Supporting information for:

Ribose Alters the Photochemical Properties of the Nucleobase in Thionated Nucleosides.[†]

Mikołaj J. Janicki,[‡,△] Corinna L. Kufner,[¶,△] Zoe R. Todd,[¶,△] Seohyun C. Kim,[§] Derek K. O'Flaherty,[§,‖] Jack W. Szostak,[§] Jiří Šponer,[⊥,#] Robert W. Góra,[*,‡] Dimitar D. Sasselov,[*,¶] and Rafał Szabla[*,@]

‡*Department of Physical and Quantum Chemistry, Wrocław University of Science and Technology, Faculty of Chemistry, Wybrzeże Wyspiańskiego 27, 50-370, Wrocław, Poland*
¶*Department of Astronomy, Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138, USA*
§*Howard Hughes Medical Institute, Department of Molecular Biology and Center for Computational and Integrative Biology, Massachusetts General Hospital, 185 Cambridge Street, Boston, MA 02114, USA*
‖*Present address: Department of Chemistry, University of Guelph, 50 Stone Rd. E., Guelph, Ontario, N1G 2W1, Canada*
⊥*Institute of Biophysics of the Czech Academy of Sciences, Královopolská 135, 61265 Brno, Czech Republic*
#*Regional Centre of Advanced Technologies and Materials, Czech Advanced Technology and Research Institute (CATRIN), Palacky University Olomouc, Slechtitelu 241/27, 783 71, Olomouc - Holice, Czech Republic*
@*EaStCHEM, School of Chemistry, University of Edinburgh, Joseph Black Building, David Brewster Road, Edinburgh, EH9 3FJ, UK*
△*These authors contributed to this work equally.*

E-mail: robert.gora@pwr.edu.pl; dsasselov@cfa.harvard.edu; rafal.szabla@ed.ac.uk

Experimental section

The 2-thiocytosine sample was purchased from Sigma Aldrich (USA) and the beta-2-thiocytidine from Berry & Associates Inc. (USA). 2-Amino-α-D-ribofurano[1',2': 4,5]-2-oxazoline and α-D-Ribofuranosyl-2,2'-anhydrouridine were synthesized according to procedures reported by Shannahoff and Sanchez[1] with slight modifications. 2-Amino-α-D-ribofurano[1',2': 4,5]-2-oxazoline was briefly filtrated and subsequently reacted with methyl propriolate in water under reflux over one hour to afford α-D-ribofuranosyl-2,2'-anhydrouridine. α-D-Ribofuranosyl-2,2'-anhydrouridine was dissolved in dimethylformamide and reacted with sodium hydrosulfide hydrate in the presence of ammonium bicarbonate to produce α-D-2-thiouridine according to Xu et al.[2] with slight modification to the purification procedure. Specifically, the product was purified by reverse phase flash chromatography (prepacked RediSep Rf Gold C18Aq 50 g columns from Teledyne Isco (Lincoln, NE)), using gradient elution between (A) aqueous 20 mM TEAB (pH 7.5) and (B) acetonitrile. The product was eluted between 0% and 30% B over 13 CVs with a flow rate of 40 mL/min. Fractions containing product were pooled and most of the solvent was removed in vacuo, followed by freeze-drying under high vacuum at room temperature. The product was characterized by ^{1}H NMR which agreed with spectra previously reported.

α-D-2-thiocytidine was obtained via a transient per-silylation methodology. First, α-D-2-thiouridine was protected using excess TBS-Cl (10 equivs), followed by convertible nucleoside with a triazoyl moiety at the C4 position. Displacement of the triazoyl group with ammonia in dioxane (0.4 M) under reflux overnight furnished silylated α-D-2-thiocytidine. Removal of the silyl groups was achieved using TEA-3HF and the product was purified using reverse phase flash chromatography, as described above (0% and 30% acetonitrile in aqueous 20 mM TEAB). The purity and identity of the α-D-2-thiocytidine was confirmed by ^{1}H NMR (in agreement with the literature Xu et al.[2]).

All samples were dissolved in 50 mM phosphate buffer (KH$_2$PO$_4$, Na$_2$HPO$_4$) in LS-MS

[†]A footnote for the title

grade H_2O. All buffer components were purchased from Sigma Aldrich (USA). The sample concentrations were ~8 mM for thioC, ~6 mM for β-thioCyd and ~2 mM for α-thioCyd. The absorbance of the samples at the excitation wavelength was between OD 0.1 and 0.3. The pump-probe experiments were performed under ambient oxygen conditions, in a flow cell with 1/16" quartz windows (sample thickness 100 μm) at a temperature of 23°C. The basic concepts of pump-probe spectroscopy can be found in the literature.[3–5] The light source for both the pump and the probe beam was a Ti:Sa based laser amplifier system (Solstice Ace, Spectra Physics) with a repetition rate of 1 kHz and an output pulse duration of 90 fs at 800 nm. An optical parametric amplifier system (Topas Prime + NirUVis, Spectra Physics) was used for the nonlinear frequency conversion to 308 nm excitation pulses (pump). The excitation pulses were then stretched by a 25 cm UV fused silica block to ~1.5 ps to reduce two-photon excitation. The excitation energy was 0.6 μJ at the sample position. The visible continuum (probe) was generated, delayed and detected in a Helios Fire system (Ultrafast Systems, USA). The diameter of the pump pulse was ~250 μm and ~100 μm for the probe pulse (90:10 level) at the sample position. Both the probe and pump beam were focused and spatially overlapped in the sample. Magic angle conditions were used in all experiments. A global fitting analysis was performed to determine the lifetimes of the transient intermediates (Surface Xplorer, Ultrafast Systems, LLC).[6–8]

Theoretical section

Computational methods

The equilibrium ground-state structures of thiocyditine and thiocytosine were located using the Møller–Plesset perturbation theory to the second-order[9] (MP2) with the cc-pVTZ basis set.[10] The abbreviation of the method is MP2/cc-pVTZ. To assess relative differences between aqueous Gibbs free energies of various arrangements of the sugar ring in thionucleosides, the Kohn-sham density functional theory (KS-DFT) was used, assuming the M06-2x

hybrid functional[11] with the def2-TZVPP basis set.[12] Solvent effects induced by bulk water were estimated by performing single-point calculations using polarizable continuum model (PCM) for the optimized gas-phase ground-state geometries located at the M06-2X/def2-TZVPP level. The MP2 and KS-DFT/PCM calculations were performed using the Turbomole 7.3[13] and Gaussian 16[14] package, respectively.

Vertical excitations energies were computed using the algebraic diagrammatic construction to the second-order method [ADC(2)][15–17] with the cc-pVTZ basis set[10] [ADC(2)/cc-pVTZ], assuming the equilibrium ground-state structures located using the MP2/cc-pVTZ method. The character of electronic transitions was assigned based on an analysis of natural transition orbitals[18] (NTOs) obtained by means of the TheoDore package.[19] Stationary points on excited-state potential energy surfaces as well as harmonic vibrational frequencies were obtained employing the ADC(2)/cc-pVTZ method. Minimum-energy crossing points (MECPs) between two different electronic states were located by means of the sequential penalty constrained optimization introduced by Levine et al.[20] in the CIOpt package. MECPs were optimized using energies and analytical gradients computed at the MP2/ADC(2)/cc-pVTZ level. To combine the CIOpt with the TURBOMOLE 7.3 package, an in-house interface between these two programs was written. To validate ADC(2) excited-state potential energy surfaces with a multiconfigurational method, the extended multi-state complete active space second-order perturbation method[21,22] (XMS-CASPT2) was used, assuming a reference wavefunction that is a state-averaged complete active space self-consistent field (SA-CASSCF) with the cc-pVDZ basis set (XMS-CASPT2/SA-CASSCF/cc-pVDZ). The selection of molecular orbitals constituting complete active spaces was made by rules proposed by Veryazov et al.[23]. They suggest that the natural orbital occupations of active orbitals should be in the range 0.02–1.98. In addition, three electronic states were included in the XMS-CASPT2 calculations and the vertical shift was adapted as 0.5. Detailed description of molecular orbitals included in the active space can be found below in the section discussing potential energy profiles calculated at the XMS-CASPT2 level. The

XMS-CASPT2/SA-CASSCF/cc-pVDZ level of theory was used to optimize excited-state minimum-energy structures. To locate MECPs between various electronic states using energies and analytical gradients obtained at the XMS-CASPT2/SA-CASSCF/cc-pVDZ level, the CIOpt program was interfaced with the BAGEL 1.2.0 package.[24,25] Schematic potential energy surfaces (PES) showing non-radiative deactivation channels were constructed using ground and excited-state energies of optimized structures, the energies were computed at the either MP2/ADC(2)/cc-pVTZ or XMS-CASPT2/SA-CASSCF/cc-pVDZ level of theory. Spin-orbit couplings (SOCs) were estimated for the $S_1 \rightarrow T_2$ and $T_1 \rightarrow S_0$ MECPs using the MS-CASPT2/SA-CASSCF/cc-pVTZ-DK method. The scalar relativistic effects were incorporated in the calculations through the second-order Douglas–Kroll–Hess Hamiltonian. The SOCs were obtained using the MOLCAS 8.4 package.[26] All XMS-CASPT2 and ADC(2) calculations were performed using the Bagel 1.2.0 and the Turbomole 7.3, respectively.

Simulation of excited-state absorption spectra - computational protocol

Simulations of excited-state absorption (ESA) spectra were performed based on the independent mode displaced harmonic oscillator[27,28] (IMDHO) model that allows for prediction of the vibrational structure of electronic spectra and includes homogeneous and inhomogeneous broadening in a single absorption band. To employ the IMDHO framework for the simulation of electronic spectra for excited-state microhydrated structures, vertical excitation energies, oscillator strengths, excited-state gradients for 11 excited states as well as harmonic vibrational frequencies for a given excited-state minimum were computed. It is worth adding that only electronic transitions in the range of 300-700 nm and having oscillator strength above 0.001 a.u. were selected to simulate ESA spectra. The homogeneous line broadening in the IMDHO model was adopted as 10 cm^{-1}. To estimate solvent effects on excited-state energies of microhydrated structures, an implicit solvent model was used in the form of the conductor-like screening model[29] (COSMO). The inhomogeneous broadening for each elec-

tronic transition was estimated based on the COSMO/ADC(2)/cc-pVTZ calculations and the Marcus theory (1).[30,31]

$$\Gamma = 2\sqrt{\lambda k_B T \ln 2} \tag{1}$$

Broadening derived from the Marcus theory (1) allows obtaining the half-width at half-maximum corresponding to Gaussian line shape (Γ), where k_B is the Boltzmann constant, T is the temperature. The solvent reorganisation energy (λ) was computed as the difference between energies calculated in the non-equilibrium and equilibrium regimes for given electronic state using the COSMO/ADC(2)/cc-pVTZ method. To estimate the solvent-induced energy shifts for each absorption band, the energy difference between corresponding excitation energies obtained at the ADC(2)/cc-pVTZ and COSMO/ADC(2)/cc-pVTZ level (in the non-equilibrium regime) was computed. The molecular orbital character of electronic transitions was inspected in the gas phase and COSMO calculations to identify the same electronic excitation. The estimated solvent shifts for electronic excitations were used to move each single absorption band. The excited-state absorption spectra were simulated employing the ORCA 4.2.1 package.[32–34] All ADC(2) and COSMO/ADC(2) calculations were performed with the Turbomole 7.3 package.

Ground-state structures

All quantum-chemical calculations involving 2-thiocytosine were performed with the amino-thion-N_1H tautomer which was found as the lowest-energy structure in the COSMO calculations[35] (see Fig. S1). To estimate preferable conformational arrangements of the sugar ring in thiocytidine, we performed Gibbs free calculations at the M06-2X/def2-TZVPP/PCM level of theory. Our results suggest that the C2'-endo arrangement of β-thiocytidine is slightly more stable (about 0.71 kcal/mol) than the C3'-endo configuration of the ribose ring. In the case of α-thiocytidine, the C3'-endo conformer is characterized by the lower relative energy of 0.48 kcal/mol in comparison of the C2'-endo structure. Consequently, we anticipate that

the C2'-endo and C3'-endo conformational arrangements of the sugar ring in both anomers of thiocytidine could coexist in a water solution. It is also worth adding that these Gibbs free energy differences are below the accuracy of the employed methods.
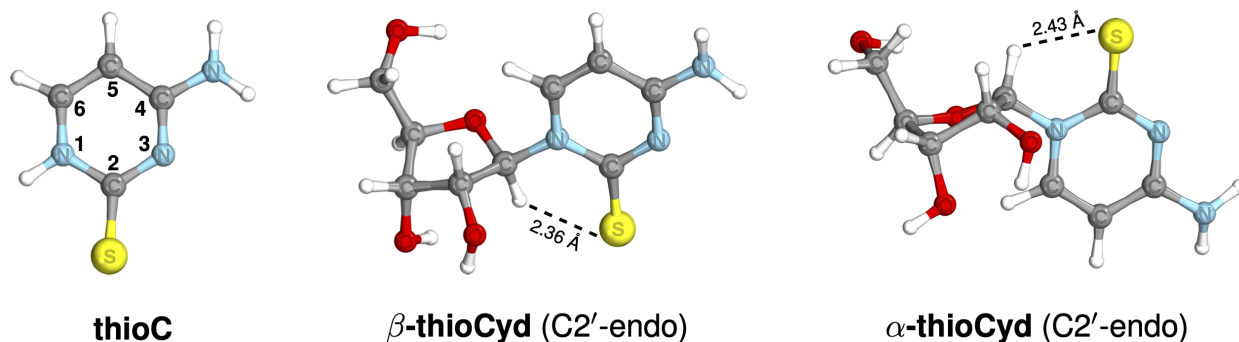
Vertical excitation energies

Figure S1: The equilibrium ground-state geometries of 2-thiocytosine, α- and β-thiocytidine found at the MP2/cc-pVTZ level of theory.
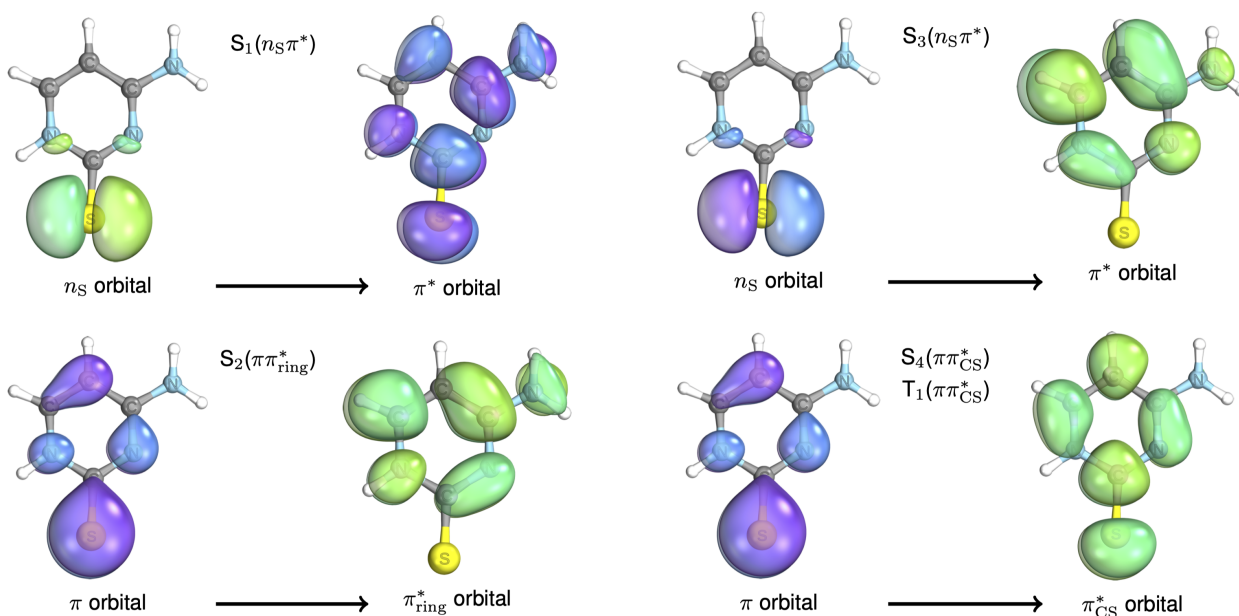

Figure S2: The molecular orbitals present a character of electronic transitions in the Franck-Condon region of thioC.

Vertical excitation energies of thioC and thioCyd, assuming the gas-phase optimized geometries located at the MP2/cc-pVTZ level (see Fig. S1), are presented in Tab. S1. We

Table S1: Vertical excitation energies of thioC and α- and β-thioCyd computed using the MP2/ADC(2)/cc-pVTZ method, assuming the equilibrium ground-state structures located at the MP2/cc-VTZ level of theory.

State / Transition		$E_{exc}[eV]$	f_{osc}	λ [nm]
	thioC			
S_1	$n_S\pi^*$	3.43	2.43×10^{-5}	361.5
S_2	$\pi\pi^*_{ring}$	3.72	3.57×10^{-2}	324.6
S_3	$n_S\pi^*$	3.83	3.66×10^{-5}	323.7
S_4	$\pi\pi^*_{CS}$	4.41	4.36×10^{-1}	281.1
S_5	$n_N\pi^*$	4.96	1.54×10^{-3}	250.0
S_6	$\pi\pi^*$	5.55	2.94×10^{-1}	223.4
T_1	$\pi\pi^*_{CS}$	3.30	-	-
T_2	$n_S\pi^*$	3.32	-	-
T_3	$\pi\pi^*_{ring}$	3.41	-	-
	β-thioCyd			
S_1	$n_S\pi^*$	3.27	8.39×10^{-4}	379.2
S_2	$\pi\pi^*_{ring}$	3.62	3.22×10^{-2}	342.5
S_3	$n_S\pi^*$	3.72	1.66×10^{-2}	333.3
S_4	$\pi\pi^*_{CS}$	4.38	3.21×10^{-1}	283.1
S_5	$n_N\pi^*$	4.83	5.58×10^{-3}	256.7
S_6	$\pi\pi^*_{ring}$	5.33	3.95×10^{-1}	232.6
T_1	$n_S\pi^*$	3.17	-	-
T_2	$\pi\pi^*_{CS}$	3.28	-	-
T_3	$\pi\pi^*_{ring}$	3.43	-	-
	α-thioCyd			
S_1	$n_S\pi^*$	3.36	7.94×10^{-4}	369.0
S_2	$\pi\pi^*_{ring}$	3.70	1.82×10^{-2}	335.1
S_3	$n_S\pi^*$	3.83	2.99×10^{-2}	323.7
S_4	$\pi\pi^*_{CS}$	4.41	3.35×10^{-1}	281.1
S_5	$n_N\pi^*$	4.94	2.36×10^{-3}	251.0
S_6	$\pi\pi^*_{ring}$	5.43	3.77×10^{-1}	228.3
T_1	$n_S\pi^*$	3.25	-	-
T_2	$\pi\pi^*_{CS}$	3.36	-	-
T_3	$\pi\pi^*_{ring}$	3.48	-	-

S$_1$($n_S\pi^*$)
T$_1$($n_S\pi^*$)

n_S orbital $\longrightarrow$ π^* orbital

S$_3$($n_S\pi^*$)

n_S orbital $\longrightarrow$ π^* orbital

S$_2$($\pi\pi^*_{ring}$)

π orbital $\longrightarrow$ π^*_{ring} orbital
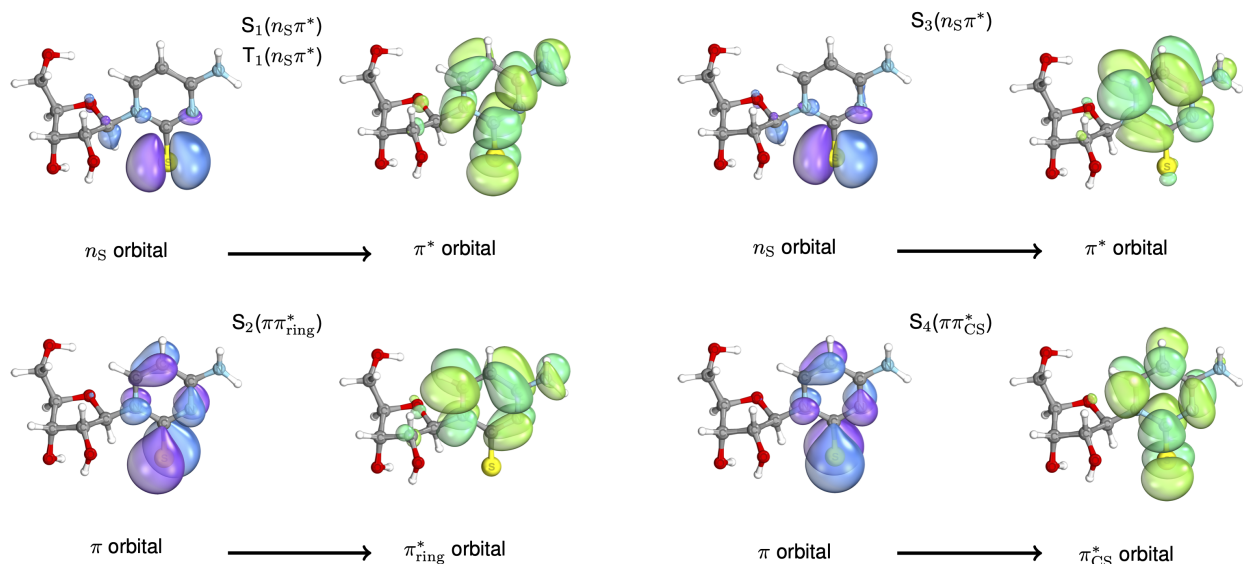
S$_4$($\pi\pi^*_{CS}$)

π orbital $\longrightarrow$ π^*_{CS} orbital

Figure S3: The molecular orbitals show a character of electronic transitions in the Franck-Condon region of β-thioCyd.

S$_1$($n_S\pi^*$)
T$_1$($n_S\pi^*$)

n_S orbital $\longrightarrow$ π^* orbital

S$_3$($n_S\pi^*$)

n_S orbital $\longrightarrow$ π^* orbital

S$_2$($\pi\pi^*_{ring}$)

π orbital $\longrightarrow$ π^*_{ring} orbital

S$_4$($\pi\pi^*_{CS}$)
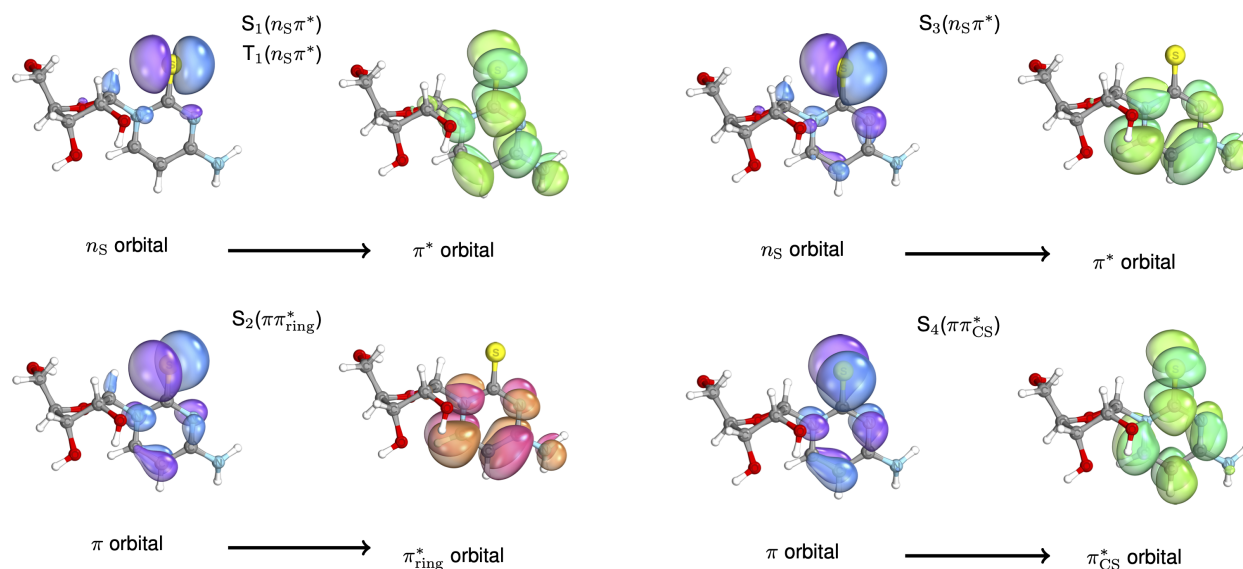
π orbital $\longrightarrow$ π^*_{CS} orbital

Figure S4: The molecular orbitals present a character of electronic transitions in the Franck-Condon region of α-thioCyd.

are particularly interested in energetically low-lying excited states because these states play a crucial role in the photodynamics of the chromophores which we tracked in our pump-probe experiments with the excitation wavelength equal to 308 nm. thioC as well as (α and β) thioCyd are characterized by similar vertical excitations of the four lowest-lying excited states. The most brightest state in thioC and in both anomers of thioCyd is the S_4 excited state (see Tab. S1). The vertical excitation energy of the S_4 state for such chromophores amounts to approximately 4.40 eV. The molecular orbital character of the S_4 excitation of thioC and thioCyd is the $\pi\pi^*_{CS}$ transition (see Fig. S2, S3 and S4). The energetically lower-lying optical state (S_2) is located at 3.70 eV in thioC and α-thioCyd whereas the same transition is slightly red-shifted by approximately 0.1 eV in β-thioCyd. Molecular orbitals participating in the $S_2(\pi\pi^*_{ring})$ transition are virtually the same for these molecules (see Fig. S2, S3 and S4). The lowest-lying $S_1(n_S\pi^*)$ dark state is energetically lower in both anomers of thioCyd (3.27 and 3.36 eV, Tab. S1) compared to thioC (3.43 eV, Tab. S1). Our results clearly show that the presence of the sugar moiety in thioCyd can lead to the stabilization of the $n_S\pi^*$ transition which could affect the photochemistry of thioCyd in comparison with thioC that does not possess the sugar part. Thiated nucleobases and nucleosides exhibit an efficient population of triplet excited states through the readily accessible intersystem crossing. Vertical excitation energies of low-lying triplet excited states in the Franck-Condon region of the molecules are presented in Tab. S1. The lowest-lying triplet excited state of thioC (see Tab. S1) is located at 3.30 eV and the corresponding molecular orbital character agrees with the S_4 ($\pi\pi^*_{CS}$) electronic excitation (see Fig. S2). Interestingly, the first triplet excited state is located energetically lower in thioCyd (at 3.17 and 3.25 eV in Tab. S1), and the T_1 state is characterized by the $^3n_S\pi^*$ transition (see Fig. S4, S3) which is the T_2 state in thioC. The observed inversion of the triplet states is associated with the excited-state C1'-H$\cdots$S interaction (see Fig. S3 and S4) in thioCyd that could result in the stabilization of the $^3n_S\pi^*$ state. The photophysical picture of thioC and thioCyd suggests that the presence of the sugar in thionucleosides could energetically stabilize the $n_S\pi^*$ electronic excitation in the

singlet and triplet manifold. This is also reflected by the differences in the photodynamics of these two molecules.

Microhydrated structures

thioC $\qquad$ β-**thioCyd** (C2'-endo) $\qquad$ α-**thioCyd** (C2'-endo)
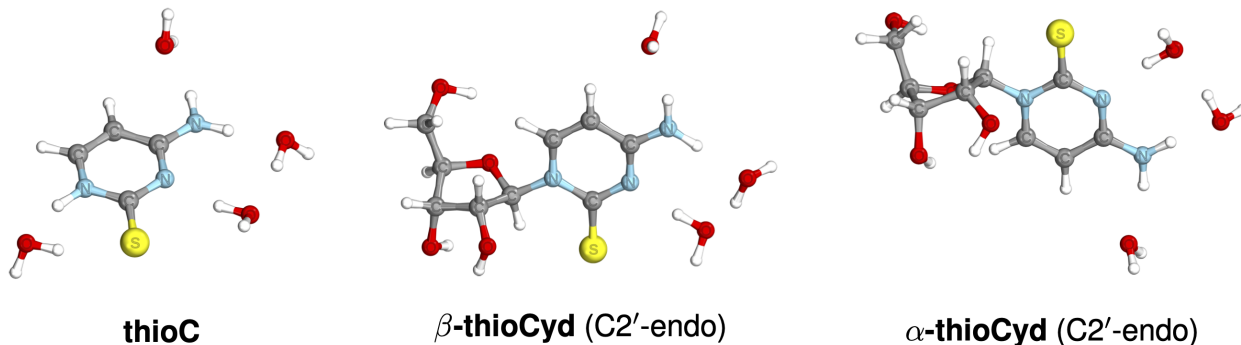
Figure S5: The equilibrium ground-state structures of thioC and thioCyd with four and three explicit water molecules, respectively.

Implicit solvent models are not able to fully reproduce quantum-mechanical interactions between solvent and solute (such as hydrogen bonds), which can significantly alter the energetics of excited states having in particular large dipole moments. Importantly, incorporating solvent effects in the simulation of excited-states absorption bands is crucial for correct assignment of experimental data. Therefore, to simulate excited-state absorption spectra (ESA) of thioC and both anomers of thioCyd in an aqueous environment, we constructed microhydrated model systems that contain thioC surrounded by four quantum-mechanical (QM) water molecule and α- and β-thioCyd with three QM water molecules (see Fig. S5). The key geometrical features of the microhydrated structures such as the C=S bond length and the C1'-H$\cdots$S=C distance are virtually the same as in the gas-phase ground-state structures in Fig. S1. We added a suitable number of water molecules to saturate all hydrogen bonding sites of the chromophores, namely four for the nucleobase and three for the nucleoside. The addition of explicit water molecules to the chromophores enables including quantum-mechanical interactions originating from hydrogen bonds in excited-states calculations. In addition, in bulk water, there are many more solvent molecules surrounding

chromophores that can affect on the energy of excited states through electrostatic interactions that can consequently also induce substantial shifts in the electronic absorption bands of excited states. Therefore, to reproduce the effects of bulk aqueous environment on our UV-excited chromophores, we decided to also include the COSMO implicit solvation model in the simulations of ESA spectra.

Table S2: Vertical excitation energies of microhydrated thioC and β-thioCyd computed using the MP2/ADC(2)/cc-pVTZ method, assuming the equilibrium ground-state structures located at the MP2/cc-VTZ level of theory.

State / Transition		$E_{exc}[eV]$	f_{osc}	λ [nm]
	thioC			
S_1	$n_S\pi^*$	4.05	2.82×10^{-4}	306.1
S_2	$\pi\pi^*_{ring}$	4.15	3.54×10^{-2}	298.8
S_3	$\pi\pi^*_{CS}$	4.46	3.61×10^{-1}	278.0
S_4	$n_S\pi^*$	4.52	1.27×10^{-2}	274.3
	β-thioCyd			
S_1	$n_S\pi^*$	3.60	8.90×10^{-4}	344.4
S_2	$\pi\pi^*_{ring}$	3.87	4.03×10^{-2}	320.4
S_3	$n_S\pi^*$	4.02	1.39×10^{-2}	308.4
S_4	$\pi\pi^*_{CS}$	4.41	2.84×10^{-1}	281.1

Tab. S2 shows vertical excitation energies of microhydrated thioC and β-thioCyd presented in Fig. S5. The lowest-lying excited states having the $n_S\pi^*$ character are destabilized about $\sim$0.60 eV and $\sim$0.30 eV in thioC and β-thioCyd, respectively, comparing to the gas-phase calculations in Tab. S1. The different extents of the energy destabilization of the S_1 dark states arise from the local chemical environment. The explicit water molecule (H_2O-HN) in microhydrated thioC directly interacts with the sulphur atom (Fig. S5) inducing a significant hypsochromic shift ($\sim$0.60 eV), whereas such a position of H_2O is not possible in β-thioCyd due to the presence of the ribose moiety. Thus, the S_1 excited state should be less destabilized ($\sim$0.30 eV) by the water environment owing to the partial shielding by the sugar. The energies of the S_2 bright states in both microhydrated model systems are blueshifted to a

similar extent as it was computed using the COSMO implicit solvent (Tab. 1 in the article). Furthermore, the $\pi\pi^*_{CS}$ excited states are only slightly destabilized (Tab. S2) owing to the similar the $\boldsymbol{\mu}$ vector of the excited state to the ground-state $\boldsymbol{\mu}$ vector which consequently enables only a modest blue-shift of the excitation energy. Due to the different orientation of the $\boldsymbol{\mu}$ vector of the higher-lying dark states in comparison of the ground-state $\boldsymbol{\mu}$ vector, the $n_S\pi^*$ excited states are strongly destabilized in the presence of explicit water molecules. In addition, the S_4 state in microhydrated thioC is noticeable stronger destabilized than the corresponding state in thioCyd molecule owing to the explicit interaction between H_2O and the sulphur atom (Fig. S5). It is worth adding that vertical excitation energies for the microhydrated structures are consistent with the COSMO calculations, performed for the gas-phase structures, presented in the main article (Tab. 1).

Excited-state stationary points

$S_1(n_S\pi^*)$ minimum $T_1(\pi\pi^*_{ring})$-rp minimum $T_1(\pi\pi^*_{CS})$-S-oop minimum
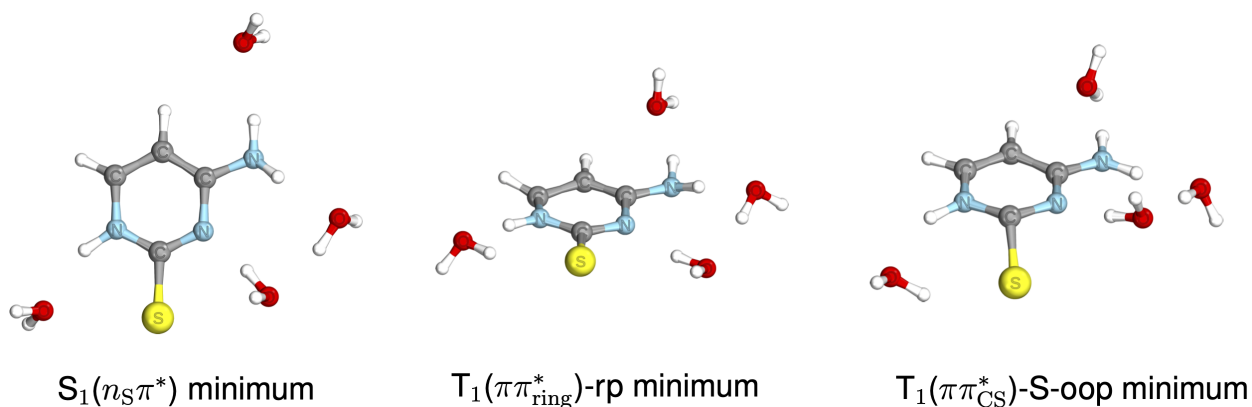
Figure S6: The excited-state minima of microhydrated thioC at the MP2/ADC(2)/cc-pVTZ level of theory.

The exploration of excited-state potential energy surfaces of thioC with four explicit water molecules allowed locating several key excited-state minima that are presented in Fig. S6. The $S_1(n_S\pi^*)$ minimum-energy structure is characterized by the rearrangement of the water cluster and the elongation of the C=S bond approximately by 0.1 Å, compared to the ground-state structure. These structural changes are caused by the population of the

$^1 n_S \pi^*$ excited state in which the n_S orbital is mostly localized on the sulphur atom. In the case of the triplet manifold, we found two different T_1 minimum-energy geometries that shown in Fig. S6. The $T_1(\pi\pi^*_{ring})$ structure has a lightly ring-puckering (rp) geometry with pyrimidalized C6 atom and slightly elongated C=S bond to 1.72 Å when compared to 1.65 Å found for the ground-state geometry. Furthermore, we also found a minimum-energy T_1 structure, having the $^3\pi\pi^*_{CS}$ molecular orbital character, that is marked by the sulphur atom tilted out of the ring plane (S-oop). In addition, this geometry is characterized by more pronounced C=S bond elongation, *i.e.* to 1.78 Å.

The excited-state optimizations performed for β-thioCyd and α-thioCyd allowed us to locate several excited-state minima of these microhydrated systems as depicted in Fig. S7 and Fig. S8. The S_1 and T_1 minimum-energy structures (Fig. S7 and S8) are associated with the $n_S\pi^*$ molecular orbital character, moreover, the length of the C=S bond (1.76 Å) and the C1′-H$\cdots$S=C distance (2.25 Å) are virtually identical in these geometries. The optimized S_2 and T_2 structures are marked by the $\pi\pi^*_{ring}$ transition for which the C1′-H$\cdots$S=C distance amounts to approximately to 2.43 Å (Fig. S7 and S8). The C=S bond length is equal to 1.76 and 1.72 Å in the S_2 and T_2 minima, respectively. These $\pi\pi^*_{ring}$ structures are characterized by slight puckering of the aromatic ring with pyramidalization on the C6 carbon atom. All of the above discussed optimized microhydrated excited-state minimum-energy geometries were used to simulate excited-state absorption spectra.

The initial position of explicit water molecules around chromophore influences the final arrangement of water molecules after the optimization step. Therefore, the relative energies of different excited-state structures could be biased by the initial arrangement of water molecules. Thus, to present plausible photochemical pathways of our model systems, we decided to demonstrate the photochemistry of the chromophores using the gas-phase structures. We would like to emphasize once again that the main motivation to consider microhydrated structures was accurate prediction of excited-state absorption spectra and solvatochromic shifts of the key bands.
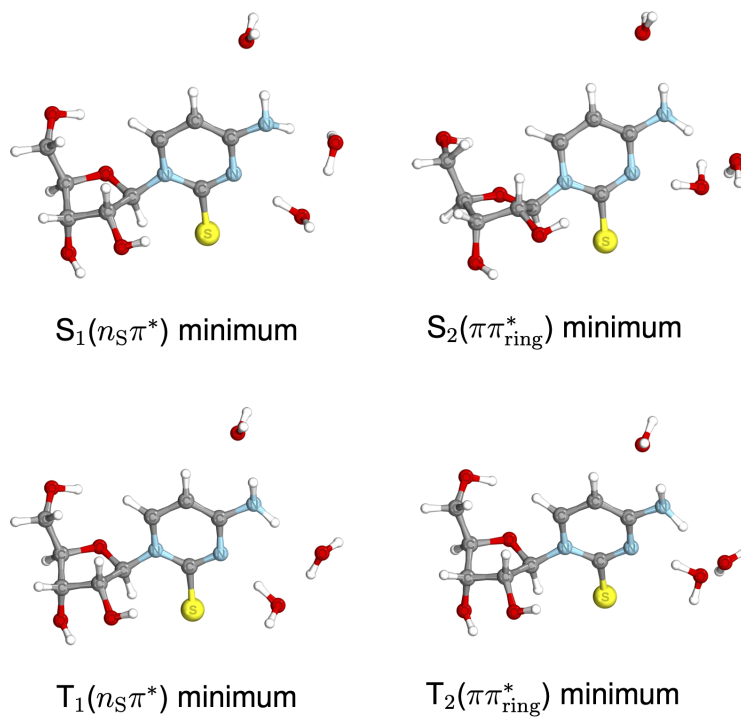
$S_1(n_S\pi^*)$ minimum $S_2(\pi\pi^*_{\mathrm{ring}})$ minimum

$T_1(n_S\pi^*)$ minimum $T_2(\pi\pi^*_{\mathrm{ring}})$ minimum

Figure S7: The excited-state minima of microhydrated β-thioCyd at the MP2/ADC(2)/cc-pVTZ level of theory.

$S_1(n_S\pi^*)$ minimum $S_2(\pi\pi^*_{\mathrm{ring}})$ minimum

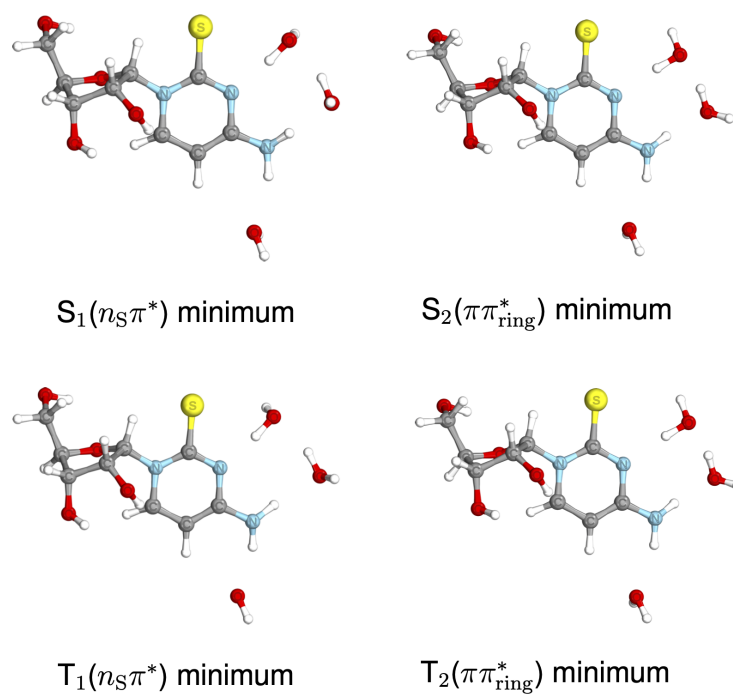$T_1(n_S\pi^*)$ minimum $T_2(\pi\pi^*_{\mathrm{ring}})$ minimum

Figure S8: The equilibrium excited-state minima of microhydrated α-thioCyd at the MP2/ADC(2)/cc-pVTZ level of theory.

Excited-state absorption spectra

FTAS spectra for α-**thioCyd**
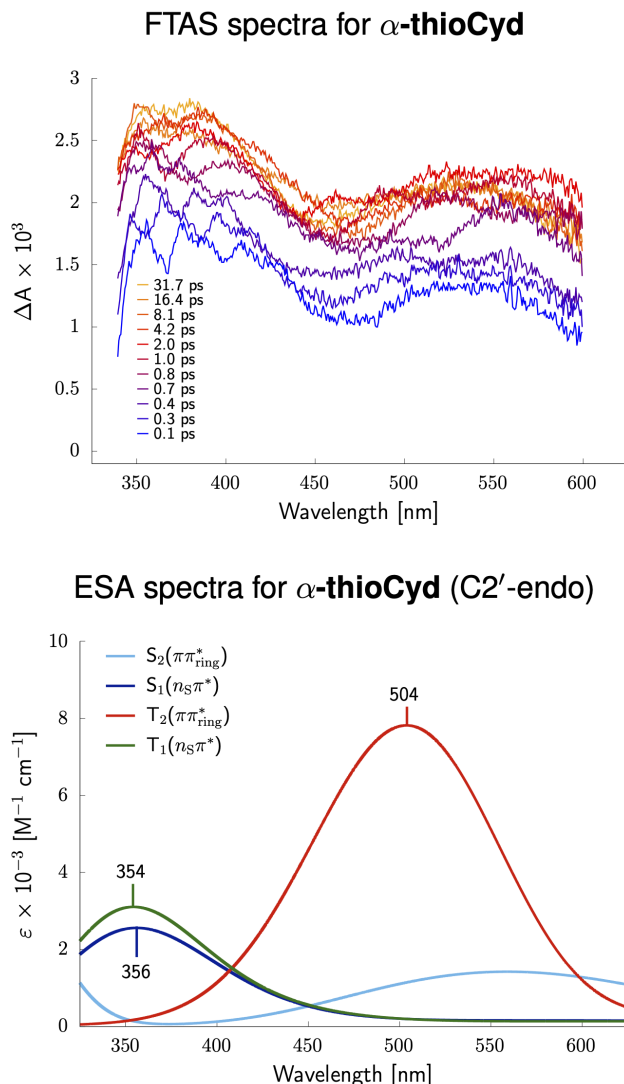

ESA spectra for α-**thioCyd** (C2′-endo)

Figure S9: The top panel presents the transient VIS absorption difference spectra of α-thioCyd following excitation at 308 nm. In the bottom panel, there are simulated excited-state absorption (ESA) spectra from specific excited-state (singlet and triplet) minima. Results for β-thioCyd are presented in the main article.

Having the excited-state minimum-energy structures of microhydrated systems and the protocol for simulations of excited-state absorption (ESA) spectra, we simulated the ESA spectra for thioC and β-thioCyd which are presented in the main article (see Fig. 3). The same computational protocol (*vide supra*) to predict ESA spectra for α-thioCyd (Fig. S9), assuming the optimized excited-state minima shown in Fig. S8. The S_1 and T_1 excited-state

structures exhibit very similar absorption bands having maxima at 354 and 356 nm (see Fig. S9, respectively. Both the S_1 and T_1 states are characterized by the $n_S\pi^*$ molecular orbital character. The T_2 minimum-energy structure associated with the $\pi\pi_{\mathrm{ring}}$ excitation has a broad absorption band in the range of 400-600 nm with the maximum at 504 nm. The measured pump-probe transient absorption spectra (TAS) of α-thioCyd (the top panel in Fig. S9) show two absorption bands marked by maxima located at 370 and 520 nm after 30 ps. Comparing the simulated ESA spectra and the collected data within the first picoseconds, we can assign the absorption band in the range of 350-450 nm to the population of the $S_1(n_S\pi^*)$ excited state. Since the photoexcitation of thioCyd enables the population of triplet excited states through the efficient intersystem crossing, the observed long-lived signals having maxima at 370 and 520 nm should originate from the triplet exited states. The long-lived absorption band at 360 nm agrees very well with the simulated ESA spectrum for the T_1 state having the $^3n_S\pi^*$ character. The other recorded band at 520 nm can be assigned to the T_2 state that is characterized by the $^3\pi\pi_{\mathrm{ring}}$ transition. Given the assignment of the collected TAS signals to the specific singlet and triplet excited states, we postulate that the dominant triplet state populated at long time-delays of the TAS measurements for α-thioCyd is the $^3n_S\pi^*$ excitation. Furthermore, our results suggest that the $^3\pi\pi_{\mathrm{ring}}$ state is also populated during the photodynamics, but it is not prevalent as in the case of thioC. It is worth noting that the photodynamics of both anomers (α and β) of thioCyd is clearly consistent.

Photorelaxation pathways of α-thioCyd

In order to propose a plausible photorelaxation pathway of α-thioCyd, we performed excited-state geometry optimizations at the ADC(2)/cc-pVTZ level that allowed finding crucial excited-state minima and minimum-energy crossing points (MECPs) between various electronic states (see Fig. S10). The description of analogous calculations performed for β-thioCyd and thioC can be found in the main article (see also Fig. 4 in the main article).

Following the excitation to the bright state at 3.70 eV, α-thioCyd can undergo vibrational relaxation towards the S_2 minimum (2.70 eV). Subsequently, having substantial excess energy, the molecule can easily reach the S_2/S_1 MECP (3.07 eV) that enables the population of the S_1 excited state which minimum can be found at 2.35 eV with respect to the ground-state geometry. $alpha$-thioCyd can subsequently undergo efficient intersystem crossing to the T_2 state which has a shallow minimum-energy structure at the 2.58 eV. This is enabled by readily accessible S_1/T_2 MECP at 2.64 eV, which is also associated with the spin-orbit coupling (SOC) between the S_1 and T_2 states equal to 149 cm^{-1}. The lowest-lying triplet excited state can be then accessed in a practically barrierless manner through the T_2/T_1 MECP (2.65 eV). Similarly as in β-thioCyd, the T_1 is associated with the $n_S\pi^*$ molecular orbital character. The deactivation of the T_1 state can occur when the system overcome the energy barrier of 0.18 eV to the $T_1 \rightarrow S_0$ intersystem crossing located as the T_1/S_0 MECP at 2.51 eV. The foregoing photochemical relaxation channel of α-thioCyd is very similar to the photochemistry of β-thioCyd which was discussed in the main article.
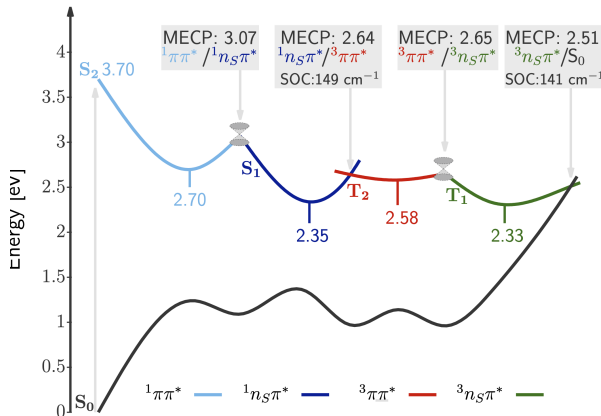

Figure S10: Excited-state potential energy surfaces calculated for α-thioCyd at the ADC(2)/cc-pVTZ level of theory.

XMS-CASPT2 excited-state profiles of thioC and β-thioCyd

In the main article, we discussed the excited-state potential energy surfaces of thioC and thioCyd calculated using the ADC(2)/cc-pVTZ method that works well for the description of photodeactivation pathways of nucleobases. However, this an approach has some limitations in the description of the state crossing between the first excited state and the electronic ground state. Therefore, we performed similar excited-state calculations employing the multiconfigurational method (XMS-CASPT2) that does not suffer from such problems. Excited-state potential energy surfaces calculated at this level of theory for thioC and β-thioCyd are presented in Fig. S13. Furthermore, Fig. S11 and Fig. S12 present molecular orbitals that were selected to build complete active spaces (CAS). In the CASSCF calculations of thioC, the following molecular orbitals were used to construct the CAS: π_{ring}, σ_{S}, π_{ring}, π, n_{S}, π_{S} and π^*_{ring}, π^*, π^*_{ring}, σ^*_{S} (12 electrons in 10 orbitals). Whereas in case of β-thioCyd the σ_{S}, π_{ring}, π, n_{S}, π_{S} and π^*_{ring}, π^*, π^*_{ring}, σ^*_{S} (10 electrons in 9 orbitals) molecular orbitals were used to build the reference CASSCF wavefunction.

The initial photoexcitation of thioC to the bright electronic state (3.63 eV) allows for efficient population of the S_1 state through the S_2/S_1 MECP at 2.80 eV. The minimum energy geometry located in the $S_1(n_{\text{S}}\pi^*)$ excited state is can be found at 2.76 eV with respect to the ground-state geometry. Optimization of the S_1/T_1 MECP (2.76 eV, SOC = 154 cm^{-1}) returned a $S_1/T_1(/T_2)$ three-state crossing that allows for efficient population of the T_1. On the $^3\pi\pi^*$ potential energy surface,

We found two different $T_1(^3\pi\pi^*)$ minima that are either characterized by lightly puckered aromatic ring (rp, 2.62 eV) or by the C=S bond tilted out of the plane of the aromatic ring (S-oop, 2.84 eV). The deactivation of the T_1 state can occur through the T_1/S_0 MECP at 3.55 eV that is characterized by the significant SOC of 125 cm^{-1}. Importantly, the system has to overcome the energy barrier of 0.93 eV in order to reach the T_1/S_0 MECP which indicates a long excited-state lifetime as demonstrated in the main article. The XMS-CASPT2 results are also qualitatively consistent with the ADC(2) excited-state calculations for thioC which
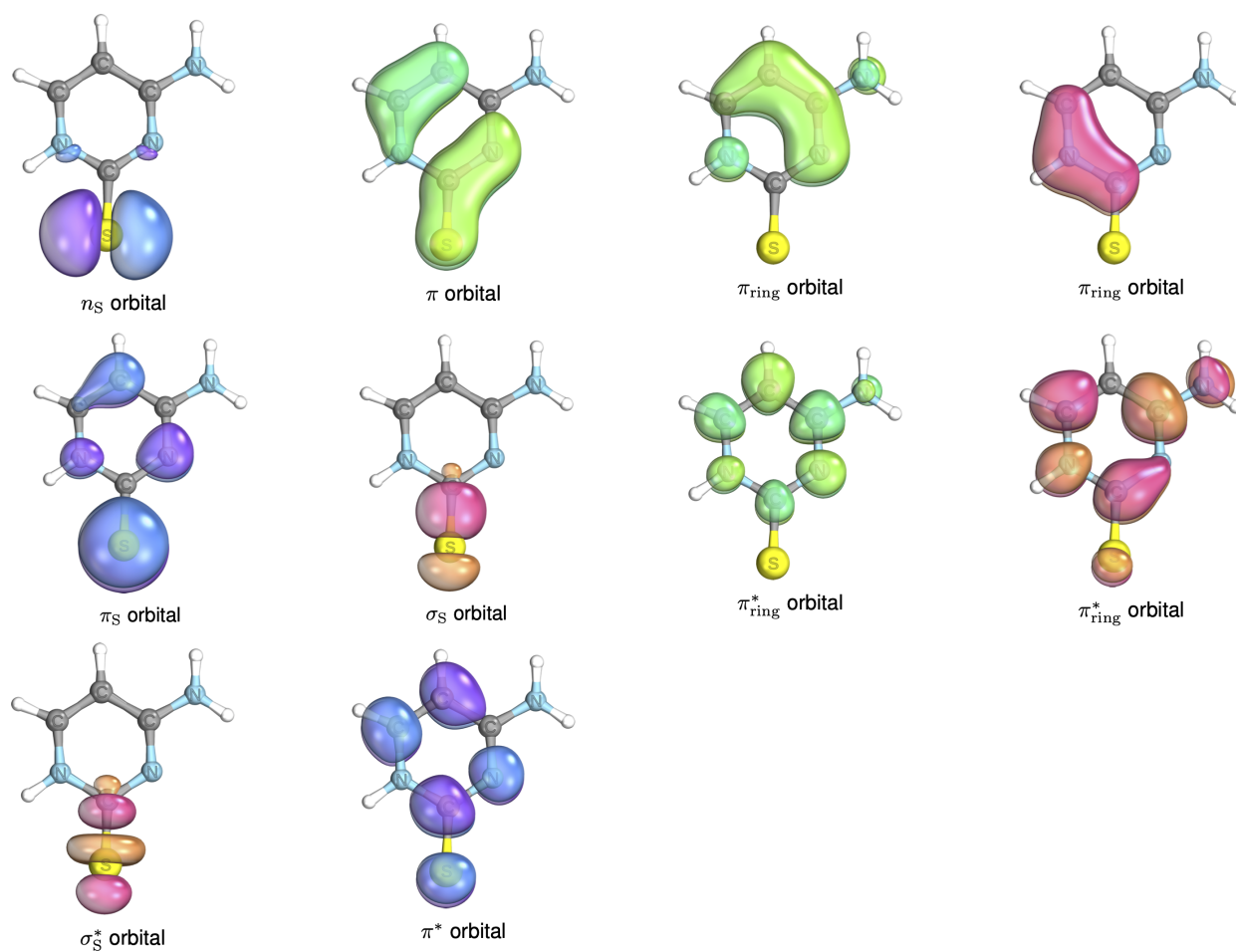
n_{S} orbital π orbital π_{ring} orbital π_{ring} orbital

π_{S} orbital σ_{S} orbital π_{ring}^{*} orbital π_{ring}^{*} orbital

σ_{S}^{*} orbital π^{*} orbital

Figure S11: The molecular orbital selected to the active space of the CASSCF wavefunction for thioC.

n_S orbital

π orbital

π_{ring} orbital

π_S orbital

σ_S orbital

π^* orbital

π^*_{ring} orbital
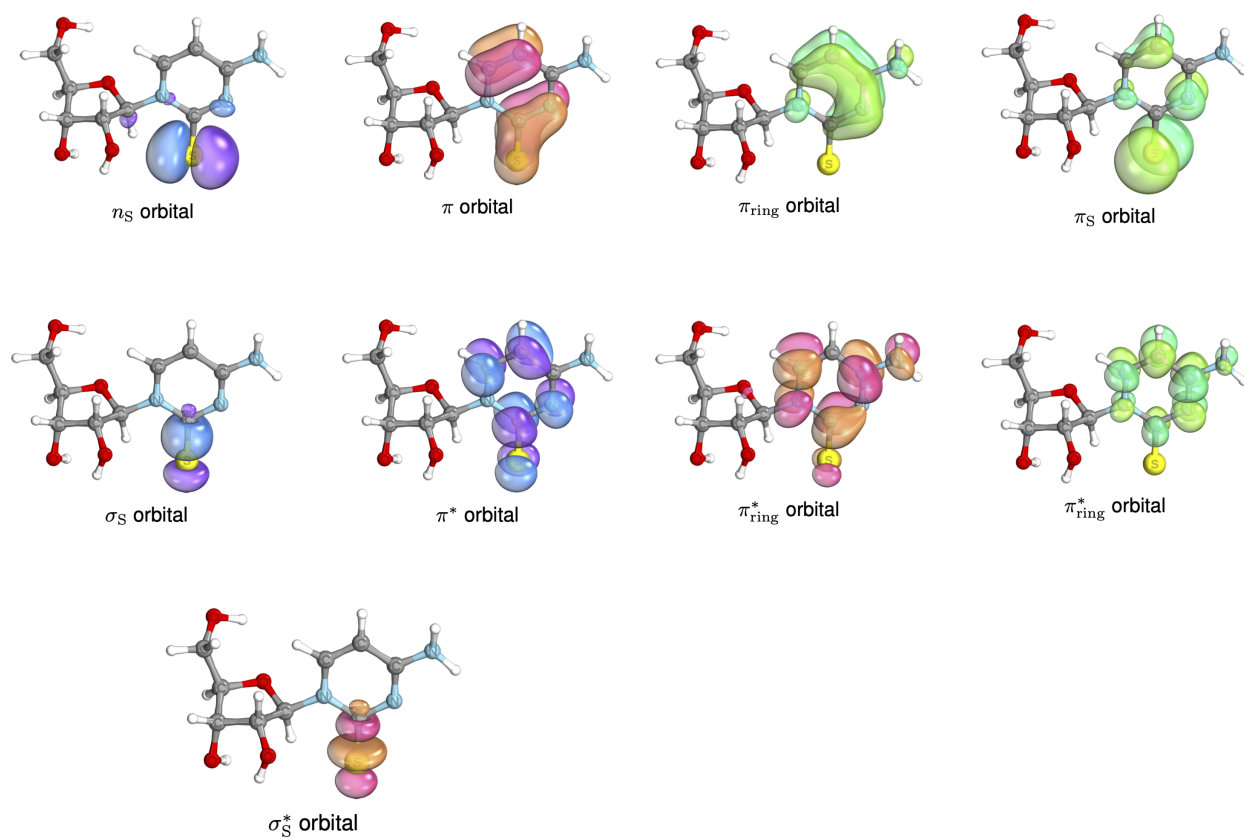
π^*_{ring} orbital

σ^*_S orbital

Figure S12: The molecular orbital selected to the active space of the CASSCF wavefunction for β-thioCyd.

are discussed in the main article. The main difference is associated with the energy barrier between the T_1 minimum and the T_1/S_0 MECP which amounts to 0.33 and 0.93 eV at the ADC(2) and XMS-CASPT2 level, respectively.

Following the excitation of β-thioCyd to the S_2 bright state at 3.70 eV (the right panel in Fig. S13), the nucleoside can undergo vibrational relaxation in the direction of the $^1\pi\pi^*$ minimum located at 2.71 eV. Having the excess energy from the population of the bright state, the system can then readily reach the S_2/S_1 MECP (2.92 eV) that enables subsequent population of the $S_1(n_S\pi^*)$ state. The minimum-energy geometry of β-thioCyd in the $S_1(n_S\pi^*)$ state has the energy of 2.47 eV. The S_1/T_2 MECP allowing for efficient intersystem crossing lies only 0.08 eV above the S_1 minimum and is associated with the SOC of 163 cm^{-1}. After intersystem crossing to the $T_2(\pi\pi^*)$ state, β-thioCyd can efficiently populate the lowest excited triplet state through the T_2/T_1 MECP, which energy is virtually identical to the energy of the T_2 minimum (2.53 eV). Similar to α-thioCyd the T_1 state of the β anomer can be characterised as an $n_S\pi^*$ excitation with the corresponding minimum on the T_1 PES at 2.46 eV. The photorelaxation of β-thioCyd from the T_1 state can occur through the $T_1 \rightarrow S_0$ intersystem crossing that we identified as the T_1/S_0 MECP having the SOC of 93 cm^{-1}. In order to reach this T_1/S_0 MECP, the system has to overcome the energy barrier that amounts to 0.62 eV. It is worth adding that this energy difference is much lower that in the case of the nucleobase, which is consistent which shorter excited-state lifetime of the nucleoside. Our XMS-CASPT2 calculations of β-thioCyd are in very good agreement with the ADC(2) calculations described in the main article. The only noticeable disparity between the XMS-CASPT2 and ADC(2) method is associated with the energy barrier for reaching the T_1/S_0 MECP which is equal to 0.62 and 0.30 eV, respectively.

The XMS-CASPT2 exploration of the excited-state potential energy surfaces of thioC and β-thioCyd provided plausible photorelaxation pathways for both molecules (Fig. S13). These results indicate that the $^3\pi\pi^*$ state plays a crucial role in the photochemistry of thioC whereas in the case of β-thioCyd, the $^3n_S\pi^*$ state seems to be the dominant excitation at

longer time delays. The stabilization of the $^3n_S\pi^*$ result from the presence of the sugar in β-thioCyd. It is wort noting that the sugar part provides a photochemically active hydrogen atom (C1′-H) that can interact with the lone electron pair of the sulphur atom (see Fig. S1) in excited states.
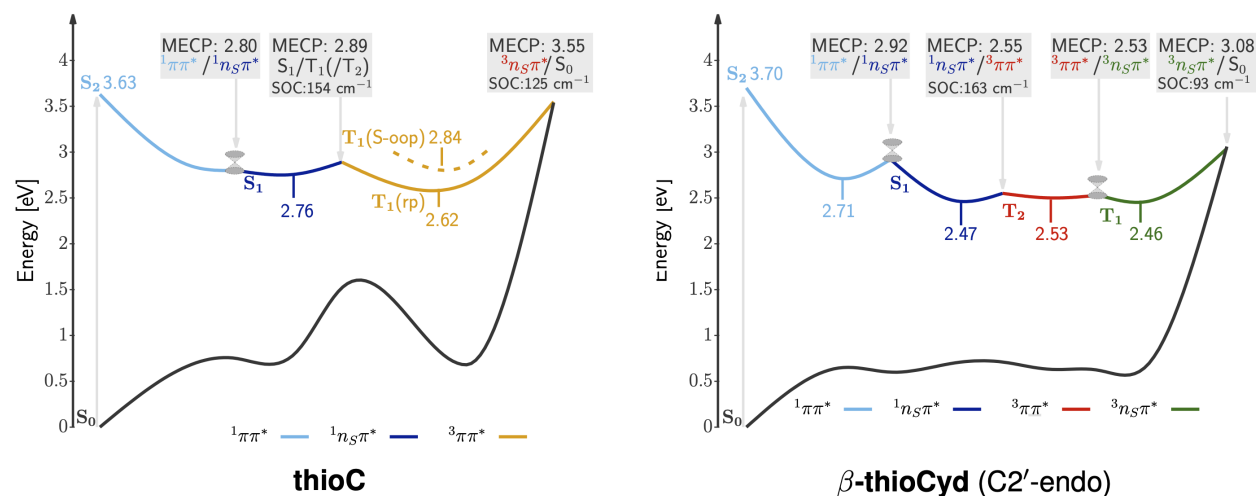

Figure S13: Excited-state potential energy surfaces calculated for thioC and β-thioCyd at the SA-CASSCF/XMS-CASPT2/cc-pVDZ level of theory.

References

(1) Shannahoff, D. H.; Sanchez, R. A. 2,2'-Anhydropyrimidine nucleosides. Novel syntheses and reactions. *J. Org. Chem.* **1973**, *38*, 593–598.

(2) Xu, J.; Tsanakopoulou, M.; Magnani, C. J.; Szabla, R.; Šponer, J. E.; Šponer, J.; Góra, R. W.; Sutherland, J. D. A prebiotically plausible synthesis of pyrimidine β-ribonucleosides and their phosphate derivatives involving photoanomerization. *Nat. Chem.* **2017**, *9*, 303–309.

(3) Schrader, T.; Sieg, A.; Koller, F.; Schreier, W.; An, Q.; Zinth, W.; Gilch, P. Vibrational relaxation following ultrafast internal conversion: comparing IR and Raman probing. *Chem. Phys. Lett.* **2004**, *392*, 358–364.

(4) Ryseck, G.; Schmierer, T.; Haiser, K.; Schreier, W.; Zinth, W.; Gilch, P. The Excited-State Decay of 1-Methyl-2(1H)-pyrimidinone is an Activated Process. *ChemPhysChem* **2011**, *12*, 1880–1888.

(5) Haiser, K.; Fingerhut, B. P.; Heil, K.; Glas, A.; Herzog, T. T.; Pilles, B. M.; Schreier, W. J.; Zinth, W.; de Vivie-Riedle, R.; Carell, T. Mechanism of UV-Induced Formation of Dewar Lesions in DNA. *Angew. Chem. Int. Ed.* **2012**, *51*, 408–411.

(6) Gutierrez-Osuna, R.; Nagle, H.; Schiffman, S. S. Transient response analysis of an electronic nose using multi-exponential models. *Sens. Actuators B Chem.* **1999**, *61*, 170–182.

(7) Satzger, H.; Zinth, W. Visualization of transient absorption dynamics – towards a qualitative view of complex reaction kinetics. *Chem. Phys.* **2003**, *295*, 287–295.

(8) Dominguez, P. N.; Himmelstoss, M.; Michelmann, J.; Lehner, F. T.; Gardiner, A. T.; Cogdell, R. J.; Zinth, W. Primary reactions in photosynthetic reaction centers of Rhodobacter sphaeroides – Time constants of the initial electron transfer. *Chem. Phys. Lett.* **2014**, *601*, 103–109.

(9) Weigend, F.; Häser, M. RI-MP2: first derivatives and global consistency. *Theor. Chem. Acc.* **1997**, *97*, 331–340.

(10) Dunning, T. H. Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *J. Chem. Phys.* **1989**, *90*, 1007–1023.

(11) Zhao, Y.; Truhlar, D. G. The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06-class functionals and 12 other function. *Theor. Chem. Acc.* **2008**, *120*, 215–241.

(12) Weigend, F.; Ahlrichs, R. Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297–3305.

(13) TURBOMOLE V7.3 201, a development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989-2007, TURBOMOLE GmbH, since 2007; available from `http://www.turbomole.com`.

(14) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H. et al. Gaussian~16 Revision A.03. 2016; Gaussian Inc. Wallingford CT.

(15) Schirmer, J. Beyond the random-phase approximation: A new approximation scheme for the polarization propagator. *Phys. Rev. A* **1982**, *26*, 2395–2416.

(16) Dreuw, A.; Wormit, M. The algebraic diagrammatic construction scheme for the polarization propagator for the calculation of excited states. *Wiley Interdiscip. Rev. Comput. Mol. Sci. 5*, 82–95.

(17) Hättig, C. Structure Optimizations for Excited States with Correlated Second-Order Methods. *Adv. Quantum Chem.* **2005**, *50*, 37–60.

(18) Martin, R. L. Natural transition orbitals. *J. Chem. Phys.* **2003**, *118*, 4775–4777.

(19) F. Plasser, "TheoDORE 1.4: A package for theoretical density, orbital relaxation, and exciton analysis". `http://theodore-qc.sourceforge.net` (2018).

(20) Levine, B. G.; Coe, J. D.; Martínez, T. J. Optimizing Conical Intersections without Derivative Coupling Vectors: Application to Multistate Multireference Second-Order Perturbation Theory (MS-CASPT2). *J. Phys. Chem. B* **2008**, *112*, 405–413.

(21) Andersson, K.; Malmqvist, P. A.; Roos, B. O.; Sadlej, A. J.; Wolinski, K. Second-order perturbation theory with a CASSCF reference function. *J. Phys. Chem.* **1990**, *94*, 5483–5488.

(22) Shiozaki, T.; Győrffy, W.; Celani, P.; Werner, H.-J. Communication: Extended multi-state complete active space second-order perturbation theory: Energy and nuclear gradients. *J. Chem. Phys.* **2011**, *135*, 081106.

(23) Veryazov, V.; Malmqvist, P. Å.; Roos, B. O. How to select active space for multiconfigurational quantum chemistry? *Int. J. Quantum Chem.* **2011**, *111*, 3329–3338.

(24) Shiozaki, T. BAGEL: Brilliantly Advanced General Electronic-structure Library. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2018**, *8*, e1331.

(25) BAGEL, Brilliantly Advanced General Electronic-structure Library. `http://www.nubakery.org` under the GNU General Public License.

(26) Aquilante, F.; Autschbach, J.; Carlson, R. K.; Chibotaru, L. F.; Delcey, M. G.; De Vico, L.; Fdez. Galván, I.; Ferré, N.; Frutos, L. M.; Gagliardi, L. et al. Molcas 8: New capabilities for multiconfigurational quantum chemical calculations across the periodic table. *J. Comput. Chem.* **2016**, *37*, 506–541.

(27) Petrenko, T.; Neese, F. Analysis and prediction of absorption band shapes, fluorescence band shapes, resonance Raman intensities, and excitation profiles using the time-dependent theoo ry of electronic spectroscopy. *J. Chem. Phys.* **2007**, *127*, 164319.

(28) Petrenko, T.; Neese, F. Efficient and automatic calculation of optical band shapes and resonance Raman spectra for larger molecules within the independent mode displaced harmonic oscillator model. *J. Chem. Phys.* **2012**, *137*, 234107.

(29) Klamt, A.; Schüürmann, G. COSMO: a new approach to dielectric screening in solvents

with explicit expressions for the screening energy and its gradient. *J. Chem. Soc., Perkin Trans. 2* **1993**, 799–805.

(30) Ferrer, F. J. A.; Improta, R.; Santoro, F.; Barone, V. Computing the inhomogeneous broadening of electronic transitions in solution: a first-principle quantum mechanical approach. *Phys. Chem. Chem. Phys.* **2011**, *13*, 17007.

(31) Zaleśny, R.; Murugan, N. A.; Tian, G.; Medved', M.; Ågren, H. First-Principles Simulations of One- and Two-Photon Absorption Band Shapes of the Bis(BF_2) Core Complex. *J. Phys. Chem. B* **2016**, *120*, 2323–2332.

(32) Neese, F. The ORCA program system. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2012**, *2*, 73–78.

(33) Neese, F. Software update: the ORCA program system, version 4.0. *WIREs Computational Molecular Science* **2018**, *8*.

(34) Neese, F.; Wennmohs, F.; Becker, U.; Riplinger, C. The ORCA quantum chemistry program package. *J. Chem. Phys.* **2020**, *152*, 224108.

(35) Mai, S.; Pollum, M.; Martínez-Fernández, L.; Dunn, N.; Marquetand, P.; Corral, I.; Crespo-Hernández, C. E.; González, L. The origin of efficient triplet state population in sulfur-substituted nucleobases. *Nat. Comm.* **2016**, *7*, 13077.