nature

https://doi.org/10.1038/s41586-020-2330-9

Supplementary information

Selective prebiotic formation of RNA pyrimidine and DNA purine nucleosides

In the format provided by the authors and unedited

Jianfeng Xu, Václav Chmela, Nicholas J. Green, David A. Russell, Mikołaj J. Janicki, Robert W. Góra, Rafał Szabla, Andrew D. Bond & John D. Sutherland✉

Selective prebiotic formation of RNA pyrimidine and DNA purine nucleosides

Jianfeng Xu[1]†, Václav Chmela[1]†, Nicholas J. Green[1], David A. Russell[1], Mikołaj J. Janicki[2], Robert W. Góra[2], Rafał Szabla[3,4], Andrew D. Bond[5] and John D. Sutherland[1]*

[1]MRC Laboratory of Molecular Biology, Francis Crick Avenue, Cambridge Biomedical Campus, Cambridge, CB2 0QH, UK.

[2]Department of Physical and Quantum Chemistry, Wrocław University of Science and Technology, Faculty of Chemistry, Wybrzeże Wyspiańskiego 27, 50-370, Wrocław, Poland.

[3]EaStCHEM, School of Chemistry, University of Edinburgh, Joseph Black Building, David Brewster Road, Edinburgh, EH9 3FJ, UK.

[4]Institute of Physics, Polish Academy of Sciences, Al. Lotników 32/46, PL-02668 Warsaw, Poland.

[5]Department of Chemistry, University of Cambridge, Lensfield Road, CB2 1EW, UK.

*Correspondence to: johns@mrc-lmb.cam.ac.uk

†These authors contributed equally to this work.

This PDF file includes:

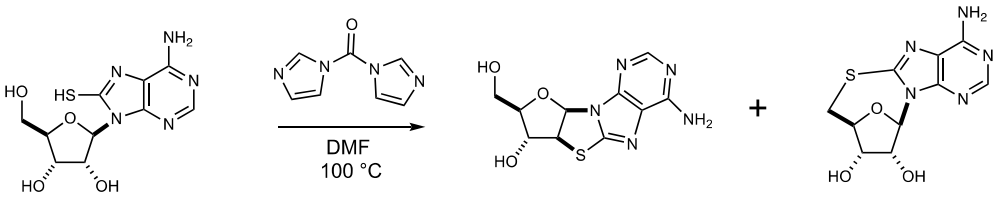
Materials and Methods

General Methods

Reagents and deuterated solvents used for reactions were purchased from Sigma-Aldrich or Acros Organics and were used without further purification. All photochemical reactions were carried out in Norell Suprasil quartz NMR tubes purchased from Sigma-Aldrich using Hg lamps with principal emission at 254 nm in a Rayonet photochemical chamber reactor RPR-200, acquired from The Southern New England Ultraviolet Company. A Mettler Toledo SevenEasy pH Meter S20 was used to monitor the pH, and deoxygenation of solutions was achieved by sparging anhydrous argon through the solution for 15-20 min. A preparative Varian Prostar HPLC System was used for the reverse phase high-pressure liquid chromatography (RP-HPLC) linked with an Atlantis T3 C18 Prep Column OBD 10 μm (19×250 mm). Analytical HPLC was used to monitor reactions, using a Thermofisher Ultimate 3000 UPLC and Waters Atlantis T3 C18 column (5 μm, 4.6x150 mm). All unknown compounds in reaction mixtures were confirmed by spiking experiments with authentic compounds either purchased from Sigma-Aldrich or synthesised in house using conventional synthetic chemistry. ^{1}H and ^{13}C NMR spectra were acquired using a Bruker Ultrashield 400 Plus NMR spectrometer operating at 400.1 MHz and 100.6 MHz, respectively. Samples consisting of H_2O/D_2O mixtures were analysed using HOD suppression to collect ^{1}H NMR data. Chemical shifts (δ) are shown in ppm. The conversion and yields were determined by relative integration of signals using a known amount of pentaerythritol as internal standard in the ^{1}H NMR spectrum. Coupling constants (*J*) are given in Hertz (Hz) and the notations s, d, and m represent the multiplicities singlet, doublet, and multiplet. Mass spectra were recorded with an Agilent Technologies 6130 Quadrupole LC-MS using positive and negative Electron Spray Ionisation. The accurate mass spectra were recorded with a Waters Vion IMS QTOF Ion Mobility Quadrupole Time-of-flight mass spectrometer using positive ESI. Single-crystal X-ray diffraction data were collected using a D8-QUEST PHOTON-100 diffractometer equipped with an Incoatec IμS Cu microsource (CuKα radiation, λ= 1.5418 Å). The temperature was held at 180(2) K using an Oxford Cryosystems N_2 cryostat. Data integration and reduction were undertaken with SAINT in the APEX3 software suite. Multi-scan corrections were applied using SADABS. Structures were solved using SHELXT[51] and refined using SHELXL[52].

Synthetic Procedures

Synthesis of N^9-8, 2'-anhydro-thioadenosine



36 **18** **37**

8-Mercaptoadenosine **36**[53] (200 mg, 0.67 mmol) and 1,1'-carbonyldiimidazole (217 mg, 1.34 mmol) were dissolved in DMF (5 mL) and the solution was heated at 100 °C for 4 days (90 hours). The reaction mixture was allowed to cool to room temperature and the solvent was evaporated under reduced pressure. The residue was applied to reverse phase HPLC to afford 51 mg (0.18 mmol, 27% yield) of N^9-8, 2'-anhydro-thioadenosine **18** and 34 mg (0.12 mmol, 18% yield) of N^9-8, 5'-anhydro-thioadenosine **37**.

***N*^{9}-8, 2'-anhydro-thioadenosine 18:** ^{1}H NMR (400 MHz, DMSO-d_6): δ = 8.05 (1H, s, H2), 7.09 (2H, s, NH_2), 6.50 (1H, d, J = 6.6 Hz, H1'), 5.88 (1H, d, J = 4.6 Hz, HO- (C2')), 4.90–4.84 (2H, m, H2', HO-(C5'), overlapping), 4.39 (1H, td, J = 4.2, 2.4 Hz, H3'), 3.97 (1H, td, J = 5.6, 3.9 Hz, H4'), 3.45 (1H, dt, J = 11.2, 5.4 Hz, H5'), 3.39–3.34 (1H, m, H5''); ^{13}C NMR (101 MHz, DMSO-d_6): δ = 154.0 (C8), 152.9 (C4), 151.6 (C2), 149.1 (C6), 124.1 (C5), 88.0 (C4'), 86.2 (C1'), 77.3 (C3'), 62.3 (C2'), 61.1 (C5'). ESI-LCMS (pos. m/z): 282.0 $[M+1]^+$.

***N*^{9}-8, 5'-anhydro-thioadenosine 37.** ^{1}H NMR (400 MHz, D_2O): δ = 7.99 (1H, s, H2), 6.17 (1H, d, J = 1.8 Hz, H1′), 4.98 (1H, dd, J = 6.2, 1.8 Hz, H2′), 4.94 (1H, t, J = 2.8 Hz, H4′), 4.58 (1H, d, J = 6.2 Hz, H3′), 3.32 (1H, dd, J = 14.8, 2.1 Hz, H5′), 3.01 (1H, dd, J = 14.9, 3.6 Hz, H5′); ^{13}C NMR (101 MHz, D_2O): δ = 154.1 (C8), 152.5 (C2), 149.8 (C6), 147.3 (C4), 117.6 (C5), 90.3 (C1′), 89.2 (C4′), 77.0 (C2′), 73.9 (C3′), 34.0 (C5′). ESI-LCMS (pos. m/z): 282.2 $[M+1]^+$.

Prebiotic Synthesis Procedures

Adenine Hydrolysis

Adenine **8** (500 mg, 3.70 mmol) and sodium phosphate buffer (1M, pH 8) were combined in a glass tube. The tube was sealed with a crimp cap and the mixture heated at 138 °C in an aluminium heating block. After ten or twelve days, the reaction was removed from heating and allowed to cool, then centrifuged for 15 minutes. The supernatant was removed and stored at 4 °C for 24 hours. The material was centrifuged again and the supernatant removed. The quantity of supernatant was measured and an aliquot of 0.200 mL removed and freeze-dried for analysis. The resulting solid was dissolved in DMSO-d_6/D_2O (3:1) (the D_2O simplifies analysis by removing exchangeable protons from the spectrum) with a known quantity of internal standard, sodium 3-(trimethylsilyl)propionate-*2,2,3,3-d_4* . Yields were determined based on internal standard integration (Fig. S1). The identity of the products was confirmed by NMR spiking experiments and analytical HPLC.

A similar reaction was conducted using 1 g of adenine **8**. Each reaction was conducted in duplicate. Since the reaction is at low conversion, yields of the ten- and twelve-day reactions were similar but the twelve-day reactions tended to be slightly higher. Yield ranges for the eight reactions are reported.

Differential solubility measurements of TAP 21 and Adenine 8

A saturated, boiling solution of adenine was added to varying quantities of TAP. Aliquots were taken whilst the mixture was still hot, diluted, and the ratio of adenine to TAP measured by ^{1}H NMR. The mixtures were stored at 4 °C overnight and an aliquot of the supernatant taken and analysed by the same procedure. The ratios before and after cooling are reported below, plus the spectra for entry 3 (Table S1).

Synthesis of 8-mercaptoadenine 16

TAP **21** (5 mg, 0.04 mmol) and NH_4SCN **24** or thiourea **25** (2-5 equiv., see Table S2) were combined in a 5 mL round-bottomed flask and dissolved in H_2O (1 mL). In some cases the pH was modified as specified with either 0.1 M HCl or NaOH. The solution was evaporated under reduced pressure and the dry material heated in an oil bath for the specified time at the specified temperature. The material was allowed to cool before it was dissolved in DMSO-d_6 and the conversion measured by integration of ^{1}H NMR signals (Table S2). When the material was dissolved in H_2O and stored at room temperature, crystals were observed to form over several weeks. The crystals were isolated and identified as pure 8-mercaptoadenine **16**.

From adenine hydrolysate:

To 100 uL of adenine hydrolysate and 1 mL of H_2O was added 28 mg of NH_4SCN. The pH was measured to be 7.5. The solvent was evaporated and the residue heated at 150 °C for 16 hours. The residue was then cooled and dissolved in DMSO-d_6 and NMR spectra were recorded. 8-Mercaptoadenine **16** was confirmed to be present by spiking the sample with authentic material.

Synthesis of α-D-ribofuranosyl-2, 2'-anhydrouridine 15 from α-thiouridine 14 in water

α-Thiouridine **14** (1 mg, 0.01 mmol) was dissolved in 0.5 mL of H_2O (containing 10% of D_2O) and the pH of mixture was adjusted to 7. The mixture was transferred to a NMR tube and heated at 80 °C for 30 h. The reaction was monitored by ^{1}H NMR periodically. α-D-Ribofuranosyl-2, 2'-anhydrouridine **15** was formed in 63% yield after 30 hours, at which point 25% α-thiouridine **14** and 12% of the hydrolysed product, α-uridine **38** were also present in the mixture.

Synthesis of α-D-ribofuranosyl-2,2'-anhydrouridine 15 from α-thiouridine 14 in formamide

α-Thiouridine **14** (1 mg, 0.01 mmol) was dissolved in 0.5 mL of formamide and the solution was heated at 80 °C for 30 hours. The reaction was monitor by ^{1}H NMR at intervals. α-D-Ribofuranosyl-2, 2'-anhydrouridine **15** was formed in 89%

yield after 30 hours, at which point 11% of the starting material, α-thiouridine **14**, remained in the mixture.

Synthesis of N^9-8, 2'-anhydro-thioadenosine 18 from 15 by dry state heating in the absence of magnesium chloride

α-D-Ribofuranosyl-2, 2'-anhydrouridine **15** (10 mg, 0.04 mmol) and 8-mercaptoadenine **16** (15 mg, 0.09 mmol) were dissolved in 6 mL of H_2O in a round bottom flask. The pH was adjusted to 8.0 before the mixture was evaporated under reduced pressure and heated at 150 °C in the dry state for 4 days. The dry residue was dissolved in DMSO-d_6 and the resulting solution was transferred to an NMR tube to record the ^{1}H NMR spectrum of the crude reaction. 50 μL of 50 mM pentaerythritol solution in D_2O was added to the mixture to enable quantification of the reaction products by ^{1}H NMR spectroscopy. The yield of N^9-8, 2'-anhydro-thioadenosine **18** was 10%, while the yield of N^7-8, 2'-anhydro-thioadenosine **19** was 12% when referring to the proton integration compared to the internal standard. To obtain a pure sample of N^7-8, 2'-anhydro-thioadenosine **19** for characterization by NMR, the reaction was repeated on a larger scale with 1.06 g (4.7 mmol) of α-D-ribofuranosyl-2,2'-anhydrouridine **15** and 1.17 g (7 mmol) of 8-mecaptoadenine **16**. The crude material obtained after reaction was subjected to preparative reverse phase HPLC. Acetonitrile in water was used as the eluent. 185 mg of pure N^9-8, 2'-anhydro-thioadenosine **18** (14% yield) and 211 mg of pure N^7-8, 2'-anhydro-thioadenosine **19** (16% yield) were obtained in this way. The ^{1}H and ^{13}C NMR spectra of the N^9-8, 2'-anhydro-thioadenosine **18** were identical to those obtained from the synthetic procedure described above.
***N*^{7}-8, 2'-Anhydro-thioadenosine 19:** ^{1}H NMR (400 MHz, DMSO-d_6): δ = 8.14 (1H, s, H2), 6.84 (1H, d, J = 6.5 Hz, H1'), 6.72 (2H, s, NH_2), 4.78 (1H, dd, J = 6.5, 3.2 Hz, H2'), 4.39 (1H, t, J = 3.7 Hz, H3'), 4.01 (1H, q, J = 5.0 Hz, H4'), 3.36 (4H, m, H5' and OH); ^{13}C NMR (101 MHz, DMSO-d_6): δ = 164.8 (C4), 157.9 (C8), 152.5 (C2), 149.5 (C6), 110.6 (C5), 89.3 (C1'), 88.0 (C4'), 77.7 (C3'), 61.1 (C5'), 60.8 (C2'). ESI-LCMS (pos. m/z): 282.0 $[M+1]^+$.

Synthesis of N^9-8, 2'-anhydro-thioadenosine 18 from 15 by dry state heating in the presence of magnesium chloride

α-D-Ribofuranosyl-2, 2'-anhydrouridine (10 mg, 0.04 mmol) **15**, 8-mercaptoadenine **16** (15 mg, 0.09 mmol) and $MgCl_2$ (4 mg, 0.04 mmol) were dissolved in 6 mL of H_2O in a round bottom flask. The pH was adjusted to 8.0 before the mixture was concentrated under reduced pressure and heated at 150 °C in the dry state for 1 day. The dry residue was dissolved in DMSO-d_6 and the resulting solution was transferred to an NMR tube to record the ^{1}H NMR spectrum of the crude reaction. 100 μL of 50 mM pentaerythritol solution in D_2O was added to the mixture to enable quantification of the reaction products by ^{1}H NMR spectroscopy. The yield of N^9-8, 2'-anhydro-thioadenosine **18** was 39%, while the yield of N^7-8, 2'-anhydro-thioadenosine **19** was 48% when referring to the proton integration compared to the internal standard. Uracil **39** (87% yield) was observed in the crude mixture as the only product deriving from nucleobase loss in the glycosidation step, alongside α-uridine **38** (13% yield) due to hydrolysis.

Photoreduction of N^9-8, 2'-anhydro-thioadenosine 18 with hydrogen sulfide as the reductant

N^9-8, 2'-Anhydro-thioadenosine **18** (2 mg, 0.007 mmol) was mixed with of $NaSH.xH_2O$ (5 mg, 0.054 mmol, containing 60% NaSH) and NaH_2PO_4 (12 mg, 0.1 mmol) in 0.5 mL of H_2O (containing 10% of D_2O) and the pH of mixture was adjusted to 7 with 1 M HCl. The mixture was transferred to a Suprasil quartz NMR tube (sample dimensions 40 mm high, 4.2 mm diameter), which was irradiated at 254 nm in a *Rayonet* photochemical chamber reactor RPR-200 for 6 hours. The sample was taken out from the reactor at intervals to record ^{1}H NMR spectra. At the end of the reaction, 20 μL of 50 mM pentaerythritol solution in H_2O was added to the mixture to enable quantification of the reaction products by ^{1}H NMR spectroscopy. The identities of products were confirmed after the completion of the experiment by spiking with authentic standards. After the quantification and spiking experiments with ^{1}H NMR, N^9-deoxyadenosine **7** (39% yield), 8-mercapto-deoxyadenosine **26**[54] (15% yield) and 8-mercaptoadenine **16** (10%) alongside adenine **8** (17%), were confirmed as products.

Photoreduction of N^9-8, 2'-anhydro-thioadenosine 18 with bisulfite as the reductant

N^9-8, 2'-Anhydro-thioadenosine **18** (2 mg, 0.007 mmol) was mixed with Na_2SO_3 (10 mg, 0.079 mmol) in 0.5 mL of H_2O (containing 10% of D_2O) and the pH of mixture was adjusted to 7 with 1 M HCl. The mixture was transferred to a Suprasil quartz NMR tube, which was irradiated at 254 nm for 6 hours. At the end of the reaction, 30 μL of 50 mM pentaerythritol solution in H_2O was added to the mixture to enable quantification of the reaction products by ^{1}H NMR spectroscopy. After the quantification and spiking experiments, N^9-deoxyadenosine **7** (51% yield), 8-mercapto-deoxyadenosine **26** (23% yield), 8-mercaptoadenine **16** (14%) and adenine **8** (8%) were confirmed as products.

Photoreduction of 8-mercapto-deoxyadenosine 26 under UV light

8-Mercapto-deoxyadenosine **26** (1 mg, 0.004 mmol) was dissoved in 0.5 mL of H_2O (containing 10% of D_2O). The pH of mixture was adjusted to 7 before irradiation. 10 μL of 50 mM pentaerythritol solution in H_2O was added to the solution to enable quantification of the reaction products by ^{1}H NMR spectroscopy. The mixture was transferred to a Suprasil quartz NMR tube, which was irradiated at 254 nm for 23 hours. 8-Mercapto-deoxyadenosine **26** was completely consumed and N^9-deoxyadenosine **7** (in 90% yield referred to the proton integration) was the only product observed in the crude ^{1}H NMR spectrum.

Photoreduction of N^7-8, 2'-anhydro-thioadenosine 19 with hydrogen sulfide as the reductant

N^7-8, 2'-Anhydro-thioadenosine **19** (1 mg, 0.004 mmol) was mixed with $NaSH.xH_2O$ (5 mg, 0.054 mmol, containing 60% NaSH) and NaH_2PO_4 (12 mg, 0.1 mmol) in 0.5 mL of H_2O (containing 10% of D_2O) and the pH of mixture was adjusted to 7 with 1 M HCl. The mixture was transferred to a Suprasil quartz NMR tube and irradiated at 254 nm for 5 hours. The sample was taken out from the reactor at intervals to record ^{1}H NMR spectra. At the end of the reaction, 20 μL of 50 mM pentaerythritol solution in H_2O was added to the mixture to enable quantification of the reaction products by ^{1}H NMR spectroscopy. The identities of products were confirmed after completion of the experiment by spiking with authentic standards. After the quantification and spiking experiments, N^7-deoxyadenosine **20** (23% yield), 8-mercaptoadenine **16** (10%) and adenine **8** (25%) were confirmed as products. N^7-deoxyadenosine **20** was synthesised according to known procedures[55, 56], and this authentic material was used for ^{1}H NMR spiking experiments here and in the hydrolysis study described below.

Photoreduction of N^7-8, 2'-anhydro-thioadenosine 19 with bisulfite as the reductant

N^7-8, 2'-Anhydro-thioadenosine **19** (1 mg, 0.004 mmol) was mixed with Na_2SO_3 (5 mg, 0.039 mmol) in 0.5 mL of H_2O (containing 10% of D_2O) and the pH of mixture was adjusted to 7 with 1 M HCl. 20 μL of 50 mM pentaerythritol solution in H_2O was added to the solution before irradiation to enable quantification of the reaction products by ^{1}H NMR spectroscopy. The mixture was transferred to a Suprasil quartz NMR tube and irradiated at 254 nm for 2 hours. At this point the sample was taken out from the reactor and the mixture was analysed by ^{1}H NMR spectroscopy, which showed that the starting material had been completely consumed, but no new nucleosides had formed.

Photoreduction of N^9-8, 2'-anhydro-thioadenosine 18 and N^7-8, 2'-anhydro-thioadenosine 19 with hydrogen sulfide as the reductant

N^9-8, 2'-Anhydro-thioadenosine **18** (1 mg, 0.004 mmol) and N^7-8, 2'-anhydro-thioadenosine **19** (1.4 mg, 0.005 mmol) was mixed with $NaSH.xH_2O$ (5 mg, 0.054 mmol, containing 60% NaSH) and NaH_2PO_4 (12 mg, 0.1 mmol) in 0.5 mL of H_2O (containing 10% of D_2O) and the pH of mixture was adjusted to 7 with 1 M HCl. 20 μL of 50 mM pentaerythritol solution in H_2O was added to the solution before irradiation to enable quantification of the reaction products by ^{1}H NMR spectroscopy. The mixture was transferred to a Suprasil quartz NMR tube, which was irradiated at 254 nm for 12 hours. The sample was taken out from the reactor at intervals to record ^{1}H NMR spectra. After the quantification and spiking experiments, N^9-deoxyadenosine **7** (31% yield) and N^7-deoxyadenosine **20** (17% yield) were confirmed to have formed. The ratio of the N^9 to N^7 regioisomers was found to have increased to 56:44 from 38:62 (the ratio in starting materials).

Photoreduction of N^9-8, 2'-anhydro-thioadenosine 18 and N^7-8, 2'-anhydro-thioadenosine 19 with bisulfite as the reductant

N^9-8,2'-Anhydro-thioadenosine **18** (0.8 mg, 0.003 mmol) and N^7-8,2'-anhydro-thioadenosine **19** (1 mg, 0.0034 mmol) were mixed with Na_2SO_3 (5 mg, 0.04 mmol) in 0.5 mL of H_2O (containing 10% of D_2O) and the pH of mixture was adjusted to 7 with 1 M HCl. 20 µL of 50 mM pentaerythritol solution in H_2O was added to the solution before irradiation to enable quantification of the reaction products by ^1H NMR spectroscopy. The mixture was transferred to a Suprasil quartz NMR tube and irradiated at 254 nm for 7 hours. The sample was taken out from the reactor at intervals to record ^1H NMR spectra. After the quantification and spiking experiments, N^9-deoxyadenosine **7** (44% yield), 8-mercapto-deoxyadenosine **26** (18% yield) and adenine **8** (11% yield) were confirmed as products. No N^7-regioisomers were detected in the crude ^1H NMR spectra after completion of the experiment.

Hydrolysis studies of N^7-deoxyadenosine 20, N^9-deoxyadenosine 7 and N^7-deoxyadenosine 20 *vs* N^9-deoxyadenosine 7 in 1 M NaOAc/ HOAc buffer (pH = 4)

The hydrolysis of 150 µM of nucleosides in 1 M NaOAc/ HOAc buffer (200 µL) at pH 4.0 was monitored by analytical HPLC, using a Thermofisher Ultimate 3000 UPLC and Waters Atlantis T3 C18 column (5 µm, 4.6x150 mm). 5 µL of solution was taken out from the reaction at intervals by the auto sampler and injected into analytical HPLC. The C18 column was eluted with acetonitrile in 25 mM TEAA (triethylammonium acetate, pH = 4.5) buffer. Plotting of the nucleoside concentration (relative peak area) against time yielded first order half-lives at pH 4.0 for N^7-deoxyadenosine **20** and N^9-deoxyadenosine **7** of 21.5 hours and 1565 hours, respectively.

Reaction of 8-mercapto-deoxyadenosine 26 with nitrous acid

8-Mercapto-deoxyadenosine **26** (2.8 mg, 0.010 mmol), sodium nitrite (6.9 mg, 0.100 mmol) and dibasic sodium phosphate (14.2 mg, 0.100 mmol) were dissolved in 1.0 mL of H_2O (containing 10% of D_2O). The pH of the solution was carefully adjusted to 4.0 with 1 M HCl to initiate the reaction. ^1H NMR spectra were acquired periodically and product ratios were calculated based on the relative integration of signals.

Reaction of N^9-deoxyadenosine 7 with nitrous acid

N^9-Deoxyadenosine **7** (2.6 mg, 0.010 mmol), sodium nitrite (6.9 mg, 0.100 mmol) and dibasic sodium phosphate (14.2 mg, 0.100 mmol) were dissolved in 1.0 mL of H_2O (containing 10% of D_2O). The pH of the solution was carefully adjusted to 4.0 with 1 M HCl to initiate the reaction. ^1H NMR spectra were acquired periodically and product ratios were calculated based on the relative integration of signals.

Reactions of a mixture of 8-mercapto-deoxyadenosine 26 and N^9-deoxyadenosine 7 with nitrous acid

N^9-Deoxyadenosine **7** (1.7 mg, 0.0066 mmol), 8-mercapto-deoxyadenosine **26** (0.9 mg, 0.0033 mmol), sodium nitrite (6.9 mg, 0.100 mmol) and dibasic sodium phosphate (14.2 mg, 0.100 mmol) were dissolved in 1.0 mL of H_2O (containing 10% of D_2O). The pH of the solution was carefully adjusted to 4.0 with 1 M HCl to initiate the reaction. ^{1}H NMR spectra were acquired periodically and product ratios were calculated based on the relative integration of signals.

Reactions of a mixture of 2'-deoxyadenosine 7 and cytidine 1 with nitrous acid

2'-Deoxyadenosine **7** (1.4 mg, 0.005 mmol), cytidine **1** (1.2 mg, 0.005 mmol), sodium nitrite (6.9 mg, 0.100 mmol) and dibasic sodium phosphate (14.2 mg, 0.100 mmol) were dissolved in 1.0 mL of H_2O (containing 10% of D_2O). The pH of the solution was carefully adjusted to 4.0 with 1 M HCl to initiate the reaction. ^{1}H NMR spectra were acquired periodically and product ratios were calculated based on the relative integration of signals.

Reactions of a mixture of 2-thiocytidine 3 and α-2-thiocytidine 13 with nitrous acid

2-Thiocytidine **3** (1.0 mg, 0.004 mmol), α-2-thiocytidine **13** (1.0 mg, 0.004 mmol), sodium nitrite (3.0 mg, 0.04 mmol) and dibasic sodium phosphate (11.0 mg, 0.05 mmol) were dissolved in 0.5 mL of H_2O (containing 10% of D_2O). The pH of the solution was carefully adjusted to 4.0 with 1 M HCl to initiate the reaction. ^{1}H NMR spectra were acquired periodically and product ratios were calculated based on the relative integration of signals compared to the added internal standard 6.0 uL of 25 mM of pentaerythritol solution. α-2-Thiocytidine **13** was fully converted to anhydrocytidine **12** (100% yield) after 8 hours at room temperature. 2-Thiocytidine **3** was slowly hydrolysed to give cytidine **1** (54% yield), thiouridine **4** (17% yield) and uridine **2** (10% yield). Anhydrocytidine **12** is stable under these conditions.

Synthesis of N^9-8, 2'-anhydro-thioadenosine 18 and N^7-8, 2'-anhydro-thioadenosine 19 in dry state with α-anhydrocytidine 12

α-D-Ribofuranosyl-2,2'-anhydrocytidine (2.6 mg, 0.01 mmol) **12**, 8-mercaptoadenine **16** (3.0 mg, 0.02 mmol) and $MgCl_2$ (2.0 mg, 0.02 mmol) were dissolved in 6 mL of H_2O in a round bottom flask. The pH was adjusted to 8.0 before the mixture was concentrated under reduced pressure. The residue was redissolved in H_2O and concentrated again. This procedure was repeated three times. The mixture was then heated at 150 °C in the dry state for 1.5 days. The dry residue was dissolved in D_2O and the resulting solution was transferred to an NMR tube to record the ^{1}H NMR spectrum of the crude reaction. 20 μL of 25 mM pentaerythritol solution in H_2O was added to the mixture to enable quantification of the reaction products by ^{1}H NMR spectroscopy. The yield of N^9-8, 2'-anhydro-thioadenosine **18** was 26%, while the yield of N^7-8, 2'-anhydro-thioadenosine **19** was 42% when referring to the proton integration compared to the internal standard. Cytosine **40** (77% yield) was observed

in the crude mixture as the only product deriving from nucleobase loss in the glycosidation step.

Stability studies of cytidine 1 or uridine 2 at 254 nm at the presence of bisulfite and $K_4Fe(CN)_6$ with or without N^9-thioanhydroadenosine 18

Cytidine **1** (3.2 mg, 0.01 mmol) or uridine **2** (3.2 mg, 0.01 mmol) was mixed with sodium sulfite (6.3 mg, 0.05 mmol) and $K_4Fe(CN)_6$ (1.0 mg) dissolved in 0.8 mL of H_2O (containing 10% D_2O). The mixture was irradiated at 254 nm for 10 hours. In parallel, the same composition of components (**1**, **2**, bisulfite and $K_4Fe(CN)_6$) with addition of N^9-thioanhydroadenosine **18** (1.5 mg, 0.005 mmol) was irradiated at 254 nm for the same period of time. 20 μL of 25 mM pentaerythritol solution in H_2O was added to the mixture before irradiation to enable quantification of the reaction products by ^{1}H NMR spectroscopy. 80% of cytidine and 82% of uridine were photochemically destroyed after 10 hours of irradiation while 95% of cytidine and 60% of uridine remained following irradiation in the presence of **18**. At the same time, 8-mercapto-deoxyadenosine **26** (88% yield) was afforded as the reduced product from **18**.

Stability studies of cytidine 1 and uridine 2 with $MgCl_2$ at 150 °C in the dry state

Cytidine **1** (5.0 mg, 0.02 mmol), uridine **2** (5.0 mg, 0.02 mmol) and $MgCl_2$ (1.0 mg, 0.01 mmol) were dissolved in 6 mL of H_2O in a round bottom flask. The mixture was concentrated under reduced pressure. The residue was then heated at 150 °C in the dry state for 1.5 day. The dry residue was dissolved in D_2O and the resulting solution was transferred to an NMR tube to record the ^{1}H NMR spectrum of the crude reaction. 9% cytosine **40** and 7% uracil **39** were detected in crude ^{1}H NMR spectrum. 91% of cytidine **1** and 93% of uridine **2** remained.

Dry state reaction of a mixture of α-anhydrouridine 15, cytidine 1, uridine 2 and 8-mercaptoadenine 16 in the presence of $MgCl_2$ at 150 °C followed by photoreduction with hydrogen sulfide and reaction with nitrous acid

α-Anhydrouridine **15** (2.5 mg, 0.01 mmol), cytidine **1** (1.3 mg, 0.005 mmol), uridine **2** (1.3 mg, 0.005 mmol), 8-mercaptoadenine **16** (5.5 mg, 0.03 mmol) and $MgCl_2$ (2.0 mg, 0.02 mmol) were dissolved in 10 mL of H_2O in a round bottom flask. The pH was adjusted to 8.0 before the mixture was concentrated under reduced pressure. The residue was then redissolved in H_2O and concentrated again. This procedure was repeated three times. The mixture was then heated at 150 °C in the dry state for 1.5 days. The dry residue was dissolved in D_2O and the resulting solution was transferred to an NMR tube to record the ^{1}H NMR spectrum of the crude reaction. 20 μL of 25 mM pentaerythritol solution in H_2O was added to the mixture to enable quantification of the reaction products by ^{1}H NMR spectroscopy. The yield of N^9-8, 2'-anhydro-thioadenosine **18** was 30%, while the yield of N^7-8, 2'-anhydro-thioadenosine **19** was 50% when referring to the proton integration compared to the internal standard. 95% of uridine **2** and 80% of cytidine **1** remained after the reaction. The sample was then lyophilized and used directly for the next step.

The crude reaction mixture from above was dissolved in 0.8 mL of degased H_2O (containing 10% D_2O), followed by addition of NaSH.xH2O (12 mg, 0.13 mmol,

containing 60% NaSH) and NaH_2PO_4 (20 mg, 0.17 mmol). The pH of the mixture was adjusted to 7 with 1 M HCl, then the mixture was irradiated at 254 nm in a Suprasil quartz NMR tube for 10 hours until thioanhydroadenosine **18** and **19** were consumed. N^7-deoxyadenosine **20**, deoxyadenosine **7** and their 8-SH intermediates **27** and **26** were observed in the crude ^{1}H NMR spectrum. The sample was then lyophilized.

The dry sample was dissolved in 0.8 mL of H_2O (containing 10% D_2O), followed by addition of $NaNO_2$ (7.0 mg, 0.09 mmol) and NaH_2PO_4 (12 mg, 0.10 mmol). The pH of mixture was adjusted to 4 with 1 M HCl. ^{1}H NMR spectra were acquired periodically. The 8-SH intermediates **27** and **26** were all converted to N^7-deoxyadenosine **20**, deoxyadenosine **7** within an hour. Deoxyadenosine **7** was gradually hydrolysed to deoxyinosine **9** while N^7-deoxyadenosine **20** was all hydrolysed within two days. The final yield (for three steps) of dA **7** and dI **9** was 6% each. The final ratio (after two days) of nucleosides in the mixture (dA:dI:C:U) was 14:14:44:28). 43% of ribosylpyrimidines (C + U) remained after 3 steps.

Dry state reaction of a mixture of α-anhydrocytidine 12, cytidine 1 and 8-mercaptoadenine 16 in the presence of $MgCl_2$ at 150 °C followed by photoreduction with bisulfite and reaction with nitrous acid

α-Anhydrocytidine **12** (3.0 mg, 0.01 mmol), cytidine **1** (3.0 mg, 0.01 mmol), 8-mercaptoadenine **16** (3.0 mg, 0.02 mmol) and $MgCl_2$ (2.0 mg, 0.02 mmol) were dissolved in 10 mL of H_2O in a round bottom flask. The pH was adjusted to 8.0 before the mixture was concentrated under reduced pressure. The residue was then redissolved in H_2O and concentrated again. This procedure was repeated three times. The mixture was then heated at 150 °C in the dry state for 1.5 day. The dry residue was dissolved in D_2O and the resulting solution was transferred to an NMR tube to record the ^{1}H NMR spectrum of the crude reaction. 40 μL of 25 mM pentaerythritol solution in H_2O was added to the mixture to enable quantification of the reaction products by ^{1}H NMR spectroscopy. The yield of N^9-8, 2'-anhydro-thioadenosine **18** was 26%, while the yield of N^7-8, 2'-anhydro-thioadenosine **19** was 34% when referring to the proton integration compared to the internal standard. 95% of cytidine **1** remained after the reaction. The sample was then lyophilized and used directly for the next step.

The crude reaction mixture from above was dissolved in 0.8 mL of degased H_2O (containing 10% D_2O), followed by addition of Na_2SO_4 (6.0 mg, 0.05 mmol) and $K_4Fe(CN)_6$ (2.0 mg, 0.005 mmol). The pH of the mixture was adjusted to 7 with 1 M HCl and the mixture was irradiated at 254 nm in a Suprasil quartz NMR tube for 6 hours until thioanhydroadenosine **18** and **19** were consumed. N^7-deoxyadenosine **20**, deoxyadenosine **7** and their 8-SH intermediates **27** and **26** were observed in the ^{1}H NMR spectrum. The sample was then lyophilized.

The dry sample above was dissolved in 0.8 mL of H_2O (containing 10% D_2O), followed by addition of $NaNO_2$ (7.0 mg, 0.09 mmol) and NaH_2PO_4 (12 mg, 0.10 mmol). The pH of mixture was adjusted to 4 with 1 M HCl. ^{1}H NMR spectra were acquired periodically. The 8-SH intermediates **27** and **26** were converted to N^7-deoxyadenosine **20** and deoxyadenosine **7** within a few hours. Deoxyadenosine **7** was gradually hydrolysed to deoxyinosine **9** while N^7-deoxyadenosine **20** was completely

hydrolysed within two days. The final yields (over three steps) of dA **7** and dI **9** were 10% and 9% respectively. The final ratio (after two days) of nucleosides in the mixture (dA:dI:C:U) was 10:9:67:14. 84% of ribosylpyrimidines (C + U) remained after 3 steps.

Hydrolysis of α-thiocytidine 13 in 0.1 M 4,5-dicyanoimidazole buffer (pH 5.2) at 60 ^{o}C

α-Thiocytidine **13** (2.6 mg, 0.010 mmol) was dissolved in 4,5-dicyanoimidazole buffer (1.0 mL of a 0.1 M solution in H_2O, containing 10% of D_2O, at pH 5.2) and the pH of the mixture was re-adjusted to 5.2. The mixture was transferred to a NMR tube and heated at 60 ^{o}C. The reaction was monitored periodically by ^{1}H NMR spectroscopy. After 24 days the mixture consisted of α-D-ribofuranosyl-2,2'-anhydrocytidine **12** (2%), α-D-ribofuranosyl-2,2'-anhydrouridine **15** (30%), α-uridine **38** (48%), α-thiouridine **14** (4%) and α-cytidine **41** (16%).

Hydrolysis of β-thiocytidine 3 in 0.1 M 4,5-dicyanoimidazole buffer (pH 5.2) at 60 ^{o}C

β-Thiocytidine **3** (2.6 mg, 0.010 mmol) was dissolved in 4,5-dicyanoimidazole buffer (1.0 mL of a 0.1 M solution in H_2O, containing 10% of D_2O, at pH 5.2) and the pH of the mixture was re-adjusted to 5.2. The mixture was transferred to a NMR tube and heated at 60 ^{o}C. The reaction was monitored periodically by ^{1}H NMR spectroscopy. After 24 days the mixture consisted of β-thiocytidine **3** (87%), cytidine **1** (7%), β-thiouridine **4** (1%) and uridine **2** (5%).

Hydrolysis of α-thiouridine 14 in 0.1 M 4,5-dicyanoimidazole buffer (pH 5.2) at 60 ^{o}C

α-Thiouridine **14** (2.6 mg, 0.010 mmol) was dissolved in 4,5-dicyanoimidazole buffer (1.0 mL of a 0.1 M solution in H_2O, containing 10% of D_2O, at pH 5.2) and the pH of the mixture was re-adjusted to 5.2. The mixture was transferred to a NMR tube and heated at 60 ^{o}C. The reaction was monitored periodically by ^{1}H NMR spectroscopy. After 24 days the mixture consisted of α-thiouridine **14** (28%), α-D-ribofuranosyl-2,2'-anhydrouridine **15** (47%) and α-uridine **38** (25%).

Hydrolysis of β-thiouridine 4 in 0.1 M 4,5-dicyanoimidazole buffer (pH 5.2) at 60 ^{o}C

β-Thiouridine **4** (2.6 mg, 0.010 mmol) was dissolved in 4,5-dicyanoimidazole buffer (1.0 mL of a 0.1 M solution in H_2O, containing 10% of D_2O, at pH 5.2) and the pH of the mixture was re-adjusted to 5.2. The mixture was transferred to a NMR tube and heated at 60 ^{o}C. The reaction was monitored periodically by ^{1}H NMR spectroscopy. No reaction took place and β-thiouridine **4** remained unchanged after 24 days.

Reaction of α-thiocytidine 13 with nitrous acid

α-Thiocytidine **13** (2.6 mg, 0.010 mmol), sodium nitrite (6.9 mg, 0.100 mmol) and dibasic sodium phosphate (14.2 mg, 0.100 mmol) were dissolved in 1.0 mL of H_2O (containing 10% of D_2O). The pH of the mixture was carefully adjusted to 4 with 1 M HCl to initiate the reaction. The mixture was transferred to a NMR tube and the reaction monitored periodically by ^1H NMR spectroscopy. Within 2 minutes the starting material was converted into α-D-ribofuranosyl-2,2'-anhydrocytidine **12** (100%), with concomitant formation of a precipitate (presumably sulfur), and no further reactions took place over the next 24 hours.

Reaction of β-thiocytidine 3 with nitrous acid

β-Thiocytidine **3** (2.6 mg, 0.01 mmol), sodium nitrite (6.9 mg, 0.1 mmol) and dibasic sodium phosphate (14.2 mg, 0.100 mmol) were dissolved in 1.0 mL of H_2O (containing 10% of D_2O). The pH of mixture was carefully adjusted to 4 with 1 M HCl to initiate the reaction. The mixture was transferred to a NMR tube and the reaction monitored periodically by ^1H NMR spectroscopy. After 24 hours the mixture consisted of β-thiocytidine **3** (56%), cytidine **1** (26%), β-thiouridine **4** (13%) and uridine **2** (5%).

Reaction of α-thiouridine 14 with nitrous acid

α-Thiouridine **14** (2.6 mg, 0.01 mmol), sodium nitrite (6.9 mg, 0.1 mmol) and dibasic sodium phosphate (14.2 were dissolved in 1.0 mL of H_2O (containing 10% of D_2O). The pH of mixture was carefully adjusted to 4 with 1 M HCl. The mixture was then transferred to a NMR tube and the reaction monitored periodically by ^1H NMR spectroscopy. No reaction took place and α-thiouridine **14** remained unchanged after 24 hours.

Reaction of β-thiouridine 4 with nitrous acid

β-Thiouridine **4** (2.6 mg, 0.01 mmol), sodium nitrite (6.9 mg, 0.1 mmol) and dibasic sodium phosphate (14.2 were dissolved in 1.0 mL of H_2O (containing 10% of D_2O). The pH of mixture was carefully adjusted to 4 with 1 M HCl. The mixture was then transferred to a NMR tube and the reaction monitored periodically by ^1H NMR spectroscopy. No reaction took place and β-thiouridine **4** remained unchanged after 24 hours.

Reaction of cytidine 1 with nitrous acid

Cytidine **1** (2.4 mg, 0.01 mmol), sodium nitrite (6.9 mg, 0.1 mmol) and dibasic sodium phosphate (14.2 mg, 0.100 mmol) were dissolved in 1.0 mL of H_2O (containing 10% of D_2O). The pH of mixture was carefully adjusted to 4 with 1 M HCl to initiate the reaction. The mixture was transferred to a NMR tube and the reaction monitored periodically by ^1H NMR spectroscopy. After 24 hours the mixture consisted of cytidine **1** (91%) and uridine **2** (9%).

Hydrolysis of a 1:1 mixture of α-thiocytidine 13 and β-thiocytidine 3 in 0.1 M 4,5-dicyanoimidazole buffer (pH 5.2) at 60 °C

α-Thiocytidine **13** (1.3 mg, 0.005 mmol) and β-thiocytidine **3** (1.3 mg, 0.005 mmol) were dissolved in 4,5-dicyanoimidazole buffer (1.0 mL of a 0.1 M solution in H_2O, containing 10% of D_2O, at pH 5.2) and the pH of the mixture was re-adjusted to 5.2. 10 μL of 25 mM pentaerythritol solution in H_2O was added to the mixture to enable quantification of the reaction products by ^{1}H NMR spectroscopy. The mixture was transferred to a NMR tube and heated at 60 °C. The reaction was monitored periodically by ^{1}H NMR spectroscopy. After 16 days the mixture consisted of α-D-ribofuranosyl-2,2'-anhydrouridine **15** (13%), α-cytidine **41** (9%), α-uridine **38** (28%), β-thiocytidine **3** (47%), cytidine **1** (2%) and uridine **2** (1%).

Reaction of a 1:3 mixture of α-D-ribofuranosyl-2,2'-anhydrouridine 15 and β-thiocytidine 3 with nitrous acid

α-D-ribofuranosyl-2,2'-anhydrouridine **15** (0.6 mg, 0.0025 mmol), β-thiocytidine **3** (2.0 mg, 0.0075 mmol), sodium nitrite (6.9 mg, 0.100 mmol) and dibasic sodium phosphate (14.2 mg, 0.100 mmol) were dissolved in 1.0 mL of H_2O (containing 10% of D_2O). The pH of the mixture was carefully adjusted to 4 with 1 M HCl to initiate the reaction. 10 μL of 25 mM pentaerythritol solution in H_2O was added to the mixture to enable quantification of the reaction products by ^{1}H NMR spectroscopy. The mixture was transferred to a NMR tube and the reaction monitored periodically by ^{1}H NMR spectroscopy. After 24 hours the mixture consisted of α-D-ribofuranosyl-2,2'-anhydrouridine **15** (21%), β-thiocytidine **3** (44%), β-thiouridine **4** (4%), cytidine **1** (28%) and uridine **2** (3%).

Crystal structure determination and refinement details

***N*^{9}-8, 2'-anhydro-thioadenosine 18** (grown in H_2O): the crystals were thin needles. A long needle (exceeding the diameter of the incident beam) was selected to maximise diffracted intensity. Data were collected and integrated to 0.84 Å resolution, and structure solution/refinement was largely straightforward. The crystal was found to be twinned by application of a 2-fold rotation axis parallel to the 3_2 axis (twin law: –1 0 0 / 0 –1 0 / 0 0 1, refined batch scale factor = 0.143 (2)). H atoms bound to C atoms were placed in idealised positions. H atoms bound to N and O atoms were located in difference Fourier maps, normalised to standard bond lengths along the O/N—H vectors, then allowed to ride on their parent atoms for subsequent refinement cycles. The resulting H-bond network is physically sensible. The absolute structure was confirmed by the Flack parameter obtained from 2098 quotients[57].

***N*^{9}-8, 5'-anhydro-thioadenosine 37** (grown in H_2O): the crystals were very thin laths and the diffracted intensity was weak. Data were collected and integrated to 1.0 Å resolution, but the intensity falls below the $3\sigma(I)$ level around 1.25 Å resolution. The overall R_{int} is large, and the resulting refinement produces correspondingly large R values. However, the structure refinement behaves well and the result is clear.

There are four independent molecules in the asymmetric unit in space group *C*2. Three of them (with suffices A, B and C) are ordered. The fourth molecule is modelled as two overlaid components (with suffices D and E), in which the adenine group occupies a consistent position but the remainder of the molecule is split over two positions. The molecular conformation is very similar in the two components (the disorder does not reflect alternative molecular structures) so the disorder appears to reflect "frustration" in the intermolecular contacts. The two molecular components are accompanied by alternative positions for lattice water molecules.

To handle the limited data resolution and quality, extensive geometrical restraints were applied. Molecules B, C, D and E were restrained to have geometry similar to that of molecule A (SAME restraint in SHELXL). The site occupancies of molecules D and E were both constrained to be 0.5 and a common isotropic displacement parameter was applied to all atoms in those molecules. The atoms in molecules A, B and C were refined with individual isotropic ADPs. Only the S atoms were refined with anisotropic ADPs. The lattice water molecules were refined with a further common isotropic ADP. H atoms were included in idealised positions, or in positions defined to produce a reasonable H-bonding network. They were subsequently refined as riding on their parent atoms with U_{iso} = 1.2 or 1.5 U_{eq}(C/N/O). H atoms were not included on the lattice water molecules. An additional solvent water correction (SWAT in SHELXL) was also applied.

Theoretical section

Computational Methods

The equilibrium ground-state geometries and the corresponding harmonic vibrational frequencies for isolated N^9-8,2'-anhydro-thioadenosine **18** and N^7-8,2'-anhydro-thioadenosine **19** were obtained using the Kohn–Sham density functional theory (KS-DFT) including the range-separated hybrid ωB97X-D functional including dispersion correction within the parametrization[58] with the ma-def2-TZVP basis set[59]. The abbreviation of the method is ωB97X-D/ma-def2-TZVP. In order to assess solvent effects exerted by bulk water on charged model systems, the polarizable continuum model (PCM) with the integral equation formalism variant (IEFPCM) was used. The last approach is further denoted as ωB97X-D/IEFPCM/ma-def2-TZVP and was used as the primary method to optimize the complexes of **18** and **19** with negatively charged HS^- anion and one water molecule which maintained the integrity of these complexes. The KS-DFT and PCM calculations were conducted using the Gaussian 16 package[60].

Adding PCM terms to the Hamiltonian of our model systems allowed including solvent effects which are exerted by bulk water on the charged systems. In Gaussian 16, the default settings in the PCM section assume that a molecular cavity is constructed using the continuous surface charge (CSC) approach[61], based on interlocking spheres centered on atoms and assuming the corresponding UFF force-field atomic radii with the scale factor of 1.1. This standard protocol led to the formation of nonphysical cavities for complexes **18** and **19** in which some of the surface charge basis functions (SCBFs) are placed in between the hydrosulfide and the chromophores (Fig. S47, default radii, centers of SCBFs marked as red dots). Therefore, for the sulfur atom in HS^-, we decided to use a larger scale factor equal to

1.4, in order to remove the artificial surface charge density from the inner part of cavities (Fig. S47, modified radii).

Vertical excitation calculations of isolated molecules were performed using the algebraic diagrammatic construction to the second order method [ADC(2)][43,44,62] with the ma-def2-TZVP basis set (ADC(2)/ma-def2-TZVP), assuming the ground-state geometries found at the ωB97X-D/ma-def2-TZVP level. Orbital character of electronic excited states was determined based on analysis of natural transition orbitals[63] (NTOs) which were generated using the TheoDore 1.7.2 package[64]. The optimization of stationary points on excited-state potential energy surfaces of isolated molecules was performed using the ADC(2)/ma-def2-TZVP method. Minimum-energy crossing points (MECPs) between the first excited state and electronic ground state were located using the sequential penalty constrained optimization approach implemented by Levine et al.[65] in the CIOpt package. MECPs were located using the energies and analytical gradients were computed at the MP2/ADC(2)/ma-def2-TZVP level of theory. The CIOpt package[65] was interfaced with the Turbomole 7.3 program[66]. The potential energy profiles of nonradiative photorelaxation pathways were constructed using the image dependent pair potential (IDPP) interpolation[67] between key minimum-energy structures, the energies of ground and excited states were obtained at the MP2/ma-def2-TZVP and ADC(2)/ma-def2-TZVP levels, respectively.

Vertical excitation calculations of charged model systems were conducted using either time-dependent density functional theory (TD-DFT) and the TD-ωB97X-D/IEFPCM/ma-def2-TZVP method or the ADC(2)/COSMO/ma-def2-TZVP approach in which conductor-like screening model (COSMO) was applied, assuming the ground-state geometries found at the ωB97X-D/IEFPCM/ma-def2-TZVP level. In vertical excitation calculations, both implicit solvation models (IEFPCM and COSMO) were used in the nonequilibrium limit. Relaxed scans on excited-state potential energy surfaces were performed using the TD-ωB97X-D/IEFPCM/ma-def2-TZVP method. The potential energy profiles of radiationless deactivation channels were prepared using the image dependent pair potential (IDPP) interpolation[67] between the key geometries and the energies of ground and excited states were obtained at the TD-ωB97X-D/IEFPCM/ma-def2-TZVP and MP2/ADC(2)/ COSMO/ma-def2-TZVP level. The relaxed scans and potential energy profiles were obtained using the equilibrium limit in implicit solvation models (IEFPCM and COSMO).

Adiabatic electron affinities were computed as the difference between free energies of neutral and radical anion of molecule, both structures were optimized at the ωB97X-D/IEFPCM/ma-def2-TZVP level. The free energies were obtained based on these minimum-energy structures using IEFPCM method in a variant of solvent reaction field self-consistent with the solute electrostatic potential, including cavitation, dispersion, electrostatic, exchange-repulsion contributions and assuming the harmonic oscillator and rigid rotor approximations.

All (TD)-DFT and MP2/ADC(2) calculation were performed using the Gaussian 16 and Turbomole 7.3 package, respectively.

The topological features of electron density in the studied complexes were analysed using the Bader's Quantum Theory of Atoms in Molecules (QTAIM). This analysis was performed using AIMALL package[68], based on the generalized one-particle electron densities obtained with the MP2/aug-cc-pVDZ method. Calculations of many-body interaction energy partitioning were performed using a modified version[69] of GAMESS (US) package[70].

Equilibrium Ground-State Geometries

The minimum energy geometries of N^9-8,2'-anhydro-thioadenosine **18** and N^7-8,2'-anhydro-thioadenosine **19** optimized at the ωB97X-D/ma-def2-TZVP are shown on the left-hand side of Fig. S48. The complexes of **18** and **19** with hydrosulfide anion were constructed by first mapping the ground-state electrostatic potential of the isolated molecules and then placing the HS- entity on the most positively charged side of the sulfur heteroatom in both **18** and **19**. It is worth noting that the ground-state geometry optimizations of **18** and **19**, each clustered with an HS^- anion, led to the formation of stable complexes with the free energies of interaction of -64.7 and -65.2 kJ/mol. These complexes were additionally stabilized by a single explicit water molecule and exhibit a strong sulfur-sulfur interaction with the interatomic S···S distance of 3.44 and 3.41 Å, for the **18** and **19** complexes, respectively. The origins of intermolecular interactions in these complexes are discussed later in greater detail (cf. section **Origins of intermolecular interactions in the studied complexes**). It is worth noting though that the formation of complexes is largely due to electrostatic and (to lesser extent) dispersion attraction forces between the hydrosulfide anion and the sulfur atom from chromophore which keeps HS^- very close to either **18** or **19**.

Vertical Excitation Energies

The vertical excitation energies (Tab. S6) of **18** and **19** were calculated in the gas phase using the ADC(2)/ma-def2-TZVP method and assuming the ground-state structures optimized at the ωB97X-D/ma-def2-TZVP level. The results of our calculations (Tab. S6) for **18** and **19** show that the first excited state in both chromophores is the optically bright state characterized by substantial oscillator strength. These electronic excited states, having $^1\pi\pi^*$ character, are at 4.969 eV (249.5 nm) and 4.783 eV (259.2 nm) for **18** and **19**, respectively. Both **18** and **19** are also characterized by higher $^1\pi\pi^*$ states at 5.064 eV (244.8 nm) and 5.334 eV (232.4 nm) associated with sizeable oscillator strengths. These results point that the both chromophores can easily absorb photons in the UV-C spectral region.

UV-excitation of **18** to the first bright state could also lead to the population of the $^1\pi\sigma^*$ state which lies energetically very close to the $^1\pi\pi^*$ excited state in the vertical spectrum. Therefore, we anticipate that radiationless relaxation processes in case of **18** could occur on both $^1\pi\pi^*$ (S_1) and $^1\pi\sigma^*$ (S_2) hypersurfaces. Similarly, absorption of a UV photon by **19** may enable population of the $^1\pi\pi^*$ (S_1) state. However, the repulsive $^1\pi\sigma^*$ (S_3) state lies somewhat higher in energy than in **18** and is additionally separated from the S_1 state by a $^1n\pi^*$ state, which could also contribute to the overall photochemistry of **19**. These findings suggest that in **19,** low energy nonradiative deactivation channels could occur on $^1\pi\pi^*$ and $^1n\pi^*$ excited state surfaces with potentially lower contribution from the $^1\pi\sigma^*$ (S_3) state.

The performed quantum-chemical calculations demonstrate that even though in both chromophores the bright states are the S_1 $^1\pi\pi^*$ states, the character of neighboring states in the Franck-Condon region is different. In addition, these low-lying $^1\pi\pi^*$ states are also associated with different redistribution of the π-bonding structure in the excited state. Therefore, UV-induced photochemical processes in the both chromophores could lead to quite different nonradiative deactivation pathways, despite their structural similarity.

The vertical excitation energies of **18** and **19** complexes with one explicit water molecule and bisulfide anion, shown in Tab. S7, were calculated employing the ADC(2)/COSMO/ma-def2-TZVP method, assuming the ground-state structures optimized at the ωB97X-D/IEFPCM/ma-def2-TZVP level of theory. The results obtained using the TD-ωB97X-D/IEFPCM/ma-def2-TZVP method are shown in parentheses. The ADC(2) vertical excitation energies of the **18** complex show that two lowest-lying excited states are bright states (4.820 and 4.955 eV) characterized by substantial oscillator strengths. Interestingly, the S_3 ($^1n\sigma^*$) excited state is also a bright state having a partial charge-transfer (CT) character. Wave-function analysis of the latter state revealed that the charge of 0.495 e^- is transferred from the hydrosulfide to the chromophore. It is worth noting that this S_3 state lies energetically very close to the optically bright excited states. Therefore, the population of the $^1n\sigma^*$ hypersurface may readily occur soon after UV absorption in the Franck-Condon region and consequently could allow for an efficient charge transfer process. The computed vertical excitations for the **19** complex show that the S_1 excited state is a typical $^1\pi\pi^*$ bright state. Interestingly, the S_2 state can be described as a bright charge-transfer state which could also lead to the electron transfer process (of 0.513 e^-) from hydrosulfide to the chromophore moiety. The latter state should also be easily accessible after photoexcitation due to large oscillator strength. Vertical excitation spectra calculations of **18** and **19** complexes at the TDDFT level indicate that energies of both $^1n\pi^*$ and $^1n\sigma^*$ excited states are generally underestimated when compared to the ADC(2) results. Nonetheless, the ordering of low-lying states remains unchanged.

Deactivation Pathways of Isolated Chromophores

Fig. S49 presents a radiationless relaxation mechanism of UV-excited **1**8 calculated at the ADC(2)/ma-def2-TZVP level of theory. Following the vertical excitation to the optically bright state S_1 ($^1\pi\pi^*$, red line), the chromophore can efficiently undergo molecular orbital change to the $^1\pi\sigma^*$ (5.057 eV, blue line) state. This is supported by our observation that the $^1\pi\sigma^*$ character is readily stabilized during the first steps of S_1 geometry optimization, even though this procedure is initiated in the $S_1(^1\pi\pi^*)$ state. Subsequently, the relaxation on the $^1\pi\sigma^*$ hypersurface results in C2'-S bond rupture of **18** leading to the peaked $S_1(\pi\sigma^*)/S_0$ state crossing. This radiationless relaxation mechanism can deactivate the UV-excited chromophore in a completely barrierless manner. In the absence of reducing agents, the photochemically broken C2'-S bond could be reformed when **18** goes back to the ground state. Therefore, we anticipate that this chromophore should be characterized by a significant photostability in UV-rich environment.

UV-excitation of isolated **19** can lead to the vibrational relaxation (see Fig. S50) on the $^1\pi\pi^*$ (red line) excited-state hypersurface towards the S_1 minimum in which the C8 and N7 atoms of the nucleobase are substantially pyramidalized indicating the localization of the biradical structure on these atoms. After passing through a modest barrier of less than 0.1 eV the photoexcited system can reach the $S_1(\pi\pi^*)/S_0$ minimum-energy crossing point associated with complete N7-C8 bond rupture. The photorelaxation mechanism of UV-excited **19** through the discussed S_1/S_0 state crossing could be responsible for the structural damages of the chromophore further occurring in the hot electronic ground state. Thus, it seems that **19** is less photostable than **18** in UV-rich prebiotic environments.

Photorelaxation Channels of Complexes with Hydrosulfide and Water

The performed photochemical calculations of **18** and **19** complexes enabled to identify the relaxation pathways which are presented in Fig. S51 and Fig. S52, respectively. We have started the exploration of the excited-state potential energy surfaces from an excitation of the equilibrium ground-state geometries of **18** and **19** complexes, which contain, beside chromophores, a hydrosulfide anion and one explicit water molecule. As discussed above, in both our model systems, there is a stabilizing interaction between HS^- and the sulfur atom from chromophore which keeps the hydrosulfide very close to either **18** or **19**. This unique arrangement of HS^- with respect to the chromophores results in the appearance of bright and charge-transfer states which were described in detail in the Vertical Excitation Energies paragraph.

Fig. S51 and Fig. S52 present the UV-induced radiationless deactivation pathways for **18** and **19** complexes obtained using the TD-ωB97X-D/IEFPCM/ma-def2-TZVP method and TDDFT energies are presented as dots. The first parts of these figures are composed of the image dependent pair potential (IDPP) interpolation between the ground-state structures and the constrained $S_1(n\sigma^*)$ geometries which are marked by a C2'-S distance equal to 2.02 and 2.08 Å for **18** and **19** complexes, respectively. The S_1 minimum-energy structures found by the constrained geometry optimization at the TD-ωB97X-D/IEFPCM/ma-def2-TZVP level of theory, for which the C2'-S distance was kept fixed, are presented in Fig. S53. For both model systems, the S_1 structure is characterized by a $^1n\sigma^*$ transition in which an electron is promoted from a lone electron pair molecular orbital (n) localized on the hydrosulfide anion to a σ^* molecular orbital of the C2'-S bond (see Fig. S53). The electron distribution analysis of these S_1 structures revealed a photoinduced charge-transfer (CT) process from the hydrosulfide anion to the chromophores. More precisely, the CT character arises from an excited-state interaction between lone electron pairs of the sulfur atoms and is promoted by the excitation from the n orbital of HS^- to the anti-bonding σ^* orbital located at the C2'-S bond what enforces the bond rupture.

The second part of Fig. S51 and S52 was obtained from the relaxed scans which were performed along the C2'-S bonds of **18** and **19** using the TD-ωB97X-D/IEFPCM/ma-def2-TZVP method. These relaxed scans allowed to explore the S_1 hypersurface and to find crucial geometries which are located near the S_1/S_0 state crossing seams in both model systems. The constrained S_1 minimum energy geometries for which the crossing seam is nearly reached are marked by the C2'-S distance equal to 2.41 and 2.38 Å for **18** and **19** complexes, respectively. These structures are 0.50 eV (**18** complex) and 0.55 eV (**19** complex) above the corresponding ground state. Since the TDDFT method cannot reliably describe potential energy surface in the vicinity of surface crossings (conical intersections) between S_1 and S_0 states, the C2'-S bond was not elongated further. However, our TDDFT quantum-chemical calculations show that UV-excitation of our model complex systems could enable the population of $^1n\sigma^*$ excited state which subsequently can trigger the charge-transfer process from the hydrosulfide anion to either of the studied chromophores.

Geometries obtained from the IDPP interpolations and the TDDFT relaxed PE scans were further used to calculate single-point energies employing the more accurate ADC(2)/COSMO/ma-def2-TZVP level of theory. As demonstrated for the vertical excitation energies, both methods provide consistent ordering of excited states. The ADC(2) potential energy profiles are presented in Fig. S51 and S52 with solid lines. These results indicate that there is indeed a S_1/S_0 surface crossings along

the C2'-S bond elongation path that could allow for the deactivation of excited states of our model systems to the ground state through peaked $^1n\sigma^*/S_0$ conical intersections in each of the cases. It is worth noting that both methods also yield qualitatively consistent shapes of potential energy surfaces.

Radical Anions

Water solutions containing bisulfite irradiated at 248 nm were suggested to photochemically generate solvated electrons[45]. These free electrons penetrating an aqueous environment may be intercepted by **18** and **19** chromophores. To understand structural changes induced by the attachment of an electron to our model systems, we performed optimizations for **18** and **19** radical anions at the ωB97X-D/IEFPCM/ma-def2-TZVP level of theory.

Fig. S54 shows the equilibrium ground-state structures of our model systems with an excess electron. The attachment of electron to **18** triggers the C2'-S bond rupture, which subsequently enables the formation of experimentally observed deoxyribonucleotides, deoxyadenosine **7** and deoxyinosine **9**. In turn, the presence of the additional electron in **19** results in an out-of-plane position of the sulfur atom in relation to the base moiety. In the latter chromophore, the excess electron is located on the C8 atom of the purine moiety. Therefore, the C2'-S bond rupture is not preferable, and **19** is unlikely to follow similar chemical routes as **18** after attachment of free electron from the aqueous environment. Moreover, we expect that the formation of **19** radical anion could initiate a fragmentation process resulting in the structural damages of the chromophore.

The QTAIM Analysis

The molecular electron density distribution ρ(r) can be obtained from a many-particle wave function as follows:

$$\rho(\mathbf{r}) = \mathrm{N}\sum_{\sigma}\int |\Psi(\mathbf{r},\mathbf{r}_2,\ldots,\mathbf{r}_\mathrm{n})|^2\, d\mathbf{r}_2,\ldots,\mathbf{dr}_\mathrm{N}$$

where the summation is performed over spin coordinates and the integration is over all, but one spatial coordinate and N stands for the total number of electrons. In the case of the intermolecular complexes studied here, we analysed the generalized MP2 densities obtained using the Z-vector approach[71].

Particularly interesting in the context of this study are the properties of bond critical points (BCPs) localized on the bond paths linking the interacting species (cf. Table S8). The sign of the Laplacian $\nabla^2\rho(\mathbf{r}_{BCP})$ at a specific point in space determines whether the negative potential energy or the positive kinetic energy is in excess of the virial ratio amounting to 2. The positive values of Laplacian, in Table S8 imply a depletion of electronic charge in the intermolecular region. This is an indication of a closed-shell nature of interactions. It is also well established that the magnitude of the $\rho(\mathbf{r}_{BCP})$ reflects the strength of interaction. If one compares the magnitudes of electronic densities at the S-S BCPs obtained for the complexes studied here with the corresponding values for S_2 (ground state $^3\Sigma_g^-$) and S_2^{-2} molecules (as well as the corresponding interatomic distances given in the last column) the weakly bonded nature of the studied complexes **18** and **19** becomes evident.

The local kinetic $G(\mathbf{r}_{BCP})$, potential $V(\mathbf{r}_{BCP})$ and total $H(\mathbf{r}_{BCP})$ energy densities, estimated at the bond critical point are also very helpful in analysis of the weak intermolecular interactions. The electronic energy density of the local charge distribution may be expressed as the sum of the local kinetic and potential energy densities: $H(\mathbf{r}_{BCP}) = G(\mathbf{r}_{BCP}) + V(\mathbf{r}_{BCP})$. The relation between the Laplacian and the components of the local energy density is given by the local expression for the virial theorem (in atomic units):

$$\frac{1}{4}\nabla^2\rho(\mathbf{r}_{\boldsymbol{BCP}}) = 2G(\mathbf{r}_{\boldsymbol{BCP}}) + V(\mathbf{r}_{\boldsymbol{BCP}})$$

The results of the QTAIM analysis are gathered in Table S8 and the structures of the analysed complexes with the assumed atomic labels (relevant only for the QTAIM Analysis paragraph) are shown in Fig. S55 and S56.

The small values of $\rho(\mathbf{r}_{BCP})$ and the positive sign of the Laplacian and of the total energy density $H(\mathbf{r}_{BCP})$ indicates a closed-shell character of the S-S interactions (possibly weakly ionic interaction but at the distances much larger than for typical S-S contacts in ionic crystals). Considering the properties of the BCPs for the relevant hydrogen bonds in the hydrated cluster, these may be characterized as typical (O23-H4 and O21-H4) or partially covalent (S1-H5 BCPs in both clusters). The latter are also slightly stronger judging by the larger magnitudes of charge density at the corresponding BCPs. In both complexes we located also bond paths for the weak CH-π contacts between the water molecule's oxygen atoms and the aliphatic hydrogens of the sugar moiety (O3-H24 and P3-H25), carrying a small positive charge. These also contribute to the overall stability of the complexes but to a rather small extent.

The results obtained for the relaxed densities computed in the presence of polarizable dielectric (using the same cavities and PCM variant as in the geometry optimizations), representing bulk water environment are virtually identical to the gas-phase data. One should note, however, that the effect of the solvent is not fully accounted for in the QAIM analysis, since it analyses only the relaxed electronic densities of interacting species in the presence of the dielectric medium and not their interactions with the environment.

Origins of intermolecular interactions in the studied complexes

In order to determine the origins of interactions in the studied complexes we used the many-body hybrid variational–perturbational interaction energy decomposition scheme which allows to analyse intermolecular interactions in many-body complexes embedded in polarizable dielectric medium (cf. Góra et al.[72] and the references cited therein). In this method, the total interaction energy of a trimer, calculated by a supramolecular approach using the second-order Møller–Plesset perturbation theory (MP2), is partitioned into the two- and three-body components. Taking into account that both the electrostatic components and the second-order dispersion energy are pairwise additive the total MP2 interaction energy, of a complex can be partitioned as follows:

$$\Delta E^{MP2} = \sum_{i=2}^{3} \Delta E_{\mathrm{i}}^{\mathrm{HF}} + \sum_{i=2}^{3} \Delta E_{\mathrm{corr,i}}^{\mathrm{MP2}} =$$

$$= \varepsilon_{\mathrm{el}}^{(10)} + \sum_{i=2}^{3} \Delta E_{\mathrm{ex,i}}^{\mathrm{HL}} + \sum_{i=2}^{3} \Delta E_{\mathrm{del,i}}^{\mathrm{HF}} +$$
$$+ \varepsilon_{\mathrm{el,r}}^{(12)} + \varepsilon_{\mathrm{disp}}^{(20)} + \sum_{i=2}^{3} \Delta E_{\mathrm{ex,i}}^{(2)}$$

The Hartree–Fock (HF) term can be partitioned into the Heitler–London (HL) and the delocalization ($\Delta E_{\mathrm{del}}^{\mathrm{HF}}$) energy components. The former involves the electrostatic interactions of unperturbed monomer charge ($\varepsilon_{\mathrm{el}}^{(10)}$) densities as well as the associated exchange repulsion ($\Delta E_{\mathrm{ex}}^{\mathrm{HL}}$). The latter collectively encompasses the induction and the associated exchange effects due to the Pauli exclusion principle.

$$\Delta E^{\mathrm{HF}} = \Delta E^{\mathrm{HL}} + \Delta E_{\mathrm{del}}^{\mathrm{HF}}$$
$$= \varepsilon_{\mathrm{el}}^{(10)} + \Delta E_{\mathrm{ex}}^{\mathrm{HL}} + \Delta E_{\mathrm{del}}^{\mathrm{HF}}$$

The second order electron correlation term ($\Delta E_{\mathrm{corr}}^{\mathrm{MP2}}$) includes the second order dispersion interaction ($\varepsilon_{\mathrm{disp}}^{(20)}$) as well as the electron correlation correction to the first order electrostatic interaction ($\varepsilon_{\mathrm{el,r}}^{(12)}$) and the remaining electron correlation effects ($\Delta E_{\mathrm{ex}}^{(2)}$). The latter term accounts mainly for the uncorrelated exchange-dispersion and electron correlation corrections to the Hartree–Fock exchange repulsion.

$$\Delta E_{\mathrm{corr}}^{\mathrm{MP2}} = \varepsilon_{\mathrm{el,r}}^{(12)} + \varepsilon_{\mathrm{disp}}^{(20)} + \Delta E_{\mathrm{ex}}^{(2)}$$

The $\varepsilon_{\mathrm{el}}^{(10)}$ and the $\varepsilon_{\mathrm{disp}}^{(20)}$ terms are obtained in the standard polarization perturbation theory[73], whereas the $\varepsilon_{\mathrm{el,r}}^{(12)}$ term is calculated using the formula proposed by Moszyński et al.[74] The indices in parenthesis denote perturbation orders in intermolecular interaction operator and intramonomer correlation operator respectively.

In all necessary calculations the complex centered basis set was consistently used and, therefore, the results are basis set superposition error free due to the full counterpoise correction[75]. In the case of trimer this procedure is equivalent to the Site-Site Function Counterpoise method[76]. The discussed interaction energy partitioning scheme can be easily generalized to the complexes interacting with polarizable dielectric medium.[72] In this case a few additional terms arise due to differences in solvation free energies of the isolated solutes and their complex. All these terms are represented here collectively as the $\Delta G^{\mathrm{HF,S}}$ component, partitioned into two-and three-body terms ($\Delta G_2^{\mathrm{HF,S}}$ and $\Delta G_3^{\mathrm{HF,S}}$, respectively).

The results of many-body interaction energy partitioning obtained for both complexes are shown in Tab. S9. The results of interaction energy analysis are consistent with the QTAIM data. The origins of intermolecular interactions in both complexes in the gas-phase are quite similar with only a small quantitative difference. Interestingly, despite the presence of highly polar and charged species the most important stabilizing contribution comes from the additive second-order dispersion terms. Although the electrostatic attraction forces are the largest in terms of magnitude, these are completely quenched by the corresponding exchange repulsion.

It should be underscored that the first-order electrostatic interactions do not correspond to properly antisymmetrized wavefunctions and the terms reflecting the effect of introduction of the proper wavefunction symmetry are the exchange-repulsion contributions. The first-order interaction of unperturbed species is thus repulsive due to large overlap of molecular wavefunctions.

The induction and exchange-induction effects reflecting the relaxation of monomeric wavefunctions upon the mutual polarization interactions are quantified by the delocalization terms. The digits in the subscripts indicate pairwise additive (2-body) or nonadditive (3-body) contributions. Even though the stabilizing polarization interactions are in excess of the repulsive first-order terms and are stabilizing the complexes already at the Hartree-Fock level the repulsive nonadditive contribution indicates that induction and charge transfer effects are, in fact, anti-cooperative in the gas-phase. In effect at the Hartree-Fock level we observe rather small stabilization, in terms of magnitude corresponding roughly to that in a water dimer. However, the total interaction energy is almost an order of magnitude larger due to electron-correlation effects, and in particular due to the additive second-order dispersion interactions.

It should be underscored though that the origins of intermolecular interactions change significantly in the presence implicit solvent model, represented here as a polarizable dielectric continuum. As expected, the electrostatic interactions are screened and significantly quenched by the dielectric medium (cf. smaller magnitudes of $\varepsilon_{\mathrm{el}}$ terms and repulsive $\Delta G^{\mathrm{HF,S}}$ term, particularly in the charged complexes). In effect, the weakly attractive interactions observed for the gas-phase trimers at the Hartree-Fock level (-21.7 and -8.0 kJ/mol for the **19** and **18** trimers, respectively) become weakly repulsive (5.6 and 6.8 kJ/mol, respectively). What is interesting, though, is that the nonadditive polarization interactions encompassed in the $\Delta E_{\mathrm{del},3}^{\mathrm{HF}}$ terms are now stabilizing the complexes, what indicates their cooperative effect. The same applies to the nonadditive contribution to the differential solvation free energy $\Delta G_3^{\mathrm{HF,S}}$ which diminishes significantly the dielectric screening effects. Since the electron-correlation correction to the first order electrostatic terms is negligible while the second-order exchange effects remain virtually unchanged and the dispersion attraction is only slightly quenched in the presence of a dielectric medium, the interaction free energies of the trimers (i.e. interaction energies corrected by the difference in solvation free energies of the complex and of the isolated species) remain significantly attractive (-58.9 and -56.5 kJ/mol, respectively for the **19** and **18** complexes), even though these were diminished in the solvent roughly by a factor of 2.

In summary, the large magnitudes of the dispersion and electrostatic interactions, which are the long-range contributions, suggest that in the solution the thioanhydronucleosides and hydrosulfide ions should reside in a close proximity and form stable complexes. The free energies of interaction obtained at the MP2/IEFPCM/aug-cc-pVDZ level (-58.9 and -56.5 kJ/mol, respectively for the **19** and **18** complexes) are consistent with the corresponding results obtained at the ωB97X-D/IEFPCM/ma-def2-TZVP level (-65.2 and -64.7 kJ/mol, respectively) which was used to optimize equilibrium geometries of the complexes.

Supplementary Figures

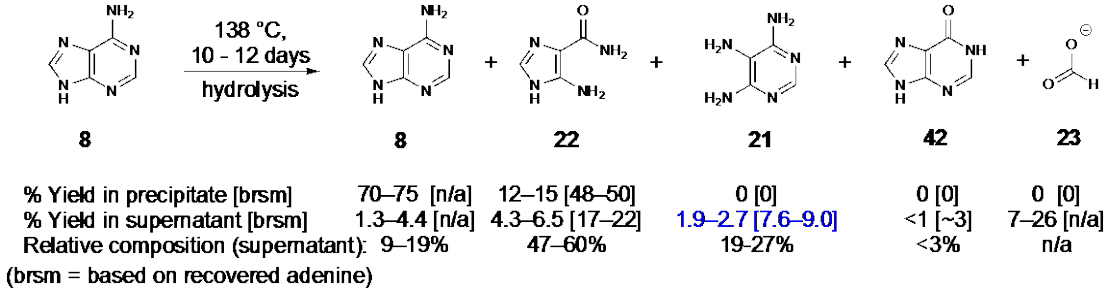


Fig. S1. Summary of distribution of products in adenine hydrolysate supernatants.

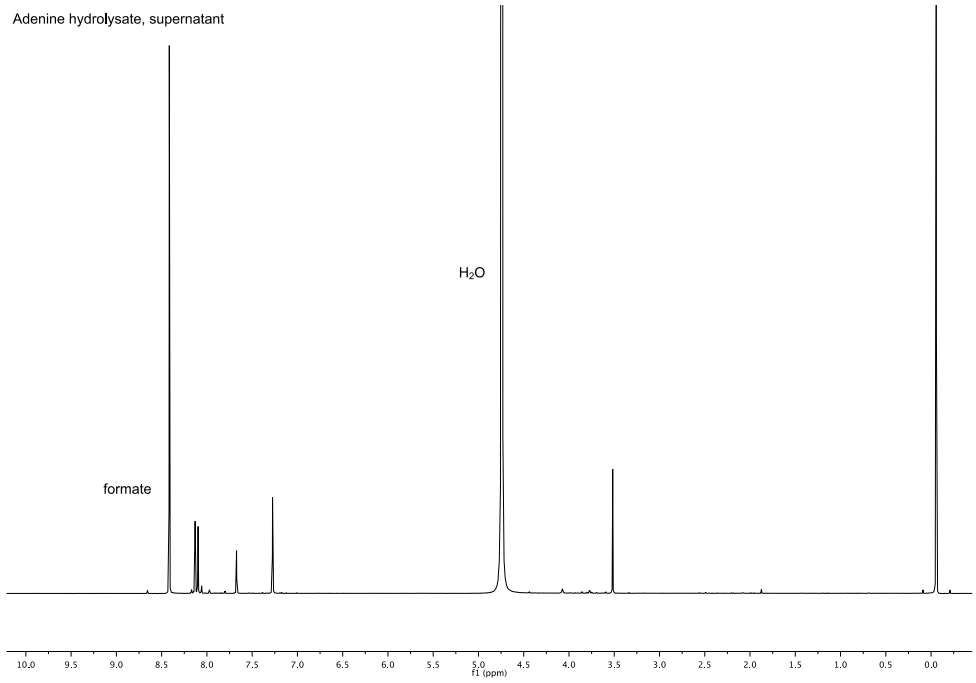


Fig. S2. ^{1}H NMR spectrum of the supernatant of adenine hydrolysis, showing formate **23**, adenine **8**, TAP **21** and AICA **22** as the constituents. The identities of these compounds were confirmed by spiking experiments and analytical HPLC.

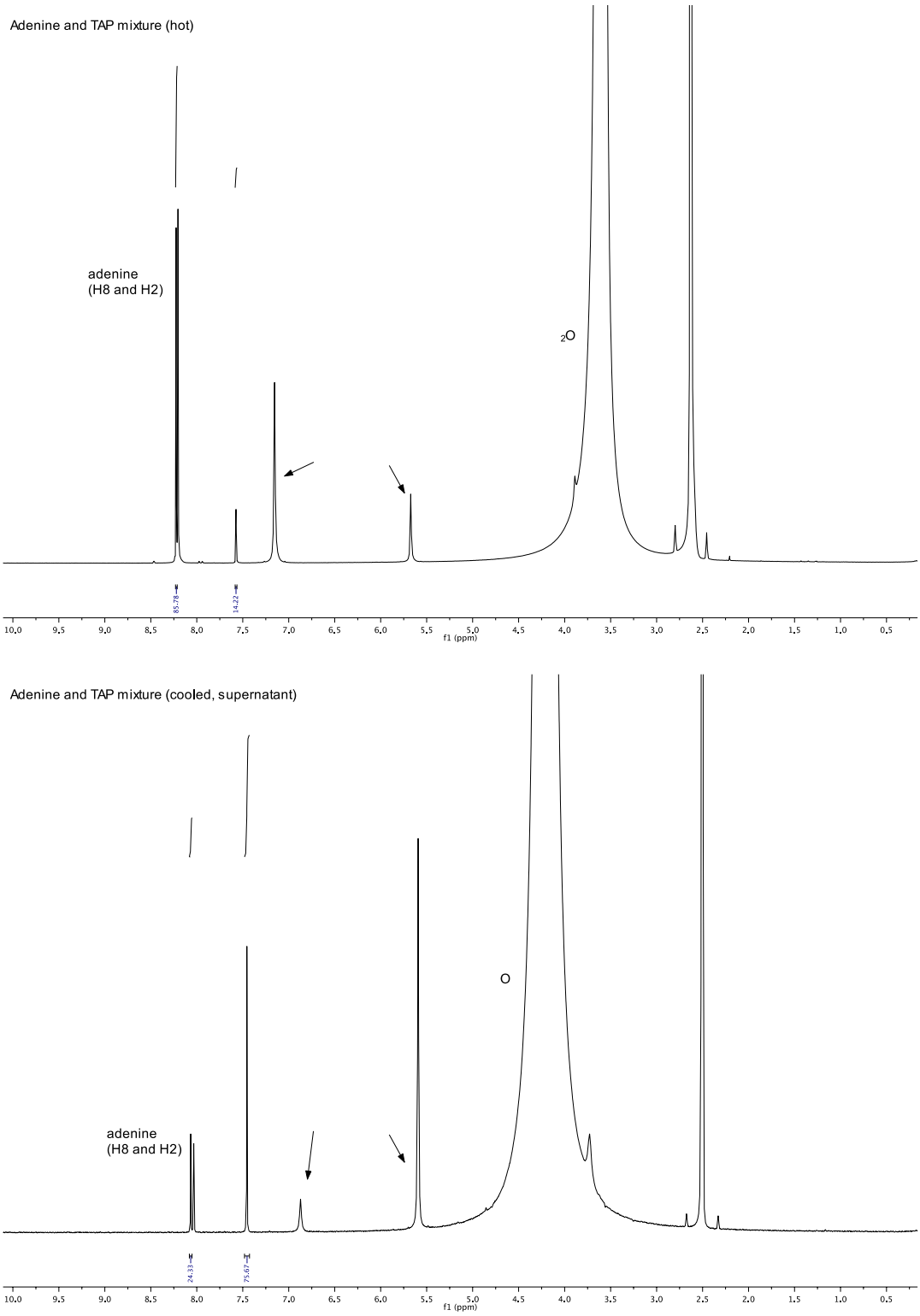


Fig. S3. ^{1}H NMR spectra for Table S1 entry 3, showing TAP **21** becoming the major compound in solution after cooling a solution which initially contained a majority of adenine.

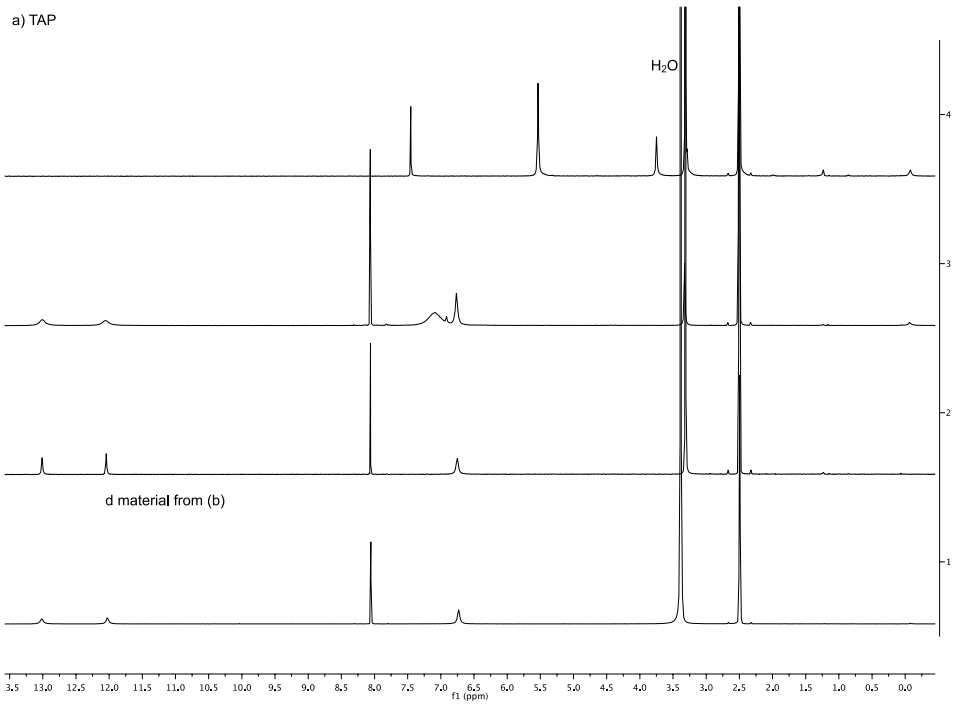


Fig. S4. ^{1}H NMR spectra of a) TAP **21**; b) dry state reaction of TAP **21** and NH_4SCN **24**; c) authentic 8-mercaptoadenine **16**; d) crystals recovered from the dry state reaction, after dissolving in water.

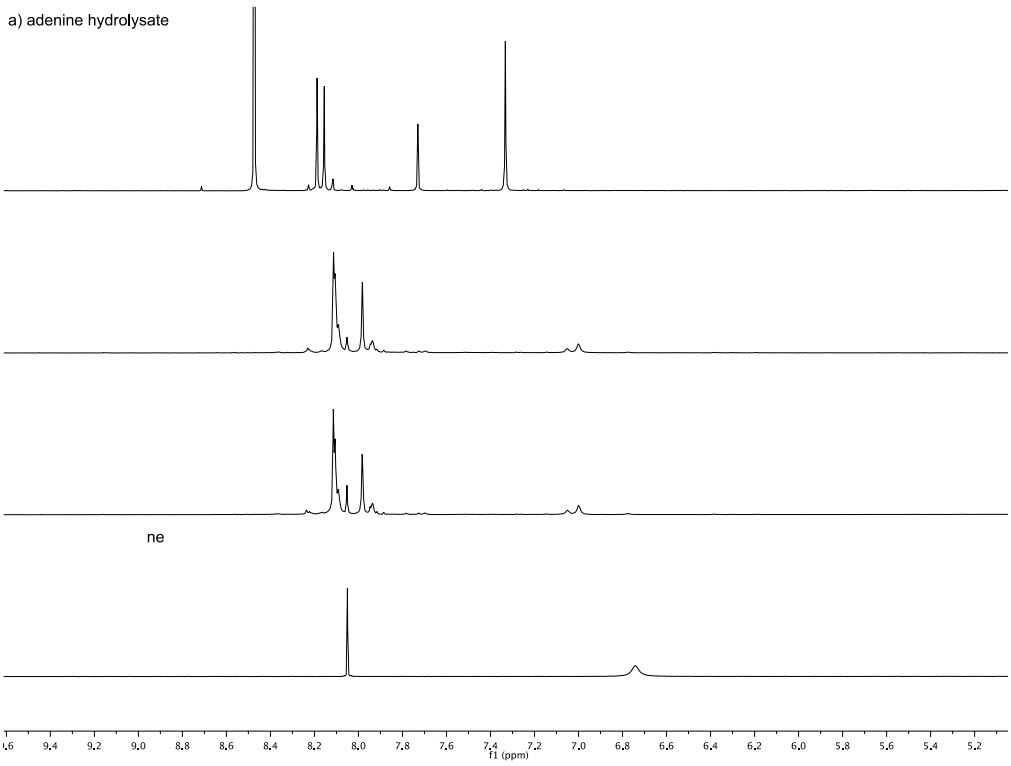


Fig. S5. ^{1}H NMR spectra of a) adenine hydrolysate; b) adenine hydrolysate after solid state reaction with NH_4SCN **24**; c) the reaction crude spiked with authentic 8-mercaptoadenine **16**; and d) 8-mercaptoadenine **16**.

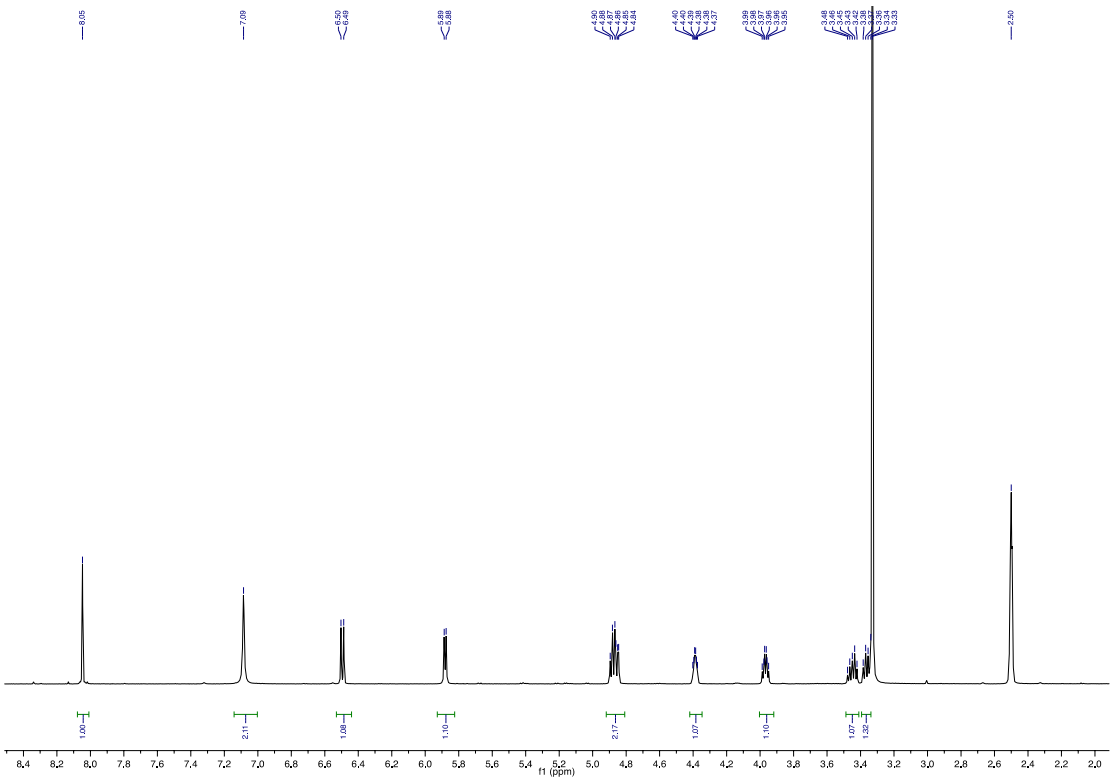


Fig. S6. ^{1}H NMR spectrum of N^{9}-8,2'-anhydro-thioadenosine **18** in DMSO-d_6.

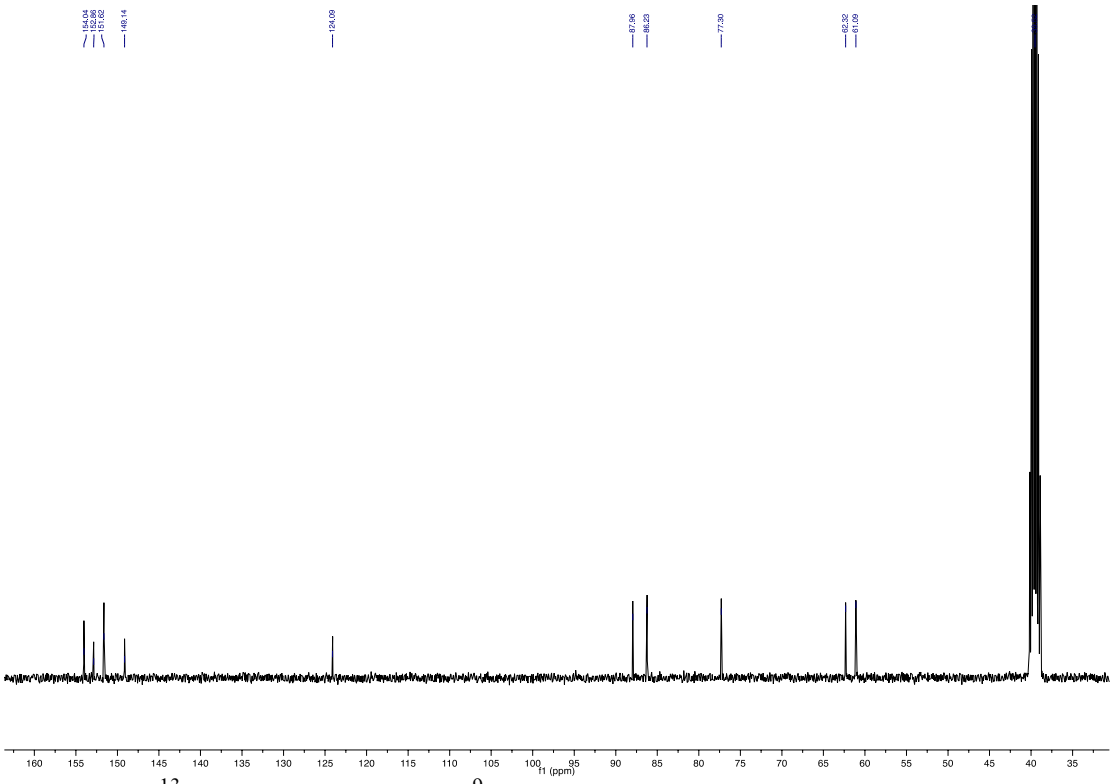


Fig. S7. ^{13}C NMR spectrum of N^{9}-8,2'-anhydro-thioadenosine **18** in DMSO-d_6.

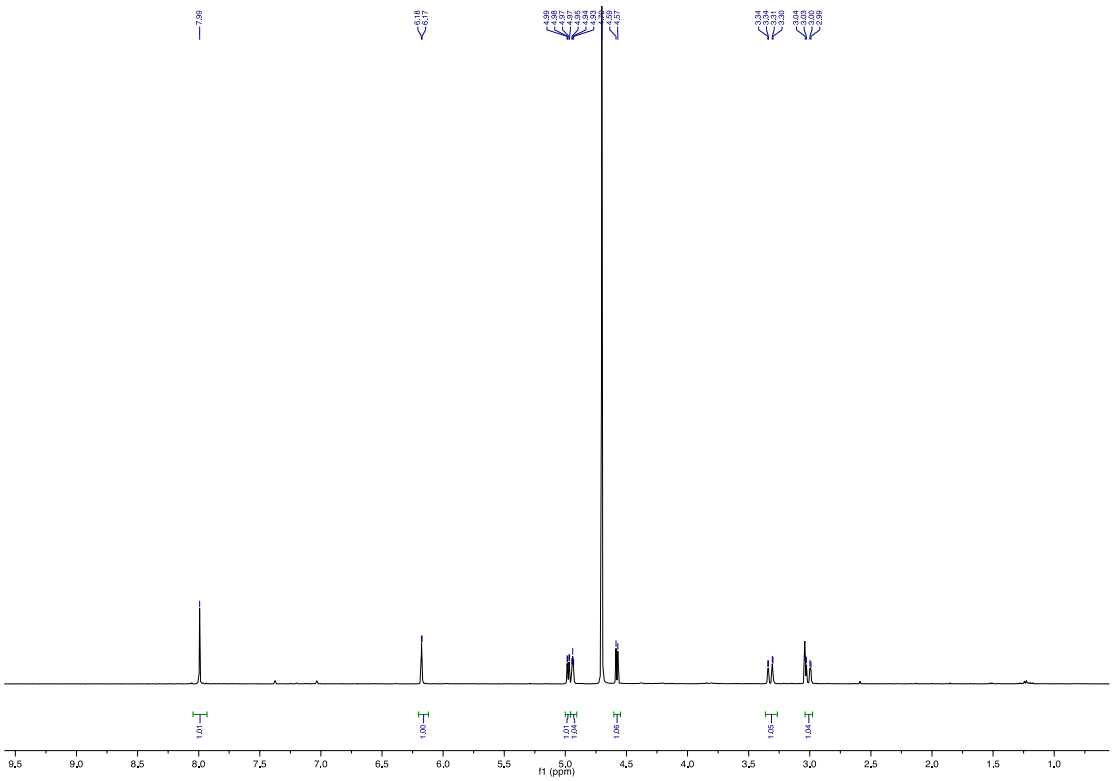


Fig. S8. ^{1}H NMR spectrum of N^{9}-8,5'-anhydro-thioadenosine **37** in D_2O.

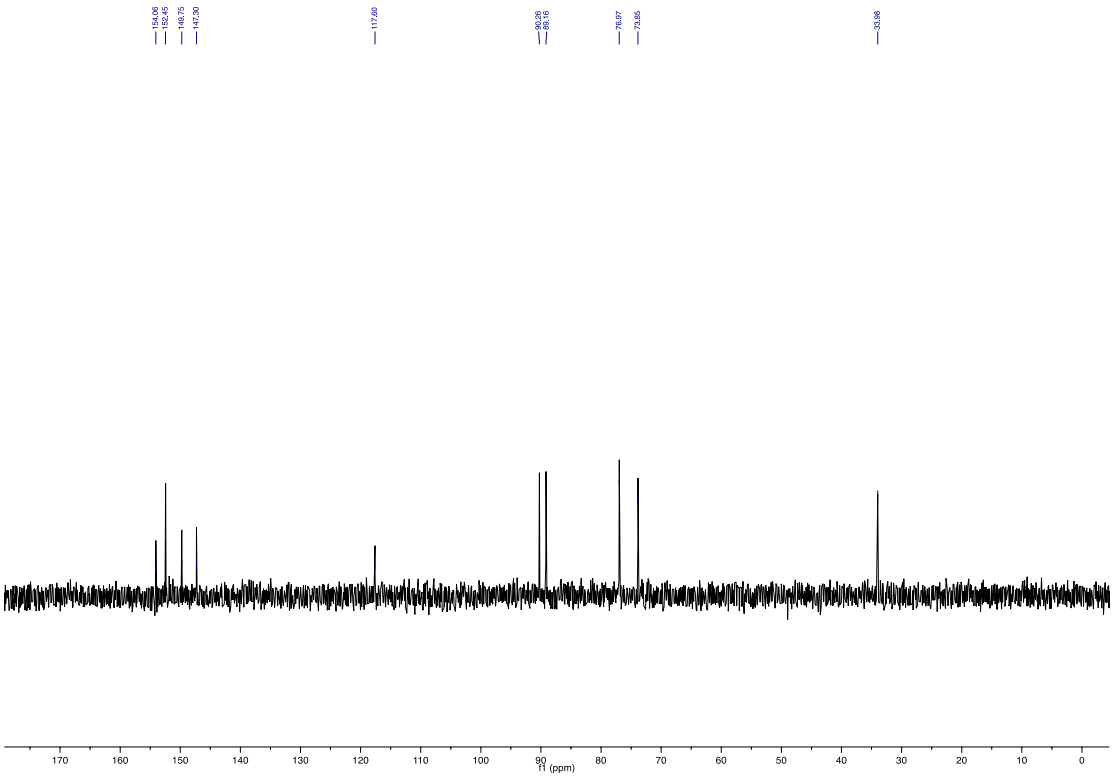


Fig. S9. ^{13}C NMR spectrum of N^{9}-8,5'-anhydro-thioadenosine **37** in D_2O.

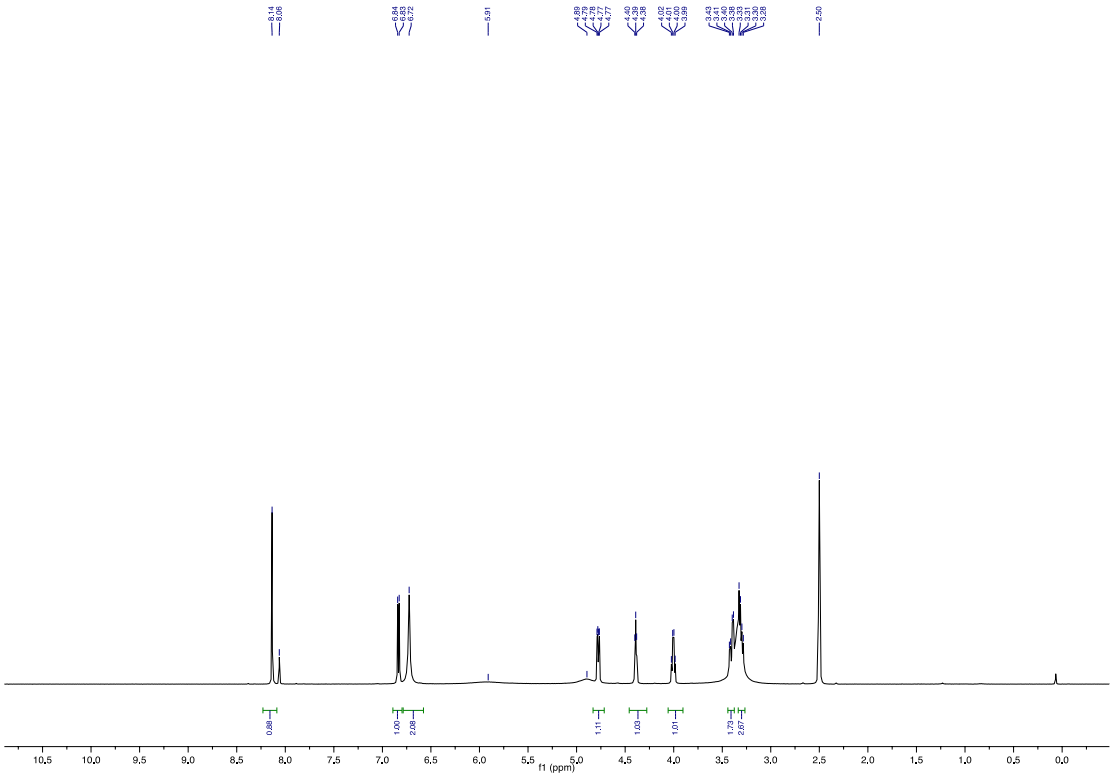


Fig. S10. ^{1}H NMR spectrum of N^{7}-8,2'-anhydro-thioadenosine **19** in DMSO-d_6.

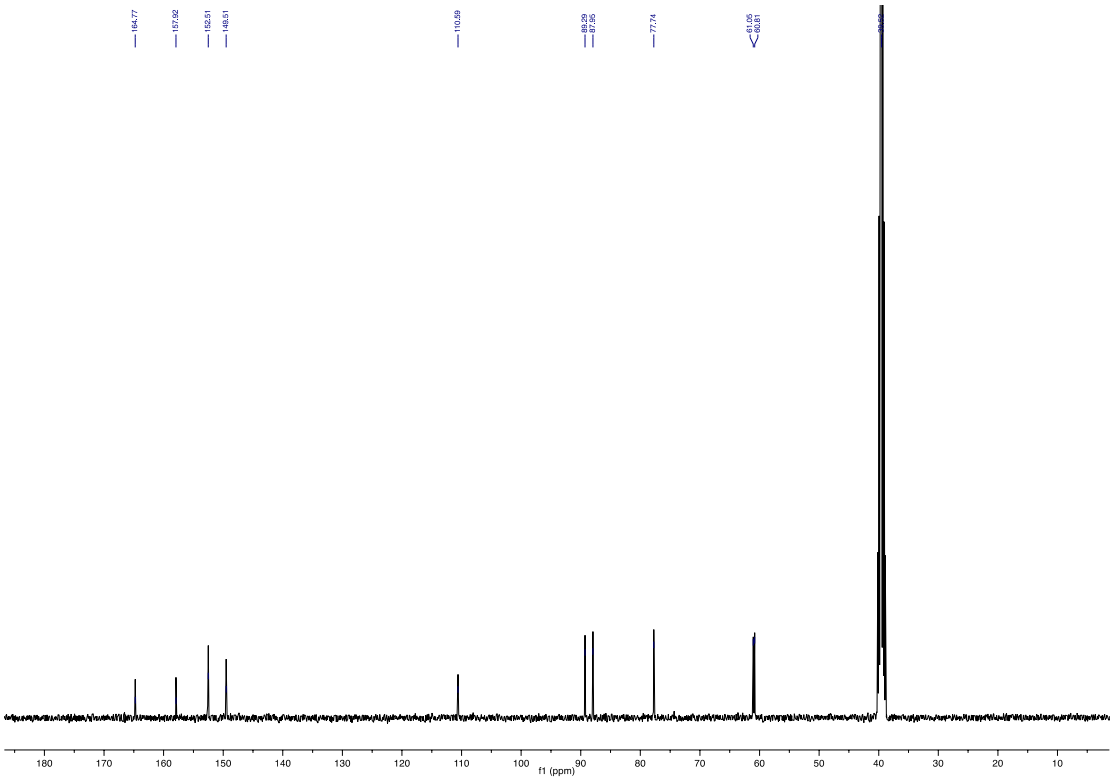


Fig. S11. ^{13}C NMR spectrum of N^{7}-8,2'-anhydro-thioadenosine **19** in DMSO-d_6.

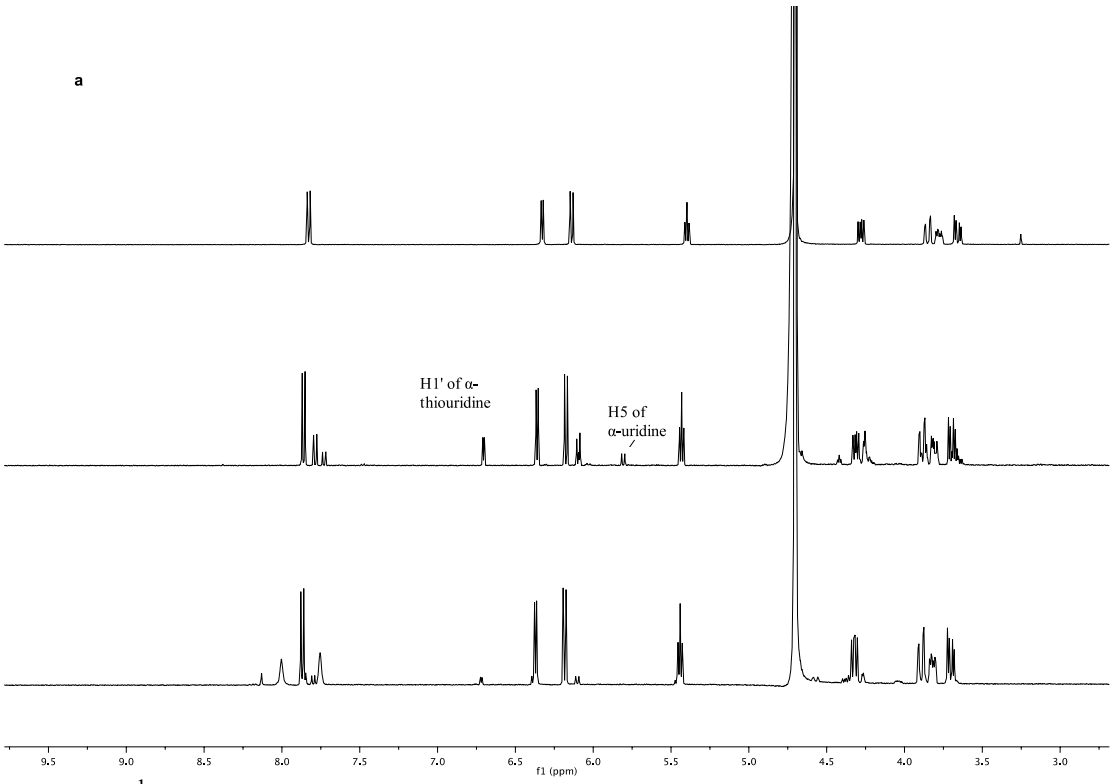


Fig. S12. ^{1}H NMR spectra of the mixture after heating α-thiouridine **14** at 80 °C. a) ^{1}H NMR spectrum of α-D-ribofuranosyl-2,2'-anhydrouridine **14**; b) ^{1}H NMR spectrum of the mixture after heating in H_2O; c) ^{1}H NMR spectrum of the mixture after heating in formamide.

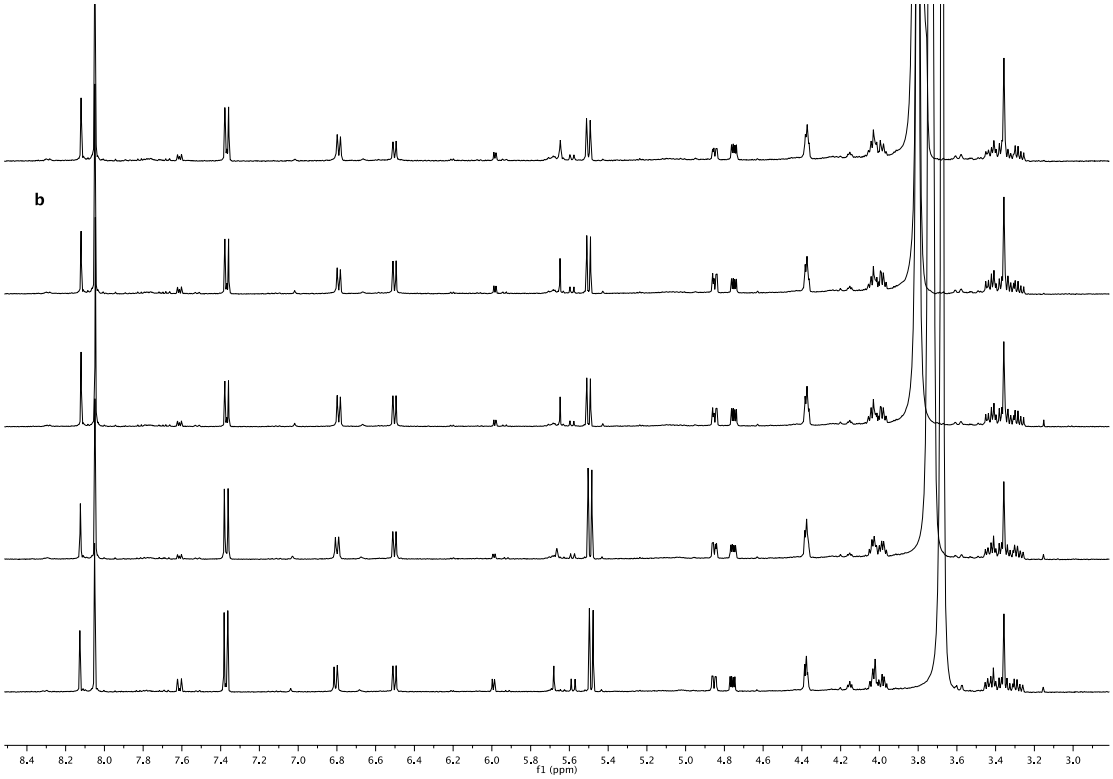


Fig. S13. ^{1}H NMR spectra of spiking experiments for the glycosidation of α-D-ribofuranosyl-2,2'-anhydrouridine **15** with 8-mercaptoadenine **16** in the presence of magnesium. a) ^{1}H NMR spectrum of the crude mixture after glycosidation; b) as a), spiked with N^{9}-8, 2'-anhydro-thioadenosine **18**; c) as b), spiked with N^{7}-8, 2'-anhydro-thioadenosine **19**; d) as c), spiked with uracil **39**; e) as d), spiked with α-uridine **38**. Signals from the spiking compounds are indicated with blue arrows, as shown above.

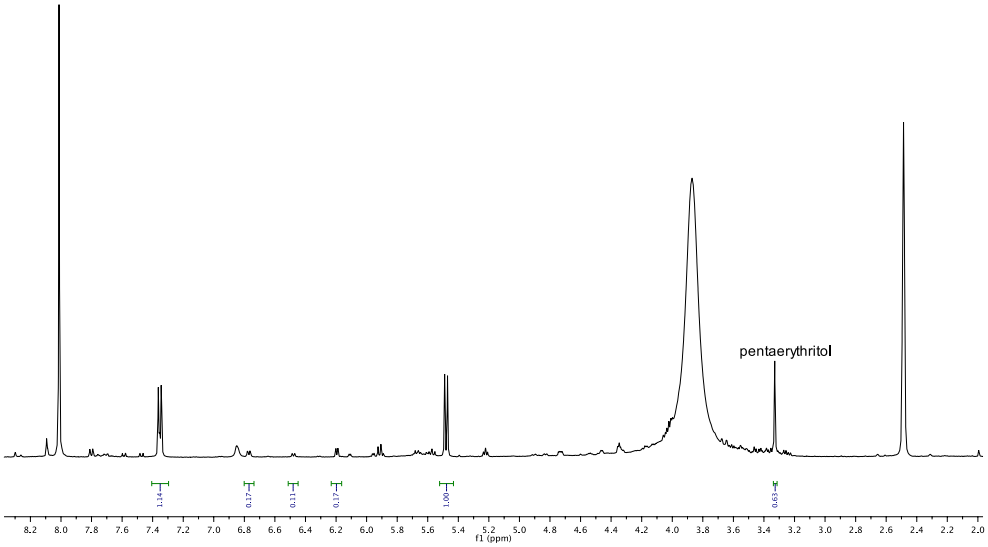


Fig. S14. ^{1}H NMR spectrum of the glycosidation of α-D-ribofuranosyl-2,2'-anhydrouridine **15** in the dry state without magnesium with added pentaerythritol (50 uL of 50 mM solution).

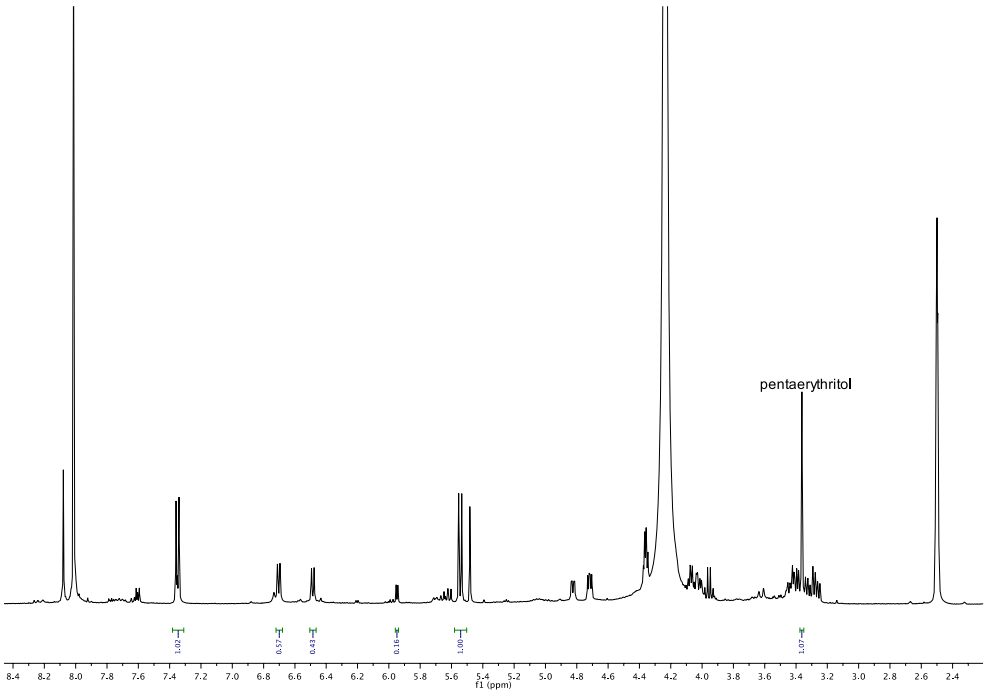


Fig. S15. ^{1}H NMR spectrum of the glycosidation of α-D-ribofuranosyl-2,2'-anhydrouridine **15** in the dry state in the presence of magnesium with added pentaerythritol (100 uL of 50 mM solution).

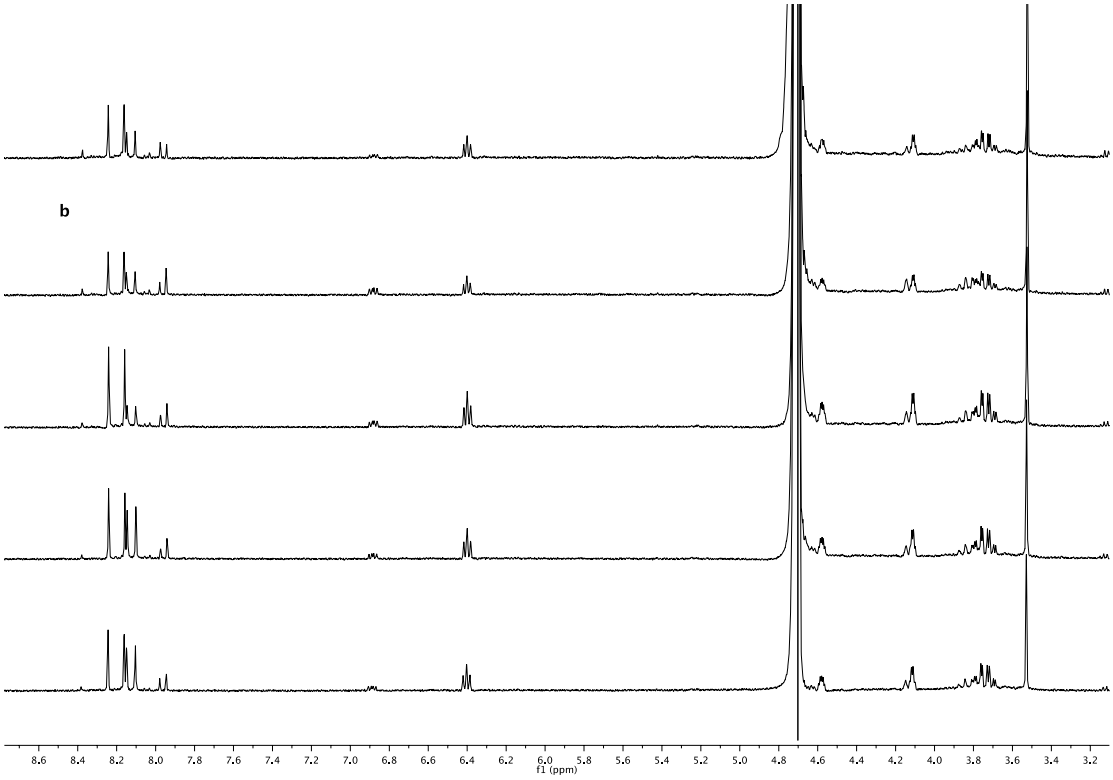


Fig. S16. ^{1}H NMR spectra of spiking experiments for the photoreduction of N^{9}-8, 2'-anhydro-thioadenosine **18** with hydrogen sulfide. a) ^{1}H NMR spectrum of the crude mixture after irradiation; b) as a), spiked with 8-mercapto-deoxyadenosine **26**; c) as b), spiked with N^{9}-deoxyadenosine **7**; d) as c), spiked with adenine **8**; e) as d), spiked with 8-mercaptoadenine **16**. Signals from the spiking compounds are indicated with blue arrows, as shown above.

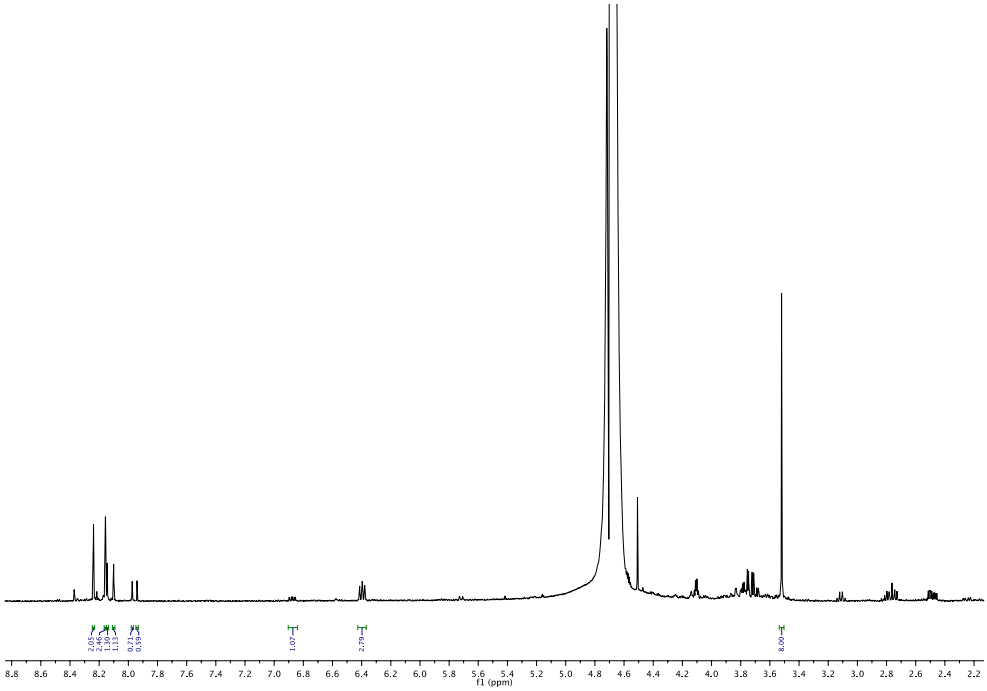


Fig. S17. ^{1}H NMR spectrum following photoreduction of N^{9}-8,2'-anhydro-thioadenosine **18** with hydrogen sulfide, with added pentaerythritol (20 uL of 50 mM solution).

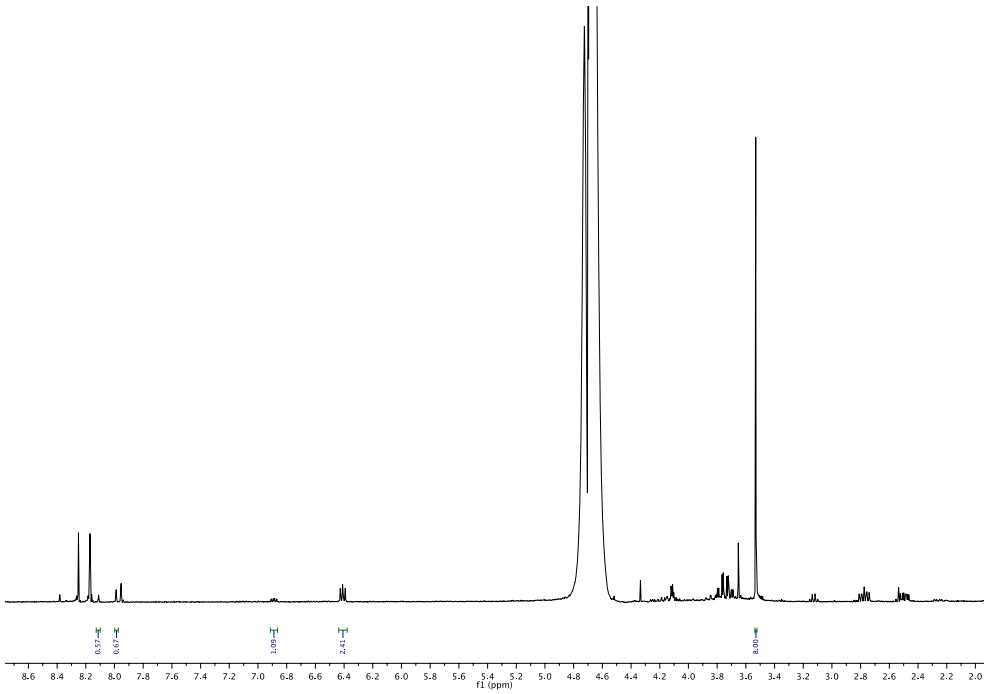


Fig. S18. ^{1}H NMR spectrum following photoreduction of N^{9}-8,2'-anhydro-thioadenosine **18** with bisulfite, with added pentaerythritol (30 uL of 50 mM solution).

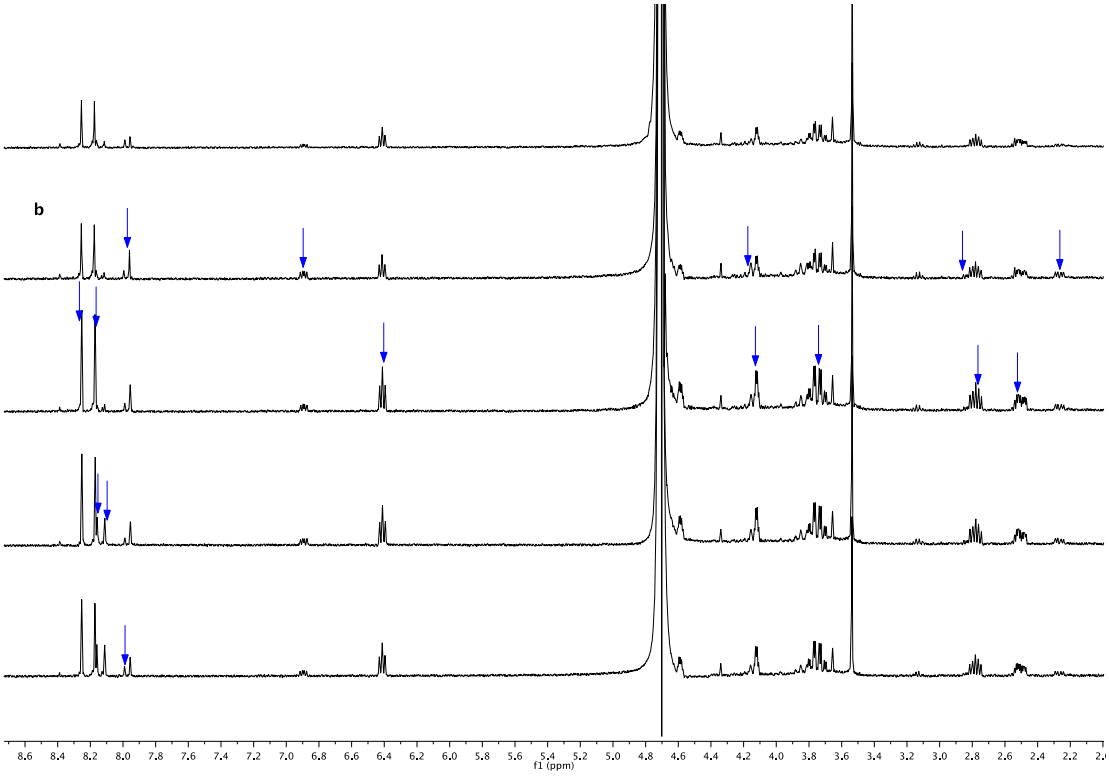


Fig. S19. ^{1}H NMR spectra of spiking experiments following photoreduction of N^{9}-8,2'-anhydro-thioadenosine **18** with bisulfite. a) ^{1}H NMR spectrum of the crude mixture after irradiation; b) as a), spiked with 8-mercapto-deoxyadenosine **26**; c) as b), spiked with N^{9}-deoxyadenosine **7**; d) as c), spiked with adenine **8**; e) as d), spiked with 8-mercaptoadenine **16**. Signals from the spiking compounds are indicated with blue arrows, as shown above.

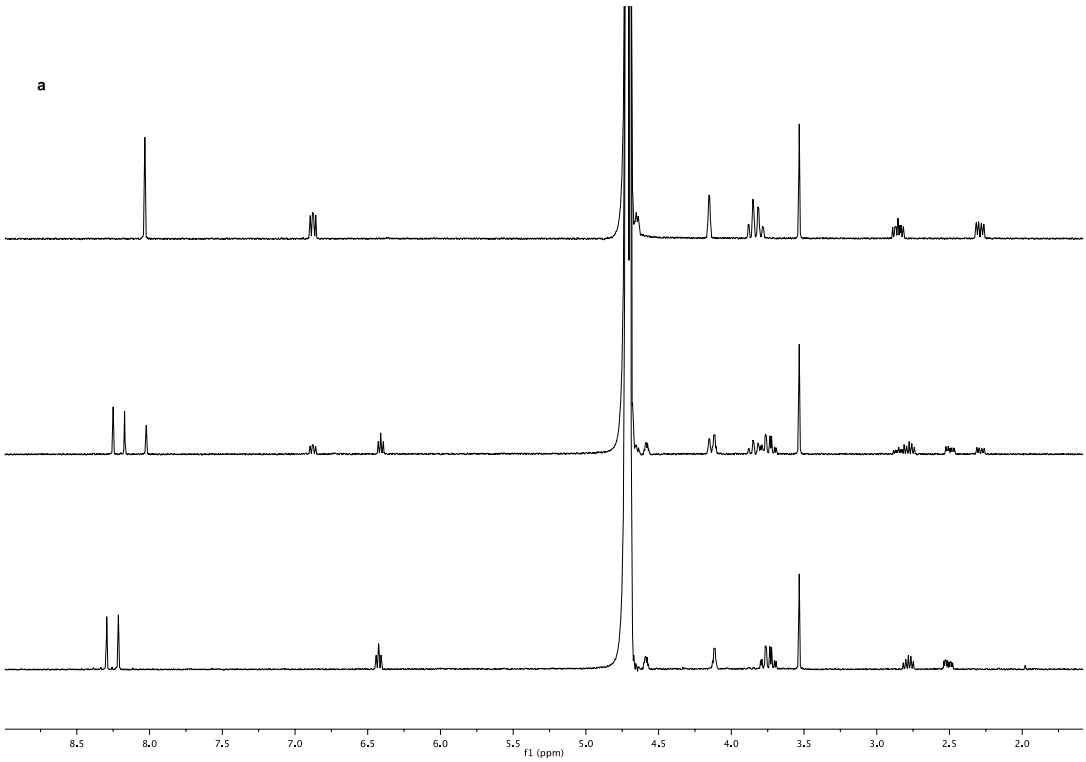


Fig. S20. ^{1}H NMR spectra following photoreduction of N^{9}-8-mercapto-deoxyadenosine **26**, with added internal standard pentaerythritol (10 uL of 50 mM solution). a) ^{1}H NMR spectrum of the crude mixture before irradiation; b) the mixture after irradiation for 4 hrs; c) the mixture after irradiation for 23 hours.

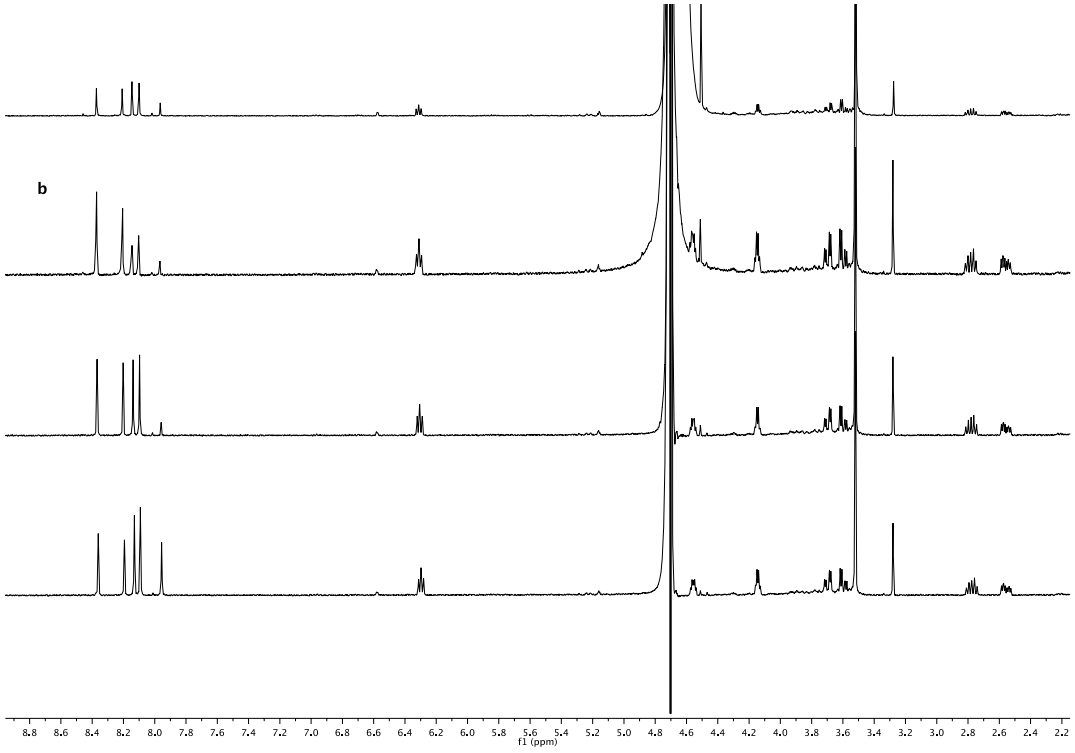


Fig. S21. ^{1}H NMR spectra of spiking experiments following photoreduction of N^7-8,2'-anhydro-thioadenosine **19** with hydrogen sulfide. a) ^{1}H NMR spectrum of the crude mixture after irradiation; b) as a), spiked with N^7-deoxyadenosine **20**; c) as b), spiked with adenine **8**; d) as c), spiked with 8-mercaptoadenine **16**. Signals from the spiking compounds are indicated with blue arrows, as shown above.

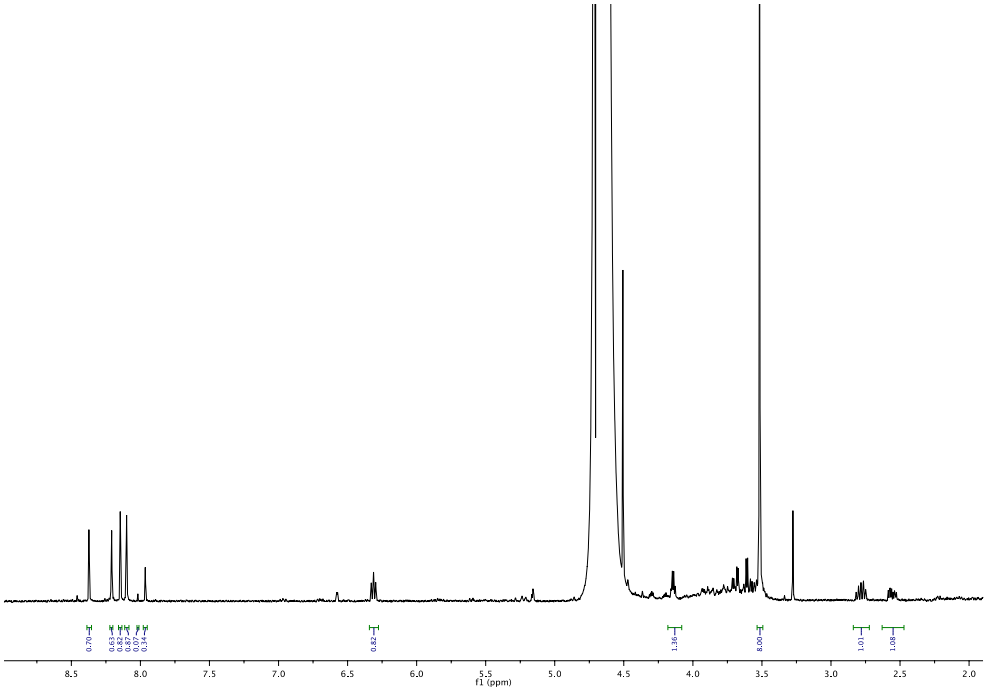


Fig. S22. ^{1}H NMR spectrum following photoreduction of N^{7}-8,2'-anhydro-thioadenosine **19** with hydrogen sulfide, with added pentaerythritol (20 uL of 50 mM solution).

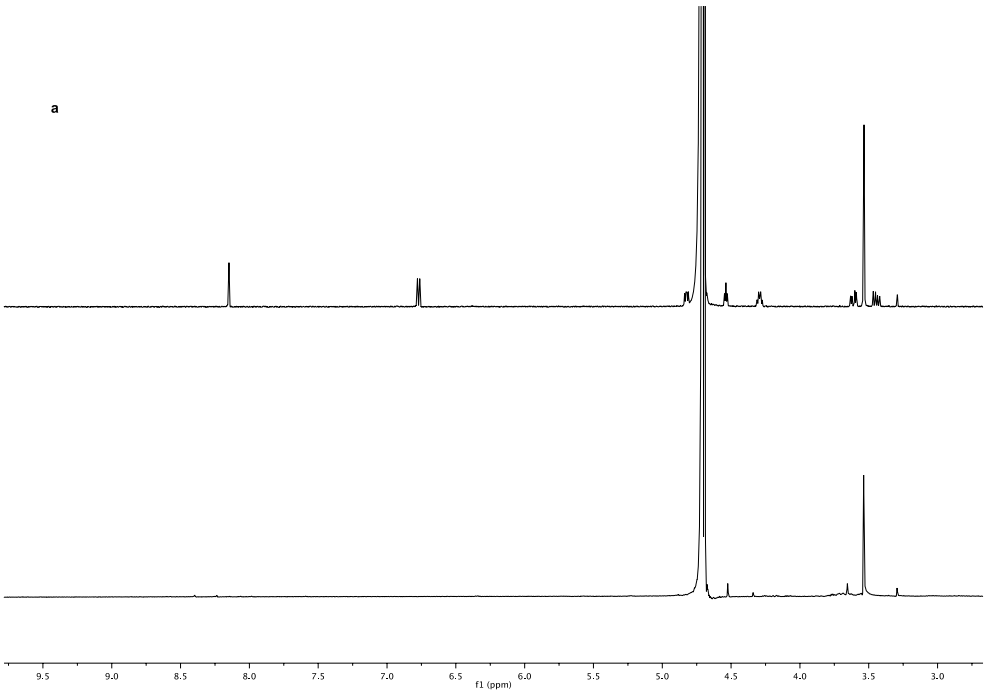


Fig. S23. ^{1}H NMR spectra following photoreduction of N^{7}-8,2'-anhydro-thioadenosine **19** with bisulfite, with internal standard pentaerythritol (20 uL of 50 mM solution). a) ^{1}H NMR spectrum of the crude mixture before irradiation; b) the mixture after irradiation for 2 hours.

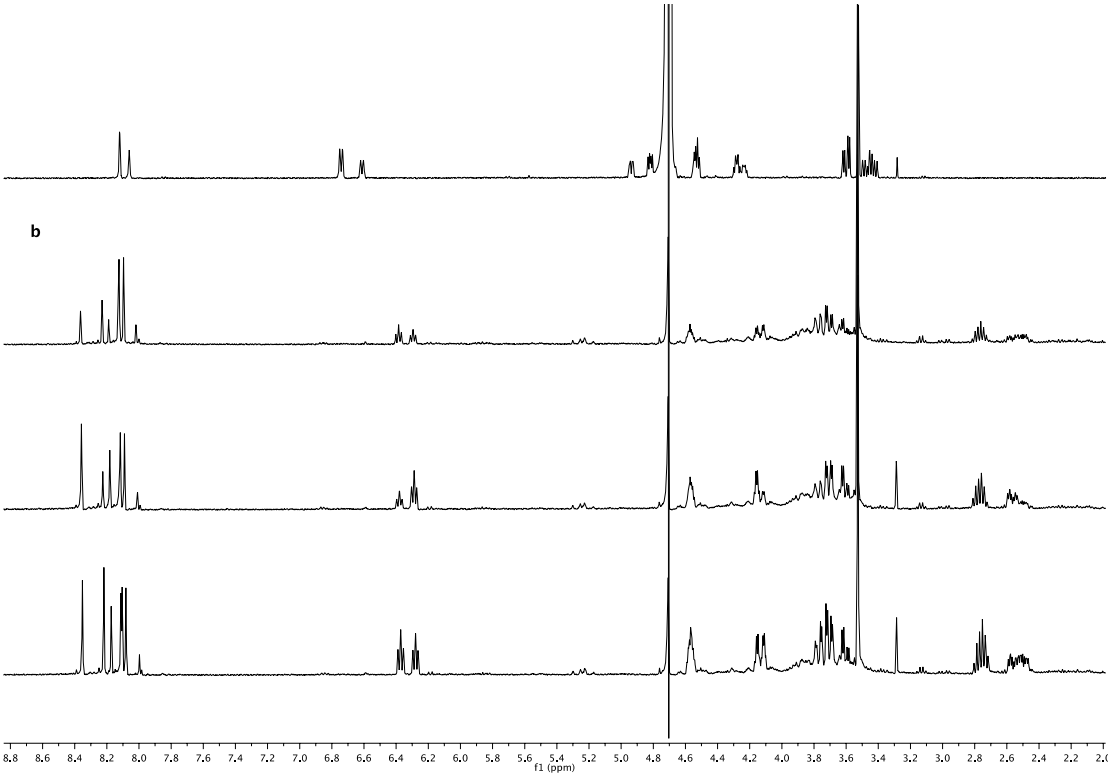


Fig. S24. ^{1}H NMR spectra following photoreduction of N^7-8,2'-anhydro-thioadenosine **19** and N^9-8,2'-anhydro-thioadenosine **18** with hydrogen sulfide, with internal standard pentaerythritol (20 uL of 50 mM solution). a) ^{1}H NMR spectrum of the crude mixture before irradiation (the N^7:N^9 isomer ratio was 62:38); b) the mixture after irradiation for 12 hrs (the N^7:N^9 isomer ratio was 44:56); c) as b), spiked with N^7-deoxyadenosine **20**; d) as c), spiked with N^9-deoxyadenosine **7**. Signals from the spiking compounds are indicated with blue arrows, as shown above.

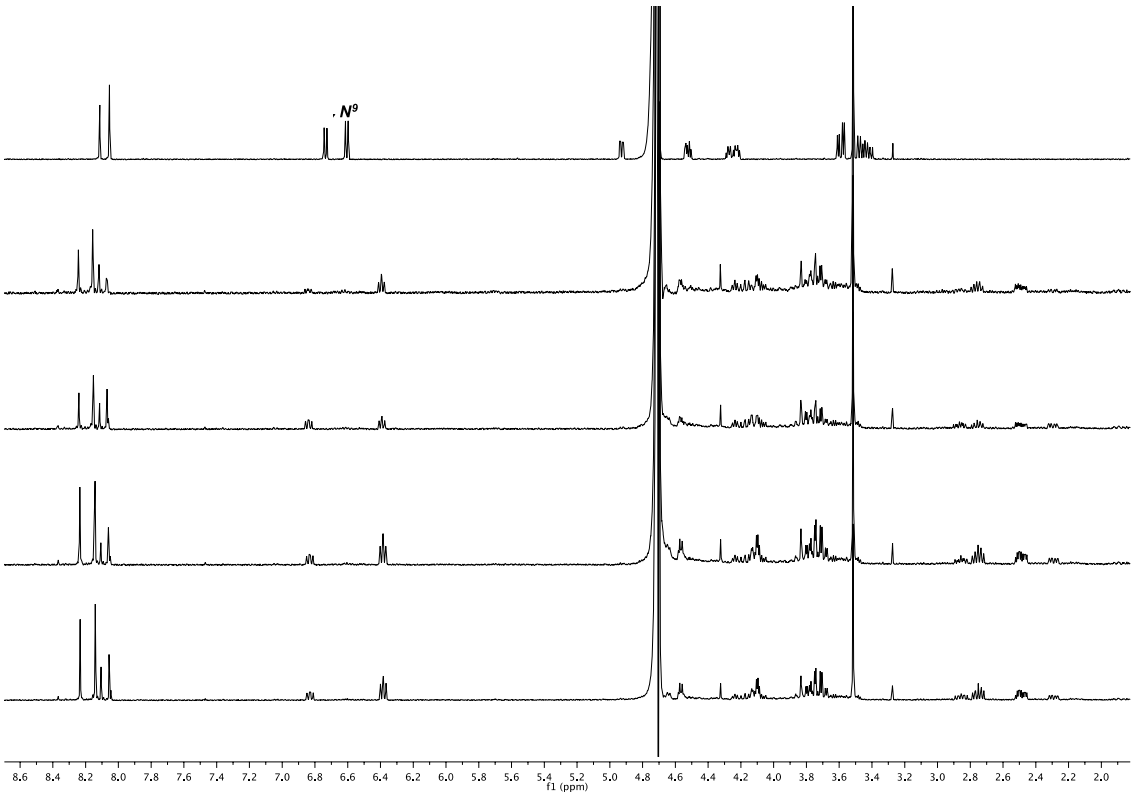


Fig. S25. ^{1}H NMR spectra of photoreduction of N^{7}-8,2'-anhydro-thioadenosine **19** and N^{9}-8,2'-anhydro-thioadenosine **18** with bisulfite, with internal standard pentaerythritol (20 uL of 50 mM solution). a) ^{1}H NMR spectrum of the crude mixture before irradiation (the N^{7}:N^{9} isomer ratio was 4:5); b) the mixture after irradiation for 7 hours (the N^{9} isomers were the only products after irradiation); c) as b), spiked with 8-mercapto-deoxyadenosine **26**; d) as c), spiked with N^{9}-deoxyadenosine **7**; e) as d), spiked with adenine **8**. Signals from the spiking compounds are indicated with blue arrows, as shown above.

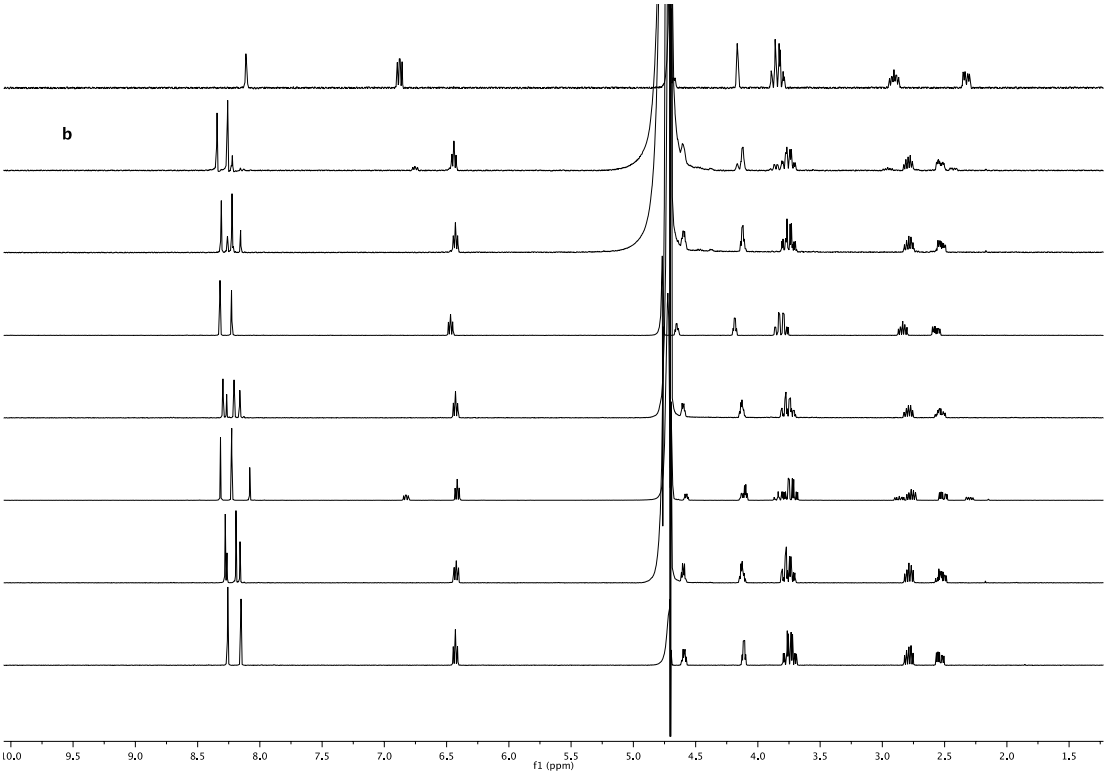


Fig. S26. ^{1}H NMR spectra for the reactions of deoxynucleosides with nitrous acid. a) ^{1}H NMR spectrum of N^9-8-mercapto-deoxyadenosine **26**; b) ^{1}H NMR spectrum of the reaction mixture of N^9-8-mercapto-deoxyadenosine **26** with nitrous acid after 2 h; c) ^{1}H NMR spectrum of the reaction mixture of N^9-8-mercapto-deoxyadenosine **26** with nitrous acid after 24 h; d) ^{1}H NMR spectrum of N^9-deoxyadenosine **7**; e) ^{1}H NMR spectrum of the reaction mixture of N^9-deoxyadenosine **7** with nitrous acid after 4 days; f) ^{1}H NMR spectrum of the mixture of N^9-8-mercapto-deoxyadenosine **26** and N^9-deoxyadenosine **7** (ratio of **26**:**7** is 1:2); g) ^{1}H NMR spectrum of the reaction mixture of N^9-8-mercapto-deoxyadenosine **26** and N^9-deoxyadenosine **7** with nitrous acid after 48 h (ratio of **7**:**9** is 68:32); h) ^{1}H NMR spectrum of deoxyinosine **9**.

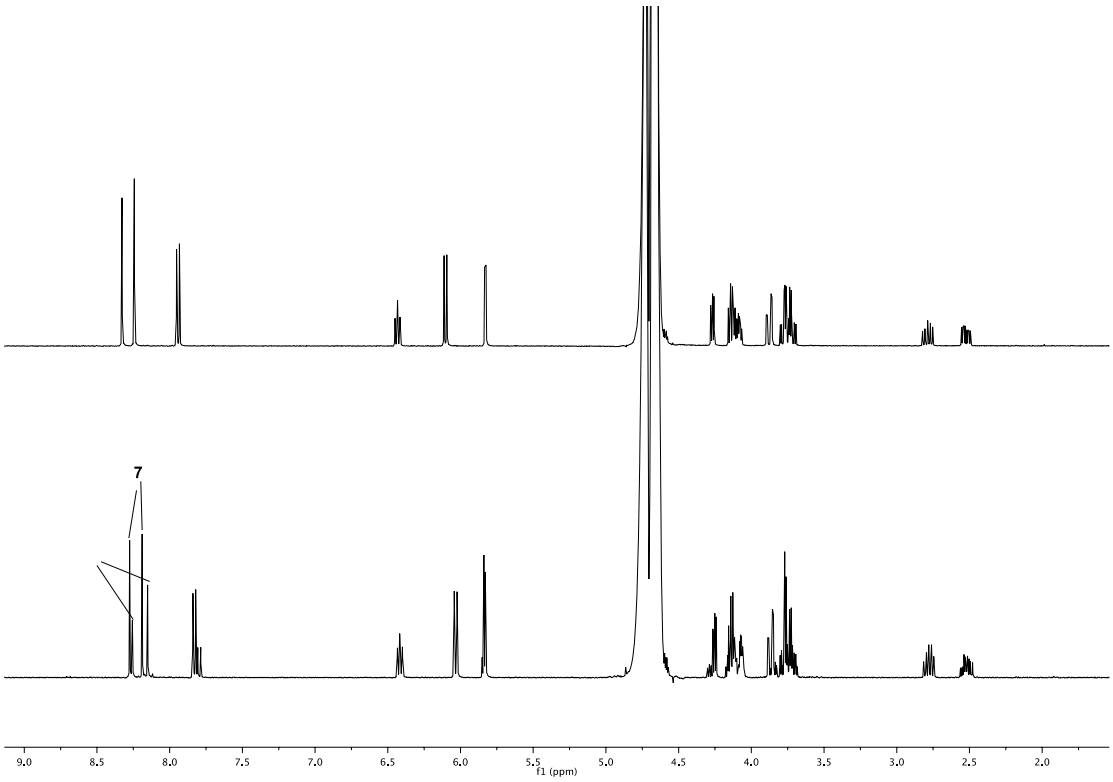


Fig. S27. ^{1}H NMR spectra for the reactions of deoxyadenosine **7** and cytidine **1** with nitrous acid. a) ^{1}H NMR spectrum of the mixture of deoxyadenosine **7** and cytidine **1**; b) ^{1}H NMR spectrum of the reaction mixture after 4 days, showing the ratio of all four (deoxy)nucleosides deoxyadenosine **7**, deoxyinosine **9**, cytidine **1**, and uridine **2** is 30:17:42:11.

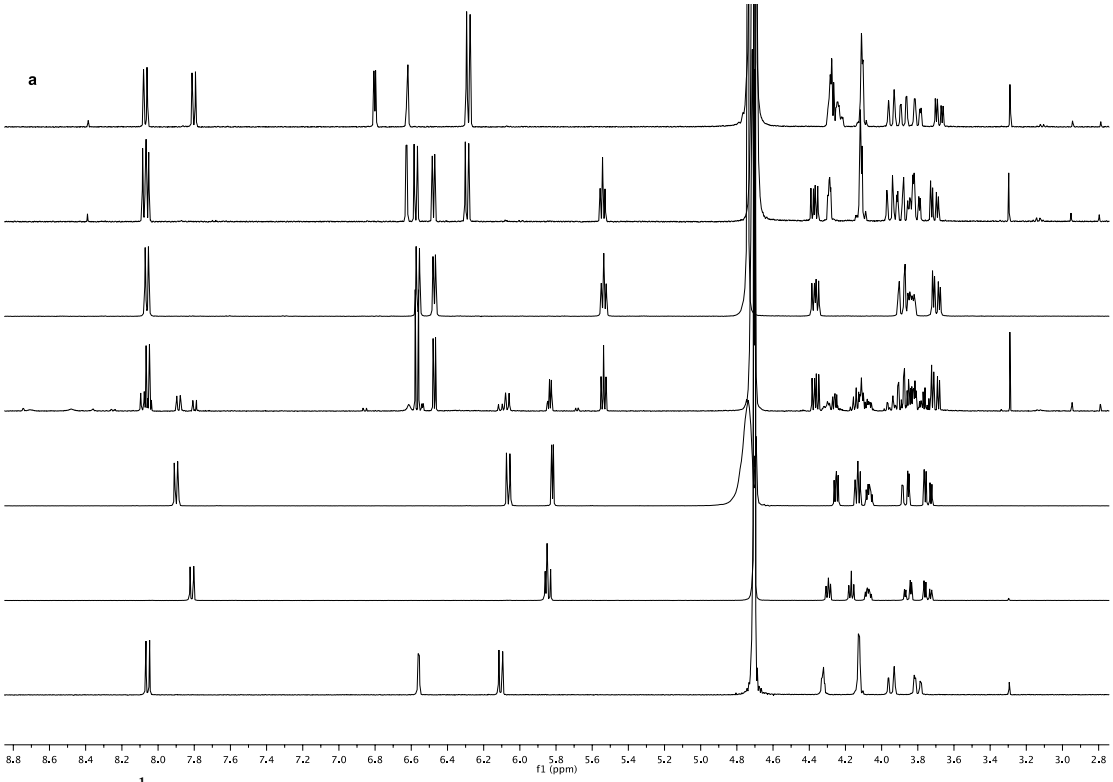


Fig. S28. ^{1}H NMR spectra for the reactions of 2-thiocytidine **3** and α-2-thiocytidine **13** with nitrous acid. a) ^{1}H NMR spectrum of the mixture of 2-thiocytidine **3** and α-2-thiocytidine **13**; b) ^{1}H NMR spectrum of the reaction mixture after 8 hours; c) ^{1}H NMR spectrum of anhydrocytidine **12**; d) ^{1}H NMR spectrum of the reaction mixture after 5.5 days; e) ^{1}H NMR spectrum of cytidine **1**; f) ^{1}H NMR spectrum of uridine **2**; g) ^{1}H NMR spectrum of 2-thiouridine **4**.

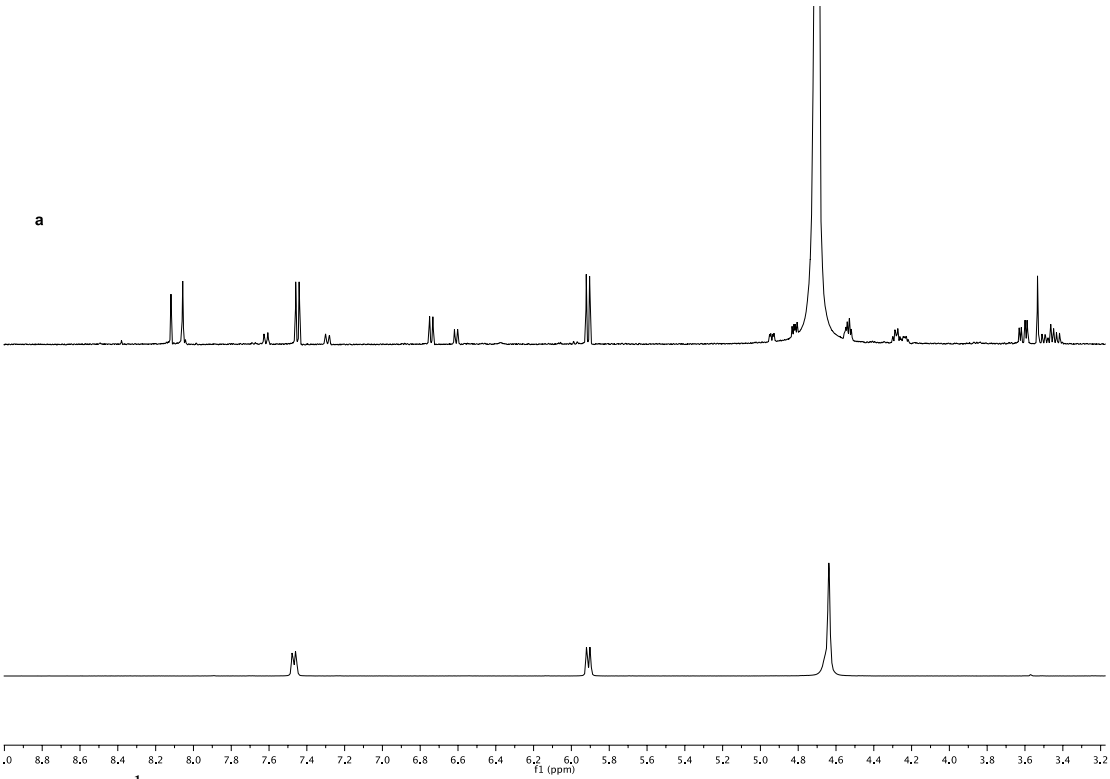


Fig. S29. ^{1}H NMR spectra for the reaction of α-D-ribofuranosyl-2,2'-anhydrocytidine **12** and 8-mercaptoadenine **16** with magnesium chloride. a) ^{1}H NMR spectrum of the mixture after heating for 1.5 days in dry state; b) ^{1}H NMR spectrum of cytosine **40**.

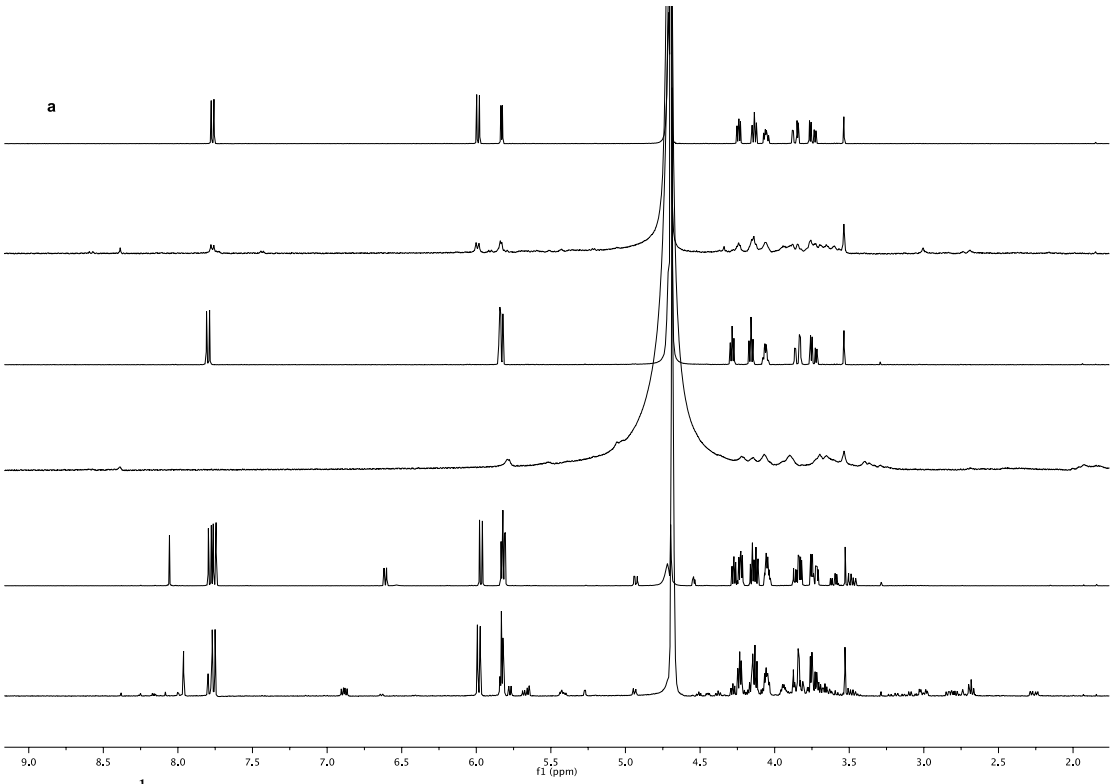


Fig. S30. ^{1}H NMR spectra for stability study of cytidine **1** and uridine **2** at 254 nm irradiation with bisulfite**.** a) ^{1}H NMR spectrum of the mixture of cytidine **1**, bisulfite and $K_4Fe(CN)_6$ in the dark; b) as a), ^{1}H NMR spectrum after 10 hours of irradiation; c) ^{1}H NMR spectrum of the mixture of uridine **2**, bisulfite and $K_4Fe(CN)_6$ in the dark; d) as c), ^{1}H NMR spectrum after 10 hours of irradiation; e) ^{1}H NMR spectrum of the mixture of cytidine **1**, uridine **2**, N^9-thioanhydroadenosine **18**, bisulfite and $K_4Fe(CN)_6$ in the dark; f) as e), ^{1}H NMR spectrum after 10 hours of irradiation.

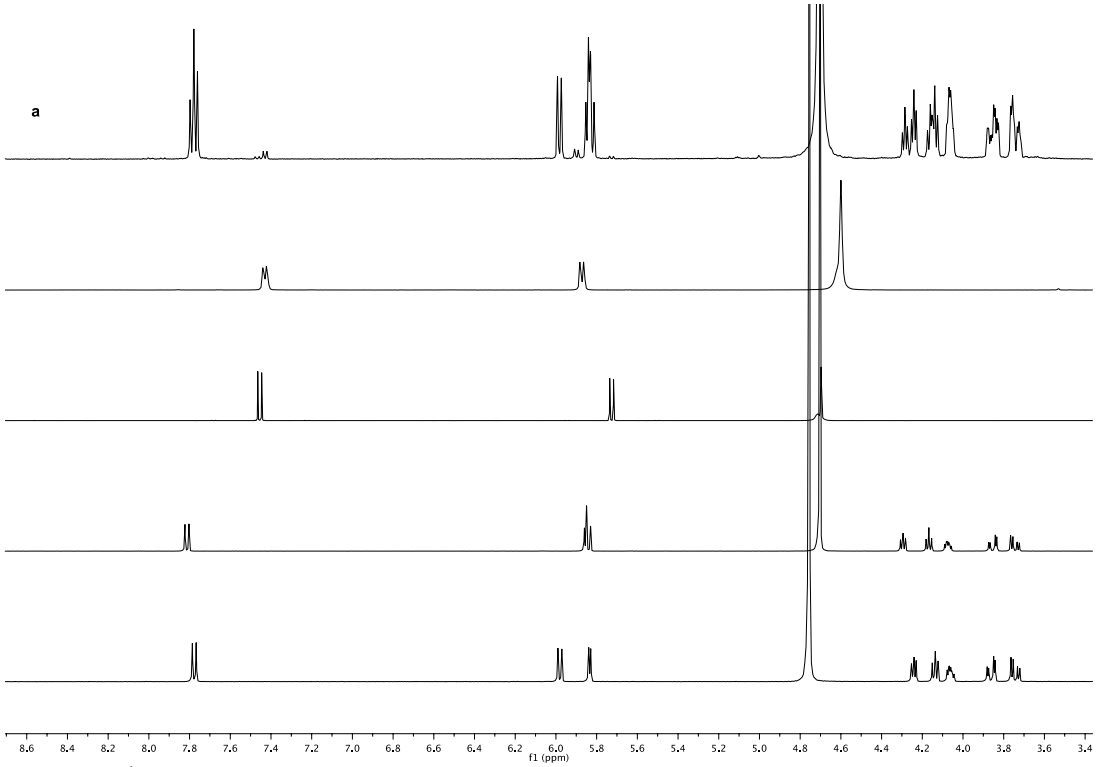


Fig. S31. ^{1}H NMR spectra for stability study of cytidine **1** and uridine **2** at 150 °C in dry state with $MgCl_2$. a) ^{1}H NMR spectrum of the mixture after heating at 150 °C for 1.5 days; b) ^{1}H NMR spectrum of cytosine **40**; c) ^{1}H NMR spectrum of uracil **39**; d) ^{1}H NMR spectrum of uridine **2**; e) ^{1}H NMR spectrum of cytidine **1**.

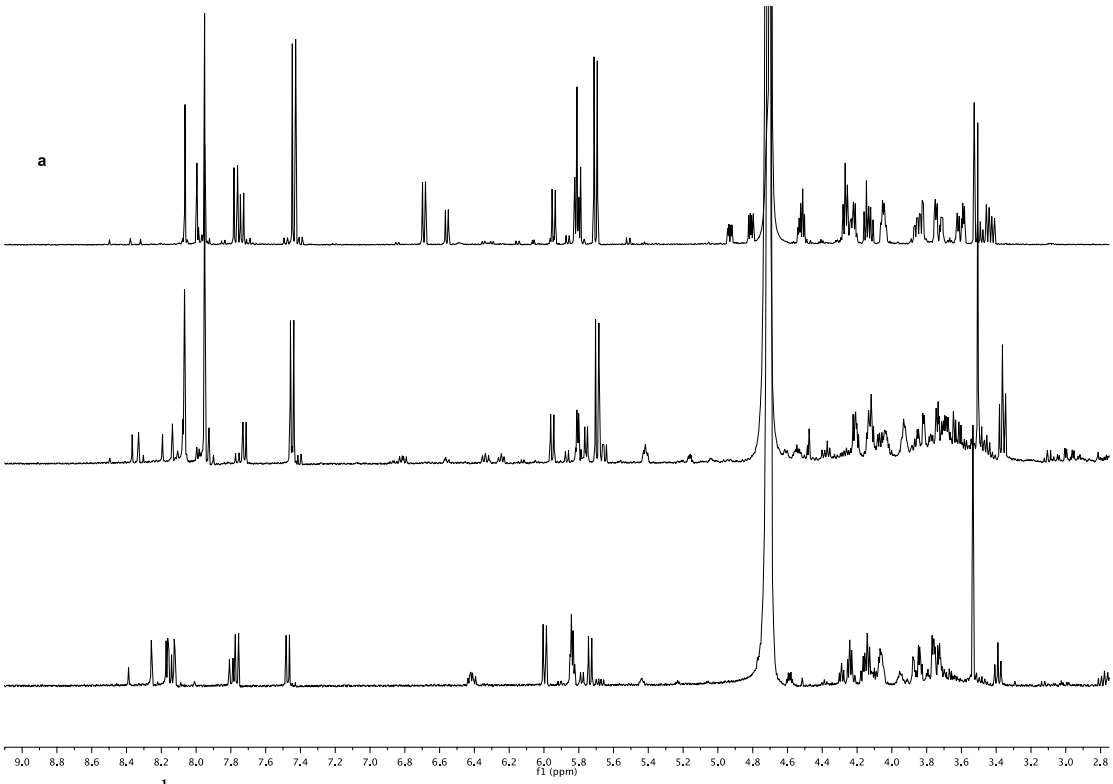


Fig. S32. ^{1}H NMR spectra for sequential reactions with the mixture of α-anhydrouridine **15,** cytidine **1** and uridine **2**. a) ^{1}H NMR spectrum of the mixture after heating with 8-mercaptoadenine **16** and magnesium chloride at 150 °C for 1.5 days; b) ^{1}H NMR spectrum of the same mixture after irradiation with hydrogen sulfide at 254 nm; c) ^{1}H NMR spectrum of the same mixture after reacting with nitrous acid for 2 days.

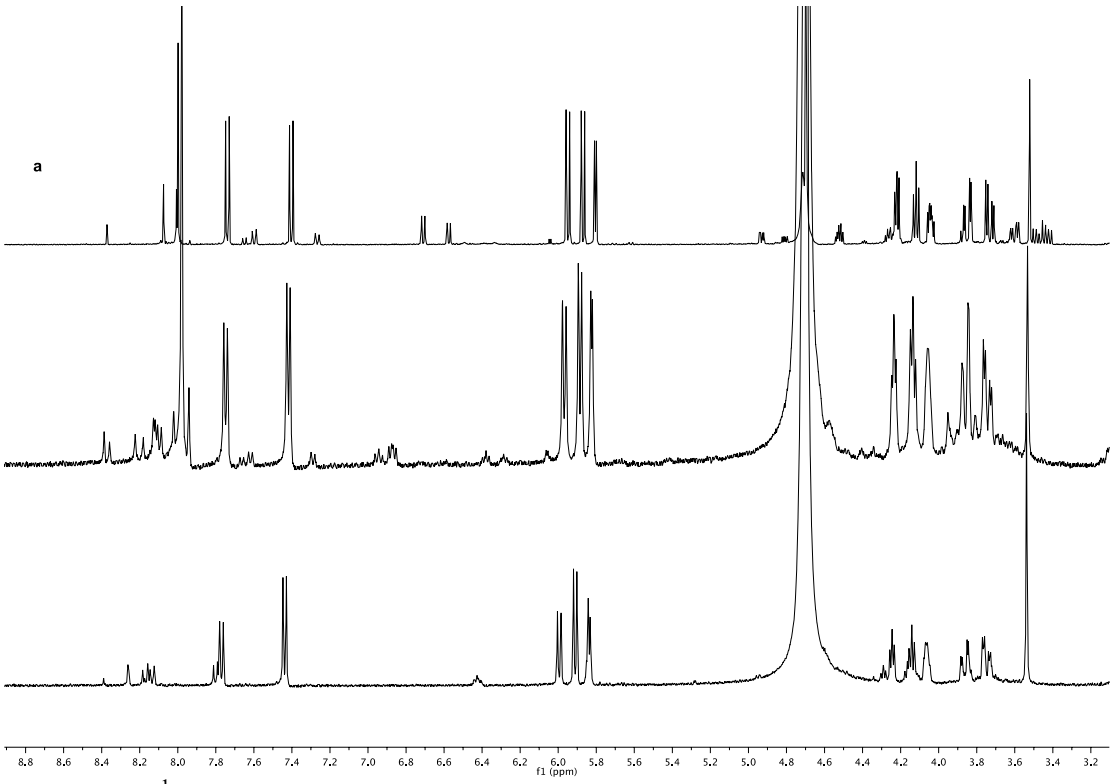


Fig. S33. ^{1}H NMR spectra for sequential reactions with the mixture of α-anhydrocytidine **12,** and cytidine **1**. a) ^{1}H NMR spectrum of the mixture after heating with 8-mercaptoadenine **16** and magnesium chloride at 150 °C for 1.5 days; b) ^{1}H NMR spectrum of the same mixture after irradiation with bisulfite and $K_4Fe(CN)_4$ at 254 nm; c) ^{1}H NMR spectrum of the same mixture after reacting with nitrous acid for 2 days.

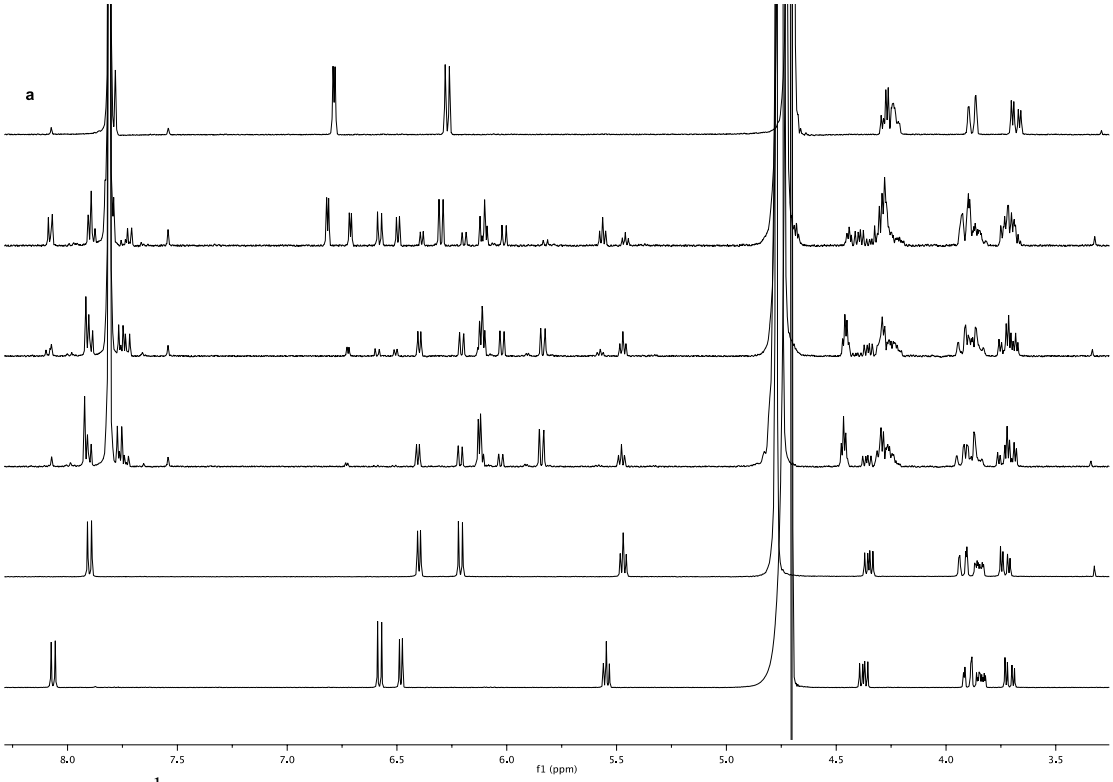


Fig. S34. ^{1}H NMR spectra for hydrolysis of α-thiocytidine **13** in 0.1 M 4,5-dicyanoimidazole (DCI) buffer (pH 5.2) at 60 °C. a) ^{1}H NMR spectrum of **13**; b) ^{1}H NMR spectrum of the mixture after 8 days of hydrolysis; c) ^{1}H NMR spectrum of the mixture after 16 days of hydrolysis; d)^{1}H NMR spectrum of the mixture after 24 days of hydrolysis; e) ^{1}H NMR spectrum of **15**; f) ^{1}H NMR spectrum of **12**.

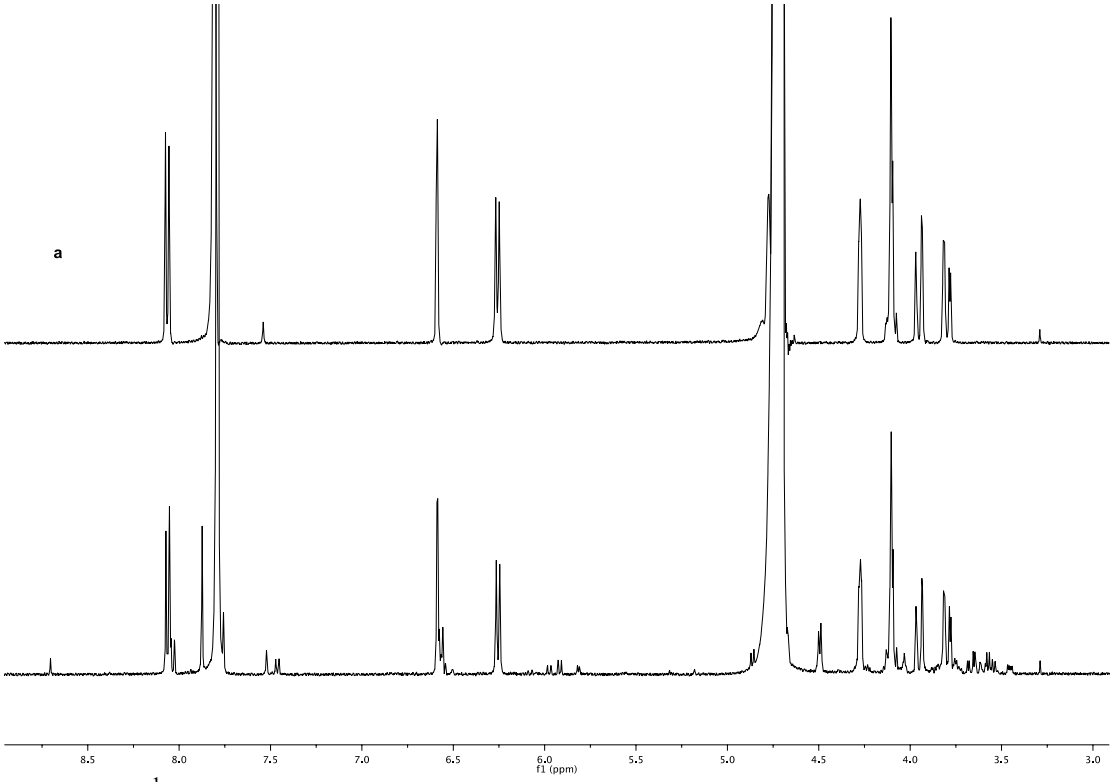


Fig. S35. ^{1}H NMR spectra for hydrolysis of β-thiocytidine **3** in 0.1 M 4,5-dicyanoimidazole (DCI) buffer (pH 5.2) at 60 °C. a) ^{1}H NMR spectrum of **3**; b) ^{1}H NMR spectrum of the mixture after 24 days of hydrolysis.

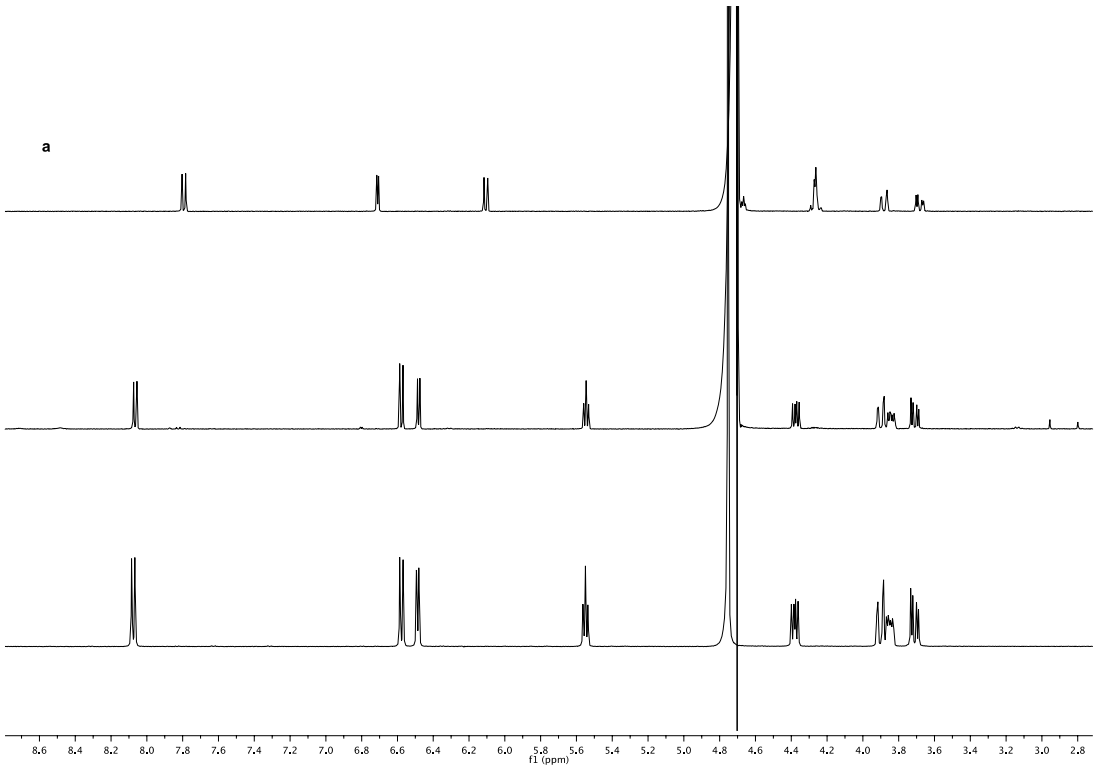


Fig. S36. ^{1}H NMR spectra for the reaction of α-thiocytidine **13** with nitrous acid. a) ^{1}H NMR spectrum of **13**; b) ^{1}H NMR spectrum of the mixture after 1 hour. c) ^{1}H NMR spectrum of **12**.

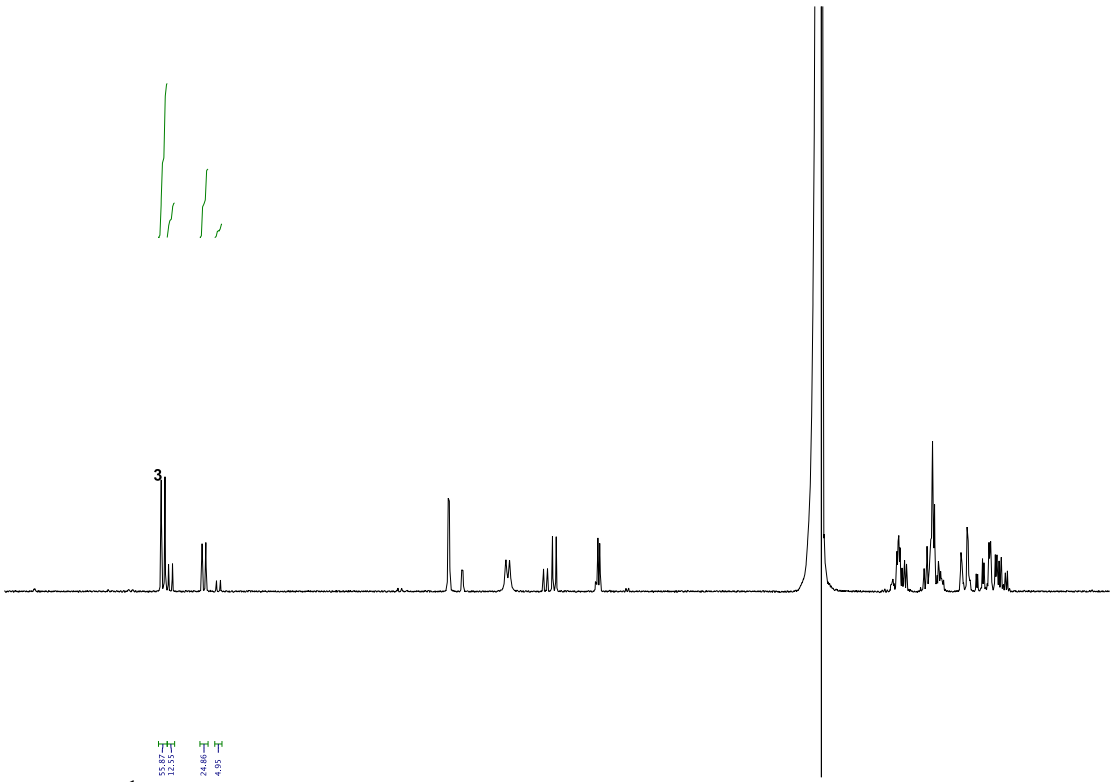


Fig. S37. ^{1}H NMR spectrum for the reaction of β-thiocytidine **3** with nitrous acid.

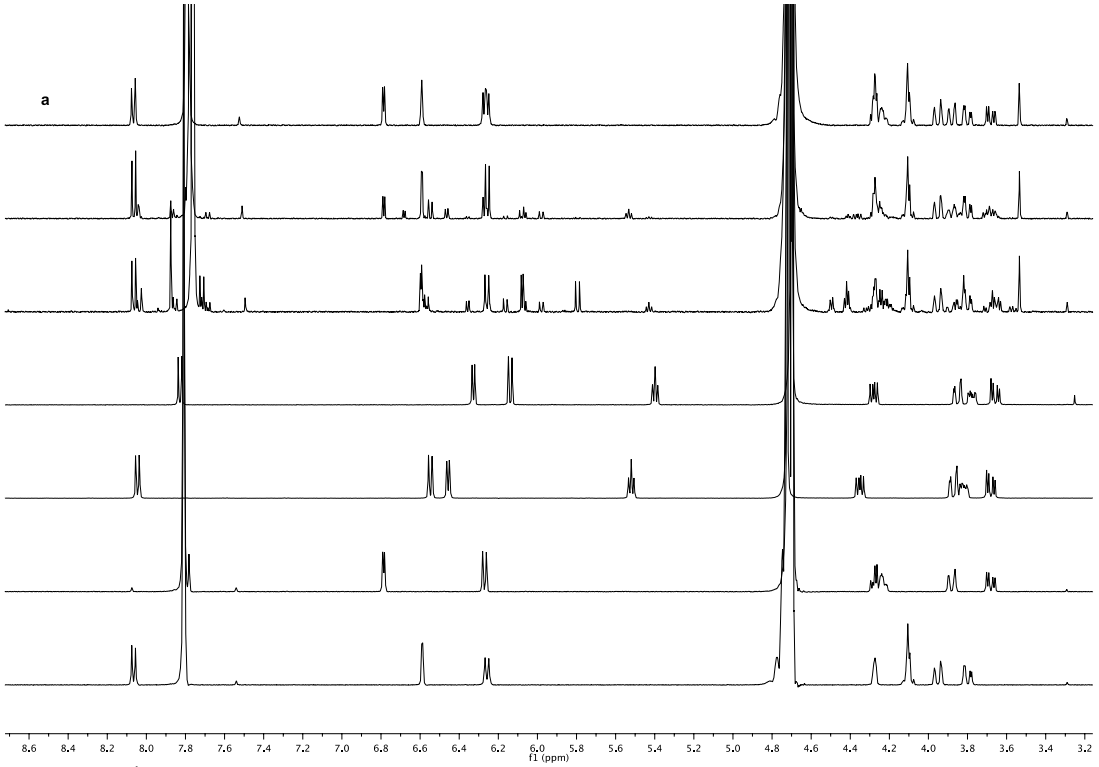


Fig. S38. ^{1}H NMR spectra for hydrolysis of α-thiocytidine **13** and β-thiocytidine **3** (1:1 mixture) in 0.1 M 4,5-dicyanoimidazole (DCI) buffer (pH 5.2) at 60 °C. a) ^{1}H NMR spectrum of **13** and **3** (1:1); b) ^{1}H NMR spectrum of the mixture after 4 days of hydrolysis; c) ^{1}H NMR spectrum of the mixture after 16 days of hydrolysis; d) ^{1}H NMR spectrum of **15**; e) ^{1}H NMR spectrum of **12**; f) ^{1}H NMR spectrum of **13**; g) ^{1}H NMR spectrum of **3**.

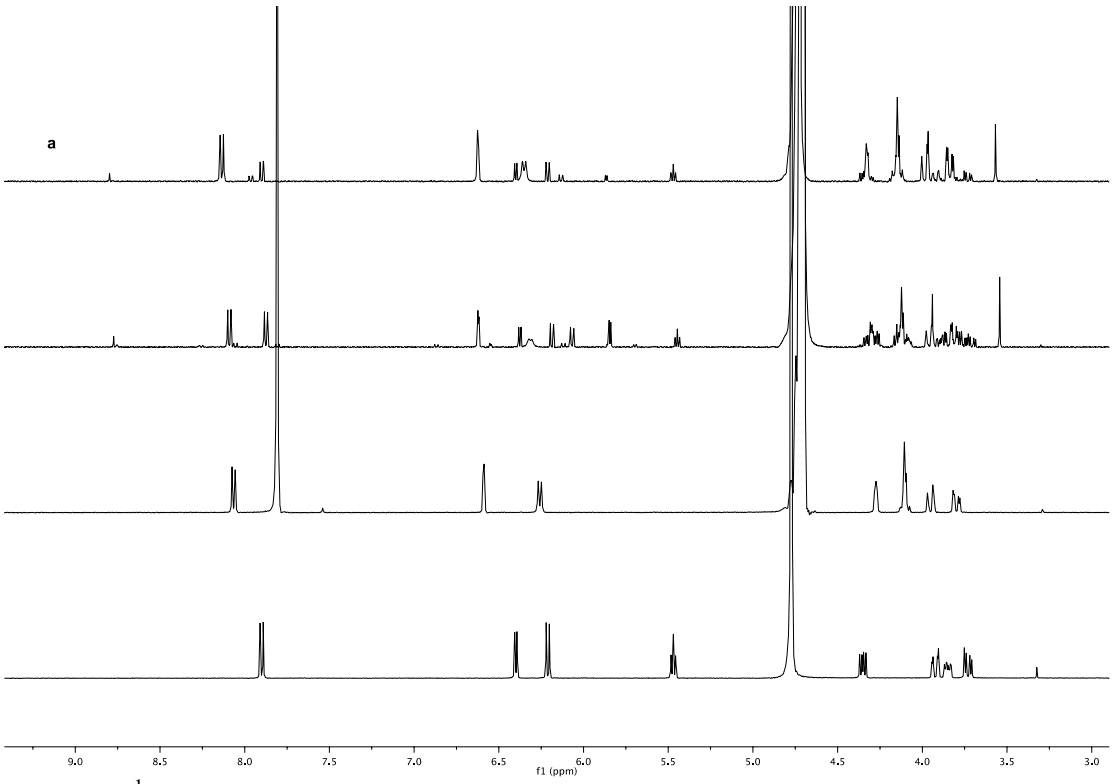


Fig. S39. ^{1}H NMR spectra for the reaction of α-anhydrouridine **15** and β-thiocytidine **3** (1:3 mixture) with nitrous acid. a) ^{1}H NMR spectrum of **15** and **3** (3:1) with nitrous acid (0 time point); b) ^{1}H NMR spectrum of the mixture after 24 hours; c) ^{1}H NMR spectrum of **3**; d) ^{1}H NMR spectrum of **15**.

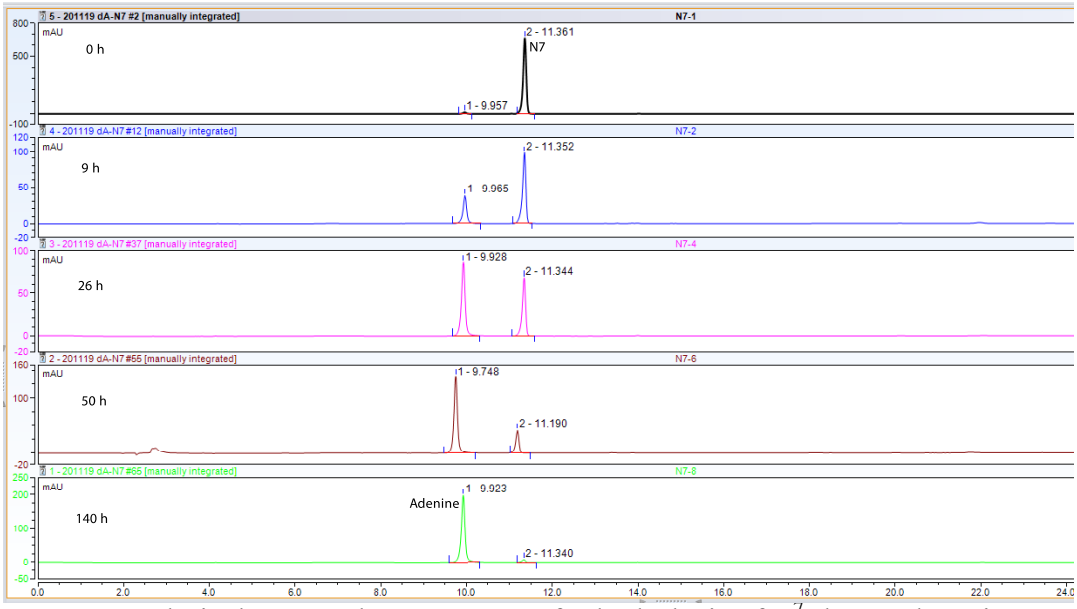


Fig. S40. Analytical HPLC chromatograms for hydrolysis of N^7-deoxyadenosine **20** in 1M NaOAc/HOAc buffer (pH = 4).

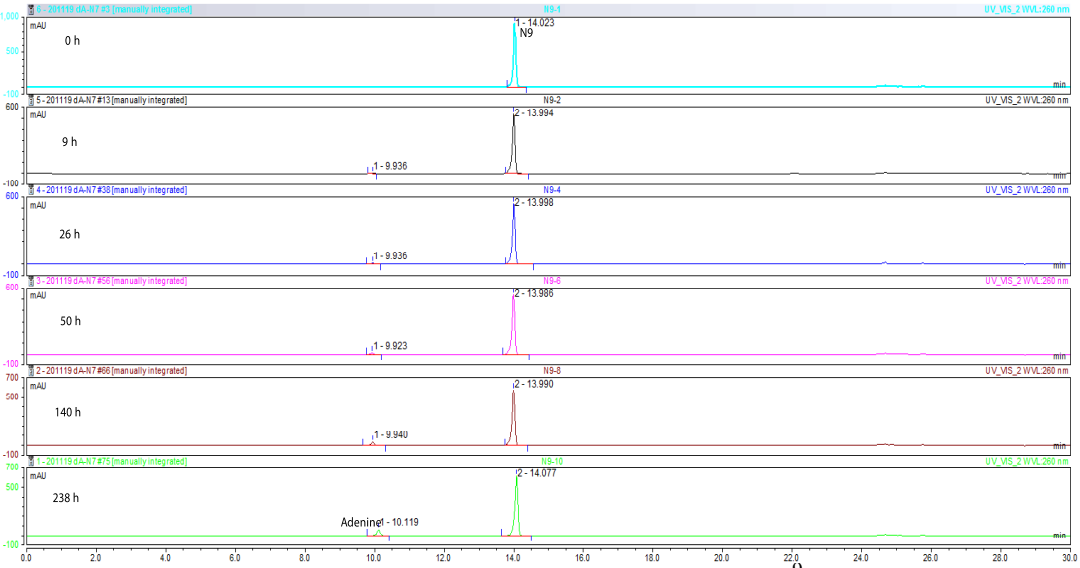


Fig. S41. Analytical HPLC chromatograms for hydrolysis of N^9-deoxyadenosine **7** in 1M NaOAc/HOAc buffer (pH = 4).

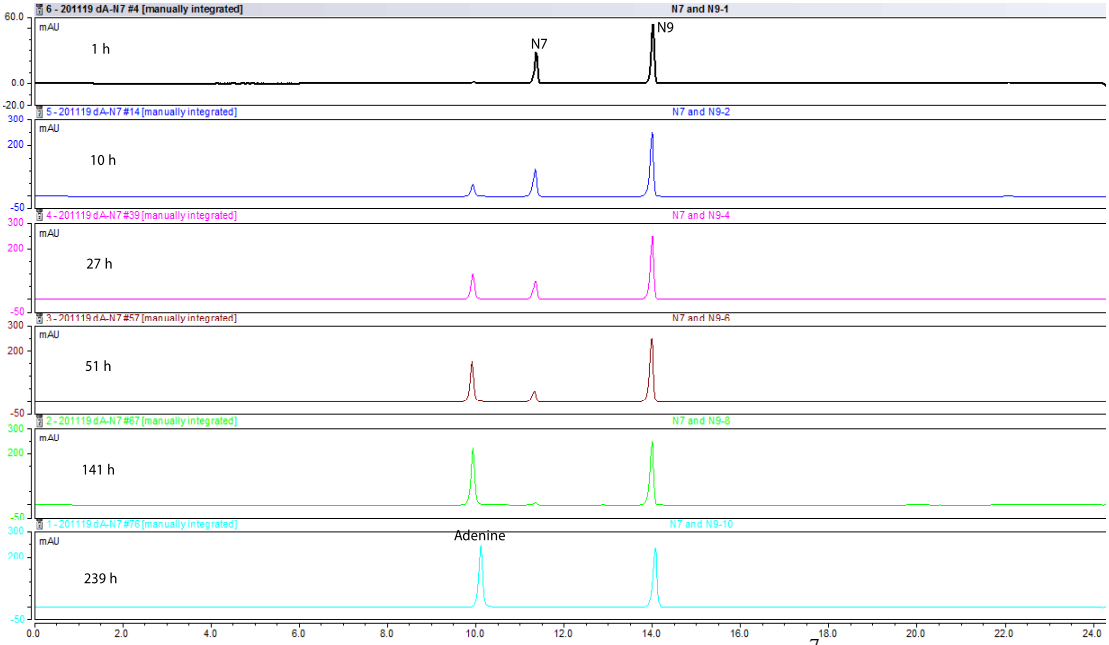


Fig. S42. Analytical HPLC chromatograms for hydrolysis of N^7-deoxyadenosine **20** and N^9-deoxyadenosine **7** (mixture) in 1M NaOAc/HOAc buffer (pH = 4).

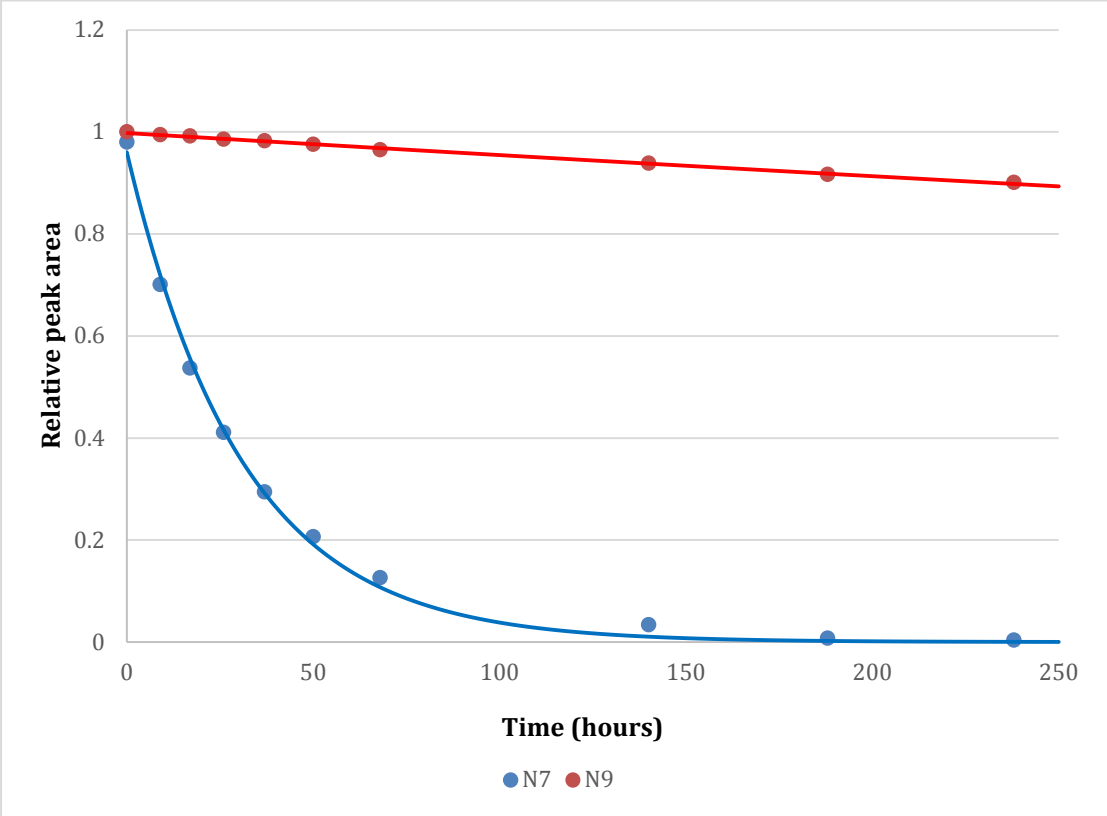


Fig. S43. Hydrolysis curves for N^7-deoxyadenosine **20** (in blue) and N^9-deoxyadenosine **7** (in red) in 1M NaOAc/HOAc buffer (pH = 4). Half lives for N^7-deoxyadenosine **20** and N^9-deoxyadenosine **7** are 21.5 hours and 1565 hours, respectively.

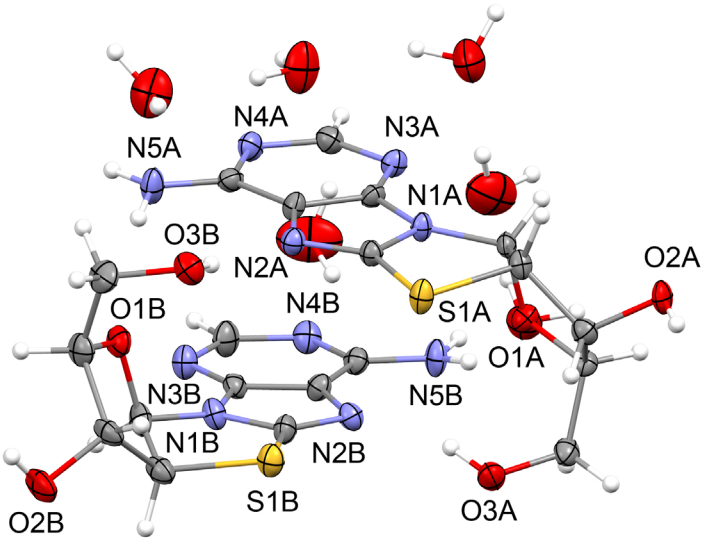


Fig. S44. Molecular structure of N^9-8, 2'-anhydro-thioadenosine **18** showing displacement ellipsoids at 50% probability for non-H atoms. The two independent molecules in the asymmetric unit have essentially identical geometry (RMSD = 0.043 Å).

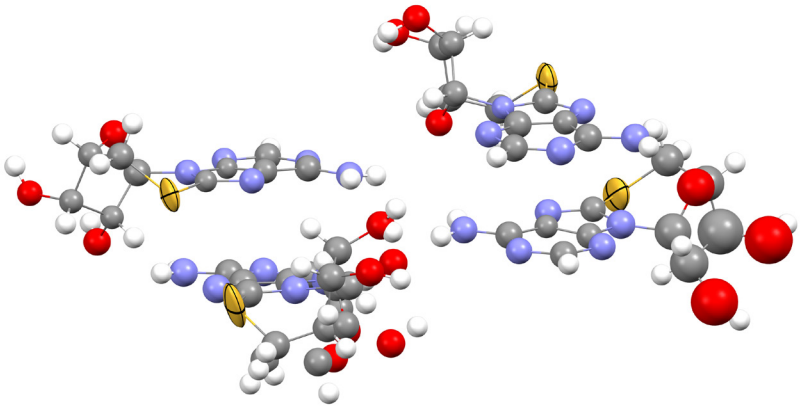


Fig. S45. Molecular structure of N^9-8, 5'-anhydro-thioadenosine **37** showing displacement ellipsoids at 50% probability for non-H atoms. Additional lattice water molecules are omitted for clarity.

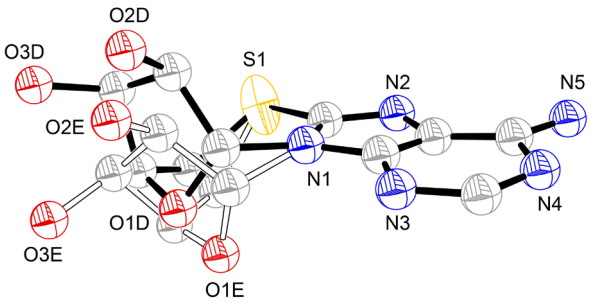


Fig. S46. Disordered molecule in N^9-8, 5'-anhydro-thioadenosine **37**. H atoms omitted. The adenine group is common to both disorder components.

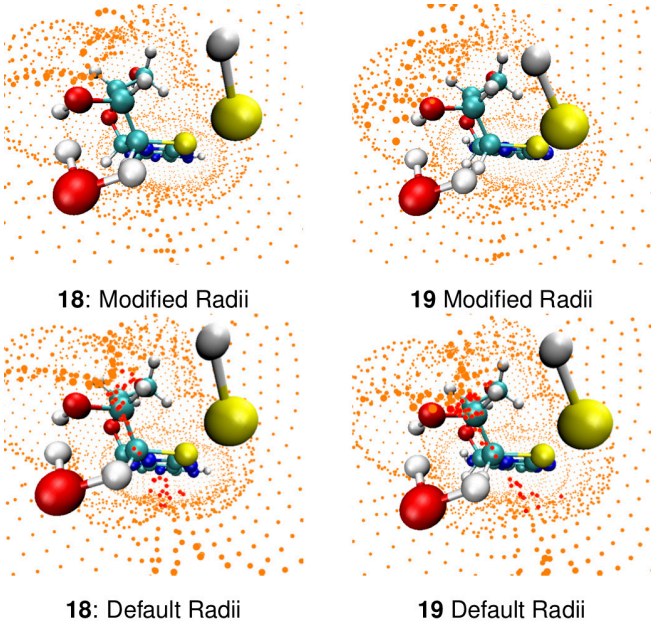


Fig. S47. The equilibrium ground-state structures of **18** and **19** complexes containing an explicit water molecule and hydrosulfide anion, located assuming PCM model of bulk water. The orange dots, which size represents the distance from the observer, indicate centering of surface charge basis functions (SCBFs).

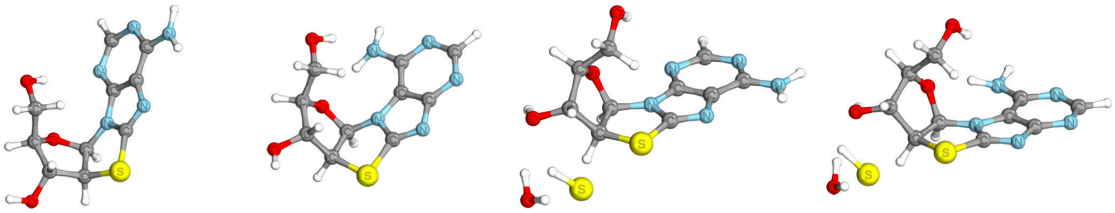


Fig. S48. The equilibrium ground-state structures of **18** and **19** and their complexes with an explicit water molecule and hydrosulfide anion located in PCM model of bulk water.

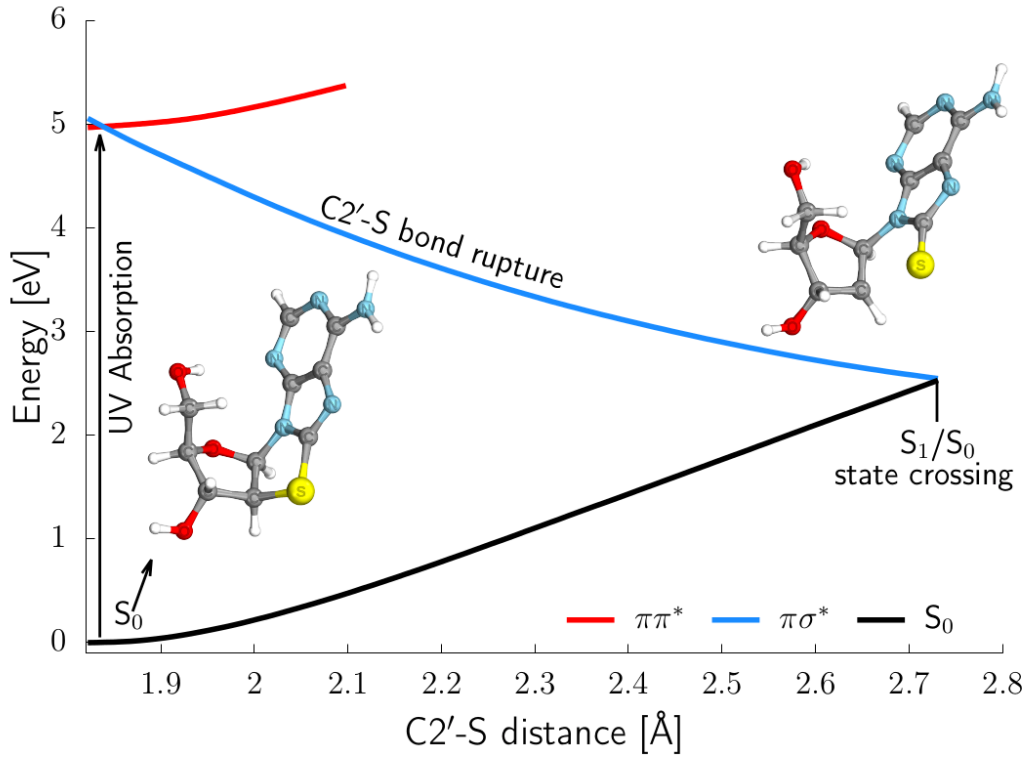


Fig. S49. The potential energy (PE) profiles of the UV-excited **18** present a barrierless deactivation pathway from the Franck-Condon region of the chromophore to the $S_1(\pi\sigma^*)/S_0$ minimum-energy crossing point. This photochemical pathway leads to the C2'-S bond rupture in the vicinity of the S_1/S_0 state crossing (2.55 eV) which can be reached in barrierless manner from the vertical excitation at 5.0 eV. The shown PE profiles were obtained based on the image dependent pair potential (IDPP) interpolation between the equilibrium ground-state structure and the S_1/S_0 minimum-energy crossing point. Energies were computed using the MP2/ADC(2)/ma-def2-TZVP level of theory.

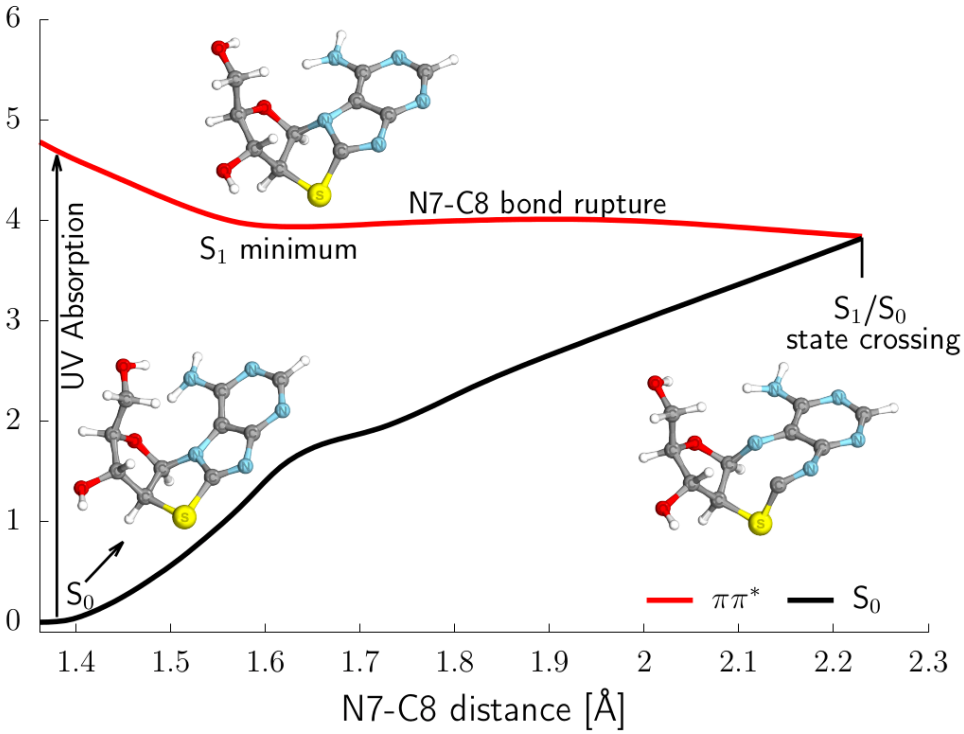


Fig. S50. The potential energy (PE) cuts of the photoexcited **19** show a radiationless relaxation channel through the $S_1(\pi\pi^*)$ minimum-energy structure, in which the sulfur atom is in the out-of-plane position in relation to the base moiety, to $S_1(\pi\pi^*)/S_0$ minimum-energy crossing point (MECP). In the S_1/S_0 state crossing the N7-C8 bond is ruptured which allows deactivation of UV-excited **19**. The presented PE profiles indicate also a small energy barrier (<0.1 eV) which could occur between the S_1 minimum and the S_1/S_0 MECP. The PE cuts were obtained from the image dependent pair potential (IDPP) interpolation between key minimum-energy structures. All energies were computed using the MP2/ADC(2)/ma-def2-TZVP level of theory.

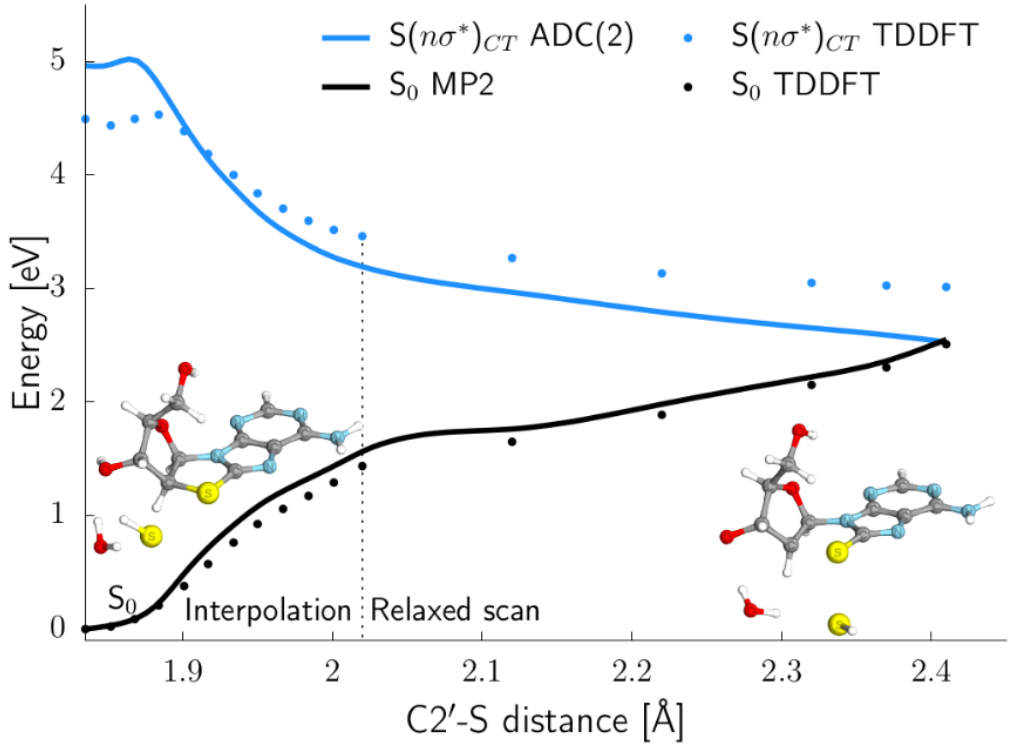


Fig. S51. The first part of the potential energy (PE) profiles of the UV-excited **18** complex is built of the image dependent pair potential (IDPP) interpolation[67] between the minimum-energy ground-state structure and the constrained S_1 geometry which is marked by the C2'-S distance equal to 2.02 Å. The second part of the PE profiles was obtained from the relaxed scan which was performed through the elongation of the C2'-S bond in the range of 2.02-2.41 Å. The TD-ωB97X-D/IEFPCM/ma-def2-TZVP method was employed to compute the relaxed scan and TDDFT single-point energies which are presented with dots. The obtained TDDFT geometries were used to compute excited and ground state energies at the MP2/ADC(2)/COSMO/ma-def2-TZVP level of theory and these results are shown as lines.

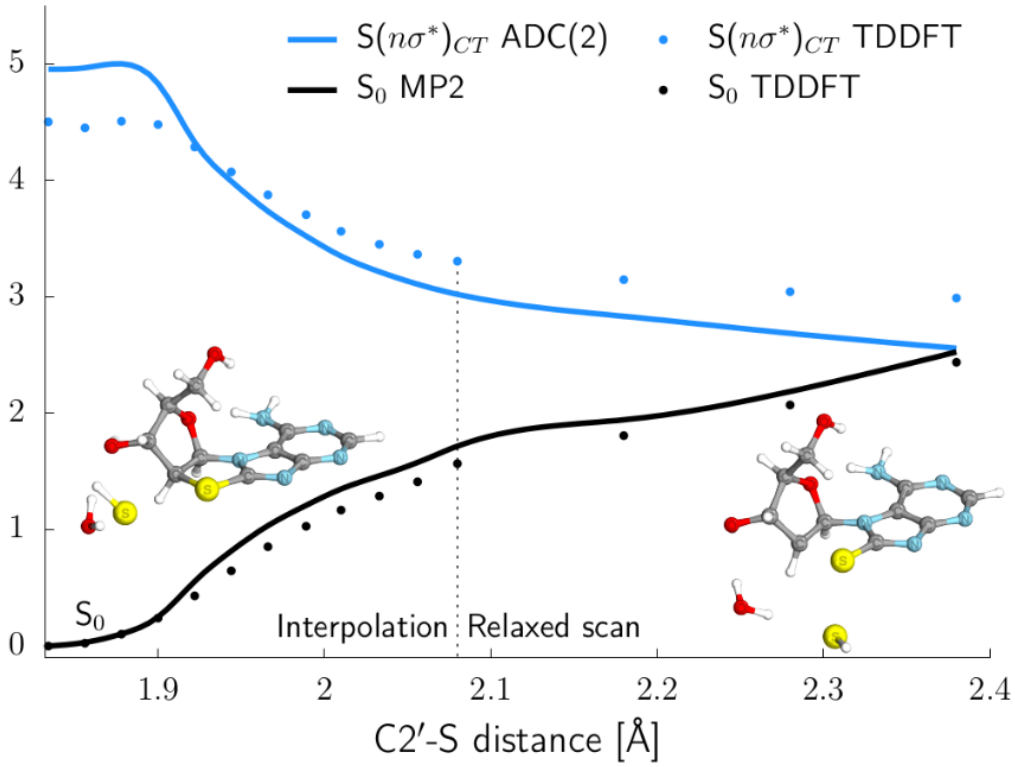


Fig. S52. The potential energy (PE) cuts of the UV-excited **19** complex are constructed of the image dependent pair potential (IDPP) interpolation[67] between the equilibrium ground-state structure and the constrained S_1 geometry which is characterized by the C2'-S distance equal to 2.08 Å. The right part of the PE profiles was obtained from the relaxed scan which was performed by the elongation of the C2'-S bond in the range of 2.08-2.38 Å. The presented single-point energies (dots) along with the relaxed scan were computed using the TD-ωB97X-D/IEFPCM/ma-def2-TZVP method. On the TDDFT geometries, single-point energies were calculated using the MP2/ADC(2)/COSMO/ma-def2-TZVP level and the results are shown as lines.

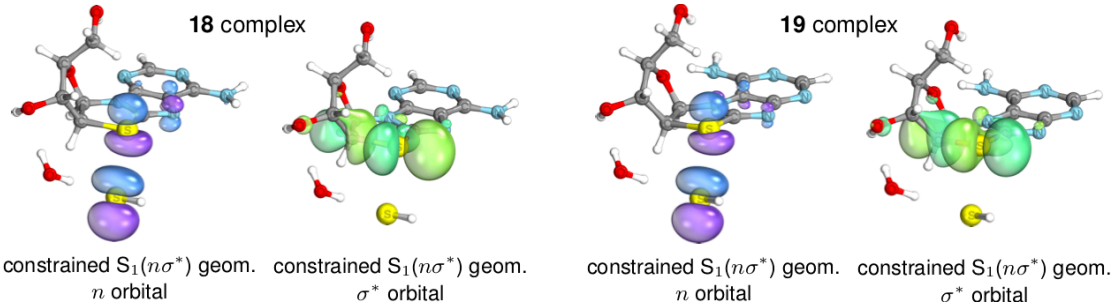


Fig. S53. The presented constrained S_1(nσ*) minimum-energy structures were found at a C2'-S distance equal to 2.02 and 2.08 Å for **18** and **19** complexes, respectively. The shown molecular orbitals depict a character of ^{1}nσ* excitation for both model systems. The constrained optimization was performed at the TD-ωB97X-D/IEFPCM/ma-def2-TZVP level of theory.

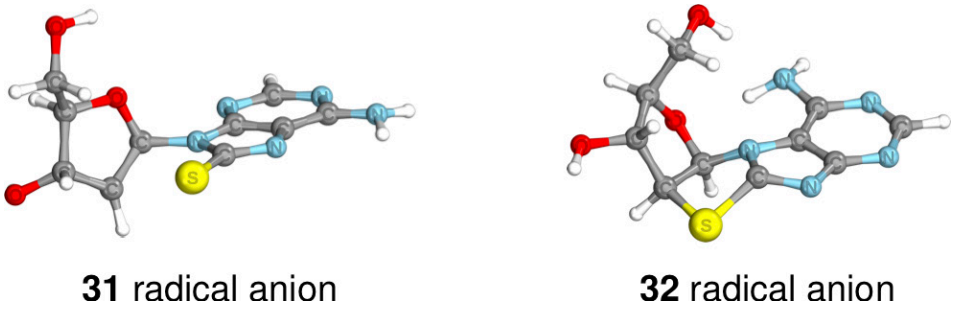


Fig. S54. The presented equilibrium ground-state geometries of **31** and **32** radical anions were found at the ωB97X-D/IEFPCM/ma-def2-TZVP level of theory. The ground-state structure of **31** radical anion is characterized by the C2'-S bond rupture. In turn, the performed geometry optimization of **32** radical anion resulted in an out-of-plane movement of the sulfur atom in relation to the purine moiety.

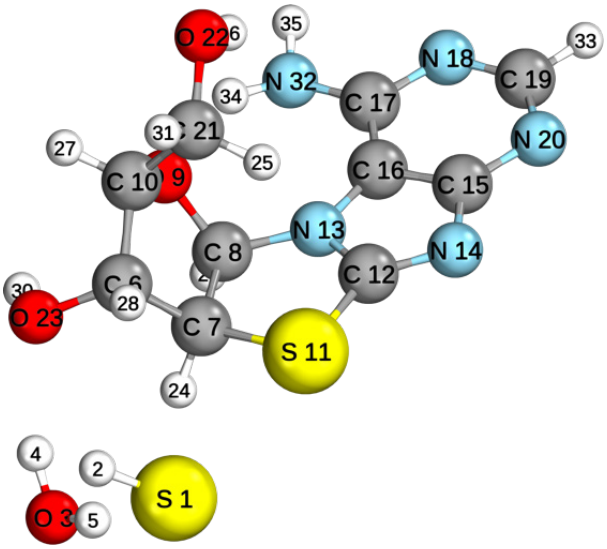


Fig. S55. Geometry of the studied **19** trimer with atomic labels.

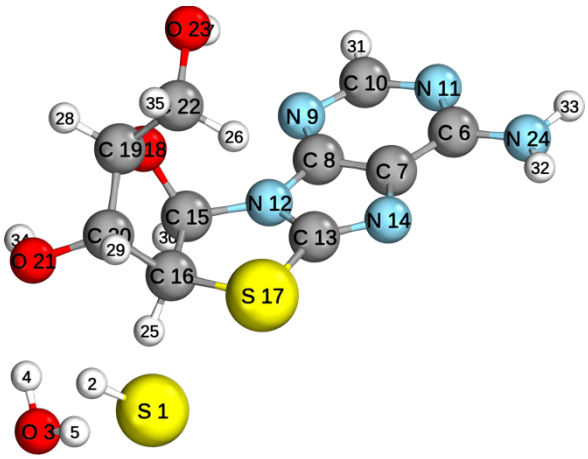


Fig. S56. Geometry of the studied **18** trimer with atomic labels.

Supplementary Tables

Entry	Initial Ratio Adenine:TAP (hot)	Ratio of supernatant (cool)
1	~99.7:0.3	~95:5
2	98:2	~60:40
3	86:14	24:76

Table S1. Ratios of Adenine to TAP before and after cooling and precipitation.

Entry	Reagent	pH	Temperature, Time	TAP : 8-mercaptoadenine
1	Thiourea (5.0 equiv.)	8.8 (not modified)	150 °C, 16 hrs	0:100
2	Thiourea (5.0 equiv.)	7.0	150 °C, 16 hrs	0:100
3	Thiourea (5.0 equiv.)	5.0	150 °C, 16 hrs	0:100
4	Thiourea (5.0 equiv.)	9.0	150 °C, 16 hrs	0:100
5	Thiourea (2.0 equiv.)	7.5	120 °C, 16 hrs	~30:70
6	NH_4SCN (2.0 equiv)	7.5	120 °C, 16 hrs	5:95
7	Thiourea (2.0 equiv.)	7.5	110 °C, 40 hrs	0:100
8	NH_4SCN (2.0 equiv)	7.5	120 °C, 40 hrs	0:100
9	Thiourea (2.0 equiv.)	7.0	100 °C, 66 hrs	0:100
10	NH_4SCN (2.0 equiv)	7.0	100 °C, 40 hrs	0:100

Table S2. Summary of dry state reaction outcomes between TAP **21** and thiourea **25** or NH_4SCN **24**.

Thioanhydronucleosides	**18**		**19**		**18** and **19** (mixture)	
	HS^-	HSO_3^-	HS^-	HSO_3^-	HS^-	HSO_3^-
Yields for deoxynucleosides	54%	74%	23%	0%	31% for N^9 isomer; 17% for N^7 isomer	62% for N^9 isomer; 0% for N^7 isomer

Table S3. Yields of deoxynucleosides from photoreduction with hydrogen sulfide or bisulfite.

Compound	**N^9-8, 2'-anhydro-thioadenosine 18**
CCDC Deposition No.	1970701
Cambridge Data No.	AB_B1_0018
Formula	$C_{10}H_{11}N_5O_3S$, $3(H_2O)$
M.w.	281.29
Crystal system	trigonal
Space group (IT no.)	$P\,3_2$ (145)
a [Å]	19.0264 (5)
b [Å]	19.0264 (5)
c [Å]	6.8590 (2)
α [°]	90
β [°]	90
γ [°]	120
Z	6
V [Å^3]	2150.33 (13)
D_x [g cm^{-3}]	1.554
Crystal size [mm]	0.60 x 0.04 x 0.04
Crystal colour, shape	colourless needle
μ [mm^{-1}]	2.391
T_{min},T_{max}	0.600, 0.753
Measured reflections	34196
Independent reflections (R_{int}[a])	4940 (0.072)
Observed reflections [I>2σ(I)]	4710
No. of parameters / restraints	398 / 1
$R1$ (observed data)[b]	0.050
$wR2(F^2)$ for all data	0.134
GOF[c]	1.08
Residual electron density [e·Å^{-3}]	–0.37, 0.40
Flack parameter	0.01 (2)

[a] $R_{int} = \Sigma \left| F_o^2 - F_{o,mean}^2 \right| / \Sigma F_o^2$

[b] $R(F) = \Sigma \left| \left| F_o \right| - \left| F_c \right| \right| / \Sigma \left| F_o \right|$; $wR(F^2) = [\Sigma(w(F_o^2 - F_c^2)^2)/(\Sigma w(F_o^2)^2)]^{½}$

[c] $\mathrm{GOF} = [\Sigma(w(F_o^2 - F_c^2)^2)/(N_{diffrs} - N_{params})]^{½}$

Table S4. Crystal data, data collection, and refinement parameters for N^9-8, 2'-anhydro-thioadenosine **18** (grown in H_2O).

Compound	**N^9-8, 5'-anhydro-thioadenosine 37**
CCDC Deposition No.	1970700
Cambridge Data No.	AB_B1_0016
Formula	$C_{10}H_{11}N_5O_3S$, 1.75(H_2O)
M.w.	281.29
Crystal system	monoclinic
Space group (IT no.)	*C*2 (5)
a [Å]	30.364 (4)
b [Å]	6.7022 (9)
c [Å]	25.496 (3)
α [°]	90
β [°]	93.364 (10)
γ [°]	90
Z	16
V [Å^3]	5179.6 (12)
D_x [g cm^{-3}]	1.605
Crystal size [mm]	0.20 x 0.04 x 0.01
Crystal colour, shape	colourless lath
μ [mm^{-1}]	2.529
T_{min},T_{max}	0.506, 0.750
Measured reflections	19497
Independent diffractions (R_{int}[a])	5089 (0.284)
Observed diffract. [I>2σ(I)]	2164
No. of parameters / restraints	365 / 561
*R*1 (observed data)[b]	0.153
*wR*2(F^2) for all data	0.378
GOF[c]	1.04
Residual electron density [e·Å^{-3}]	–0.47, 0.41
Flack parameter	0.07 (6)

[a] $R_{int} = \Sigma \left| F_o^2 - F_{o,mean}^2 \right| / \Sigma F_o^2$

[b] $R(F) = \Sigma \left| \left| F_o \right| - \left| F_c \right| \right| / \Sigma \left| F_o \right|$; $wR(F^2) = [\Sigma(w(F_o^2 - F_c^2)^2)/(\Sigma w(F_o^2)^2)]^{½}$

[c] GOF = $[\Sigma(w(F_o^2 - F_c^2)^2)/(N_{diffrs} - N_{params})]^{½}$

Table S5. Crystal data, data collection, and refinement parameters for N^9-8, 5'-anhydro-thioadenosine **37** (grown in H_2O).

State / Transition		E_{exc} [eV]	f_{osc}	λ [nm]
		18		
S_1	ππ*	4.969	0.309	249.5
S_2	πσ*	5.057	0.005	245.2
S_3	ππ*	5.064	0.158	244.8
S_4	πσ*	5.492	0.000	225.8
		19		
S_1	ππ*	4.783	0.123	259.2
S_2	nπ*	4.879	0.004	254.1
S_3	πσ*	5.120	0.001	242.2
S_4	ππ*	5.334	0.034	232.4

Table S6. Vertical excitations energies (in eV) of isolated **18** and **19** were obtained at the ADC(2)/ma-def2-TZVP level.

State / Transition		E_{exc} [eV]	f_{osc}	CT [e^-]	λ [nm]
		18 complex			
S_1	ππ* (nπ*)	4.820 (4.494)	0.617 (0.264)	0.057 (0.574)	257.2 (275.9)
S_2	ππ* (nπ*)	4.955 (4.631)	0.108 (0.110)	0.059 (0.546)	250.2 (267.7)
S_3	nσ* (ππ*)	4.976 (4.978)	0.095 (0.517)	0.495 (0.272)	249.2 (249.1)
S_4	πσ* (ππ*)	5.070 (5.060)	0.021 (0.124)	0.167 (0.490)	244.5 (245.0)
		19 complex			
S_1	ππ* (nσ*)	4.653 (4.503)	0.269 (0.272)	0.029 (0.558)	266.5 (275.3)
S_2	nσ* (nσ*)	4.969 (4.640)	0.218(0.113)	0.513 (0.565)	249.5 (267.2)
S_3	nσ* (ππ*)	5.075 (4.898)	0.116(0.268)	0.506 (0.078)	244.3 (253.1)
S_4	ππ* (nπ*)	5.199 (5.012)	0.004 (0.017)	0.141 (0.756)	238.5 (247.7)

Table S7. Vertical excitation energies (in eV) of **18** and **19** complexes obtained at the ADC(2)/COSMO/ma-def2-TZVP level of theory. In parentheses, the results computed using the TD-ωB97X-D/IEFPCM/ma-def2-TZVP method are shown.

BCP(A-B)	q_A/e	q_B/e	ρ	$\nabla^2\rho$	V x 10^3	G x 10^3	H x 10^3	R_{AB} in A
S_2X^3			0.20547					1.889
S_2^{-2}			0.10481					2.03 - 2.36
19 ···HS^-···H_2O (**19** complex, gas phase data)								
Sl-Sll	-0.519	0.462	0.0113	0.0263	-5.12	5.85	0.73	3.410
S1-H5	-0.519	0.619	0.0283	0.0579	-15.93	15.20	-0.73	2.217
O23-H4	-1.126	0.600	0.0225	0.0761	-17.02	18.02	1.00	2.004
O3-H24	-1.312	0.089	0.0076	0.0261	-4.92	5.72	0.80	2.672
18 ···HS^-···H_2O (**18** complex, gas phase data)								
S1-S17	-0.525	0.436	0.0107	0.0249	-4.82	5.53	0.71	3.437
S1-H5	-0.525	0.618	0.0281	0.0578	-15.77	15.11	-0.66	2.219
O21-H4	-1.128	0.602	0.0229	0.0782	-17.31	18.43	1.12	1.994
O3-H25	-1.313	0.083	0.0070	0.0249	-4.45	5.34	0.89	2.719
19 ···HS^-··· H_2O (**19** complex, PCM bulk water data)								
Sl-Sll	-0.580	0.368	0.0108	0.0276	-5.23	6.06	0.83	3.410
S1-H5	-0.580	0.610	0.0280	0.0596	-15.98	15.44	-0.54	2.217
O23-H4	-1.154	0.627	0.0226	0.0745	-16.78	17.71	0.93	2.004
O3-H24	-1.331	0.078	0.0074	0.0263	-4.88	5.73	0.85	2.672
18···HS^- ···H_2O (**18** complex, PCM bulk water data)								
S1-S17	-0.581	0.345	0.0103	0.0261	-4.93	5.73	0.80	3.437
S1-H5	-0.581	0.610	0.0279	0.0594	-15.84	15.34	-0.50	2.219
O21-H4	-1.154	0.628	0.0230	0.0766	-17.08	18.12	1.04	1.994
O3-H25	-1.332	0.074	0.0069	0.0251	-4.40	5.34	0.94	2.719

Table S8. MP2/aug-cc-pVDZ generalized electron density properties evaluated at the bond critical points (BCPs) of the selected complexes. In the second and third column the charges on atoms connected via a particular bond path (indicated in the first column) are given. All the results were calculated for the equilibrium geometries located using ωB97X-D/IEFPCM/ma-def2-TZVP approach. The gas-phase data were obtained for the same geometries and are shown for the reference. All results are in a.u. Values for S_2 molecule and anion[77].

	$\varepsilon_{\mathrm{el}}^{(10)}$	$\Delta E_{\mathrm{ex},2}^{\mathrm{HL}}$	$\Delta E_{\mathrm{ex},3}^{\mathrm{HL}}$	$\Delta E_{\mathrm{del},2}^{\mathrm{HF}}$	$\Delta E_{\mathrm{del},3}^{\mathrm{HF}}$	$\Delta G_2^{\mathrm{HF,S}}$	$\Delta G_3^{\mathrm{HF,S}}$	ΔG^{HF}	$\varepsilon_{\mathrm{el}}^{(12)}$	$\varepsilon_{\mathrm{disp}}^{(20)}$	$\Delta E_{\mathrm{ex},2}^{(2)}$	$\Delta E_{\mathrm{ex},3}^{(2)}$	ΔG^{MP2}
19···HS^- ··H_2O (**19** complex, bulk water PCM)													
HS^- ·H_2O·**19**	-217.5	317.7	-2.7	-83.9	-25.5	41.0	-23.6	5.6	1.0	-113.1	49.0	-1.4	-58.9
HS^-·H_2O	-45.2	69.4	-	-19.6	-	12.9	-	17.6	-3.5	-23.0	10.7	-	1.8
H_2O·**19**	-32.1	33.4	-	-7.7	-	-0.3	-	-6.8	-2.1	-15.5	8.4	-	-16.0
HS^-·**19**	-140.2	215.0	-	-56.6	-	28.4	-	46.6	6.7	-74.6	29.9	-	8.6
18···HS^- ··H_2O (**18** complex, bulk water PCM)													
HS^-·H_2O·**18**	-206.0	308.3	-2.6	-79.6	-20.9	32.4	-24.8	6.8	-1.6	-111.9	50.3	-0.1	-56.5
HS^-·H_2O	-50.0	74.4	-	-20.2	-	12.4	-	16.6	-3.3	-23.9	11.5	-	0.9
H_2O·**18**	-33.2	33.4	-	-8.2	-	0.5	-	-7.5	-1.8	-15.4	8.2	-	-16.5
HS^-·**18**	-122.8	200.5	-	-51.3	-	19.5	-	46.0	3.6	-72.6	30.6	-	7.6
	$\varepsilon_{\mathrm{el}}^{(10)}$	$\Delta E_{\mathrm{ex},2}^{\mathrm{HL}}$	$\Delta E_{\mathrm{ex},3}^{\mathrm{HL}}$	$\Delta E_{\mathrm{del},2}^{\mathrm{HF}}$	$\Delta E_{\mathrm{del},3}^{\mathrm{HF}}$			ΔE^{HF}	$\varepsilon_{\mathrm{el}}^{(12)}$	$\varepsilon_{\mathrm{disp}}^{(20)}$	$\Delta E_{\mathrm{ex},2}^{(2)}$	$\Delta E_{\mathrm{ex},3}^{(2)}$	ΔE^{MP2}
19···HS^- ··H_2O (**19** complex, gas phase)													
HS^- H_2O·**19**	-245.0	340.5	-0.9	-129.2	12.8			-21.7	-27.1	-128.0	43.0	21.4	-112.4
HS^-·H_2O	-132.1	178.1	-	-57.7	-			-11.7	1.9	-48.2	22.8	-	-35.2
H_2O·**19**	-31.0	33.8	-	-9.4	-			-6.6	-2.2	-15.9	9.0	-	-15.7
HS^-·**19**	-81.9	128.6	-	-62.1	-			-15.3	-26.8	-63.9	11.1	-	-94.9
18···HS^- ··H_2O (**18** complex, gas phase data)													
HS^-·H_2O·**17**	-232.8	336.4	-0.9	-122.3	11.6			-8.0	-27.6	-127.1	53.4	18.9	-90.4
HS^-·H_2O	-132.0	177.7	-	-57.6	-			-11.9	1.9	-48.0	22.8	-	-35.2
H_2O·**17**	-32.0	33.9	-	-9.7	-			-7.9	-1.9	-15.8	9.0	-	-16.6
HS^-·**17**	-68.8	124.9	-	-55.0	-			1.1	-27.7	-63.2	21.7	-	-68.1

Table S9. Results of the many-body hybrid variational-perturbational interaction energy decomposition for the studied trimers. Results were calculated at the MP2/(IEFPCM)/aug-cc-pVDZ level for equilibrium geometries located using ωB97X-D/IEFPCM/ma-def2-TZP approach. The results obtained for each of the dimers within a trimer are also reported. The gas-phase data were obtained for the same geometries and are shown for the reference. All results are in kJ/mol.

References

51. Sheldrick, G. M. SHELXT – Integrated space-group and crystal-structure determination. *Acta Cryst*. **A71**, 3-8 (2015).

52. Sheldrick, G. M. Crystal structure refinement with SHELXL. *Acta Cryst*. **C71**, 3-8 (2015).

53. Gendron, F. *et al*. Novel inhibitors of nucleoside triphosphate diphosphohydrolases: chemical synthesis and biochemical and pharmacological characterizations. *J. Med. Chem.* **43**, 2239–2247 (2000).

54. Trevisiol, E., Renard, A. Defrancq, E. & Lhomme, J. Fluorescent labelling of oligodeoxyribonucleotides by the oxyamino-aldehyde coupling reaction. *Nucleosides Nucleotides Nucleic Acids*. **19**, 1427-1439 (2000).

55. Kazimierczuk, Z. & Seela, F. Synthesis of 5-amino-4-cyano-1-imidazolyl 2-deoxy-β-D-ribofuranoside by photolysis of 2-aza-2'-deoxyadenosine or glycosylation. *Liebigs Ann. Chem.* **1991**, 695-697 (1991).

56. Seela, F. & Winter, H. N^7- DNA: synthesis and base pairing of oligonucleotides containing 7- (2- deoxy- β - D- erythro- pentofuranosyl)adenine (N^7A_d). *Helv. Chim. Acta.* **77**, 597-607 (1994).

57. Parsons, S., Flack, H. D. & Wagner, T. Use of intensity quotients and differences inabsolute structure refinement. *Acta Cryst*. **B69**, 249-259 (2013).

58. Chai, J.-D. & Head-Gordon, M. Long-range corrected hybrid density functionals with damped atom–atom dispersion corrections. *Phys. Chem. Chem. Phys.* **10**, 6615-6620 (2008).

59. Zheng, J., Xu, X. & Truhlar, D. G. Minimally augmented Karlsruhe basis sets. *Theor. Chem. Acc.* **128**, 295–305 (2011).

60. 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., Li, X., Caricato, M., Marenich, A. V., Bloino, J., Janesko, B. G., Gomperts, R., Mennucci, B., Hratchian, H. P., Ortiz, J. V., Izmaylov, A. F., Sonnenberg, J. L., Williams-Young, D., Ding, F., Lipparini, F., Egidi, F., Goings, J., Peng, B., Petrone, A., Henderson, T., Ranasinghe, D., Zakrzewski, V. G., Gao, J., Rega, N., Zheng, G., Liang, W., Hada, M., Ehara, M., Toyota, K., Fukuda, R., Hasegawa, J., Ishida, M., Nakajima, T., Honda, Y., Kitao, O., Nakai, H., Vreven, T., Throssell, K., Montgomery, J. A., Peralta, J. E., Ogliaro, F., Bearpark, M. J., Heyd, J. J., Brothers, E. N., Kudin, K. N., Staroverov, V. N., Keith, T. A., Kobayashi, R., Normand, J., Raghavachari, K., Rendell, A. P., Burant, J. C., Iyengar, S. S., Tomasi, J., Cossi, M., Millam, J. M., Klene, M., Adamo, C., Cammi, R., Ochterski, J. W., Martin, R. L., Morokuma, K., Farkas, O., Foresman, J. B., Fox, D. J. Gaussian 16, Revision B.01, Gaussian, Inc., Wallingford CT (2016).

61. Scalmani, G. & Frisch, M. J. Continuous surface charge polarizable continuum models of solvation. I. General formalism. *J. Chem. Phys.* **132**, 114110-114115 (2010).

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

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

64. Plasser, F. "TheoDORE 1.7.2: A package for theoretical density, orbital relaxation, and exciton analysis". http://theodore-qc.sourceforge.net (2019).

65. 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*, **112**, 405–413 (2008).

66. TURBOMOLE V7.3 2018, a development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989-2007, TURBOMOLE GmbH, since 2007; available from http://www.turbomole.com.

67. Smidstrup, S., Pedersen, A., Stokbro, K. & Jónsson, H. Improved initial guess for minimum energy path calculations. *J. Chem. Phys*. **140**, 214106 (2014).

68. T. A. Keith, AIMAll, 2011, http:/ /aim.tkgristmill.com.

69. R. W. Góra, EDS package, revision 2.8.4, 1998, Wroclaw University of Science and Technology.

70. Schmidt, M. W. *et al.* General atomic and molecular electronic structure system. *J. Comput. Chem.* **14**, 1347-1363 (1993).

71. Handy, N. C. & Schaefer, H. F. On the evaluation of analytic energy derivatives for correlated wave functions. *J. Chem. Phys.* **81**, 5031-5033(1984).

72. Góra, R. W., Bartkowiak, W., Roszak, S. & Leszczyński, J. Intermolecular interactions in solution: Elucidating the influence of the solvent. *J. Chem. Phys.* **120**, 2802-2813 (2004).

73. Hirschfelder, J. Perturbation theory for exchange forces, I. *Chem. Phys. Lett*. **1**, 325-329 (1967).

74. Moszyński, R., Rybak, S., Cybulski, S. & Chałasiński, G. Correlation correction to the Hartree-Fock electrostatic energy including orbital relaxation. *Chem. Phys. Lett*. **166**, 609-614 (1990).

75. Boys, S. & Bemardi, F. The calculation of small molecular interactions by the

differences of separate total energies. Some procedures with reduced errors. *Mol. Phys.* **19**, 553-566 (1970).

76. Wells, B. H. & Wilson, S. Van der Waals interaction potentials: Many-body basis set superposition effects. *Chem. Phys. Lett.* **101**, 429-434 (1983).

77. Knop, O., Boyd, R. J. & Choi, S. C. Sulfur-sulfur bond lengths, or can a bond length be estimated from a single parameter? *J. Am. Chem. Soc.* **110**, 7299-7301 (1988).