ChemPhotoChem

Research Article
doi.org/10.1002/cptc.202400093

Chemistry Europe
European Chemical Societies Publishing

www.chemphotochem.org

Preparation of Light-responsive Unnatural RNA Bases via a Chromogenic Morita-Baylis-Hillman Adduct Path

Matteo Lami,[a] Leonardo Barneschi,[a, b] Mario Saletti,[a] Massimo Olivucci,[a, c] Andrea Cappelli,[a] and Marco Paolino*[a]

RNA-based tools for biological and pharmacological research are raising an increasing interest. Among these, RNA aptamers whose biological activity can be controlled via illumination with specific wavelengths represent an important target. Here, we report on a proof-of-principle study supporting the viability of a systematic use of Morita-Baylis-Hillman adducts (MBHAs) for the synthesis of light-responsive RNA building blocks. Accordingly, a specific acetylated MBHA derivative was employed in the functionalization of the four natural RNA bases as well as two unnatural bases (5-aminomethyl uracil and 5-methylamino-methyl uracil). The results reveal a highly selective functionalization for both unnatural bases. The conjugation products were then investigated spectroscopically, photochemically and computationally. It is shown that when a single light-responsive unit is present (i.e. when using 5-methylaminomethyl uracil), the generated unnatural uracil behaves like a cinnamic-framework-based photochemical switch that isomerizes upon illumination through a biomimetic light-induced intramolecular charge transfer mechanism driving a barrierless and, therefore, ultrafast reaction path.

Introduction

The five nitrogenous bases adenine, guanine, thymine, cytosine, and uracil characterize the nucleotides of DNA and RNA.[1] Of these bases, uracil is only present in RNA.[2] While over the years, DNA has represented the target of numerous therapeutic agents (in particular antineoplastic agents),[3–5] RNA has been recognized to be a more difficult target,[6] and only recently has started to play an important role in the pharmaceutical world.[7] Furthermore, RNA is no longer considered only a "target" but, above all, a precise and specific "magic bullet".[8,9] In this regard, antisense oligonucleotides, siRNA, mRNA, and RNA aptamers are being tested and used in the treatment of various diseases when conventional pharmacological agents fail or when a more personalized medicine practice is desirable.[10–12]

In the above context, some research groups are moving to implement RNA therapies with systems for controlling the pharmacological effect.[13] This can occur through controlled and targeted delivery,[14] or external stimuli such as changes in pH or temperature,[15,16] or RNA-ligand interactions.[17,18] In the last case, if the ligand corresponds to a molecular photoswitch, it is possible to obtain a pharmacological response modulated by illumination.[19–21] This concept has led to promising designs where, in principle, the activation of the pharmacological effect is highly specific and minimally invasive.[20] However, the limitation of such RNA-ligand approach lies in the achievable affinity constants dictated by the high instability of the single strand of RNA.[22] Accordingly, a different approach where the ligand is covalently bound to the RNA is highly desirable.

During the last few years, we have developed acetylated Morita Baylis Hillman adduct (MBHA) derivatives (Figure 1) capable of reacting with nucleophilic residues to afford covalently linked chromophoric ligands (i.e. chromophores) that respond to illumination.[23–27] In particular, the reaction of MBHA derivatives with imidazole[23] or butylamine[27,28] (model compounds of histidine and lysine residues, respectively) allowed us to generate, through the nucleophilic addition to the acrylic double bond followed by the elimination of the acetyl group, chromophores with push-pull electronic structures.[29] These are characterized by an electron-withdrawing cinnamic ester derivative where a carbonyl is conjugated through an exocyclic double bond with an electron-donating aromatic system. Upon light irradiation, some of these products showed interesting fluorescence properties that, in some instances, were amplified in the aggregated states (AIE property),[23,24,30] while others were found to be capable of undergoing a photoinduced trans-cis (or *E*/*Z*) isomerization.[27,28]

Differently from many other conjugation techniques of light-responsive chromophores, the use of MBHA derivatives represents an *in situ* chromogenic strategy compatible with the biological environment that occurs through an addition/elimination reaction definable as a "click"[31] and does not require catalysts. This makes MBHA derivatives extremely versatile tools for chemical biology.

[a] *Dr. M. Lami, Dr. L. Barneschi, Dr. M. Saletti, Prof. Dr. M. Olivucci, Prof. Dr. A. Cappelli, Prof. Dr. M. Paolino*
Department of Biotechnology, Chemistry and Pharmacy
University of Siena
Via A. Moro 2, 53100 Siena, Italy
E-mail: paolino3@unisi.it

[b] *Dr. L. Barneschi*
Department of Chemistry, Biology and Biotechnology
University of Perugia
Via Elce di Sotto 8, 06123 Perugia, Italy

[c] *Prof. Dr. M. Olivucci*
Chemistry Department
Bowling Green State University
Overmann Hall, Bowling Green, OH 43403, USA

Supporting information for this article is available on the WWW under https://doi.org/10.1002/cptc.202400093

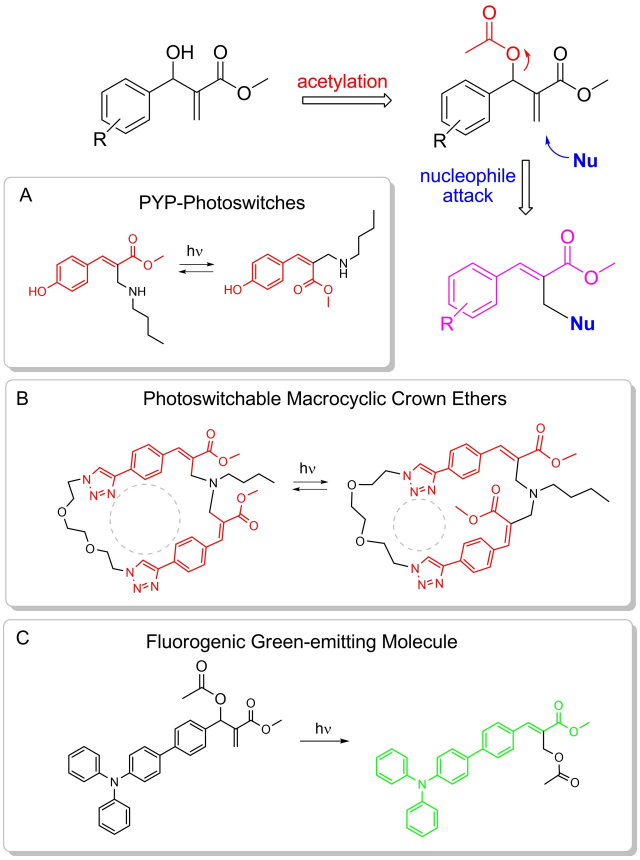


Figure 1. Nucleophilic attack to a generic MBHA derivative and application of this conjugation reaction leading to cinnamic-based photoswitches (A and B) or fluorescent molecules (C).

Recently we have shown that MBHA derivatives can be used to functionalize proteins in specific sites.[26,27,32] However, while proteins contain nucleophilic amino acid side chains (e.g., cysteine, lysine, arginine, and histidine), the four RNA bases (guanine, adenine, cytosine, and uracil) have a weaker nucleophilic character and, therefore, it is still not clear if MBHA derivatives can be employed in their direct functionalization. For this reason, in the present work, we explore the reactivity of a specific acetylated MBHA derivative **1**[33] towards the RNA bases as well as towards the unnatural nitrogenous bases 5-aminomethyl uracil (**5-AMU**)[34] and 5-methylaminomethyl uracil (**5-MAMU**)[35] already used to prepare synthetic RNA (Figure 2).[36,37]

We find that while **1** shows a poor reactivity with natural bases, it reacts in mild conditions with **5-AMU** and **5-MAMU** forming the expected cinnamic chromophore demonstrating a realistic, although indirect, route towards achievement of light-responsive RNAs. The obtained products, featuring a light-responsive double-bond as part of a cinnamic chromophore, were spectroscopically and photochemically studied both experimentally and computationally.

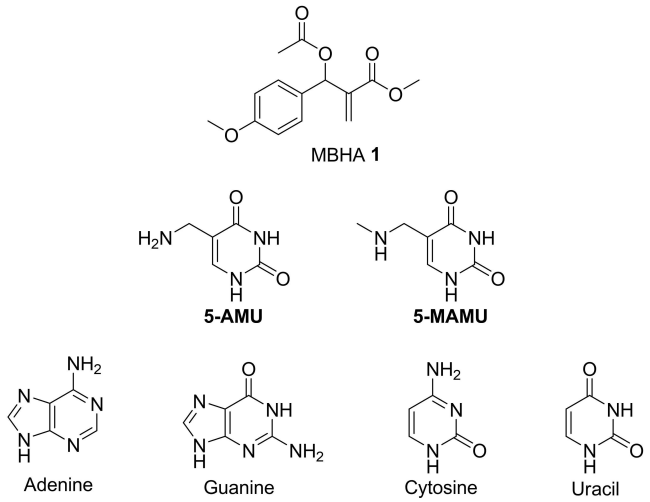
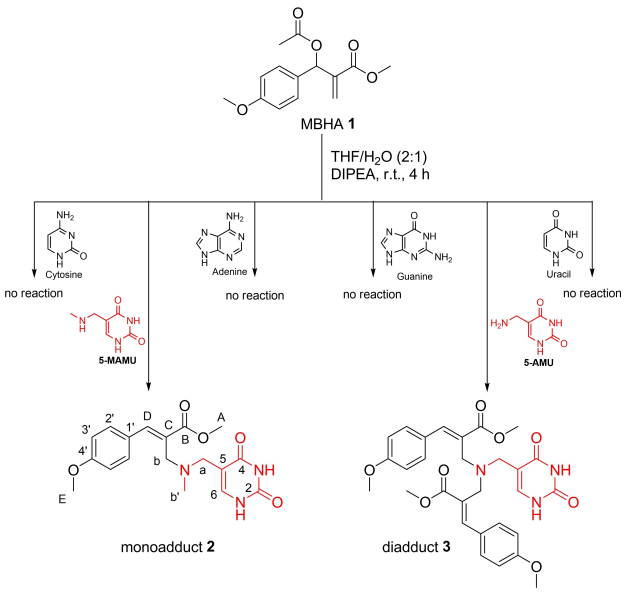


Figure 2. Structure of the selected MBHA derivative **1**, unnatural bases **5-AMU**, **5-MAMU** and natural bases adenine, guanine, cytosine, and uracil.

Results and Discussion

Design, Synthesis, and Chemical Characterization

Compound **1** was prepared by condensation of methyl acrylate and *p*-anisaldehyde[38] and subsequent acetylation of the corresponding MBHA derivative following the synthetic procedure reported in the literature (see supporting information for details).[33] The reactivity of **1** towards each of the four natural RNA bases was evaluated by solubilizing both reagents in a water and THF (1:2) mixture in order to ensure their solubility. Each reaction was monitored at room temperature for five days without observing any reactivity even after the addition of *N,N*-diisopropylethylamine (DIPEA) as the base (Scheme 1). The



Scheme 1. Reactivity of MBHA derivative **1** with the natural RNA (adenine, guanine, cytosine, and uracil) and unnatural (**5-MAMU** and **5-AMU**) bases.

confirmation of the poor nucleophilic character of natural RNA bases towards MBHA derivative **1** in mild conditions could represent the starting point for the optimization of site-specific derivatization procedures on RNA strands. In fact, similarly to what has already been developed in the past on proteins,[39–41] we assumed that unnatural bases containing nucleophilic functional groups could open a more viable route to the selective modification of synthetic RNAs by MBHA derivatives.

Therefore, we explored the reactivity of **1** towards **5-MAMU**, bearing a secondary alkyl amino group, and **5-AMU** bearing a primary alkyl amino group. Compound **5-MAMU** was prepared starting from commercially available 5-chloromethyluracil and methylamine, while **5-AMU** was prepared as hydrochloride salts using the same starting material and hexamethylenetetramine followed by acid hydrolysis of the quaternary ammonium salt according to the Delépine reaction (see supporting information for the details).[42] The reactivity of **1** towards **5-MAMU** and **5-AMU** was then studied, similarly to what was done previously, at room temperature in a mixture of water and THF (1 : 2) and in the presence of DIPEA as a base. As hoped, **1** reacted with **5-MAMU** and **5-AMU** in four hours producing monoadduct **2** on the secondary amine of **5-MAMU** or the diadduct **3** on the primary amine of **5-AMU** with the apparent absence of side reaction on the uracil ring (Scheme 1).

The structures of **2** and **3** were characterized by NMR spectroscopy (Figure 3, S1 and S2) and HRMS (Figure S3 and S4) confirming the formation of cinnamic chromophores covalently linked to the aminic function. The 1D and 2D NMR analysis of **2** supported an *E* configuration of the cinnamic residue. In particular, NOESY experiments highlighted dipolar contacts between H-b (3.43 ppm) and H-2′ (7.63 ppm) possible only for the *E* diastereoisomer (see Scheme 1 for the numbering). In addition, ^{1}H NMR spectrum of **2** shows the signal attributed to H–D significantly shifted to down-field (7.78 ppm) (Figure 3). This was in agreement with the *E* configuration of the exocyclic double bond in which the H–D proton was found in the de-shielding cone of the carbonyl of the ester function. The same type of down-field shift of the H–D protons is present in the ^{1}H NMR spectrum of **3** (Figure 3). However, despite the structural similarity, the signals of the aromatic protons are significantly shifted up-field compared to the analogous protons of the **2**, a sign of a possible intramolecular π-π interaction (Figure 3). Furthermore, the ^{1}H NMR spectrum clearly revealed the presence of only one diastereoisomeric form for the two cinnamic residues, allowing the attribution of the *E* configuration for both its chromophores. In addition, NOESY experiments showed contacts between H-b (3.55 ppm) and H-2′ (7.60 ppm) also in **3** (Figure 3).

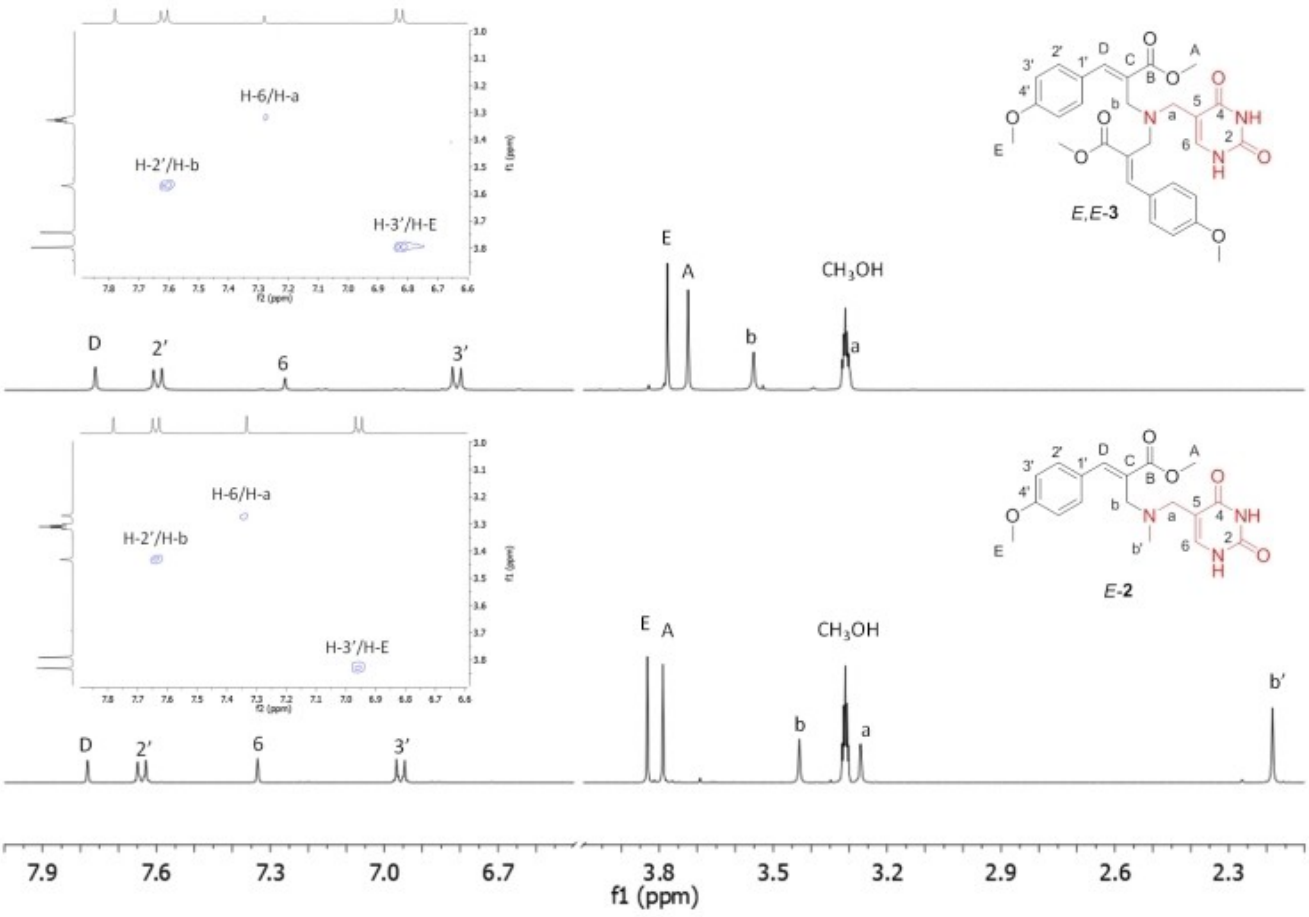


Figure 3. Comparison of ^{1}H NMR (400 MHz, CD_3OD) spectra of *E*-**2** and *E,E* -**3**. The insets show the NOESY H-b/H-2′ contacts, used to attribute the *E* configuration of the exocyclic double bonds.

To further confirm the selectivity of **1** towards the **5-MAMU** and **5-AMU** amine nitrogen, the conjugation reaction was carried out under the same previously optimized reaction conditions but adding a stoichiometric amount of each natural base to the reaction mixture. After 4 hours, only the formation of the **2** (when using **5-MAMU**) and **3** (when using **5-AMU**) with the *E* configuration were observed, excluding any possible electrophilic attack towards the natural RNA bases.

Photophysical and Photochemical Characterization

We demonstrated the responsiveness to light of the achieved chromophores experimentally and elucidate the corresponding geometrical structures and double-bond isomerization mechanism computationally. As shown in Figure 4A, the UV-vis absorption spectrum of *E*-**2** in methanol (1×10^{-5} M) showed the presence of a dominant band centered at ca. 305 nm. Similarly, the spectrum of *E,E*-**3** showed, in the same conditions, a similar absorption band this time showing a non-linear increase of the molar extinction coefficient. Such non-linear increase supports the hypothesis of π-π interaction between the chromophoric residues of **3** (Figure 4).[43–46]

The two compounds were irradiated with monochromatic light centered at 330 nm or 310 nm to probe their reactivity. This was followed via both UV-Vis and/or ^{1}H NMR spectroscopies.

Firstly, *E*-**2** in deuterated methanol (ca. 1×10^{-2} M) was irradiated in a 5 mm NMR Pyrex tube until a photostationary state was reached. The irradiation produced *Z*-**2** which constituted the 62% and 54% of the sample when using 330 or 310 nm light, respectively (Figure 5). This *Z*/*E* ratio was obtained from the integral values of the signals attributed to the same protons in the two diastereomeric species. The formation of *Z*-**2** is supported by a H–D signal shifted to 6.72 ppm as a consequence of the displacement out from the de-shielding cone of the ester carbonyl. The *Z* configuration was also confirmed by the appearance of a new contact between H–D and H-b evident in the NOESY spectrum registered after irradiation (Figure 5).

As shown in Figure 4B, the photoconversion of *E*-**2** was also monitored by UV-vis spectroscopy. A solution of *E*-**2** in methanol (1×10^{-5} M) was constantly irradiated in a quartz cuvette using an optical fiber emitting monochromatic UV light at 310 nm and the absorption spectrum was measured at fixed four minutes intervals. The *E*-**2** to *Z*-**2** interconversion is in-line with the progressive decrease of the band centered at 305 nm and the appearance of a new band centered at ca. 277 nm.

To demonstrate the reversibility of the photoconversion without net loss of *E*-**2**, the photostationary state obtained at 310 nm and monitored via NMR was irradiated with a monochromatic light centered at 275 nm producing the decrease of the *Z*-**2** and an increase *E*-**2** signals respectively, until a new photostationary state consisting of 62% of the latter was obtained (Figure 5).

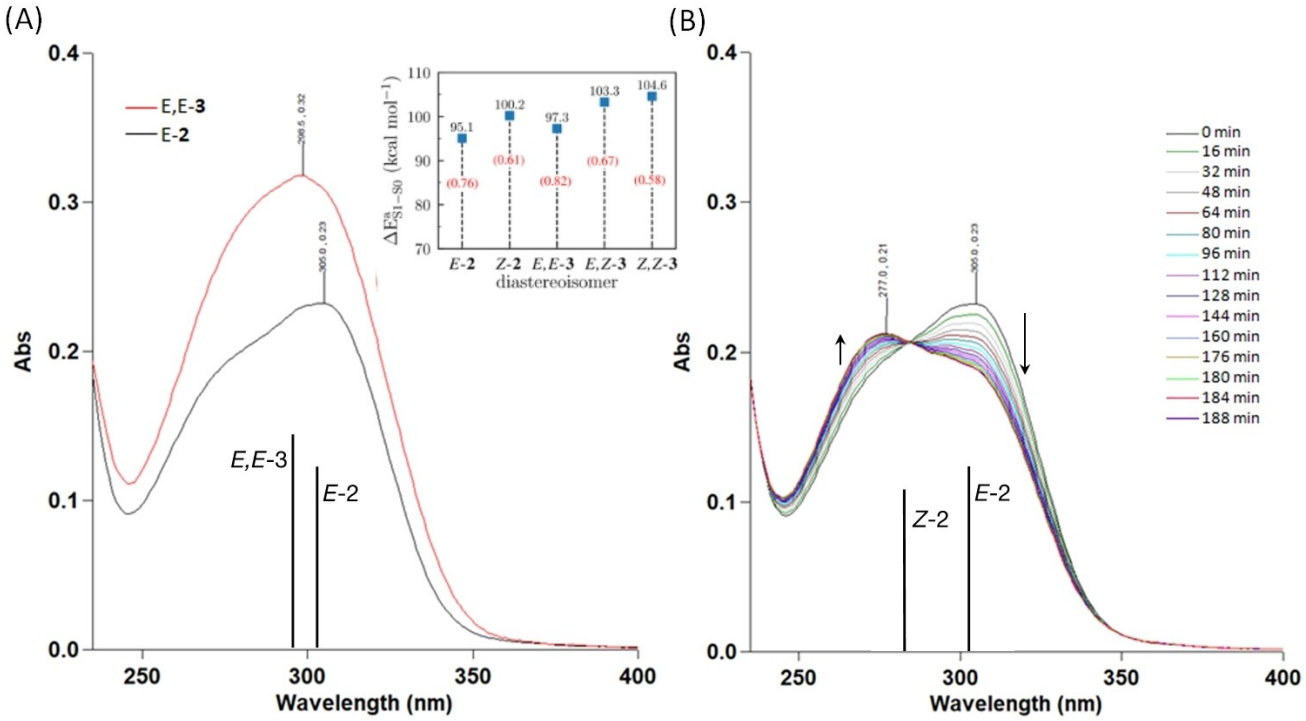


Figure 4. (A) UV-Vis absorption spectra of compounds *E*-**2** and *E,E*-**3** (1×10^{-5} M in methanol). Vertical excitation energies (ΔE^{a}_{S1-S0}, the values in black are given in kcal/mol) and oscillator strengths (f, values in red) obtained at the TDDFT/M06-2X/def2-TZVP level for the five molecules studied in the present work are given in the inset. The calculated excitation energy position in nm and the relative intensity for *E*-**2** and *E,E*-**3** are also displayed below the corresponding experimental bands. (B) Evolution of the absorption spectra of compound *E*-**2** (1×10^{-5} M in methanol) under continuous irradiation with a light at 310 nm. The calculated excitation energy position in nm and the relative intensity for *E*-**2** and *Z*-**2** are also displayed below the corresponding experimental bands.

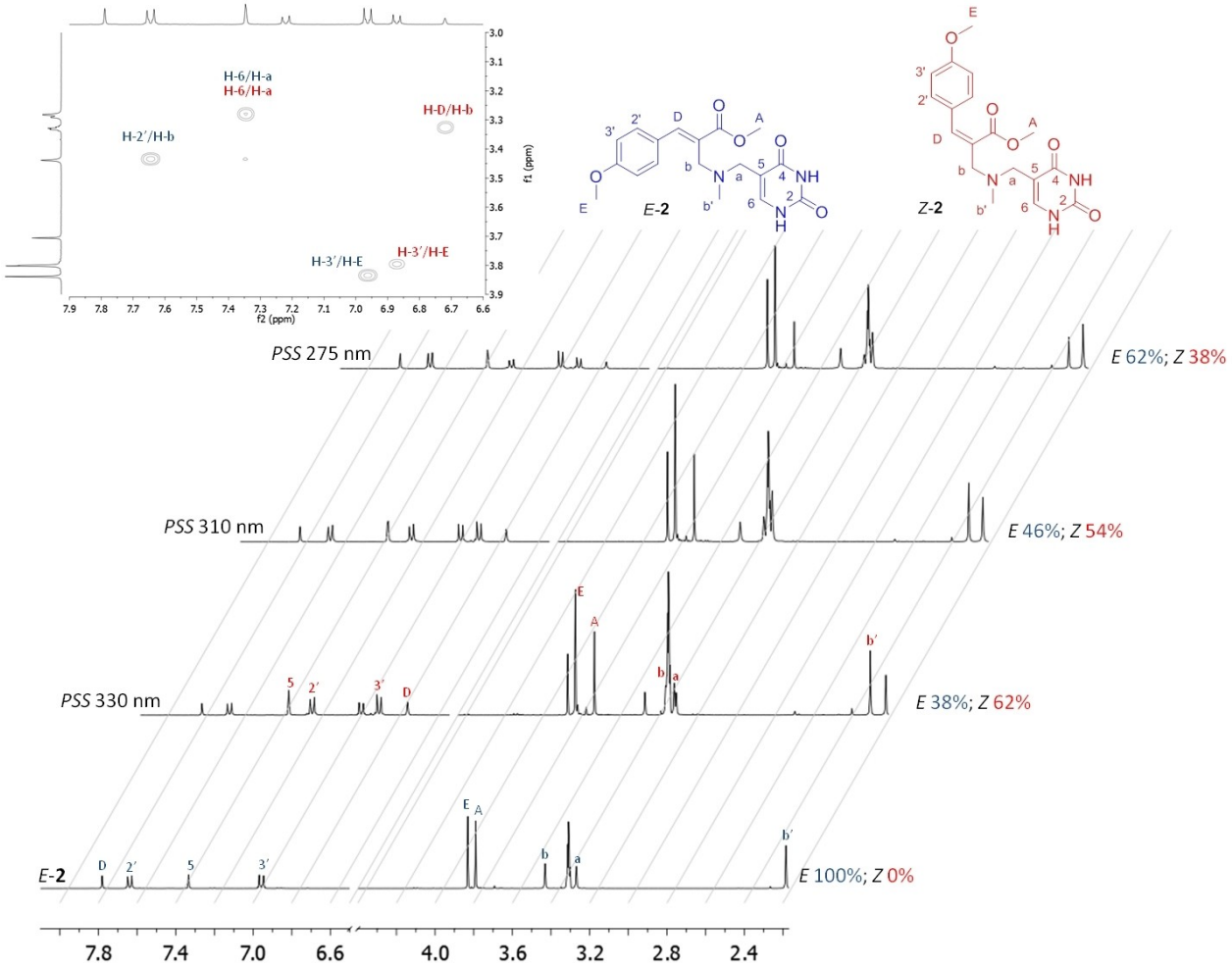


Figure 5. Comparison of ^{1}H NMR spectrum of *E*-**2** with those recorded with different PSS obtained by irradiation at 330, 310, and 275 nm. The insert show the NOESY contacts, used in the attribution of the *E* and *Z* configuration of the double bond.

Contrary to the high photostability shown by the mono-adduct **2**, the irradiation of the diadduct *E*,*E*-**3** in the same conditions led to the formation of several chemical species. In particular, as suggested by ^{1}H NMR spectra of the kinetic experiment reported in Figure S5, the photoconversion of the chromophores from *trans* to *cis* configuration is accompanied by the formation of a complex mixture (here not characterized) of chemical species compatible with a [2 + 2] photocycloaddition reaction which is known to occur in cinnamate esters for specific setups.[47] This poor photostability of the diadduct *E*,*E*-**3** could be motivated to the presence of the π-π intramolecular interaction between the aromatic moieties of the two chromophoric residues capable of bringing the acrylic double bonds closer together in space, making them readily available for the photocycloaddition reaction.[48]

In order to get geometric and mechanistic information on the observed photoreaction of *E*-**2**, we calculated the relative Gibbs free energies of the two diastereoisomers of **2** and three diastereoisomers (*E*,*E*, *E*,*Z*, and *Z*,*Z*) of **3** in isolated conditions. As reported in the Computational Methods section, for each isomer we performed a sampling of the conformational space using the Conformer-Rotamer Ensemble Sampling Tool (CREST)[49] based on the fast semi-empirical extended tight-binding method GFN2-xTB.[50] The lowest energy conformer produced by CREST was then relaxed using density functional theory (DFT),[51] using with the B3LYP functional, empirical dispersion correction D3BJ[52,53] and the def2-SVP basis set.[54] The located energy minima were confirmed via frequency calculation at the same level of theory. Such calculation was also used to determine Gibbs free energies differences. For **2**, we found that the identified *E*-**2** isomer is 4.4 kcal mol^{-1} more stable than *Z*-**2**, while for **3** we observe that the *E*,*E*-**3** diastereoisomer is 6.3 and 4.7 kcal mol^{-1} more stable than the *E*,*Z*-**3** and *Z*,*Z*-**3** isomers, respectively. For the five located energy minima we calculated the vertical excitation energy (ΔE^{a}_{S1-S0}) and the corresponding oscillator strengths (f) with time-dependent density functional theory (TDDFT)[55] with the M06-2X functional[56] and def2-TZVP basis set[54] (see inset in Figure 4A).

Remarkably, the relative maximum absorption wavelengths (λ^{a}_{S1-S0}) obtained from the ΔE^{a}_{S1-S0} values agree with those observed experimentally. In fact, we compute for *E*-**2**, *Z*-**2** and

E,E-**3** a λ^{a}_{S1-S0} of 300.7 nm, 285.3 nm and 293.9 nm, respectively that reproduces the observed trend of 305.0 nm, 277.0 nm and 296.5 nm and thus support the isomer assignment (see also Figure 4A and 4B).

While the computational results support the experimental assignment of the most stable structures, they don't provide information on the mechanism allowing the spectral progression of Figure 4B. In fact, this would require complex non-adiabatic dynamics simulation of the *Z*-**2**→*E*-**2** and *E*-**2**→*Z*-**2** photoisomerizations[57] as well as the calculation of the activation $\Delta G^{\neq}$ for the two corresponding thermal isomerizations that are beyond the scope of the present work. However, as we discuss in the next paragraph, approximate mechanistic information could already be obtained by studying the computed equilibrium geometries and, most importantly by a mapping of the S_1 potential energy profile along the isomerization coordinate. Accordingly, and in contrast with the expected structures for *E*-**2**, the dominant *E,E*-**3** isomer (see Figure 6A) clearly exhibits a π-stacking interaction of the aromatic moieties, suggesting that this spatial orientation, which brings the two cinnamic double bonds in close vicinity, might lead to photoinduced reactions such as a $2\pi_s + 2\pi_s$ cycloaddition. This justifies the complexity observed in the corresponding NMR data presented above. No further investigation has been carried out on this isomer.

Finally, to qualitatively characterize the observed photochemical interconversion of *E*-**2** driving the spectral evolution and the formation of the stationary states (see Figure 4B) we computed the potential energy profile of the first excited state (S_1) along the corresponding initial relaxation dominated by a stretching coordinate (compare the geometrical parameters in Figure 6A, left with those at the top-right of Figure 4B) followed by the torsional deformation of the cinnamic double bonds assumed to be the reaction coordinates. Notice that the initial stretching relaxation corresponds to a limited displacement of the reaction coordinate. Since TDDFT (at least in its linear response formalism) lacks the multireference character necessary to produce the correct potential energy surface in the regions of S_1/S_0 conical intersection (photochemical transition state or funnel; $Coln_{S1/S0}$) region, we employed the recently reported mixed reference spin-flip time dependent density functional theory (MRSF-TDDFT or MRSF), which has been shown to generate $Coln_{S1/S0}$ topologies and topographies with accuracy comparable to that of higher level multireference methods.[58,59] As shown in Figure 6B, it has been possible to optimize the conical intersection $Coln_{S1/S0}$ driving the isomerization reaction (left) and featuring a fully broken (ca. 97 degree twisted) cinnamic π-bond. Accordingly, and consistently with the energy profiles reported in the right panel of the same figure, the relaxation starting from the FC point of *E*-**2** initially generate a geometrically unlocked structure where the cinnamic double bond is now a single bond as also described by the initial changed in geometry described above. The geometrical evolution (see the methods section for details on the construction of the reaction coordinates and energy profiles) then continues along the torsional coordinate associated to the same bond until the $Coln_{S1/S0}$ is entered and, after decay to S_0,

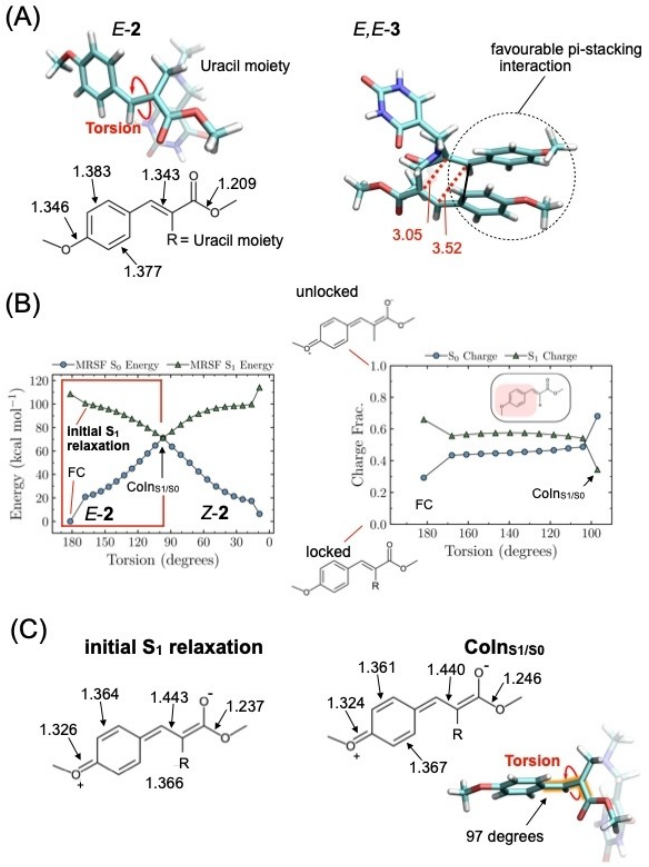


Figure 6. (A) Computed geometry of the dominating monoadduct *E*-**2** (left) and diadduct *E,E*-**3** (right) optimized at the B3LYP-D3BJ/def2-SVP level of theory. Relevant geometrical parameters (in Å, values computed at the MRSF/BH&HLYP/6-31G* level) are given in the skeletal formula at the bottom of structure *E*-**2**. The distances between the cinnamic double-bond carbons of the two chromophoric moieties in *E,E*-**3** are given in red. (B) Left. S_1 and S_0 energy profiles along the *E*-**2**→*Z*-**2** photoisomerization coordinate (right) computed at the MRSF/BH&HLYP/6-31G* level of theory (see also text). Right. Charge progression of the red-framed moiety along the computed *E*-**2**→$Coln_{S1/S0}$ branch of the same *E*-**2**→*Z*-**2** photoisomerization path (see framed region of the energy profile diagram on the left part). Notice the rapid exchange of charges expected in the vicinity of the degeneracy at a conical intersection. The resonance formulas showing the limiting electronic structures (resonance formulas) dominating the ground and reactive excited state are also given. (C) Left. Skeletal formula illustrating the change (compare the values with the structure of part A, left corresponding to the structure of the Franck-Condon (FC) point) in geometrical parameters (in Å) occurring during the initial bond-lengths relaxation from FC after a limited torsional deformation along the S_1 potential energy surface. Right. $Coln_{S1/S0}$ geometry associated to *E*-**2**→*Z*-**2** photoisomerization and corresponding relevant geometrical parameters (in Å and degrees). Both geometries correspond to the geometries marked of the energy profile in part B, left and were computed at the MRSF/BH&HLYP/6-31G* level of theory.

the *Z*-**2** product is generated. The S_1 energy profile is barrierless and the reaction is thus predicted to occur in a subpicosecond time-scale. This qualitatively agrees with previously reported studies on the *E* to *Z* gas-phase photoisomerization of similar chromophores.[60,61] Most notably, Nomoto et al. performed very accurate multireference calculation for *E* to *Z* photoisomerization of a *para*-methoxy methylcinnamate (pMMC) model in methanol, reporting a quasi-barrierless isomerization path.[62] As reported in the supporting information a similar geometrical

evolution and energy profiles have been documented for the opposite *Z*-**2**→*E*-**2** photoisomerization process (Figure S6). These findings indicates that the *E* and *Z* configurations of monoadduct **2** rapidly interconvert upon illumination to generate the stationary state and behaves like a light-driven molecular switch.

The agreement with previous simulations suggests that the uracil moiety only marginally interact with the aromatic moiety of monoadduct **2**. This was also confirmed by recomputing the *E*-**2**→*Z*-**2** isomerization energy profile after substitution of the uracil group with a methyl group (Figure S7), resulting in a negligible difference with the original path.

As reported in Figure 6C, the evolution of the charge on the aromatic moiety along the S_1 branch of the *E*-**2**→*Z*-**2** isomerization coordinate indicates that the spectroscopically bright state features a charge-transfer character associated to a "quinone-like" resonance formula, which is necessary to unlock and elongate the cinnamic double bond (see above) and promote isomerization.

Conclusions

In conclusion, we have shown that the conjugation of specific MBHA adducts to unnatural nitrogenous base represents a viable, although indirect, synthetic pathway for the preparation of unnatural RNA aptamers responding efficiently to a light stimulus and, for instance, potentially useful in the control of gene expression.[20,22] In fact, the demonstrated cinnamic chromophore generated after the electrophilic attack of the investigated MBHA derivative could act as a permanent and ultrafast on-off molecular switch. Much still needs to be done to assess the limits of the presented approach and, most importantly, to achieve a prototype library of light-responsive RNA aptamer that represents one of the main targets in our lab currently.

Experimental Section

Materials and Methods

Synthesis. All chemicals used were of reagent grade. Yields refer to purified products and are not optimized. Merck silica gel 60 (230–400 mesh) was used for column chromatography. Merck TLC plates and silica gel 60 F_{254} were used for TLC. Melting points were determined in open capillaries in a Gallenkamp apparatus and were uncorrected. NMR spectra were obtained with a Bruker 400 AVANCE spectrometer in the indicated solvents. The chemical shifts are referenced to the residual not deuterated solvent signal (CHD_2OD: δ (^1H) = 3.31 ppm, δ (^{13}C) = 49.86 ppm). The values of the chemical shifts are expressed in ppm and the coupling constants (*J*) in Hz. An Agilent 1100 LC/MSD operating with an electrospray source was used in mass spectrometry experiments, and a Bruker TimsTOF apparatus was used in recording HR-MS spectra. The absorption spectra were recorded with a Agilent Cary-60 in the indicated solvent. A tunable light source Zolix (TLS2-X300PU–G, 300 W UV Xenon Light Source with monochromator Omni-λ2047i) equipped with a fiber bundle (200–1100 nm, fibers 19 of 200 μn) was used for the irradiation of the solutions.

General Procedure for the Reaction of MBHA 1 with Natural and Unnatural Nitrogenous Bases

The appropriate nitrogenous bases (1 equivalent) was dissolved in a mixture of THF-H_2O (2:1 v/v). Then, MBHA acetate **1** (0.10 g, 0.38 mmol) and DIPEA (0.13 mL, 0.76 mmol) were added and the resulting mixture was stirred at room temperature for 4–24 h. The reaction mixture was monitored by TLC and ESI-MS observing the conversion of the MBHA **1** only in presence of the unnatural nitrogenous bases **5-MAMU** and **5-AMU**. With these, the solvents were concentrated under reduced pressure and the resulting residue was purified by flash chromatography to obtain pure amine derivative *E*-**2** or *E,E*-**3**.

(E)-Methyl 2-((((2,4-dioxo-1,2,3,4-tetrahydropyrimidin-5-yl) methyl) (methyl)amino) methyl)-3-(4-methoxyphenyl) acrylate (E-2)

The title compound was obtained starting from **5-MAMU** (0.072 g, 0.38 mmol) as white solid (85 mg, yield 62%, m.p. 126.0–127.0 °C) using ethyl acetate as the eluent. ^1H NMR (400 MHz, CD_3OD): 2.18 (s, 3H), 3.29 (s, 2H), 3.43 (s, 2H), 3.79 (s, 3H), 3.83 (s, 3H), 6.96 (d, *J* = 8.9, 2H), 7.33 (s, 1H), 7.64 (d, *J* = 8.7, 2H), 7.78 (s, 1H). ^{13}C NMR (100 MHz, CD_3OD): 43.1, 53.4, 54.2, 54.9, 56.7, 111.9, 115.8, 129.3, 129.7, 134.5, 143.1, 145.3, 154.4, 163.0, 167.5, 171.7. MS (ESI): *m/z* 360.1 ($M+H^+$). HR-MS: *m/z* ($M+H^+$) calcd for $C_{18}H_{22}N_3O_5^+$ 360.15540; found 360.15435.

(Z)-Methyl 2-((((2,4-dioxo-1,2,3,4-tetrahydropyrimidin-5-yl) methyl) (methyl)amino) methyl)-3-(4-methoxyphenyl) acrylate (Z-2)

NMR data of the title compound were derived after UV light irradiation of the corresponding *E* diastereoisomer. ^1H NMR (400 MHz, CD_3OD): 2.28 (s, 3H), 3.30 (s, 2H), 3.34 (s, 2H), 3.71 (s, 3H), 3.81 (s, 3H), 6.74 (s, 1H), 6.88 (d, *J* = 8.8, 2H), 7.23 (d, *J* = 8.7, 2H), 7.35 (s, 1H). ^{13}C NMR (100 MHz, CD_3OD): 43.5, 53.1, 53.7, 56.6, 63.4, 111.9, 115.6, 130.0, 131.8, 132.5, 136.2, 142.9, 154.3, 162.1, 167.5, 172.7.

(2E,2′E)-Dimethyl 2,2′-((2,4-dioxo-1,2,3,4-tetrahydropyrimidin-5-yl) methylazanediyl) bis (methylene) bis (3-(4-methoxyphenyl) acrylate) (E,E-3)

The title compound was obtained starting from **5-AMU** (0.067 g, 0.38 mmol) as white solid (63 mg, yield 60%, m.p. 187.6–189.0) using petroleum ether-ethyl acetate (8:2) as the eluent. ^1H NMR (400 MHz, CD_3OD): 3.30 (overlapped with CHD_2OD, s, 2H), 3.55 (s, 4H), 3.72 (s, 6H), 3.78 (s, 6H), 6.81 (d, *J* = 8.8, 4H), 7.26 (s, 1H), 7.60 (d, *J* = 8.7, 4H), 7.76 (s, 2H). ^{13}C NMR (100 MHz, CD_3OD): 51.8, 52.8, 53.4, 56.6, 112.6, 115.8, 129.3, 129.6, 134.6, 143.1, 145.0, 154.3, 163.0, 167.2, 171.8. MS (ESI): *m/z* 550.2 ($M+H^+$). HR-MS: *m/z* ($M+H^+$) calcd for $C_{29}H_{32}N_3O_8^+$ 550.21839; found 550.21661.

Computational Methods

The conformer search for the five isomers was performed fully automatically using the CREST[49] program with default parameters, based on the semiempirical quantum chemistry method GFN2-xTB,[50] including the methanol as solvent with the analytic linearized Poisson-Boltzmann formalism. As input for each molecule, we provided three-dimensional structures generated with Open Babel[63] from the corresponding SMILES (Simplified Molecular Input Line Entry System) strings.[64] While CREST produces numerous conformers and rotamers, we specifically opted for the lowest

energy conformer as the sole representative of the global conformational minimum for each molecule. The five structures so obtained were subsequently relaxed on S_0 at the B3LYP/def2-SVP level, with empirical dispersion correction D3BJ.[53] Solvent effects (methanol) were accounted for via polarizable continuum model (PCM). At these geometries, vertical excitation energies were calculated at the TDDFT/M06-2X/def2-TZVP level of theory. All DFT and TDDFT calculations were perform using the Gaussian09 suite.

The *E*-**2** to *Z*-**2** photoisomerization energy profile was constructed as follows. Starting from the minimum energy conformer from the CREST protocol, the FC structure of *E*-**2** was optimized on S_0 with DFT at the BH&HLYP/6-31G* level of theory. From here, an S_1 geometry optimization constraining the exocyclic double bond torsion was performed at the MRSF/BH&HLYP/6-31G* level of theory, as no minima could be optimized on the S_1 potential energy surface. After the optimization of $CoIn_{S1/S0}$, we used linear interpolation in internal coordinates (LIIC) to interpolate the geometries between the S_1 constrained structure and $CoIn_{S1/S0}$, computing MRSF/BH&HLYP/6-31G* excitation energies at each point. From $CoIn_{S1/S0}$, we performed a DFT relaxation on S_0 to obtain the *Z*-**2** FC geometry. The path from *Z*-**2** to $CoIn_{S1/S0}$ was constructed in the same fashion as for *E*-**2**. LIIC interpolation was performed with the python package chemcoord.[65] The $CoIn_{S1/S0}$ geometry was obtained via MRSF/BH&HLYP/6-31G* optimization using the adaptive penalty function approach.[66] Solvent effects were included also in this case with PCM. All MRSF calculations were performed using the GAMESS 2023 R1 package.[67]

Supporting Information

Experimental details for the synthesis and the characterization of compounds **1**, **5**-**AMU** and **5**-**MAMU**; Additional NMR and HR-MS spectra of compounds **2** and **3**; Computed isomerization profiles for *Z*-**2**→*E*-**2** photoisomerization, and *E*-**2**→*Z*-**2** with a simplified model of monoadduct **2**. Cartesian coordinates and Gibbs Free Energies for the 5 molecules investigated optimized at the B3LYP–D3BJ/def2-SVP level of theory, $CoIn_{S1/S0}$ geometries for the *E*-**2**→*Z*-**2** and *Z*-**2**→*E*-**2** photoisomerization reaction of monoadduct **2**, and *E*-**2**→*Z*-**2** photoisomerization of monoadduct **2** after substitution of the uracil moiety with a methyl group, at the MRSF/BH&HLYP/6-31G* level of theory.

Acknowledgements

The authors acknowledge the MUR for the financial support under the project CN00000041–"National Center for Gene Therapy and Drugsbased on RNA Technology"–CUP B63C2200061 0006 Mission 4 Component 2 (M4C2)–investment 1.4 [CN3] of the National Recovery and Resilience Plan (PNRR) funded by the European Union "Next Generation EU". M. P. acknowledge the University of Siena for the financial support of the project Chromo-GEnUP through the F-CUR2022 funding line (2265-2022-PM-CONRICMIUR_PC-FCUR2022_003). M. O. is also grateful for a NSF CSDM-A grant no. 2102619.

Conflict of Interests

The authors declare no conflict of interest.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Keywords: covalent derivatization · Morita-Baylis-Hillman adduct · photoswitches · photochemistry · RNA bases

[1] S. Minchin, J. Lodge, *Essays Biochem.* **2019**, *63*, 433–456.
[2] B. G. Vértessy, J. Tóth, *Acc. Chem. Res.* **2009**, *42*, 97–106.
[3] L. H. Hurley, F. L. Boyd, *Trends Pharmacol. Sci.* **1988**, *9*, 402–407.
[4] J.-L. Mergny, C. Hélène, *Nat. Med.* **1998**, *4*, 1366–1367.
[5] W. C. Reinhold, A. Thomas, Y. Pommier, *Trends in Cancer* **2017**, *3*, 2–6.
[6] S. Kovachka, M. Panosetti, B. Grimaldi, S. Azoulay, A. Di Giorgio, M. Duca, *Nat. Chem. Rev.* **2024**, *8*, 120–135.
[7] J. L. Childs-Disney, X. Yang, Q. M. R. Gibaut, Y. Tong, R. T. Batey, M. D. Disney, *Nat. Rev. Drug Discovery* **2022**, *21*, 736–762.
[8] T. R. Damase, R. Sukhovershin, C. Boada, F. Taraballi, R. I. Pettigrew, J. P. Cooke, *Front. Bioeng. Biotechnol.* **2021**, *9*, 628137.
[9] R. Feng, S. Patil, X. Zhao, Z. Miao, A. Qian, *Front. Mol. Biosci.* **2021**, *8*, 710738.
[10] J. M. Sasso, B. J. B. Ambrose, R. Tenchov, R. S. Datta, M. T. Basel, R. K. DeLong, Q. A. Zhou, *J. Med. Chem.* **2022**, *65*, 6975–7015.
[11] L. A. Rojas, Z. Sethna, K. C. Soares, C. Olcese, N. Pang, E. Patterson, J. Lihm, N. Ceglia, P. Guasp, A. Chu, *Nature* **2023**, *618*, 144–150.
[12] C. Sheridan, *Nat. Biotechnol.* **2022**, *40*, 1160–1162.
[13] H. Zhou, D. S. Chen, C. J. Hu, X. Hong, J. Shi, Y. Xiao, *Adv. Sci.* **2023**, *10*, 2303597.
[14] K. Paunovska, D. Loughrey, J. E. Dahlman, *Nat. Rev. Genet.* **2022**, *23*, 265–280.
[15] X. Fu, T. Chen, Y. Song, C. Feng, H. Chen, Q. Zhang, G. Chen, X. Zhu, *Small* **2021**, *17*, 2101224.
[16] J. Wang, E. Ayano, Y. Maitani, H. Kanazawa, *Int. J. Pharm.* **2017**, *523*, 217–228.
[17] A. M. Abdelaal, A. L. Kasinski, *NAR Cancer* **2021**, *3*, zcab030.
[18] N. Uddin, D. W. Binzel, D. Shu, T.-M. Fu, P. Guo, *Acta Pharm. Sin. B* **2023**, *13*, 1383–1399.
[19] C. Dohno, M. Kimura, Y. Fujiwara, K. Nakatani, *Nucleic Acids Res.* **2023**, *51*, 9533–9541.
[20] D. V. Berdnikova, *Chem. Commun.* **2021**, *57*, 10819–10826.
[21] A. Jäschke, *FEBS Lett.* **2012**, *586*, 2106–2111.
[22] K. A. Rotstan, M. M. Abdelsayed, L. F. M. Passalacqua, F. Chizzolini, K. Sudarshan, A. R. Chamberlin, J. Míšek, A. Luptak, *eLife* **2020**, *9*, e51737.
[23] V. Razzano, M. Paolino, A. Reale, G. Giuliani, R. Artusi, G. Caselli, M. Visintin, F. Makovec, A. Donati, F. Villafiorita-Monteleone, et al., *ACS Omega* **2017**, *2*, 5453–5459.
[24] V. Razzano, M. Paolino, A. Reale, G. Giuliani, A. Donati, G. Giorgi, R. Artusi, G. Caselli, M. Visintin, F. Makovec, *RSC Adv.* **2018**, *8*, 8638–8656.
[25] M. Paolino, A. Reale, V. Razzano, G. Giuliani, A. Donati, C. Bonechi, G. Caselli, M. Visintin, F. Makovec, C. Scialabba, et al., *New J. Chem.* **2019**, *43*, 6834–6837.
[26] M. Paolino, M. Visintin, E. Margotti, M. Visentini, L. Salvini, A. Reale, V. Razzano, G. Giuliani, G. Caselli, F. Tavanti, *New J. Chem.* **2019**, *43*, 17946–17953.
[27] G. Tassone, M. Paolino, C. Pozzi, A. Reale, L. Salvini, G. Giorgi, M. Orlandini, F. Galvagni, S. Mangani, X. Yang, *ChemBioChem* **2022**, *23*, e202100449.
[28] M. Saletti, J. Venditti, M. Paolino, A. Zacchei, G. Giuliani, G. Giorgi, C. Bonechi, A. Donati, A. Cappelli, *RSC Adv.* **2023**, *13*, 35773–35780.
[29] M. Paolino, A. Reale, V. Razzano, G. Giorgi, G. Giuliani, F. Villafiorita-Monteleone, C. Botta, C. Coppola, A. Sinicropi, A. Cappelli, *New J. Chem.* **2020**, *44*, 13644–13653.
[30] M. Saletti, M. Paolino, J. Venditti, C. Bonechi, G. Giuliani, A. Boccia, C. Botta, A. Cappelli, *Dyes Pigm.* **2023**, *219*, 111571.
[31] N. K. Devaraj, M. G. Finn, *Chem. Rev.* **2021**, *121*, 6697–6698.

[32] M. Saletti, M. Paolino, J. Venditti, C. Bonechi, G. Giuliani, S. Lamponi, G. Tassone, A. Boccia, C. Botta, L. Blancafort, *ChemBioChem* **2024**, *n/a*, e202300862.
[33] C. R. Mateus, W. P. Almeida, F. Coelho, *Tetrahedron Lett.* **2000**, *41*, 2533–2536.
[34] T. B. Johnson, A. Litzinger, *J. Am. Chem. Soc.* **1936**, *58*, 1940–1942.
[35] J. Carbon, H. David, M. H. Studier, *Science (80)* **1968**, *161*, 1146–1147.
[36] J. N. Singer, F. M. Müller, E. Węgrzyn, C. Hölzl, H. Hurmiz, C. Liu, L. Escobar, T. Carell, *Angew. Chem. Int. Ed.* **2023**, *62*, e202302360.
[37] D. Pearson, T. Carell, *Nucleic Acids Res.* **2011**, *39*, 4818–4826.
[38] Y. Guo, G. Shao, L. Li, W. Wu, R. Li, J. Li, J. Song, L. Qiu, M. Prashad, F. Y. Kwong, *Adv. Synth. Catal.* **2010**, *352*, 1539–1553.
[39] K. J. Lee, D. Kang, H.-S. Park, *Mol. Cells* **2019**, *42*, 386–396.
[40] D. Hadar, S. Gelkop, L. Vaserman, M. Amiram, *ChemBioChem* **2021**, *22*, 1379–1384.
[41] T. Violo, A. Lambert, A. Pillot, M. Fanuel, J. Mac-Béar, C. Broussard, C. Grandjean, E. Camberlein, *Chem. A Eur. J.* **2023**, *29*, e202203497.
[42] A. Jordan, S. Huang, H. F. Sneddon, A. Nortcliffe, *ACS Sustainable Chem. Eng.* **2020**, *8*, 12746–12754.
[43] S. K. Pagire, A. Hossain, L. Traub, S. Kerres, O. Reiser, *Chem. Commun.* **2017**, *53*, 12072–12075.
[44] T. B. Nguyen, A. Al-Mourabit, *Photochem. Photobiol. Sci.* **2016**, *15*, 1115–1119.
[45] A. Cappelli, M. Paolino, P. Anzini, G. Giuliani, S. Valenti, M. Aggravi, A. Donati, R. Mendichi, L. Zetta, A. C. Boccia, *J. Polym. Sci. Part A* **2010**, *48*, 2446–2461.
[46] T. Nakano, K. Takewaki, T. Yade, Y. Okamoto, *J. Am. Chem. Soc.* **2001**, *123*, 9182–9183.
[47] F. D. Lewis, S. L. Quillen, P. D. Hale, J. D. Oxman, *J. Am. Chem. Soc.* **1988**, *110*, 1261–1267.
[48] H. Zhang, X. Li, Y. Lin, F. Gao, Z. Tang, P. Su, W. Zhang, Y. Xu, W. Weng, R. Boulatov, *Nat. Commun.* **2017**, *8*, 1147.
[49] P. Pracht, F. Bohle, S. Grimme, *Phys. Chem. Chem. Phys.* **2020**, *22*, 7169–7192.
[50] C. Bannwarth, S. Ehlert, S. Grimme, *J. Chem. Theory Comput.* **2019**, *15*, 1652–1671.
[51] W. Kohn, L. J. Sham, *Phys. Rev.* **1965**, *140*, A1133–A1138.
[52] A. D. Becke, *J. Chem. Phys.* **1992**, *96*, 2155–2160.
[53] S. Grimme, S. Ehrlich, L. Goerigk, *J. Comput. Chem.* **2011**, *32*, 1456–1465.
[54] F. Weigend, R. Ahlrichs, *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297–3305.
[55] E. Runge, E. K. U. Gross, *Phys. Rev. Lett.* **1984**, *52*, 997–1000.
[56] Y. Zhao, N. E. Schultz, D. G. Truhlar, *J. Chem. Phys.* **2005**, *123*, 161103.
[57] M. Filatov(Gulak), M. Paolino, R. Pierron, A. Cappelli, G. Giorgi, J. Léonard, M. Huix-Rotllant, N. Ferré, X. Yang, D. Kaliakin, *Nat. Commun.* **2022**, *13*, 6433.
[58] S. Lee, E. E. Kim, H. Nakata, S. Lee, C. H. Choi, *J. Chem. Phys.* **2019**, *150*, 184111.
[59] Y. Horbatenko, S. Sadiq, S. Lee, M. Filatov, C. H. Choi, *J. Chem. Theory Comput.* **2021**, *17*, 848–859.
[60] X.-P. Chang, C.-X. Li, B.-B. Xie, G. Cui, *J. Phys. Chem. A* **2015**, *119*, 11488–11497.
[61] M. Promkatkaew, S. Suramitr, T. Karpkird, S. Wanichwecharungruang, M. Ehara, S. Hannongbua, *Photochem. Photobiol. Sci.* **2014**, *13*, 583–594.
[62] A. Nomoto, N. Inai, T. Yanai, Y. Okuno, *J. Phys. Chem. A* **2022**, *126*, 497–505.
[63] N. M. O'Boyle, M. Banck, C. A. James, C. Morley, T. Vandermeersch, G. R. Hutchison, *J. Cheminf.* **2011**, *3*, 33.
[64] D. Weininger, *J. Chem. Inf. Comput. Sci.* **1988**, *28*, 31–36.
[65] O. Weser, B. Hein-Janke, R. A. Mata, *J. Comput. Chem.* **2023**, *44*, 710–726.
[66] Y. S. Baek, S. Lee, M. Filatov, C. H. Choi, *J. Phys. Chem. A* **2021**, *125*, 1994–2006.
[67] F. Zahariev, P. Xu, B. M. Westheimer, S. Webb, J. Galvez Vallejo, A. Tiwari, V. Sundriyal, M. Sosonkina, J. Shen, G. Schoendorff, *J. Chem. Theory Comput.* **2023**, *19*, 7031–7055.

Manuscript received: March 15, 2024
Revised manuscript received: April 17, 2024
Accepted manuscript online: April 23, 2024
Version of record online: June 3, 2024