomerisation pathway not involving inversion or concerted inversion. While both CNN angles are changing from 120° to 145° for the reactive trajectories, they are oscillating around 120° for the unreactive ones.

Figure S12 (for Azo-RNA1) and Figure S13 (for Azo-RNA2) show the averaged changes of the two dihedrals *C*1-*C*2-*N*3-*N*4 and *N*3-*N*4-*C*5-*C*6, which describe the motion of the two phenyl rings of azobenzene relative to the azo bridge. The other complementary dihedrals are automatically determined. As further explained in Section S6 and in the main text, these angles also inform about the helicity (i.e., P-helical or M-helical) of the *cis* form of the chromophore. Notably, the initial *cis* form of Azo-RNA1 shown in Figures S12c/d is of M-helical type (since -90°$<$ *N*3-*N*4-*C*5-*C*6 $<$ 0°), while the initial *cis* form of Azo-RNA2 shown in Figures S13c/d is of P-helical type (since 0°$<$ *N*3-*N*4-*C*5-*C*6 $<$ 90°).

In Figure S12c for the reactive trajectories starting from *cis*-Azo-RNA1, the dihedral *N*3-*N*4-*C*5-*C*6 starts from a negative value and increases to a positive value. Meanwhile, the dihedral *C*1-*C*2-*N*3-*N*4 also increases from a negative value to a positive value, but to a lesser extent. For the unreactive trajectories (Figure S10d), the changes of these dihedrals are negligible. For both cases, the dihedral *N*3-*N*4-*C*5-*C*6 is changing significantly during the isomerisation, which implies that the distant phenyl ring (seen from the perspective of the linker) undergoes more extensive reorientation during the isomerisation, which includes the rotation around the *N*4-*C*5 bond. This is consistent with the observation that the distant phenyl ring is free to reorient, differently from the ring directly bound to the linker.

In a complementary fashion, Figure S13 shows the changes of the *C*1-*C*2-*N*3-*N*4 and *N*3-*N*4-*C*5-*C*6 dihedrals for Azo-RNA2, starting from the *trans* and *cis* isomers and averaged over the reactive and unreactive subensembles. In Figure S13c and Figure S13d for the initial *cis*-Azo-RNA2 form, the dihedrals *C*1-*C*2-*N*3-*N*4 and *N*3-*N*4-*C*5-*C*6 start from a positive value (i.e., the P-helical form, as discussed above). Then both dihedrals decrease and approach to a negative value via 0° for the reactive trajectories, while they remain almost constant for the unreactive trajectories.

From the above observations, we summarize that (i) the initial *cis* forms of Azo-RNA1 and Azo-RNA2 are of M-helical and P-helical type, respectively, (ii) in the course of the dynamics starting from the respective *cis* isomers, the chromophores undergo a chirality conserving pathway, as can be inferred from the change in the NNCC dihedral as well as the CNNC dihedral. In the main text, a complementary discussion is given.

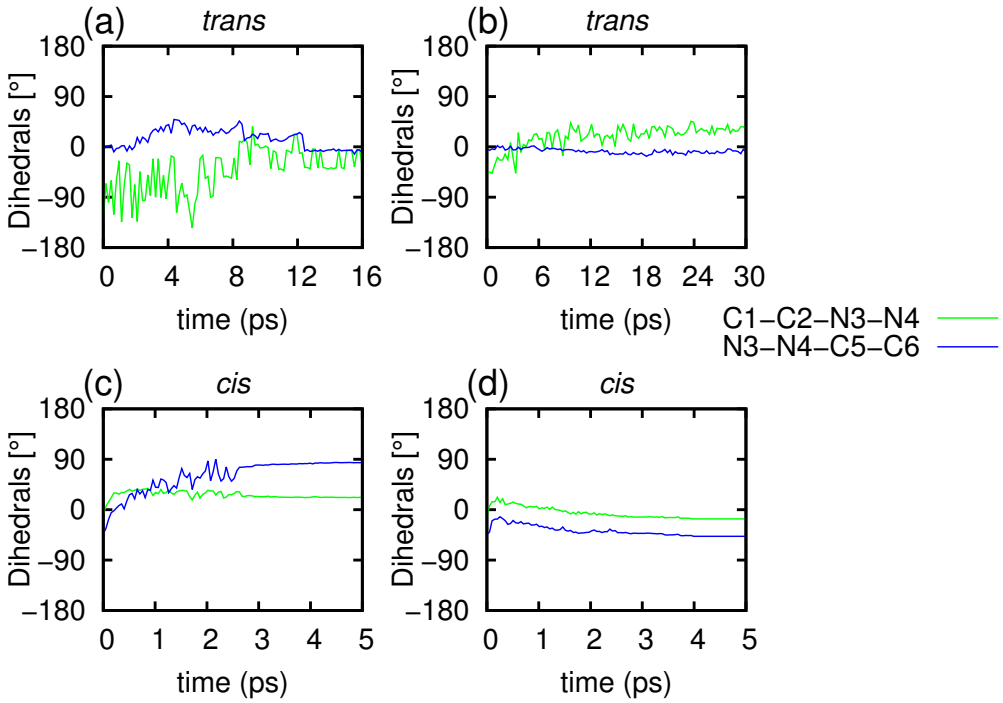


Figure S12: Changes of *C*1-*C*2-*N*3-*N*4 and *N*3-*N*4-*C*5-*C*6 dihedral angles for (a, c) reactive and (b, d) unreactive trajectories for Azo-RNA1 starting from the *trans* and *cis* isomers.

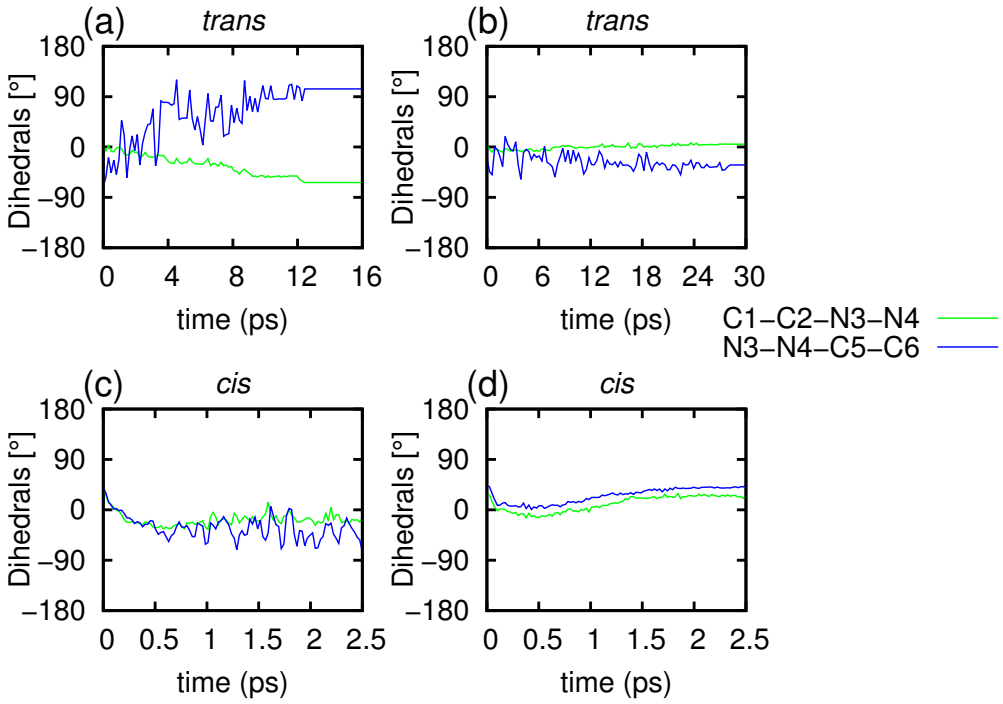


Figure S13: Changes of *C*1-*C*2-*N*3-*N*4 and *N*3-*N*4-*C*5-*C*6 dihedrals for (a, c) reactive and (b, d) unreactive trajectories for Azo-RNA2 starting from the *trans* and *cis* isomers.

References

(1) Goldau, T.; Murayama, K.; Brieke, C.; Steinwand, S.; Mondal, P.; Biswas, M.; Burghardt, I.; Wachtveitl, J.; Asanuma, H.; Heckel, A. *Chem. Eur. J.* **2015**, *21*, 2845–2854.

(2) Mondal, P.; Biswas, M.; Goldau, T.; Heckel, A.; Burghardt, I. *J. Phys. Chem. B.* **2015**, *119*, 11275–11286.

(3) Ponder, J. W.; Richards, F. M. *J. Comput. Chem.* **1987**, *8*, 1016–1024.

(4) Rastädter, D.; Biswas, M.; Burghardt, I. *J. Phys. Chem. B* **2014**, *118*, 8478–8488.

(5) Creatini, L.; Cusati, T.; Granucci, G.; Persico, M. *Chem. Phys.* **2008**, *347*, 492–502.

(6) Cantatore, V.; Granucci, G.; Persico, M. *Comput. Theor. Chem.* **2014**, *1040-1041*, 126–135.

(7) Conti, I.; Garavelli, M.; Orlandi, G. *J. Am. Chem. Soc.* **2008**, *130*, 5216–5230.

(8) Cusati, T.; Granucci, G.; Martínez-Núñez, E.; Martini, F.; Persico, M.; Vázquez, S. *J. Phys. Chem. A.* **2012**, *116*, 98–110.

(9) Sponer, J.; Sponer, J.; Mladek, A.; Jurecka, P.; Banas, P.; Otyepka, M. *Biopolymers* **2013**, *99*, 978–988.

(10) Granucci, G.; Persico, M.; Toniolo, A. *J. Chem. Phys.* **2001**, *114*, 10608–10615.

(11) Ciminelli, C.; Granucci, G.; Persico, M. *Chem. Eur. J.* **2004**, *10*, 2327–2341.