The Transcription Factor CarH Safeguards Use of Adenosylcobalamin as a Light Sensor by Altering the Photolysis Products

Supporting Information

Marco Jost[†], Jeffrey H. Simpson[†], and Catherine L. Drennan*[,†,‡]

[†]Department of Chemistry and [‡]Department of Biology and Howard Hughes Medical Institute, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139, USA

Materials and Methods

Sample preparation

A vector encoding *Thermus thermophilus* CarH, pET15b-CarH, was a generous gift from S. Padmanabhan (Instituto de Química Física "Rocasolano", Consejo Superior de Investigaciones Científicas, Madrid, Spain). AdoCbl-bound (holo) CarH was purified as described previously[1]. AdoCbl (Sigma) for photolysis experiments was dissolved in CarH buffer (100 mM NaCl, 50 mM Tris, pH 8). All AdoCbl-containing samples were handled in a dark room under red light until photolysis experiments.

Photolysis and spectroscopic analysis

For experiments under aerobic conditions, samples of either AdoCbl or holo CarH in CarH buffer were exposed to air for 2 hrs in the dark prior to photolysis to ensure the presence of oxygen. For photolysis, samples were transferred into transparent tubes and exposed to ambient light for 1 hr on ice.

For experiments under anaerobic conditions, photolysis was performed in an MBraun chamber under an atmosphere of 100% N_2 at 20 °C. Buffers and solutions were de-gassed by sparging with Ar prior to transfer into the MBraun chamber. Protein samples were transferred while frozen and allowed to de-gas within the chamber over the course of 4-12 hrs in the dark. Because photolysis of alkylcobalamins is slower under anaerobic conditions (see for example ref. [2]), samples were exposed to ambient light for 2 hrs (CarH) or 4-8 hrs (AdoCbl).

UV-Vis spectra after photolysis under aerobic conditions were recorded at 25 °C on a SpectraMax Plus 384 (Molecular Devices) using SoftMax Pro 5 software (Molecular Devices) and a 1 cm path length quartz cuvette (Starna). The spectrum of pure CarH buffer was used for

background subtraction. For UV-Vis spectroscopy after photolysis under anaerobic conditions, 120 µM holo CarH (4.0 mg/mL) or 120 µM AdoCbl in CarH buffer were photolyzed as described above. The samples were transferred to 1 cm path length quartz cuvettes (Starna) under anaerobic conditions. The cuvettes were sealed with septa and removed from the anaerobic environment, and spectra were collected as described above. The resulting spectra had the same features as UV-Vis spectra recorded within the MBraun anaerobic chamber on a NanoDrop 2000c (Thermo Scientific), but were of higher quality. After acquisition of the spectra, the samples were removed from the cuvettes and exposed to molecular oxygen for 20 min before another set of UV-Vis spectra were collected using the same experimental parameters.

For EPR spectroscopy, 400 µL of 240 µM holo CarH (8.0 mg/mL) or 400 µL of 1 mM free AdoCbl in CarH buffer supplemented with 5% (v/v) glycerol were photolyzed under anaerobic conditions as described above. The resulting samples were transferred to 3.8 mm thin wall precision quartz EPR tubes (Wilmad-LabGlass). The tubes were sealed with septa and removed from the anaerobic environment, and the samples were immediately flash-frozen in liquid nitrogen. Continuous wave X-band EPR spectra were recorded at 77 K in the MIT Department of Chemistry Instrumentation Facility on a Bruker ESP-300 X-band spectrometer equipped with a quartz finger Dewar filled with liquid nitrogen. Experimental conditions were as follows: microwave frequency, 9.45GHz; modulation amplitude, 5.00 G; modulation frequency, 100 kHz; time constant, 5.12 ms; and scan time, 41.9 s. A microwave power of 1 mW and an average of 100 scans were used.

LC-MS analysis

Samples were prepared and photolyzed as described above, under either aerobic or anaerobic conditions. After photolysis, the photolysis product was separated from CarH and CarH-bound cobalamin species using a 10 kDa molecular weight cutoff centrifugal concentrator. AdoCbl samples were used directly. The resulting samples were placed in HPLC vials under aerobic or anaerobic conditions, sealed, and then analyzed by LC-MS on an Agilent 1100 LC-MSD instrument (Agilent Technologies). The chromatography separation was carried out on a 150 × 3.0 mm Poroshell 120 SB-C_{18} (2.7 µm pore size) reversed-phase column (Agilent Technologies) at 25 °C and a flow of 200 µL min^{-1}. The solvent system consisted of H_2O + 0.1% (v/v) acetic acid (solvent A) and CH_3CN (solvent B). The following elution program was used: isocratic 0% B for 7 min, linear gradient from 0% to 95% B over 28 min, 100% B for 20 min, re-equilibration at 0% B for 15 min. Absorbance was monitored continuously at 260 nm, 350 nm, and 522 nm. An in-line ESI-TOF single quadrupole mass spectrometer (Agilent Technologies) in positive ion mode was used to determine mass-to-charge ratios (*m*/*z*) with the following settings: ESI capillary voltage, 3000 V; gas temperature, 350 °C; drying gas flow, 11 L/min; nebulizer pressure, 25 psi; fragmentor voltage, 70 V; *m*/*z* scan range, 120-1500. For treatment with trifluoroacetic acid (TFA), TFA was added to photolyzed samples to a final concentration of 8% (v/v) and the samples were incubated on ice for 10 min. For treatment with methanol, methanol was added to photolyzed samples to a final concentration of 70% (v/v) and the samples were shaken for 1 min. All samples were then centrifuged at 15000 × *g* for 10 min to remove precipitated protein and the supernatant was used for LC-MS experiments. All LC-MS experiments were performed at the MIT Center for Environmental Health Sciences Bioanalytical Sciences core facility.

Purification of CarH photolysis product

For purification of the CarH photolysis product, holo CarH was purified as described previously[1], with the following modifications: cell pellets from 2 L of culture were resuspended in 70 mL lysis buffer (50 mM sodium phosphate, pH 8, 300 mM NaCl, 1 mM PMSF, 1 mM benzamidine-HCl, supplemented with 2 cOmplete protease inhibitor tablets (Roche) and 200 U benzonase nuclease (EMD Millipore)). Ni-NTA chromatography was performed on a 5 mL HisTrap HP Ni-NTA column (GE Healthcare) and the wash step after binding contained 40 mM imidazole. The final size exclusion chromatography step was performed in 20 mM Tris, pH 8, 100 mM NaCl to reduce the amount of Tris in the sample. Fractions containing about 100 mg of holo CarH were merged and exposed to ambient light for 2 hrs until all holo CarH was photolyzed, as assessed by UV-Vis spectroscopy. Protein was precipitated by addition of methanol to a final concentration of 70% (v/v) followed by shaking for 1 min. Precipitated protein was removed by centrifugation at 20,000 × *g* for 20 min. Methanol was then removed in a rotary evaporator at 70 torr and 30 °C. The remaining aqueous solution was flash-frozen in liquid nitrogen and lyophilized. The resulting product, a pale red crystalline solid, was dissolved in 4 mL H_2O and purified by successive HPLC runs over a 3.0 × 150 mm Poroshell 120 SB-C_{18} (2.7 µm pore size) reversed-phase column (Agilent). Samples were injected in 200 µL aliquots using an auto-sampler and separated using the following elution program: flow 200 µL min^{-1}, solvent A H_2O + 0.01% (v/v) TFA, solvent B CH_3CN + 0.01% (v/v) TFA, isocratic 0% B for 5 min, ramp to 80% B over 25 min, wash with 100% B for 20 min, re-equilibrate at 0% B for 20 min. The product was collected in 40 µL fractions into tubes pre-filled with 500 µL H_2O to dilute the TFA and to prevent hydrolysis of the acid-sensitive compound. The fractions were merged

immediately after collection, flash-frozen in liquid nitrogen, and lyophilized, yielding a white powder. Fractions originating from 25 mg of protein yielded 0.22 µmol of the compound (30% yield), as judged by absorbance at 260 nm. ESI-MS (*m*/*z*): $[M+H]^+$ calculated for $C_{10}H_{11}N_5O_3$, 250.0935; found, 250.0962.

NMR-spectroscopic characterization of 4′,5′-anhydroadenosine

For NMR, 0.22 µmol of the CarH photolysis product (4′,5′-anhydroadenosine) were dissolved in 230 µL $(CD_3)_2SO$ (960 µM) and placed in a high precision NMR tube (Shigemi). All spectra were collected on a Varian INOVA 500 MHz NMR spectrometer at 25 °C. Chemical shift referencing was done using the solvent, referencing the residual ^1H shift to 2.50 ppm and the ^{13}C shift to 39.51 ppm.

The 1-D ^1H NMR spectrum (Figure S5) was collected at a field strength of 11.47 T. The final spectrum was obtained by averaging 64 scans. Each scan used a 2 s relaxation delay followed by a 9.2 µs 90° excitation pulse, and a time-domain acquisition of 3 s. The time-domain data consisted of 26,678 points (complex). No apodization function was applied to the time domain data for this spectrum. The size of the transformed data set was 256,000 (complex) to improve the fineness of the digital resolution. The spectral window was 8.9 ppm wide (4446.2 Hz) and was centered at 4.67 ppm. Baseline correction was applied prior to obtaining integrals of resonances. Chemical shifts: δ 8.37 (1H, s), δ 8.14 (1H, s), δ 7.32 (2H, s), δ 6.16 (1H, d), δ 5.75 (1H, d), δ 5.58 (1H, d), δ 4.82 (1H, q), δ 4.71 (1H, t), δ 4.28 (1H, broadened singlet), δ 4.20 (1H, broadened singlet). Assignments to 4′,5′-anhydroadenosine were made based on 2-D spectra, described below, and are shown in Figure 3.

The 2-D ^{1}H-^{1}H double quantum correlation spectroscopy (DQCOSY) NMR spectrum (Figure S6) used essentially the same ^{1}H pulse widths and spectral window as the 1-D ^{1}H NMR spectrum described above. For each t_1 time increment, 16 transients were collected. The number of t_1 time increments was 200, resulting in 400 free induction decays (FIDs) being collected for this phase-sensitive experiment. Each pass through the NMR pulse sequence consisted of a 1 s relaxation delay, the two 90° ^{1}H rf pulses separated by the t_1 evolution time, and 231 ms of t_2 time domain data acquisition (corresponding to 1024 complex points at the 4439.5 Hz sweep width). Following data acquisition, the t_2 time domain data were apodized with a shifted, squared sine bell function. Linear prediction was used to extend the size of the data set from 200 complex points along the t_1 axis to 512. The t_1 time domain data was then apodized with a shifted, squared sine bell function. The size of the final 2-D data matrix was 2048 × 2048 (1024 × 1024 complex).

The 2-D ^{1}H-^{13}C heteronuclear single quantum correlation (HSQC) NMR spectrum (Figure S7) was collected with 64 transients for each t_1 time increment. A total of 400 FIDs were collected, corresponding to two FIDs for each t_1 time increment. The ^{1}H spectral window and pulse durations were the same as those used for the DQCOSY 2-D NMR spectrum. Each pass through the pulse sequence entailed a 1 s relaxation delay and 100 ms of t_2 time domain data acquisition (the acquisition time was kept short to prevent overheating of the sample and probe due to the demands associated with ^{13}C decoupling during data acquisition). The short t_2 acquisition time limited the number of the data points in the FIDs to 444 complex points. The ^{13}C dimension (t_1/F_1) was 250 ppm wide and centered at 125 ppm. The ^{13}C hard pulses were 14.6 ms in duration. The t_2 time domain data was apodized with a Gaussian function. The number of points in the t_1 time domain was extended from 200 complex points to 512 complex points. The

t_1 time domain data was apodized with a Gaussian function. The size of the final 2-D data matrix was 2048 × 2048 (1024 × 1024 complex).

The 2-D ^{1}H-^{13}C gradient-selected heteronuclear multiple bond correlation (gHMBC) NMR spectrum (Figure S8) was collected with 192 transients for each t_1 time increment. A total of 400 FIDs were collected. The ^{1}H spectral window and pulses were the same as above for the DQCOSY and HSQC 2-D NMR spectra. Each pass through the pulse sequence employed a 1 s relaxation delay and ended with 231 ms of t_2 time domain data acquisition, corresponding to the collection of 1024 data points. The ^{13}C dimension (t_1/F_1) and rf pulses were the same as those described above for the 2-D ^{1}H-^{13}C HSQC NMR spectrum. The t_2 time domain data was apodized with a sine bell function. Linear prediction extended the size of the t_1 time domain data from 400 points to 512 points (absolute value, not complex). The t_1 time domain data was apodized using a sine bell function. The size of the final 2-D data matrix was 1024 × 1024.

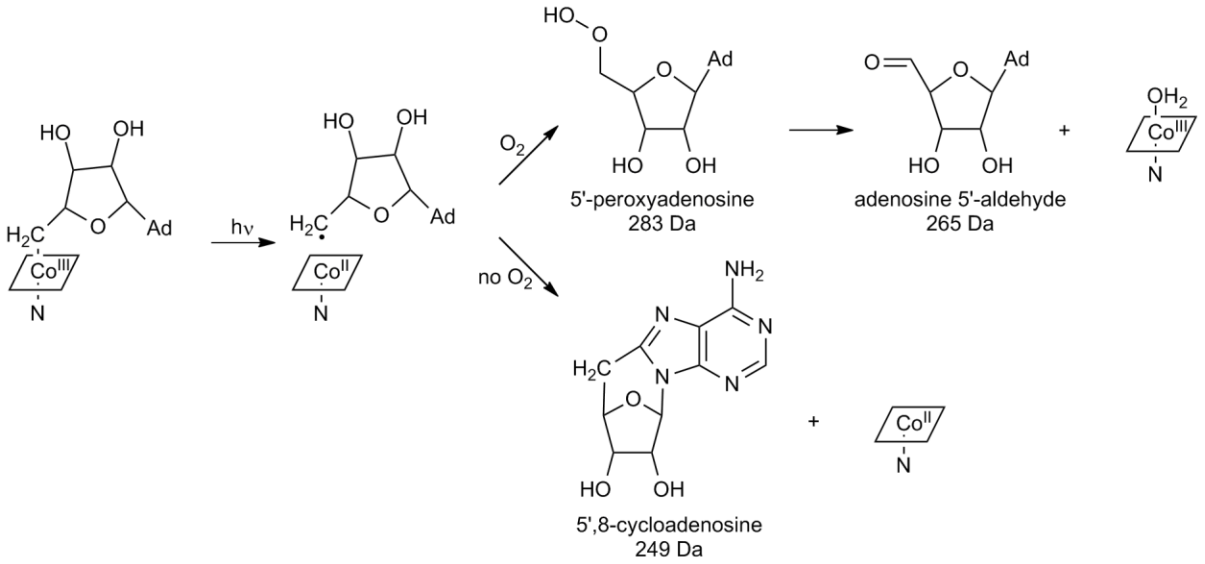

Figure S1. Fates of cobalamin and the 5′-deoxyadenosyl group upon photolysis of free AdoCbl. Light exposure causes homolytic cleavage of the covalent Co-C bond, generating the 5′-deoxyadenosyl radical and cob(II). Molecular oxygen rapidly reacts with the 5′-deoxyadenosyl radical, forming 5′-peroxyadenosine and then adenosine-5′-aldehyde as the major products. Cob(II) gets oxidized to cob(III)alamin. In the absence of molecular oxygen, the 5′-deoxyadenosyl radical undergoes a cyclization reaction to form 5′,8-cycloadenosine. Cob(II) is stable under these conditions. Molecular weights of products resulting from the 5′-deoxyadenosyl group are indicated. Cobalamin is shown as a rhombus, N denotes the lower axial base (dimethylbenzimidazole for free AdoCbl). Adenine base is abbreviated as Ad.

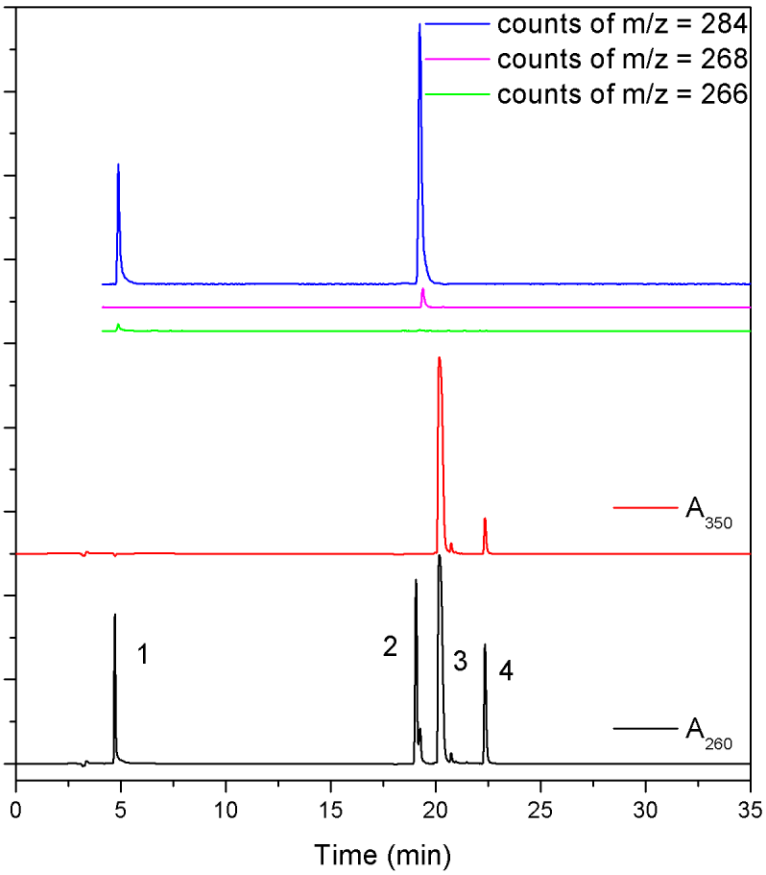

Figure S2. Photolysis products of pure AdoCbl. Shown are A_{260} (black) and A_{350} traces (red) from LC of AdoCbl exposed to light under aerobic conditions. Shown at the top are counts of specific ions, as measured by an in-line mass spectrometer. Peak 1 is adenosine-5′-aldehyde, as indicated by the associated ions with *m*/*z* = 266 (free aldehyde) and *m*/*z* = 284 (aldehyde hydrate). Peak 2 is 5′-peroxyadenosine, as indicated by the associated ion with *m*/*z* = 284. The peak furthermore has a shoulder that corresponds to adenosine, as indicated by the associated ion with *m*/*z* = 268. Peak 3 is cob(III), as indicated by the additional large absorbance at 350 nm and the associated ion with *m*/*z* = 665 (2+) (not shown). Peak 4 is unphotolyzed AdoCbl, as indicated by the additional absorbance at 350 nm and the associated ions with *m*/*z* = 790 (2+) and *m*/*z* = 527 (3+) (not shown). Note that acquisition of mass spectra started at t = 4 min.

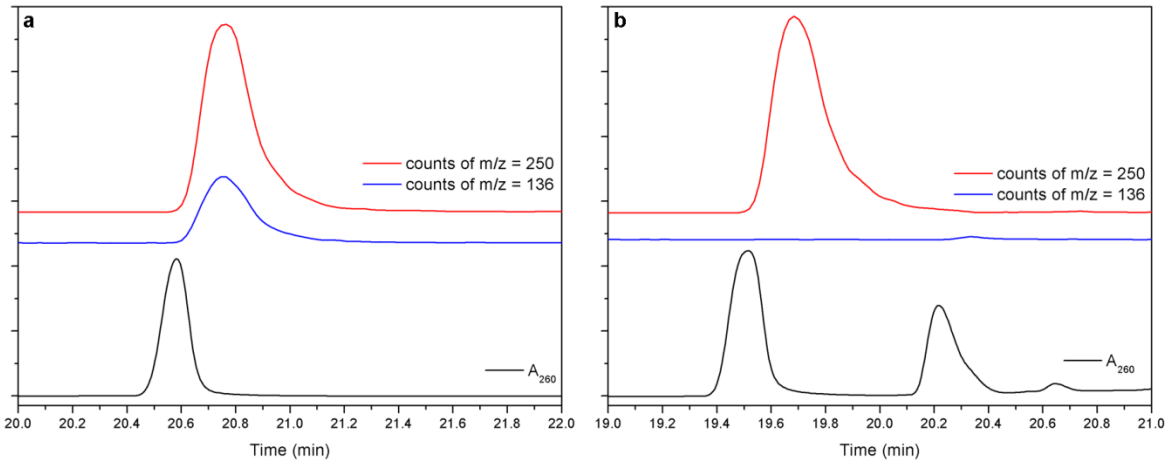

Figure S3. Comparison of fragmentation patterns of CarH photolysis product and AdoCbl photolysis product 5′,8-cycloadenosine. (**a**) Close-up of A_{260} trace (black) from LC of CarH exposed to light under anaerobic conditions. Shown at the top are counts of ions with *m*/*z* = 250 (full product, red trace) and *m*/*z* = 136 (adenine fragment, blue trace), as measured by an in-line mass spectrometer. (**b**) Close-up of A_{260} trace (black) from LC of AdoCbl exposed to light under anaerobic conditions, which generates 5′,8-cycloadenosine. Ion traces are shown as in (**a**). Note that 5′,8-cycloadenosine (*m*/*z* = 250) elutes at a different time than the CarH photolysis product and does not exhibit an adenine fragment (*m*/*z* = 136). Later peaks are cobalamin species. Full LC traces are shown in Figure 2. The ~0.2 min time delay between the A_{260} trace and the mass counts reflects the time delay between the UV cell and the mass spectrometer inlet.

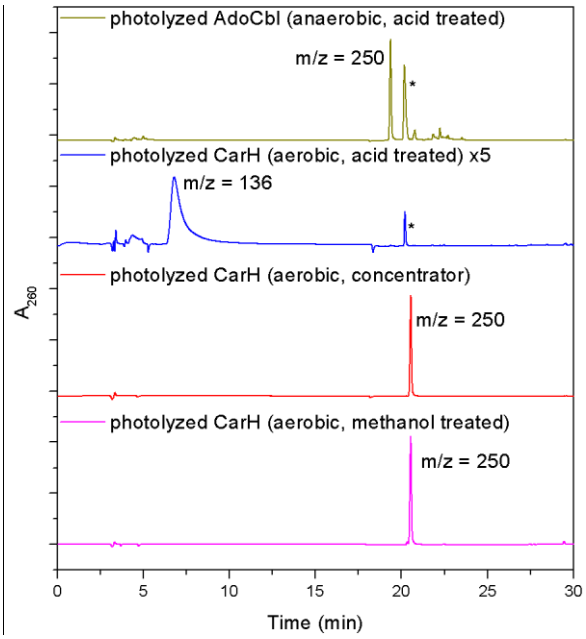

Figure S4. Stability of the anaerobic AdoCbl photolysis product and the CarH photolysis product in the presence of acid. Shown are A_{260} traces from LC of different light-exposed CarH and free AdoCbl samples. The CarH photolysis product is sensitive to treatment with 8% (v/v) trifluoroacetic acid (blue), whereas the anaerobic photolysis product of free AdoCbl is not (dark yellow). Control experiments in which the CarH photolysis product was not treated with acid but instead separated from the protein using a concentrator (red) or by treatment with methanol (magenta) show no degradation of the product, suggesting that degradation was specific to the addition of acid. Mass-to-charge ratios of peaks, as determined by an in-line mass spectrometer, are indicated. Peaks denoted with an asterisk are cobalamin species, as indicated by additional absorbance at 350 nm and 520 nm. TFA treatment induces release of a small amount of cobalamin from CarH. Note that the blue trace is magnified five-fold.

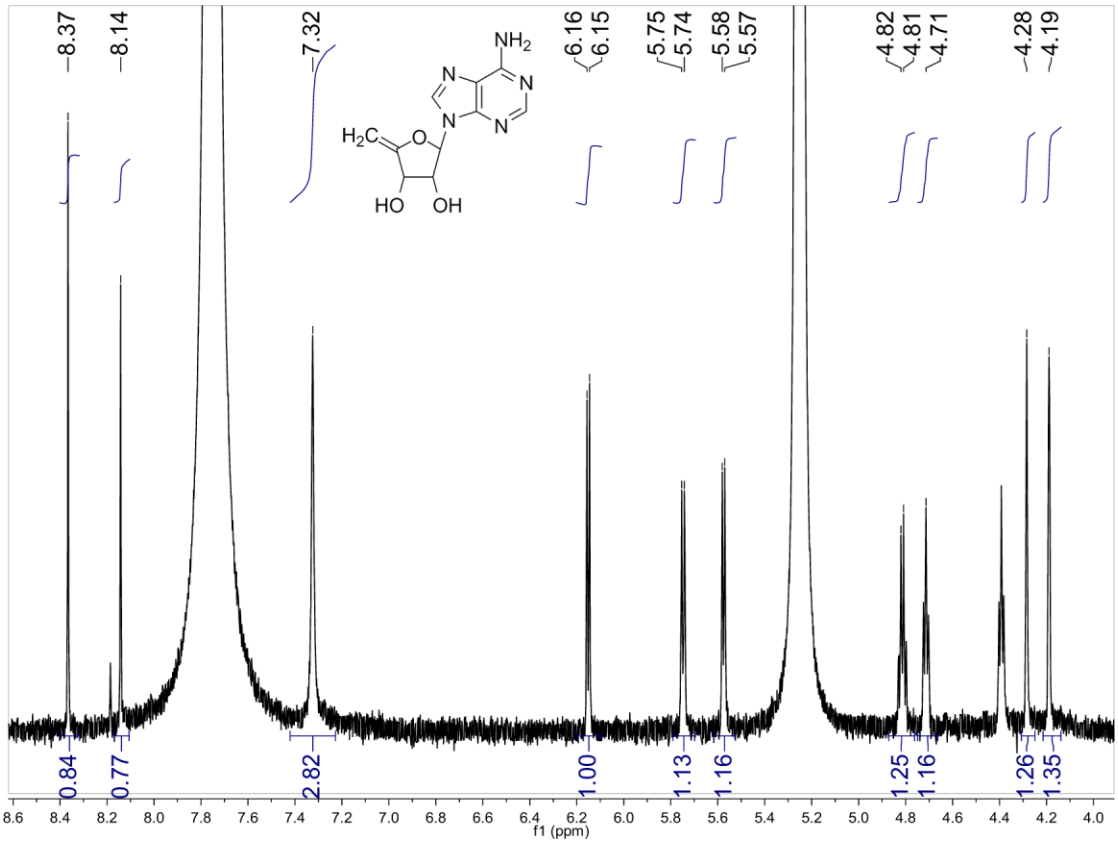

Figure S5. 1-D ^{1}H NMR spectrum of 4′,5′-anhydroadenosine collected in $(CD_3)_2SO$ at a field strength of 11.74 T with peak chemical shifts (top) and peak integrations (bottom) shown. The large peaks at 7.7 ppm and 5.25 ppm arise from contaminating Tris that is left in the sample.

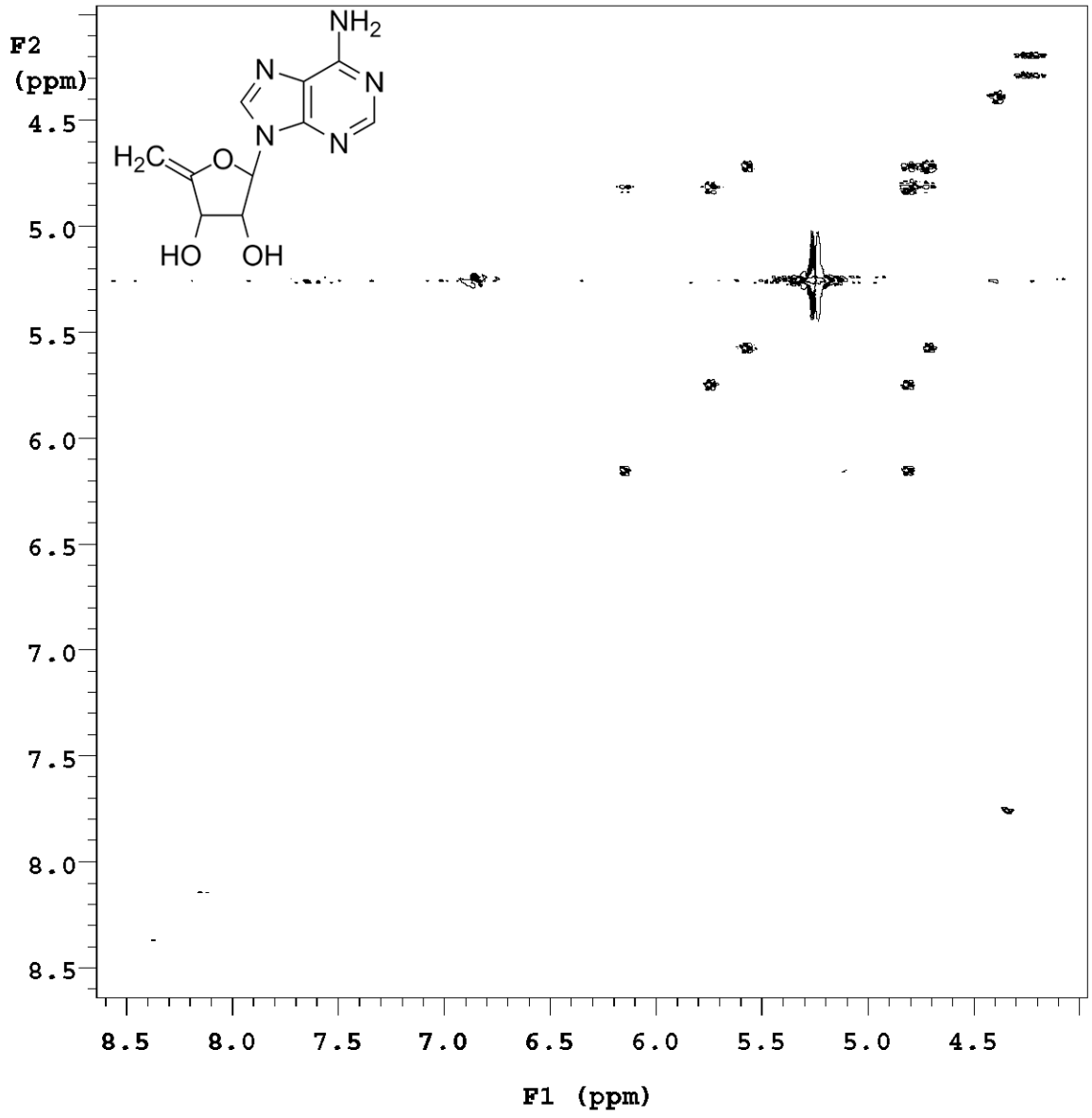

Figure S6. 2-D ^{1}H-^{1}H DQCOSY NMR spectrum of 4′,5′-anhydroadenosine collected in $(CD_3)_2SO$ at a field strength of 11.74 T. Selected chemical shift range shown to emphasize signals from 4′,5′-anhydroadenosine. Signals around 5.25 ppm arise from contaminating Tris.

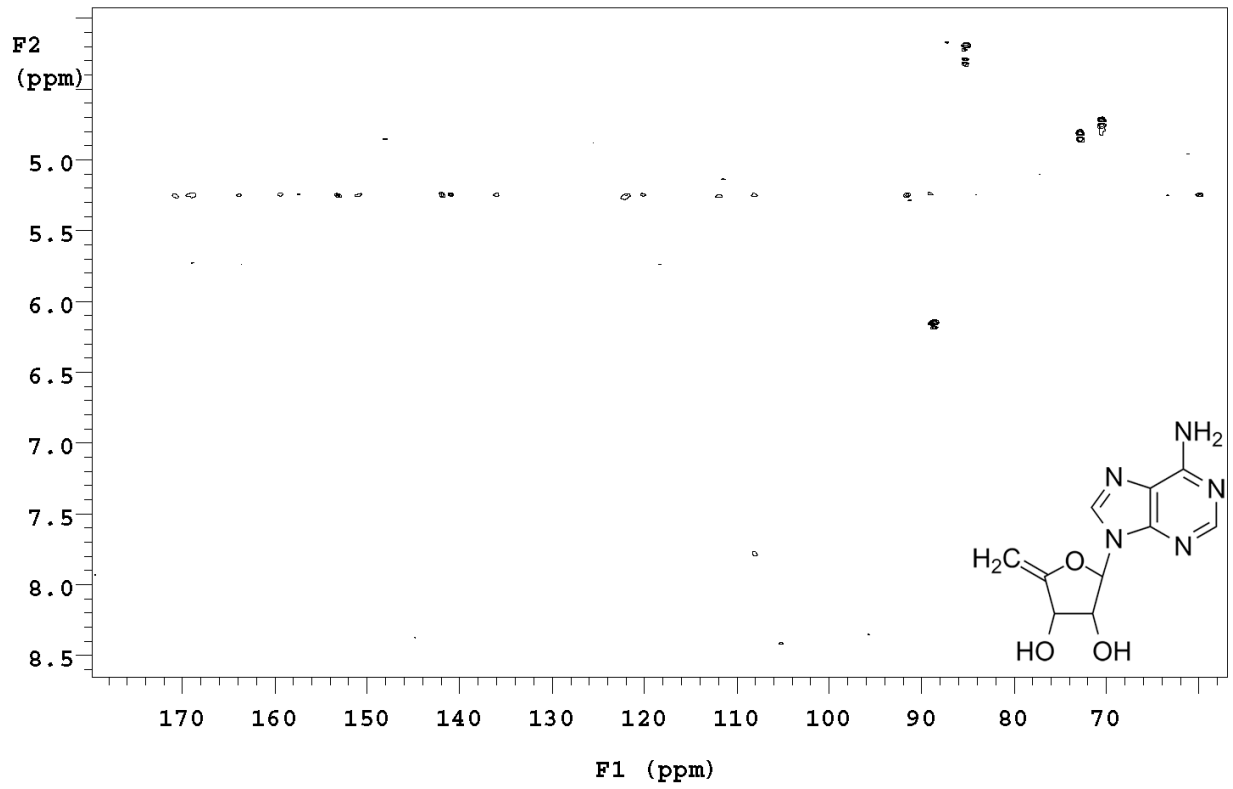

Figure S7. 2-D ^{1}H-^{13}C HSQC NMR spectrum of 4′,5′-anhydroadenosine collected in $(CD_3)_2SO$ at a field strength of 11.74 T. Selected chemical shift range shown to emphasize signals from 4′,5′-anhydroadenosine.

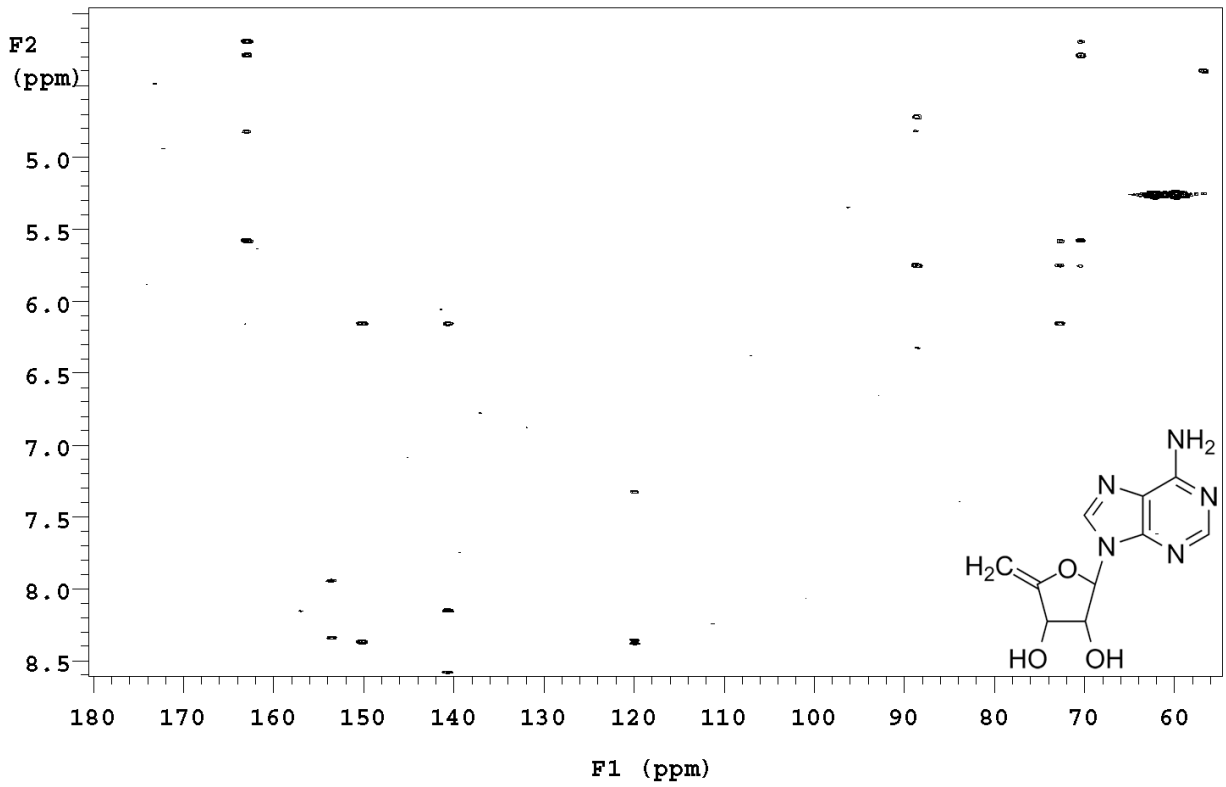

Figure S8. 2-D ^{1}H-^{13}C gHMBC NMR spectrum of 4′,5′-anhydroadenosine collected in $(CD_3)_2SO$ at a field strength of 11.74 T. Selected chemical shift range shown to emphasize signals from 4′,5′-anhydroadenosine.

Supporting References

1. Ortiz-Guerrero, J. M., Polanco, M. C., Murillo, F. J., Padmanabhan, S., and Elias-Arnanz, M. (2011), *Proc. Natl. Acad. Sci. U. S. A. 108*, 7565-7570.

2. Schrauzer, G. N., Lee, L. P., and Sibert, J. W. (1970), *J. Am. Chem. Soc. 92*, 2997-3005.