Angewandte Chemie
Eine Zeitschrift der Gesellschaft Deutscher Chemiker

Supporting Information

Proto-Urea-RNA (Wöhler RNA) Containing Unusually Stable Urea Nucleosides

Hidenori Okamura[+], Antony Crisp[+], Sarah Hübner, Sidney Becker, Petra Rovó, and Thomas Carell**

anie_201911746_sm_miscellaneous_information.pdf

Author Contributions

H.O. Conceptualization: Equal; Data curation: Equal; Formal analysis: Equal; Investigation: Equal
A.C. Conceptualization: Equal; Data curation: Equal; Formal analysis: Equal; Investigation: Equal; Methodology: Equal.

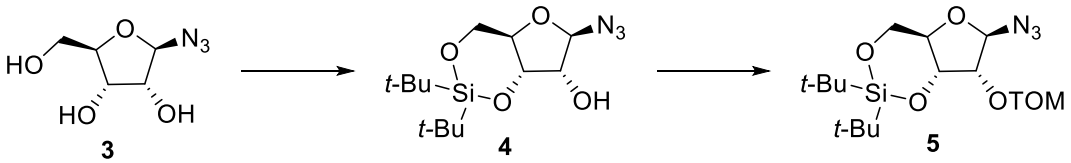
Supporting information

1. General Experimental Methods

Chemicals were purchased from Sigma-Aldrich, TCI, Fluka, ABCR, Carbosynth or Acros organics and used without further purification. The solvents were of reagent grade or purified by distillation. Reactions and chromatography fractions were monitored by qualitative thin-layer chromatography (TLC) on silica gel F_{254} TLC plates from Merck KGaA. Flash column chromatography was performed on Geduran® Si60 (40-63 μm) silica gel from Merck KGaA. Reactions were conducted under a positive pressure of dry nitrogen in oven-dried glassware, and at ambient room temperature, unless otherwise specified. NMR spectra were recorded on Bruker AVIIIHD 400 spectrometers (400 MHz). ^{1}H NMR shifts were calibrated to the residual solvent resonances: DMSO-d_6 (2.50 ppm), CD_3OD (4.87 ppm), C_6D_6 (7.16 ppm), D_2O (4.79 ppm). ^{13}C NMR shifts were calibrated to the residual solvent: DMSO-d_6 (39.52 ppm), CD_3OD (49.00 ppm), C_6D_6 (128.06 ppm). All NMR spectra were analysed using the program MestRE NOVA 10.0.1 from Mestrelab Research S. L. Normal resolved mass spectra were measured on a LTQ FT-ICR by Thermo Finnigan GmbH. High resolution mass spectra were measured by the analytical section of the Department of Chemistry of the Ludwigs-Maximilians-Universität München on the following spectrometers (ionization mode in brackets): MAT 95 (EI) and MAT 90 (ESI) from Thermo Finnigan GmbH, unless otherwise specified. IR spectra were recorded on a PerkinElmer Spectrum BX II FT-IR system. All substances were directly applied as solids or on the ATR unit. Analytical RP-HPLC was performed on an analytical HPLC Waters Alliance (2695 Separation Module, 2996 Photodiode Array Detector) equipped with the column Nucleosil 120-2 C18 from Macherey Nagel using a flow of 0.5 mL/min, a gradient of 0-25% of buffer B in 30 min was applied, unless otherwise mentioned. Preparative RP-HPLC was performed on a HPLC Waters Breeze (2487 Dual λ Array Detector, 1525 Binary HPLC Pump) equipped with the column VP 250/32 C18 from Macherey Nagel. A flow rate of 5 mL/min with varying gradients of between of 0-0.8% and 0-70% of buffer B in 45 min was applied for the purifications, unless otherwise specified. Oligonucleotides were purified using the following buffer system: buffer A: 100 mM NEt_3/HOAc (pH 7.0) in H_2O and buffer B: 100 mM NEt_3/HOAc in 80% (v/v) acetonitrile. The pH value of buffers were adjusted using a MP 220 pH-meter (Metter Toledo). Oligonucleotides were detected at wavelength: 260 nm. biuret- and triuret-nucleosides were purified using the following buffer system: buffer A: pure H_2O and buffer B: pure acetonitrile. The compounds were detected at the wavelength: 210 nm. UV spectra, melting profiles and the concentrations of purified oligonucleotides were measured on a JASCO V-650 spectrometer. Calculation of concentrations was assisted using the software OligoAnalyzer 3.0 (Integrated DNA Technologies: https://eu.idtdna.com/calc/analyzer). For strands containing artificial bases, the extinction coefficient of their corresponding canonical-only strand was employed without corrections. Matrix-assisted laser desorption/ionization-time-of-flight (MALDI-TOF) mass spectra were recorded on a Bruker Autoflex II. For MALDI-TOF measurements, the samples were desalted on a 0.025 μm VSWP filter (Millipore) against ddH_2O and co-crystallized in a 3-hydroxypicolinic acid matrix (HPA).

2. Synthesis and Characterisation of the Phosphoramidite Building-Blocks

3,5-*O*-(di-*tert*-butylsilanediyl)-2-*O*-triisopropylsilyloxymethyl-1-β-D-ribofuranosyl-azide (5)



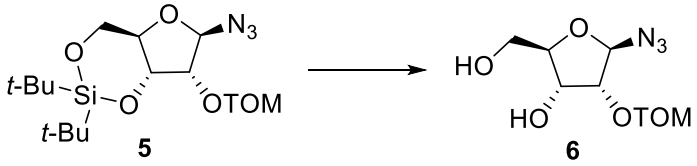
A solution of β-azidoribose[1] **3** (4.80 g, 27.4 mmol) in DMF (35 mL) was treated with di-*tert*-butylsilyl bis(trifluoromethanesulfonate) (9.40 mL, 28.8 mmol) at 0°C and the resulting mixture was stirred for 1 h. The reaction was subsequently quenched by the addition of imidazole (3.90 g, 57.6 mmol), and then suspended in water. The organic layer was washed with brine, dried over $MgSO_4$, and then concentrated to dryness to give the crude compound **4** as an oil that was used directly for the subsequent reactions. A small portion of the residue (20 mg) was purified *via* flash chromatography (silica gel, Hexane:EtOAc = 100:0 → 200:1 → 100:1) to give the compound **4** as a colorless oil, for the purposes of characterisation.

^{1}H NMR (400 MHz, CD_3OD) δ 5.17 (s, 1H), 4.40 (dd, *J* = 9.0, 5.1 Hz, 1H), 4.17-4.09 (m, 1H), 4.00-3.96 (m, 2H), 3.92 (dd, *J* = 10.6, 9.0 Hz, 1H), 1.07 (s, 9H), 1.03 (s, 9H); **^{13}C NMR** (101 MHz, CD_3OD) δ 96.8, 78.2, 75.7, 75.6, 69.3, 27.9, 27.6, 23.6, 21.3; **IR** (ν_{max}) 3480, 2931, 2107, 1471 cm^{-1}; **HRMS** (EI): calculated for $C_{13}H_{25}O_4N_3Si^+$ $[M]^+$: 315.1609; found 315.1614.

A solution of the crude compound **4** in THF (135 mL) was treated with sodium hydride (1.64 g, 40.9 mmol) at 0°C. After stirring for 30 min at the same temperature, TOMCl (9.50 mL, 40.9 mmol) was added, and the mixture was stirred overnight at room temperature. The reaction was then concentrated under reduced pressure, and the residue was taken up in EtOAc. The organic layer was washed with brine, dried over $MgSO_4$ and concentrated. The residue thus obtained was purified by flash chromatography (silica gel, Hexane:EtOAc = 100:0 → 200:1 → 100:1) to give the compound **5** as white, amorphous solid (55%, 7.6 g, 15.2 mmol).

^{1}H NMR (400 MHz, CD_3OD) δ 5.37 (s, 1H), 5.15 (d, *J* = 5.1 Hz, 1H), 5.10 (d, *J* = 5.1 Hz 1H), 4.40 (dd, *J* = 8.9, 4.1 Hz, 1H), 4.16 (d, *J* = 4.1 Hz, 1H), 4.13-4.03 (m, 2H), 3.97-3.87 (m, 1H), 1.12-1.09 (m, 21H), 1.07 (s, 9H), 1.01 (s, 9H); **^{13}C NMR** (101 MHz, CD_3OD) δ 95.4, 90.6, 79.4, 78.0, 76.2, 69.3, 27.8, 27.6, 23.6, 21.2, 18.3, 18.3, 13.2; **IR** (ν_{max}) 2934, 2863, 2108, 1463 cm^{-1}; **HRMS** (EI): calculated for $C_{23}H_{46}O_5N_3Si^+$ $[M]^+$: 472.2909; found 472.3039, calculated for $C_{23}H_{46}O_5Si^+$: 485.2878; found 458.2615; For examples of characteristic MS-fragmentations of organic azides, please see Pinto, Olariu *et al*.[2]

2-*O*-triisopropylsilyloxymethyl-1-β-D-ribofuranosyl-azide (6)

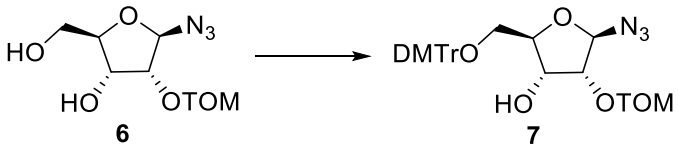


A solution of **5** (7.60 g, 15.2 mmol) in CH_2Cl_2 (150 mL) was treated at 0°C with pyridine (10 mL) followed by HF-pyridine (70% w/w, 1.50 mL, 47.7 mmol). The resulting mixture was stirred at the same temperature for 1 h. The reaction was subsequently quenched by the addition of methoxytrimethylsilane (3 mL), and the organic phase was washed with brine, dried over $MgSO_4$ and concentrated under reduced pressure. The residue thus obtained was purified by flash chromatography (silica gel, CH_2Cl_2:CH_3OH = 100:0 → 100:1 → 20:1) to give the compound **6** as

a colorless oil (61%, 3.4 g, 9.4 mmol).

^{1}H NMR (400 MHz, CD_3OD) δ 5.41 (d, *J* = 1.7 Hz, 1H), 5.15 (d, *J* = 5.1 Hz, 1H), 5.02 (d, *J* = 5.1 Hz, 1H), 4.14 (dd, *J* = 7.1, 4.8 Hz, 1H), 3.99-3.90 (m, 2H), 3.77 (dd, *J* = 12.1, 3.1 Hz, 1H), 3.60 (dd, *J* = 12.1, 5.5 Hz, 1H), 1.14-1.07 (m, 21H); **^{13}C NMR** (101 MHz, CD_3OD) δ 95.3, 91.5, 85.8, 82.6, 71.3, 63.5, 18.3, 13.1; **IR** (ν_{max}) 3405, 2942, 2866, 2108, 1463 cm^{-1}; **HRMS** (EI and ESI) neither the molecular ion, nor any discernible fragments were observed in the mass-spectra for this compound.

5-*O*-(4,4'-dimethoxytrityl)-2-*O*-triisopropylsilyloxymethyl-1-β-D-ribofuranosyl-azide (7)

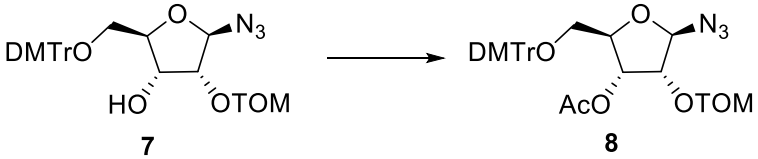


6 **7**

A solution of **6** (1.56 g, 4.15 mmol) in pyridine (94 mL) was treated with 4,4'-dimethoxytrityl chloride (1.7 g, 5.0 mmol) at 0°C and the reaction mixture was stirred at room temperature overnight. The reaction was then quenched by the addition of H_2O (0.5 mL), and the resulting mixture concentrated under reduced pressure. The residue thus obtained was purified *via* flash chromatography chromatography (silica gel, CH_2Cl_2:Hexane = 9:1 → 5:1) to give the compound **7** as colorless oil (77%, 2.12 g, 3.20 mmol).

^{1}H NMR (400 MHz, DMSO-d_6) δ 7.46-7.40 (m, 2H), 7.34-7.18 (m, 7H), 6.91-6.85 (m, 4H), 5.52 (d, *J* = 1.5 Hz, 1H), 5.10 (d, *J* = 6.8 Hz, 1H), 5.05 (d, *J* = 5.0 Hz, 1H), 4.97 (d, *J* = 5.0 Hz, 1H), 4.21-4.11 (m, 1H), 4.03-3.94 (m, 1H), 3.89 (dd, *J* = 4.5, 1.6 Hz, 1H), 3.73 (s, 6H), 3.17 (dd, *J* = 10.4, 2.5 Hz, 1H), 3.03 (dd, *J* = 10.4, 5.0 Hz, 1H), 1.13-0.94 (m, 21H); **^{13}C NMR** (101 MHz, DMSO-d_6) δ 158.5, 145.4, 140.7, 136.1, 136.1, 130.1, 129.4, 128.8, 128.1, 127.9, 127.1, 113.6, 113.2, 93.7, 89.7, 85.7, 82.8, 80.4, 79.9, 69.9, 63.7, 55.5, 18.2, 18.2, 11.8; **IR** (ν_{max}) 3327, 2941, 2865, 2106, 1607, 1507; **HRMS** (EI): calculated for $C_{36}H_{49}O_7N_3Si^+$ $[M]^+$: 663.3334; found 663.3339.

5-*O*-(4,4'-dimethoxytrityl)-3-*O*-acetyl-2-*O*-triisopropylsilyloxymethyl-1-β-D-ribofuranosyl-azide (8)

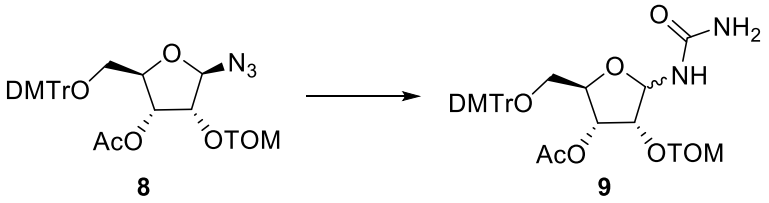


7 **8**

A solution of compound **7** (1.32 g, 1.99 mmol) in pyridine (21 mL) was treated with acetic anhydride (0.80 mL, 8.33 mmol) and DMAP (12.2 mg, 0.1 mmol) and the reaction mixture was stirred at room temperature overnight. The reaction was then quenched by the addition of MeOH (2 mL), and the solvent was evaporated under reduced pressure. The residue was taken up in EtOAc, washed with brine, dried over $MgSO_4$ and concentrated once more under reduced pressure. The residue obtained was purified by flash chromatography (silica gel, CH_2Cl_2:Hexane = 9:1 → 5:1) to give the compound **8** as a colorless oil (91%, 1.27 g, 1.81 mmol).

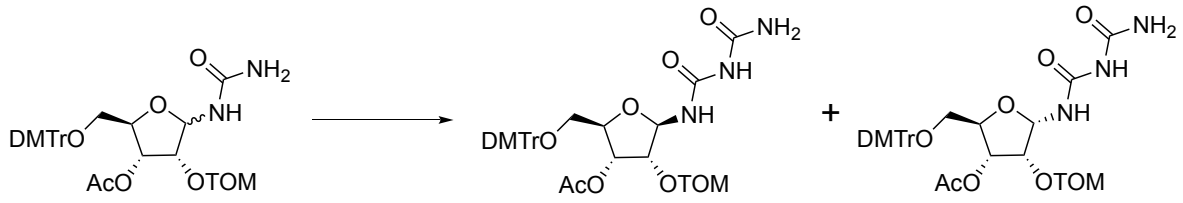
^{1}H NMR (400 MHz, DMSO-d_6) δ 7.44-7.36 (m, 2H), 7.33-7.21 (m, 7H), 6.91-6.84 (m, 4H), 5.63 (d, *J* = 2.7 Hz, 1H), 5.19 (dd, *J* = 6.1, 4.8 Hz, 1H), 4.98 (d, *J* = 5.1 Hz, 1H), 4.87 (d, *J* = 5.1 Hz, 1H), 4.29-4.12 (m, 2H), 3.73 (s, 6H), 3.26 (dd, *J* = 10.8, 3.3 Hz, 1H), 3.05 (dd, *J* = 10.8, 3.9 Hz, 1H), 1.99 (s, 3H), 1.06-0.94 (m, 21H); **^{13}C NMR** (101 MHz, DMSO-d_6) δ 169.6, 158.1, 158.1, 144.7, 135.5, 135.3, 129.7, 129.6, 127.9, 127.6, 126.7, 113.2, 92.9, 88.9, 85.6, 80.6, 77.4, 71.2, 62.3, 55.0, 20.5, 17.6, 11.3; **IR** (ν_{max}) 2941, 2864, 2107, 1744, 1606, 1507; **HRMS** (EI): calculated for $C_{38}H_{51}O_8N_3Si^+$ $[M]^+$: 705.3440; found 705.3421.

5-*O*-(4,4'-dimethoxytrityl)-3-*O*-acetyl-2-*O*-triisopropylsilyloxymethyl-1-α,β-D-ribofuranosyl-urea (9)



A solution of compound **8** (3.4 g, 4.8 mmol) in THF (24 mL) was treated with 10% Palladium on carbon (340 mg), and the resulting suspension was stirred at room temperature for 2 h under an atmosphere of H_2. The reaction mixture was filtered over a pad of $MgSO_4$, the resulting cake of which was washed with another THF (5 mL). To the filtrate was added trimethylsilyl isocyanate (3.80 mL, 24.1 mmol), and the solution was stirred at room temperature overnight. The reaction was quenched by the addition of MeOH (4 mL), and the solvent was removed by rotary evaporation. The residue was taken up in CH_2Cl_2, washed with brine, dried over $MgSO_4$ and concentrated under reduced pressure. The residue obtained was purified by flash chromatography (silica gel, CH_2Cl_2:CH_3OH = 100:0 → 99:1) to give compound **9** as (76%, 2.7 g, 3.7 mmol) as an inseparable mixture of the α- and β-anomers. **^{1}H NMR** (400 MHz, DMSO-d_6) diastereomeric mixture; δ 7.46-7.18 (m, 10H), 6.99-6.84 (m, 4H), 6.22-6.12 (m, 1H), 6.02 (s, 1H), 5.78-5.70 (m, 1H), 5.22-5.13 (m, 1H), 4.96-4.86 (m, 2H), 4.52-4.04 (m, 1H), 4.01-3.81 (m, 1H), 3.79-3.67 (m, 6H), 3.15-2.94 (m, 2H), 2.18-1.96 (m, 3H), 1.10-0.90 (m, 21H); **^{13}C NMR** (101 MHz, DMSO-d_6) diastereomeric mixture; δ 169.7, 158.1, 157.7, 157.5, 144.8, 135.6, 135.6, 135.4, 135.4, 129.8, 129.8, 129.7, 129.7, 127.9, 127.8, 127.7, 126.7, 113.2, 113.2, 89.6, 89.1, 85.6, 82.8, 80.2, 79.3, 79.0, 77.2, 73.0, 72.7, 72.6, 68.6, 63.9, 63.6, 55.9, 55.0, 55.0, 20.8, 17.7, 17.7, 17.6, 17.6, 17.5, 11.4, 11.4; **IR** (ν_{max}) 3355, 2939, 2864, 1740, 1660; **HRMS** (ESI): calculated for $C_{39}H_{54}O_9N_2NaSi^+$ $[M + Na]^+$: 745.3491; found 745.3482.

5-*O*-(4,4'-dimethoxytrityl)-3-*O*-acetyl-2-*O*-triisopropylsilyloxymethyl-1-α,β-D-ribofuranosyl-biuret (α-10/β-10)



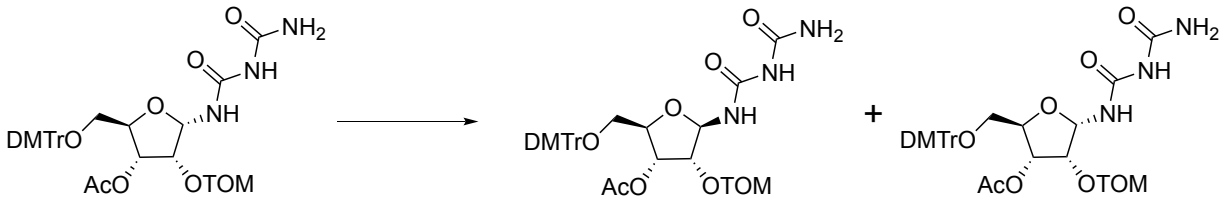
A solution of **9** (2.5 g, 3.5 mmol) in 70 mL THF was treated with pyridine (1.70 mL, 20.8 mmol), trichloroacetyl isocyanate (1.60 mL, 13.8 mmol), and the reaction mixture was stirred at room temperature. After 1 h, an excess of Al_2O_3 and MeOH (1 mL) were added, and the mixture was stirred for another 1 h. The Al_2O_3 was filtered over a pad of $MgSO_4$, and the filtrate was concentrated to dryness under reduced pressure. The residue obtained was purified by flash chromatography (silica gel, CH_2Cl_2:CH_3OH = 100:0 → 98.5:1.5 → 98:2) to give the compound **β-10** as a white foam (35%, 920 mg, 1.20 mmol), as well as the α-anomer **α-10** (55%, 1.47 g, 1.92 mmol), also as a white foam.

β-10: **^{1}H NMR** (400 MHz, DMSO-d_6) δ 8.81 (s, 1H), 8.53-8.31 (m, 1H), 7.43-7.37 (m, 2H), 7.35-7.17 (m, 8H), 7.01-6.63 (m, 5H), 5.45 (dd, J = 9.7, 7.0 Hz, 1H), 5.10 (dd, J = 5.2, 2.7 Hz, 1H), 4.89 (d, J = 5.1 Hz, 1H), 4.86 (d,

J = 5.1 Hz, 1H), 4.24 (dd, *J* = 7.0, 5.2 Hz, 1H), 4.02-3.91 (m, 1H), 3.73 (s, 6H), 3.16 (dd, *J* = 10.4, 3.6 Hz, 1H), 3.02 (dd, *J* = 10.4, 4.0 Hz, 1H), 2.02 (s, 3H), 1.08-0.91 (m, 21H); **^{13}C NMR** (101 MHz, DMSO-d_6) δ 169.7, 158.1, 158.1, 155.3, 154.3, 144.6, 135.6, 135.3, 129.8, 129.7, 127.9, 127.8, 126.7, 113.2, 89.3, 85.7, 82.1, 79.9, 77.2, 72.1, 68.5, 63.5, 55.9, 55.0, 20.7, 17.6, 11.3; **IR** (ν_{max}) 3279, 2942, 2865, 1740, 1692, 1507; **HRMS** (ESI): calculated for $C_{40}H_{55}O_{10}N_3NaSi^+$ $[M + Na]^+$: 788.3549; found 788.3538.

α-10: **^{1}H NMR** (400 MHz, DMSO-d_6) δ 8.92 (s, 1H), 7.50-7.14 (m, 9H), 7.00-6.79 (m, 4H), 5.79 (dd, *J* = 9.4, 6.1 Hz, 1H), 5.17 (dd, *J* = 5.3, 1.6 Hz, 1H), 4.89 (s, 2H), 4.52 (dd, *J* = 5.7 Hz, 1H), 4.11-4.00 (m, 1H), 3.72 (s, 6H), 3.11 (dd, *J* = 10.3, 3.8 Hz, 1H), 2.98 (dd, *J* = 10.3, 3.8 Hz, 1H), 2.07 (s, 3H), 1.13-0.76 (m, 21H); **^{13}C NMR** (101 MHz, DMSO-d_6) δ 169.7, 158.1, 155.7, 154.0, 144.7, 135.6, 135.3, 129.7, 129.6, 127.9, 127.7, 127.5, 126.8, 113.2, 89.3, 85.7, 80.4, 80.2, 79.2, 73.3, 72.6, 63.6, 55.1, 55.0, 20.5, 17.6, 17.5, 11.6; **IR** (ν_{max}) 3277, 2942, 2864, 1697, 1506; **HRMS** (ESI): calculated for $C_{40}H_{56}O_{10}N_3Si^+$ $[M + H]^+$: 766.3729; found 766.3732.

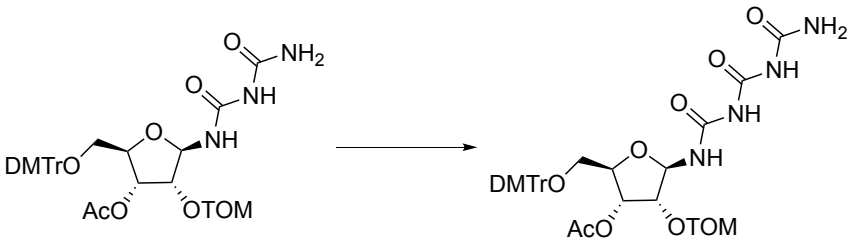
Racemization of α-10 to β-10



α–**10** β–**10** α–**10**

A solution of **α-10** (1.90 g, 2.48 mmol) in THF (25 mL) was treated with DBU (185 μL, 1.24 mmol), and the reaction mixture was stirred at 50°C overnight. The solvent was removed by rotary evaporation, and the residue was purified by flash chromatography (silica gel, CH_2Cl_2:CH_3OH = 100:0 → 98.5:1.5 → 98:2) to give compound **β-10** as a white foam (28%, 0.53 g, 0.69 mmol), as well as the α-anomer **α-10** (68%, 1.30 g, 1.69 mmol), also as a white foam.

5-*O*-(4,4'-dimethoxytrityl)-3-*O*-acetyl-2-*O*-triisopropylsilyloxymethyl-1-β-D-ribofuranosyl-triuret (11)

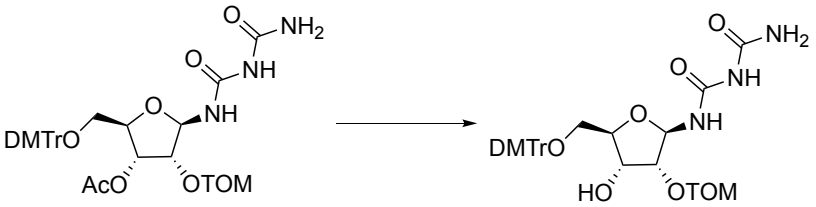


β–**10** **11**

A solution of **β-10** (800 mg, 1.04 mmol) in THF (20 mL) was treated with pyridine (11 μL, 0.10 mmol) and trichloroacetyl isocyanate (495 μL, 4.18 mmol), and the reaction mixture was stirred at room temperature. After 1 h, an excess of Al_2O_3 and MeOH (1 mL) were added, and the mixture was stirred for another 1 h. The Al_2O_3 was filtered over a pad of $MgSO_4$, and the filtrate was concentrated to dryness under reduced pressure. The residue obtained was purified by flash chromatography (silica gel, CH_2Cl_2:CH_3OH = 100:0 → 98:2) to give the compound **11** as a white foam (58%, 486 mg, 0.60 mmol).

^{1}H NMR (400 MHz, DMSO-d_6) δ 10.10 (s, 1H), 9.61 (s, 1H), 8.46 (d, J = 9.7 Hz, 1H), 7.46 – 7.17 (m, 9H), 6.94 – 6.84 (m, 4H), 6.81 – 6.66 (m, 1H), 5.46 (dd, J = 9.5, 7.0 Hz, 1H), 5.07 (dd, J = 5.2, 2.5 Hz, 1H), 4.88 (d, J = 5.2 Hz, 1H), 4.83 (d, J = 5.2 Hz, 1H), 4.32 (dd, J = 7.0, 5.2 Hz, 1H), 4.00 (m, 1H), 3.73 (s, 6H), 3.18 (dd, J = 10.4, 3.6 Hz, 1H), 3.03 (dd, J = 10.4, 3.9 Hz, 1H), 2.02 (s, 3H), 1.10 – 0.90 (m, 21H); **^{13}C NMR** (101 MHz, DMSO) δ 169.7, 158.1, 154.3, 152.8, 144.6, 135.6, 135.3, 129.8, 129.7, 127.9, 127.8, 126.8, 113.2, 89.2, 85.8, 82.3, 80.2, 77.2, 72.0, 72.0, 68.6, 63.5, 63.5, 55.9, 55.0, 20.7, 17.6, 11.3; **IR** (ν_{max}) 3313, 2942, 2866, 1701, 1607; **HRMS** (ESI): calculated for $C_{41}H_{55}O_{11}N_4Si^-$ [M – H]$^-$: 807.3642; found 807.3644.

5-*O*-(4,4'-dimethoxytrityl)-2-*O*-triisopropylsilyloxymethyl-1-β-D-ribofuranosyl-biuret (12)

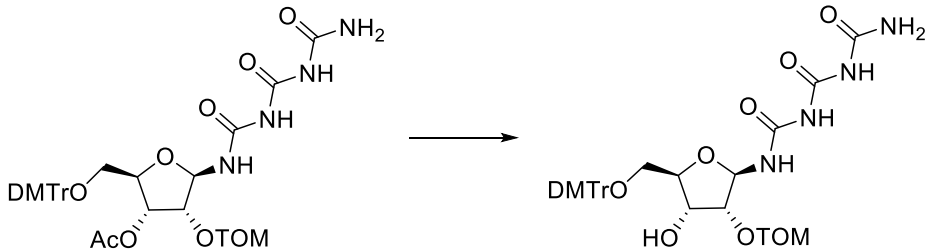


β−**10** **12**

A solution of **β-10** (400 mg, 0.522 mmol) in methanolic ammonia (7 N, 20 mL) was stirred at room temperature for 4 h. The solvent was carefully removed by rotary evaporation with a water-bath temperature of 30°C, and the residue was purified by flash chromatography (silica gel, CH_2Cl_2:CH_3OH = 99:1 → 98:2) to give compound **12** as a colorless oil (74%, 280 mg, 0.386 mmol).

^{1}H NMR (400 MHz, DMSO-d_6) δ 8.70 (s, 1H), 8.25 (d, J = 9.0 Hz, 1H), 7.50-7.10 (m, 9H), 6.96-6.68 (m, 4H), 6.78-6.62 (m, 1H), 5.45 (dd, J = 9.6, 4.1 Hz, 1H), 5.02-4.96 (m, 2H), 4.93 (d, J = 5.1 Hz, 1H), 4.02-3.92 (m, 2H), 3.84 (ddd, J = 4.0 Hz, 1H), 3.73 (s, 6H), 3.09 (dd, J = 10.2, 3.4 Hz, 1H), 2.95 (dd, J = 10.1, 5.0 Hz, 1H), 1.11-0.95 (m, 21H); **^{13}C NMR** (101 MHz, DMSO-d_6) δ 158.1, 158.0, 155.3, 154.1, 144.9, 135.7, 135.6, 129.8, 129.7, 127.8, 127.8, 126.6, 113.2, 88.8, 85.4, 82.4, 81.8, 78.9, 69.8, 68.6, 64.2, 55.0, 55.0, 17.7, 17.7, 11.4; **IR** (ν_{max}) 3278, 2939, 2864, 1697, 1507; **HRMS** (ESI): calculated for $C_{38}H_{52}O_9N_3Si^-$ [M – H]$^-$: 722.3478; found 722.3516.

5-*O*-(4,4'-dimethoxytrityl)-2-*O*-triisopropylsilyloxymethyl-1-β-D-ribofuranosyl-triuret (13)



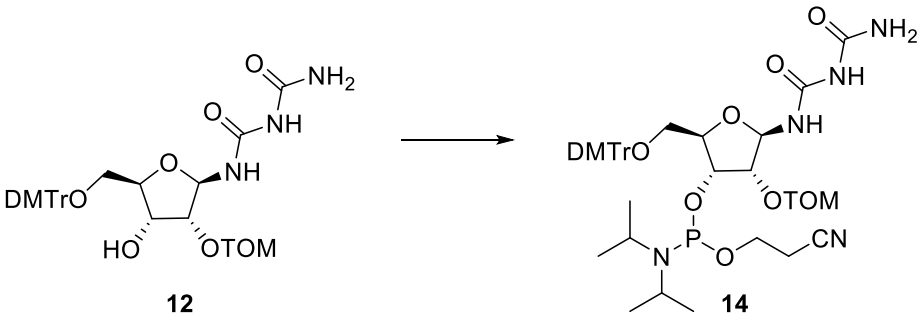
11 **13**

A solution of **11** (400 mg, 0.494 mmol) in methanolic ammonia (7 N, 20 mL) was stirred at room temperature for 4 h. The solvent was carefully removed by rotary evaporation with a water-bath temperature of 30°C, and the residue was purified by flash chromatography (silica gel, CH_2Cl_2:CH_3OH = 99:1 → 98:2) to give compound **13** as a colorless oil (76%, 290 mg, 0.378 mmol).

^{1}H NMR (400 MHz, DMSO-d_6) δ 10.00 (s, 1H), 9.57 (s, 1H), 8.40-8.15 (m, 1H), 7.46-7.15 (m, 9H), 6.95-6.80 (m, 4H), 6.78 (s, 1H), 5.46 (dd, J = 9.4, 5.4 Hz, 1H), 5.04-4.96 (m, 2H), 4.90 (d, J = 5.2 Hz, 1H), 4.04 (dd, J = 5.2 Hz, 1H), 3.96 (ddd, J = 4.9 Hz, 1H), 3.86 (ddd, J = 4.1 Hz, 1H), 3.77-3.68 (m, 6H), 3.12 (dd, J = 10.3, 3.4 Hz, 1H), 2.95

(dd, J = 10.3, 4.7 Hz, 1H), 1.08-0.92 (m, 21H); **^{13}C NMR** (101 MHz, DMSO-d_6) δ 158.1, 158.0, 152.8, 144.9, 135.7, 135.5, 129.8, 129.7, 127.8, 127.8, 126.6, 113.2, 113.1, 88.7, 85.5, 82.6, 82.3, 78.9, 69.6, 68.5, 64.0, 55.9, 55.0, 55.0, 17.7, 17.7, 11.4; **IR** (ν_{max}) 3409, 3301, 2942, 2864, 1693, 1507; **HRMS** (ESI): calculated for $C_{39}H_{53}O_{10}N_4Si^-$ [M – H]$^-$: 765.3536; found 765.3565.

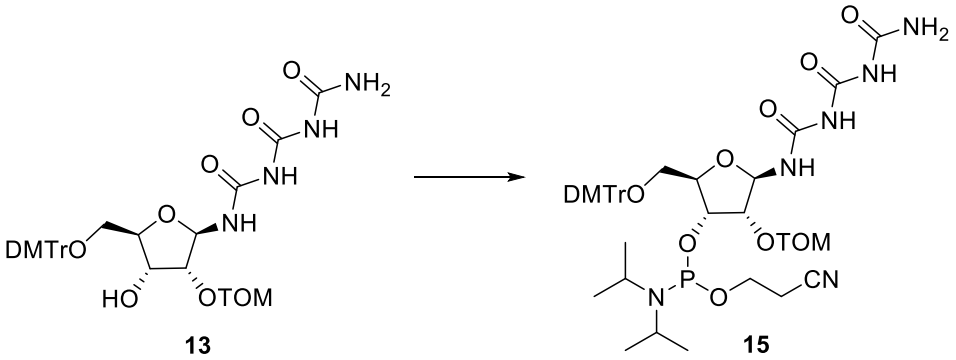
5-*O*-(4,4'-dimethoxytrityl)-3-*O*-[2-cyanoethoxy(diisopropylamino)phosphino]-2-*O*-triisopropylsilyloxy-methyl-1-β-D-ribofuranosyl-biuret (14)



A solution of **12** (240 mg, 0.331 mmol) in CH_3CN (7 mL) was treated with bis(2-cyanoethyl)-*N,N*-diisopropylphosphoramidite (158 μL, 0.50 mmol) followed by diisopropylamine-tetrazole salt (57 mg, 0.33 mmol), and the reaction mixture was stirred at room temperature overnight. The reaction mixture was diluted with CH_2Cl_2, washed with brine, dried over $MgSO_4$ and concentrated under reduced pressure. The residue was purified by flash chromatography (silica gel, CH_2Cl_2:Acetone = 100:0 → 70:30, HPLC grade solvents) to give compound **14** as a white foam (70%, 213 mg, 0.231 mmol).

^{1}H NMR (400 MHz, C_6D_6) diastereomeric mixture; δ 9.25-8.83 (m, 1H), 7.80-7.65 (m, 2H), 7.61-7.48 (m, 4H), 7.40-6.78 (m, 7H), 6.36-6.13 (m, 1H), 5.30-5.08 (m, 2H), 4.80-4.53 (m, 2H), 4.52-4.33 (m, 1H), 3.76-2.98 (m, 12H), 1.25-0.43 (m, 35H); **^{31}P NMR** (162 MHz, C_6D_6) diastereomeric mixture; δ 149.9, 149.6; **IR** (ν_{max}) 3355, 2939, 2864, 2359, 1697, 1508; **HRMS** (ESI): calculated for $C_{47}H_{70}O_{10}N_5NaSi^+$ [M + Na]$^+$: 946.4522; found 946.4524.

5-*O*-(4,4'-dimethoxytrityl)-3-*O*-[2-cyanoethoxy(diisopropylamino)phosphino]-2-*O*-triisopropylsilyloxy-methyl-1-β-D-ribofuranosyl-triuret (15)



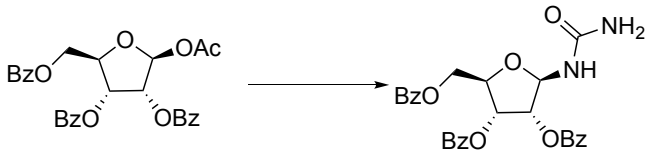
A solution of **13** (320 mg, 0.417 mmol) in CH_3CN (8 mL) was treated with bis(2-cyanoethyl)-*N,N*-diisopropylphosphoramidite (199 μL, 0.63 mmol) followed by diisopropylamine-tetrazole salt (72 mg, 0.42 mmol), and the reaction mixture was stirred at room temperature overnight. The reaction mixture was diluted with CH_2Cl_2, washed with brine, dried over $MgSO_4$ and concentrated under reduced pressure. The residue was purified by flash chromatography (silica gel, CH_2Cl_2:Acetone = 100:0 → 70:30, HPLC grade solvents) to give compound **15** as a

white foam (54%, 204 mg, 0.231 mmol).

^{1}H NMR (400 MHz, C_6D_6) diastereomeric mixture; δ 10.59-9.67 (m, 1H), 9.17-8.05 (m, 2H), 7.82-7.66 (m, 2H), 7.60-7.49 (m, 4H), 7.39-6.68 (m, 7H), 6.38-6.22 (m, 1H), 5.40-5.27 (m, 1H), 5.27-5.15 (m, 1H), 4.83-4.63 (m, 2H), 4.61-4.41 (m, 1H), 3.84-3.07 (m, 12H), 1.28-0.52 (m, 35H); **^{31}P NMR** (162 MHz, C_6D_6) diastereomeric mixture; δ 150.0, 149.6; **IR** (ν_{max}) 3305, 2939, 2866, 2359, 1701, 1508; **HRMS** (ESI): calculated for $C_{48}H_{71}O_{11}N_6NaSi^+$ [*M* + Na]$^+$: 989.4580; found 989.4582.

3. Synthesis and Characterisation of Standards for Prebiotic Reactions

2,3,5-tri-*O*-benzoyl-1-β-D-ribofuranosyl-urea (S2)

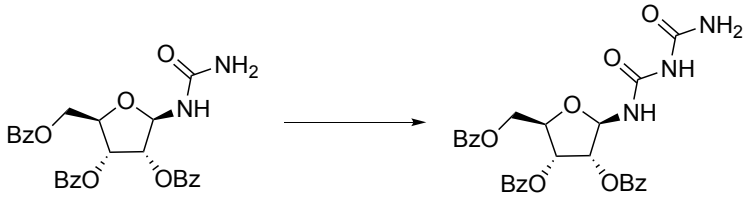


S1 **S2**

A solution of 1-*O*-acetyl-2,3,5-tri-*O*-benzoyl-1-β-D-ribose **S1** (6.00 g, 12.0 mmol) in CH_2Cl_2 (120 mL) was treated dropwise with a solution of $TiCl_4$ in CH_2Cl_2 (1.0 M, 16.6 mL, 16.6 mmol), and the reaction mixture was stirred at that temperature for 2 h. The reaction was quenched by the addition of H_2O, partitioned and extracted with CH_2Cl_2, and the organic phase dried over $MgSO_4$. Concentration under reduced pressure gave a white residue. The residue was then taken up in toluene (120 mL), added AgNCO (2.20 g, 14.2 mmol), and the mixture was heated at 110 ºC for 2 h. After cooling to room temperature, the mixture was filtered over a pad of celite. NH_3 gas was bubbled into the filtrate at 0°C for 30 min, noting the formation of a white precipitate. The solvent was subsequently evaporated under reduced pressure, and the residue subjected to flash column chromatography (silica gel, CH_2Cl_2:MeOH = 100:0 → 95:5) to give **S2** as a white amorphous solid (67%, 4.12 g, 8.17 mmol).

^{1}H NMR (400 MHz, DMSO-d_6) δ 8.17-7.99 (m, 2H), 7.94-7.82 (m, 4H), 7.75-7.60 (m, 3H), 7.58-7.38 (m, 6H), 7.24 (d, J = 9.7 Hz, 1H), 5.90 (s, 2H), 5.79-5.66 (m, 2H), 5.49 (dd, J = 5.9 Hz, 1H), 4.65-4.41 (m, 3H); **^{13}C NMR** (101 MHz, DMSO-d_6) δ 165.6, 164.8, 164.8, 157.5, 133.9, 133.8, 133.6, 129.4, 129.3, 129.3, 129.3, 128.9, 128.8, 128.8, 128.7, 128.6, 83.1, 77.2, 73.4, 71.2, 64.5; **IR** (ν_{max}) 3483, 3452, 3325, 1719, 1658; **HRMS** (ESI): calculated for $C_{27}H_{24}O_8N_2Na^+$ [M + Na]$^+$: 989.1425; found 527.1422.

2,3,5-tri-*O*-benzoyl-1-β-D-ribofuranosyl-biuret (S3)

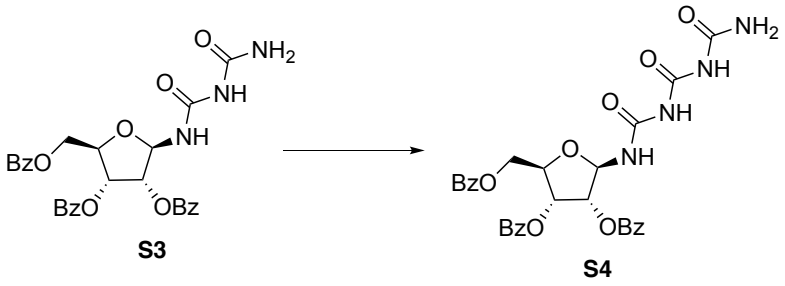


S2 **S3**

A solution of **S2** (300 mg, 0.595 mmol) in a 2:1 mixture of THF/CH_2Cl_2 (18 mL) was treated with trichloroacetyl isocyanate (130 µL, 1.96 mmol), and the reaction mixture was stirred overnight at room temperature. An excess of Al_2O_3 and MeOH (1 mL) were added, and the mixture was stirred for another 1 h. The Al_2O_3 was filtered over a pad of celite, and the filtrate was concentrated to dryness under reduced pressure. The residue obtained was purified by flash chromatography (silica gel, CH_2Cl_2:CH_3OH = 100:0 →100:1 → 50:1) to give **S3** as a white amorphous solid (96%, 313 mg, 0.571 mmol).

^{1}H NMR (400 MHz, DMSO-d_6) δ 8.89 (s, 1H), 8.57 (d, J = 9.0 Hz, 1H), 8.13-7.97 (m, 2H), 7.94-7.81 (m, 4H), 7.81-7.60 (m, 3H), 7.60-7.33 (m, 6H), 7.08-6.65 (m, 2H), 5.86-5.70 (m, 2H), 5.65 (dd, J = 5.7 Hz, 1H), 4.68-4.46 (m, 3H); **^{13}C NMR** (101 MHz, DMSO-d_6) δ 165.5, 164.7, 164.7, 155.0, 154.2, 134.0, 133.9, 133.6, 129.4, 129.3, 129.3, 129.3, 128.9, 128.8, 128.8, 128.7, 128.5, 82.8, 77.8, 73.6, 71.1, 64.2; **IR** (ν_{max}) 3270, 2988, 1715, 1697, 1543; **HRMS** (ESI): calculated for $C_{28}H_{25}O_9N_3Na^+$ [M + Na]$^+$: 570.1483; found 570.1481.

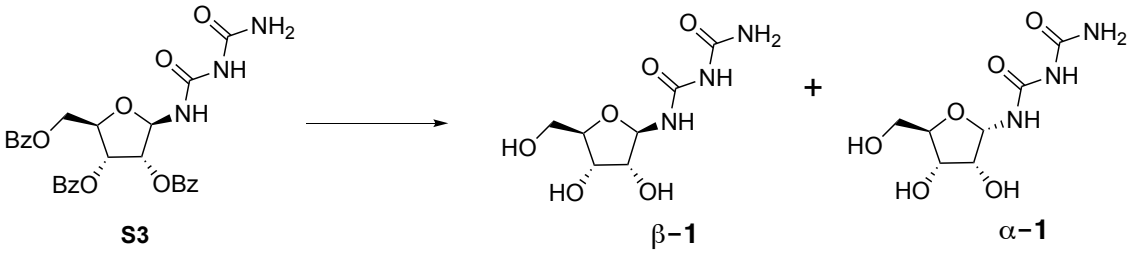
2,3,5-tri-*O*-benzoyl-1-β-D-ribofuranosyl-triuret (S4)



A solution of **S3** (300 mg, 0.548 mmol) in a 2:1 mixture of THF/CH_2Cl_2 (18 mL) was treated with trichloroacetyl isocyanate (130 µL, 1.96 mmol), and the reaction mixture was stirred overnight at room temperature. An excess of Al_2O_3 and MeOH (1 mL) were added, and the mixture was stirred for another 1 h. The Al_2O_3 was filtered over a pad of celite, and the filtrate was concentrated to dryness under reduced pressure. The residue obtained was purified by flash chromatography (silica gel, CH_2Cl_2:CH_3OH = 100:0 →100:1 → 50:1) to give **S4** as a white amorphous solid (85%, 277 mg, 0.466 mmol).

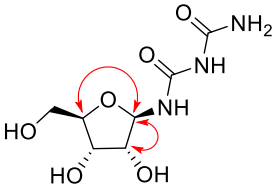
^{1}H NMR (400 MHz, DMSO-d_6) δ 10.10 (s, 1H), 9.55 (s, 1H), 8.64 (d, J = 8.5 Hz, 1H), 8.12-7.97 (m, 2H), 7.93-7.83 (m, 4H), 7.74-7.58 (m, 3H), 7.57-7.32 (m, 7H), 7.32 (s, 1H), 6.80 (s, 1H), 5.87-5.69 (m, 3H), 4.76-4.42 (m, 3H); **^{13}C NMR** (101 MHz, DMSO-d_6) δ 165.5, 164.7, 164.7, 154.2, 152.8, 152.5, 134.0, 133.9, 133.6, 129.4, 129.3, 129.3, 129.3, 128.9, 128.8, 128.8, 128.7, 128.5, 83.1, 78.0, 73.7, 71.0, 64.0; **IR** (ν_{max}) 3419, 3266, 1714, 1692, 1483; **HRMS** (ESI): calculated for $C_{29}H_{26}O_{10}N_4Na^+$ [M + Na]$^+$: 613.1541; found 613.1542.

1-α,β-D-ribofuranosyl-biuret (α-1/β-1)



A solution of **S3** (15 mg, 0.027 mmol) in methanolic ammonia (7 N, 2 mL) was stirred overnight at room temperature. The solvent was carefully removed by rotary evaporation with a water-bath temperature of 30°C, and the residue was purified by reverse-phase HPLC (H_2O:CH_3CN = 100:0 → 99:1 over 45 min) to give compound **β-1** as a white solid (15%, 1 mg, 4.25 µmol) as well as compound **α-1**, also as a white solid (46%, 3 mg, 12.76 µmol). Yields of the purified nucleosides were measured by integration of the ^{1}H NMR spectra in D_2O, with sodium 4,4-dimethyl-4-silapentane-1-sulfonate (DSS) added as an internal reference. **β-1** and **α-1** were distinguished from their related pyranoside-isomers by the distinctive ^{13}C shifts associated with their C4'- and C5'- atoms.[3] The β-furanoside **β-1** was able to be distinguished from its α-isomer **α-1** by a characteristic H1' to H4' NOE.

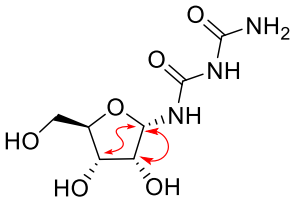
Significant NOEs:



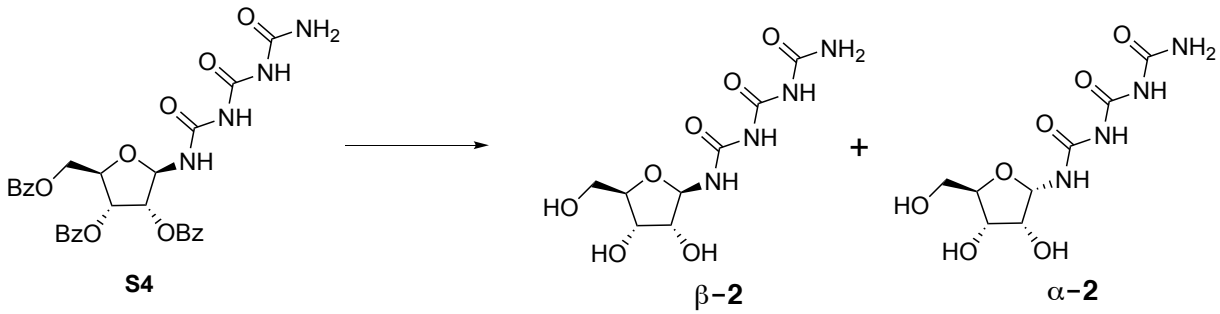
β-1: **^{1}H NMR** (400 MHz, D_2O) δ 5.36 (d, J = 5.4 Hz, 1H), 4.1 (dd, J = 5.4, 4.1 Hz, 1H), 4.1 (dd, J = 5.4 Hz, 1H), 3.95 (ddd, J = 4.1, 4.1, 3.5 Hz, 1H), 3.69 (dd, J = 12.5, 3.5 Hz, 1H), 3.61 (dd, J = 12.5, 4.1 Hz, 1H); **^{13}C NMR** (101 MHz, D_2O) δ 84.0, 70.1, 73.7, 83.3, 61.2 (measured *via* a HSQC experiment); **IR** (ν_{max}) 3319, 2924, 2490, 1663, 1534; **HRMS** (ESI): calculated for $C_7H_{14}O_6N_3Na^+$ [M + H]$^+$: 236.0877; found 236.0889.

α-1: **^{1}H NMR** (400 MHz, D_2O) δ 5.63 (d, *J* = 4.4 Hz, 1H), 4.22 (dd, *J* = 4.4, 4.7 Hz, 1H), 4.17 (dd, *J* = 6.2, 4.7 Hz, 1H), 3.97 (ddd, *J* = 6.2, 4.4, 3.0 Hz, 1H), 3.72 (dd, *J* = 12.6, 3.0 Hz, 1H), 3.60 (dd, *J* = 12.6, 4.4 Hz, 1H); **^{13}C NMR** (101 MHz, D_2O) δ 80.6, 70.2, 70.4, 81.9, 60.6 (measured *via* a HSQC experiment); **IR** (ν_{max}) 3448, 2971, 1483, 1691, 1543; **HRMS** (ESI): calculated for $C_7H_{14}O_6N_3Na^+$ [*M* + H]$^+$: 236.0877; found 236.0889.

Significant NOEs:



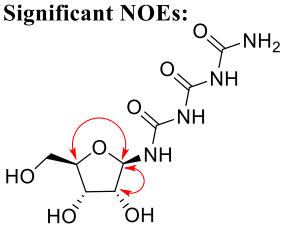
1-α,β-D-ribofuranosyl-triuret (α-2/β-2)



S4 β**−2** α**−2**

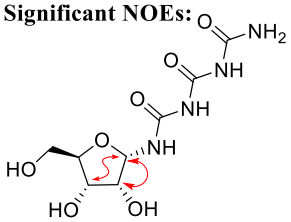
A solution of **S4** (15 mg, 0.025 mmol) in methanolic ammonia (7 N, 2 mL) was stirred overnight at room temperature. The solvent was carefully removed by rotary evaporation with a water-bath temperature of 30°C, and the residue was purified by reverse-phase HPLC (H_2O:CH_3CN = 100:0 → 99:1 over 45 min) to give compound **β-2** as a white solid (14%, 1 mg, 3.59 μmol) as well as compound **α-2**, also as a white solid (28%, 2 mg, 7.19 μmol). Yields of the purified nucleosides were measured by integration of the ^{1}H NMR spectra in D_2O, with sodium 4,4-dimethyl-4-silapentane-1-sulfonate (DSS) added as an internal reference. **β-2** and **α-2** were distinguished from their related pyranoside-isomers by the distinctive ^{13}C shifts associated with their C4'- and C5'- atoms.[3] The β-furanoside **β-2** was able to be distinguished from its α-isomer **α-2** by a characteristic H1' to H4' NOE.

Significant NOEs:



^{1}H NMR (400 MHz, D_2O) δ 5.43 (d, *J* = 5.1 Hz, 1H), 4.20 (m, 1H), 4.14 (dd, *J* = 5.1 Hz, 1H), 4.00 (ddd, *J* = 4.1, 4.1, 3.4 Hz 1H), 3.74 (dd, *J* = 12.5, 3.4 Hz, 1H), 3.66 (dd, *J* = 12.5, 4.1 Hz, 1H); **^{13}C NMR*** (101 MHz, D_2O) δ 70.0, 73.8, 83.5, 61.0 (measured *via* a HSQC experiment); **IR** (ν_{max}) 3319, 2924, 2490, 1663, 1600; **HRMS** (ESI): calculated for $C_8H_{15}O_7N_4^+$ [*M* + H]$^+$: 279.0935; found 279.0947.

Significant NOEs:

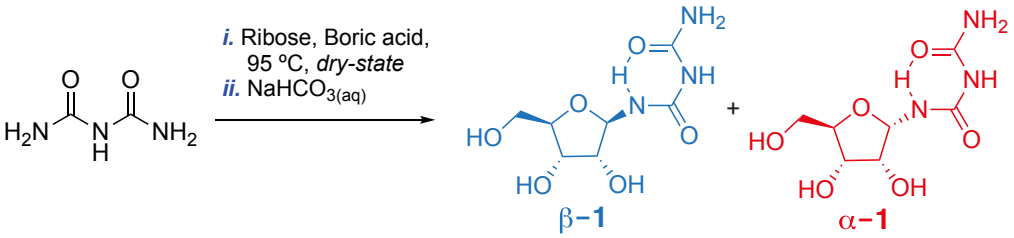


^{1}H NMR (400 MHz, D_2O) δ 5.68 (d, *J* = 4.7 Hz, 1H), 4.27 (dd, *J* = 4.7, 4.7 Hz, 1H), 4.19 (dd, *J* = 6.0, 4.7 Hz, 1H), 4.01 (ddd, *J* = 6.0, 4.4, 3.1 Hz, 1H), 3.74 (dd, *J* = 12.6, 3.1 Hz, 1H), 3.62 (dd, *J* = 12.6, 4.4 Hz, 1H); **^{13}C NMR** (101 MHz, D_2O) δ 80.6, 70.1, 70.5, 82.2, 60.6 (measured *via* a HSQC experiment); **IR** (ν_{max}) 3448, 2971, 1483, 1691, 1543; **HRMS** (ESI): calculated for $C_8H_{15}O_7N_4^+$ [*M* + H]$^+$: 279.0935; found 279.0947.

4. Prebiotic Reactions

Prebiotic reactions were carried out and analysed by LC-ESI-MS on a Thermo Finnigan LTQ Orbitrap XL, and were chromatographed using a Dionex Ultimate 3000 HPLC system with a flow of 0.15 mL/min over an Interchim Uptisphere120A-3µm-HDO C18 column. The column temperature was maintained at 30°C. The eluting buffers were: buffer A (2 mM $HCOONH_4$ in H_2O, pH 5.5) and buffer B (2 mM $HCOONH_4$ in H_2O/MeCN 20/80). The gradient for all samples was from 100% to 99.2% buffer A over 15 min. The elution was monitored at 210 nm (Dionex Ultimate 3000 Diode Array Detector). The chromatographic eluent was directly injected into the ion source without prior splitting. Ions were scanned by use of a positive polarity mode over a full-scan range of m/z 50-1000 with a resolution of 30000. The retention time of each nucleoside was confirmed by the HRMS spectrum and also by co-injection of a chemically-synthesised standard together with the reaction mixture (please see section 3 for the syntheses).

Prebiotic synthesis of biuret-nucleosides



An open glass vial containing a mixture of biuret* (16 mg, 0.155 mmol), ribose (6 mg, 0.040 mmol) and boric acid (10 mg, 0.162 mmol) in 500 µL H_2O was heated to 95°C in an oven for 18 h. The brown residue was subsequently taken up in sodium carbonate buffer (100 mM, 0.5 mL, pH 9.5) and heated again at 95°C for 1 h. The mixture was then analysed by reverse phase HPLC-MS. Formation of the furanosides was thus observed in a total combined yield of 33% (13.2 µmol, dr α:β = 3.3:1). The identities of the desired products were confirmed by co-injection of chemically synthesised standards (red: **α-1**; blue: **β-1**).

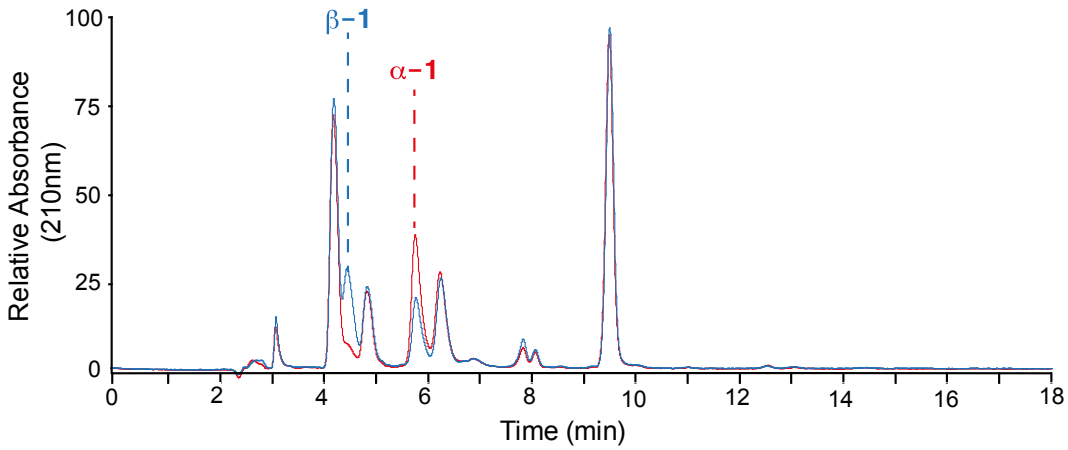
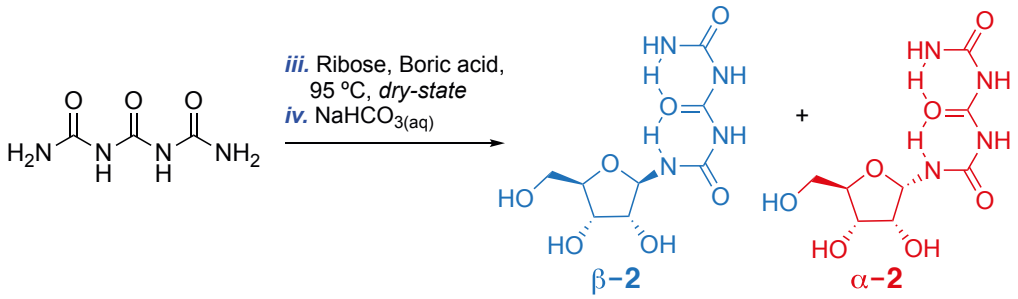


Figure S1. LC-MS chart of the mixture obtained from reacting biuret and ribose

*The biuret starting material (purchased from Sigma Aldrich) contained a small amount of triuret, which absorbs strongly at 210nm, and can be observed in the HPLC chromatogram (t = 9.5 min).

Prebiotic synthesis of triuret-nucleosides



An open glass vial containing a mixture of triuret (4 mg, 27.4 µmol), ribose (41 mg, 0.273 mmol) and boric acid (10 mg, 0.162 mmol) in 500 µL H_2O was heated to 95°C in an oven for 18 h. The brown residue was subsequently taken up in sodium carbonate buffer (100 mM, 0.5 mL, pH 9.5) and heated again at 95°C for 1 h. The mixture was then analysed by reverse phase HPLC-MS. Formation of the furanosides was thus observed in a total combined yield of 2% (0.548 µmol, dr α:β = 1.9:1). The identities of the desired products were confirmed by co-injection of chemically synthesised standards (red: **α-2**; blue: **β-2**).

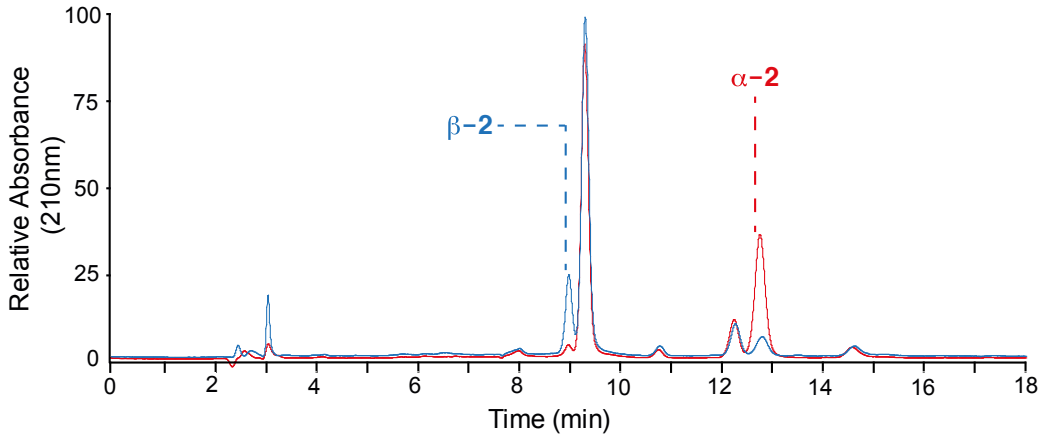


Figure S2. LC-MS chart of the mixture obtained from reacting triuret and ribose

5. Calibration Curves

In order to determine the percentage yield of a given nucleoside by HPLC analysis, solutions of the respective compound with known concentrations were prepared and their UV absorbance at 210 nm was measured three times using the same HPLC-MS setup as for the prebiotic reactions. The integrals were calculated, and the average area under the curve from three independent experiments were plotted against the concentration. A regression line was generated and its equation was extracted using the software Origin 6.0.

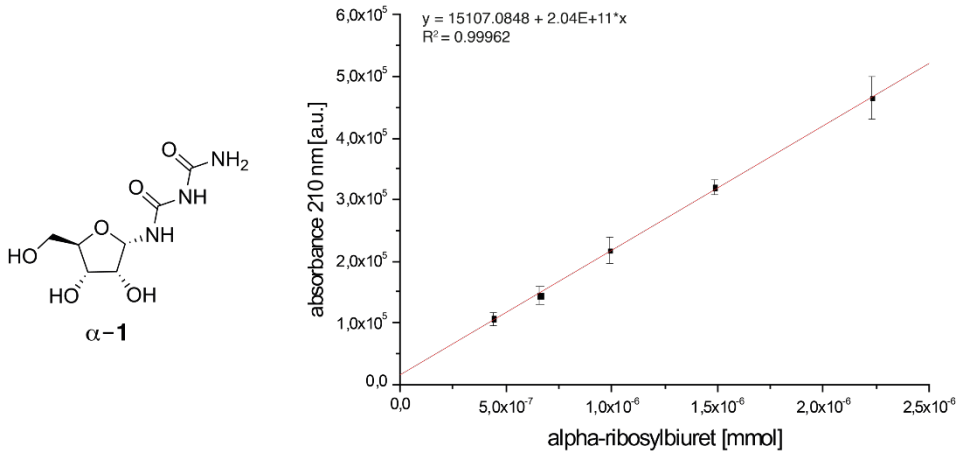


Figure S3. Calibration curve associated with compound **α-1**

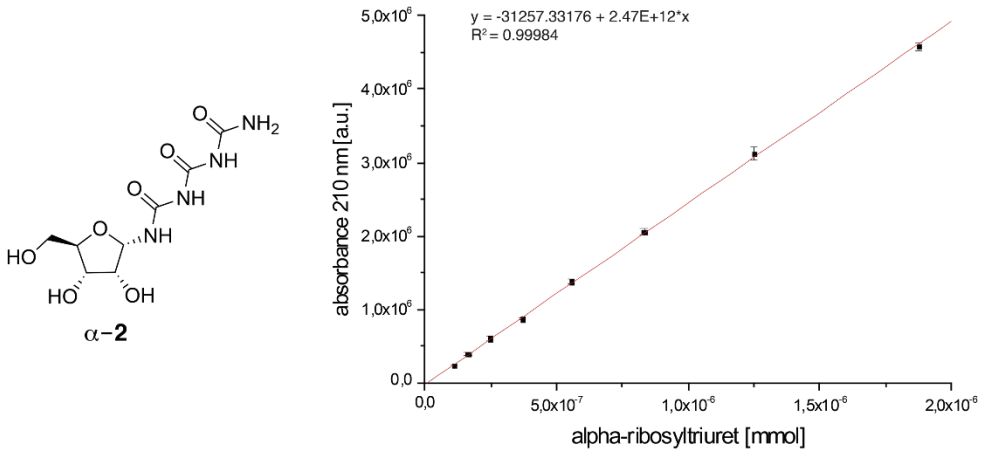


Figure S4. Calibration curve associated with compound **α-2**

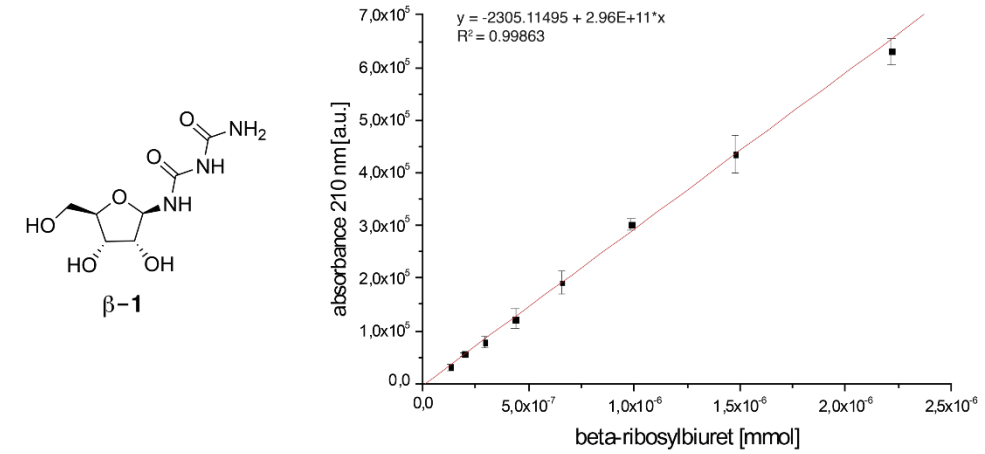


Figure S5. Calibration curve associated with compound **β-1**

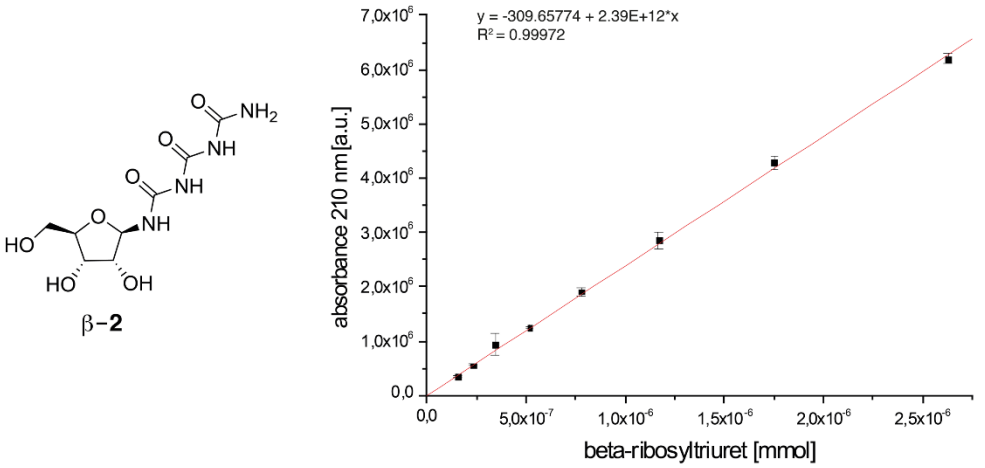


Figure S6. Calibration curve associated with compound **β-2**

6. Synthesis and Purification of Oligonucleotides

All of the oligonucleotides used in this study were synthesized on a 1 μmol scale using a DNA automated synthesizer (Applied Biosystems 394 DNA/RNA Synthesizer) with standard phosphoramidite chemistry. The phosphoramidites of canonical ribonucleosides and of inosine were purchased from Glen Research and Sigma-Aldrich. Oligonucleotides containing ribosylurea nucleosides were synthesized in DMT-OFF mode using ultramild phosphoramidites (Pac-A, Tac-G, Ac-C, U, I) with BTT in CH_3CN as an activator, DCA in CH_2Cl_2 as a deblocking solution and Pac_2O in pyridine/THF as a capping reagent. Oligonucleotides containing a 3'-triuret moiety were synthesized using Universal Support III, purchased from Sigma Aldrich. Incorporation of Cy3 was achieved using the Cyanine-3-CE Phosphoramidite (Cyanine 540), purchased from Linktech. The cleavage and deprotection of the CPG bound oligonucleotides were performed with ammonia in MeOH (7 N, 1 mL) at room temperature for 4 h. The resin was removed by filtration and the solution was evaporated at room temperature under reduced pressure. The residue was subsequently heated with a solution of trimethylamine trihydrofluoride (98 μL) and triethylamine (60 μL) in DMSO (120 μL) at 65°C for 2 h. Upon cooling in an ice bath, NaOAc (3.0 M, 25 μL) and *n*-BuOH (1 mL) were added. The resulting suspension was vortexed and cooled in a freezer (-20°C) for 30 min. After the centrifugation, supernatant was removed and the remaining oligonucleotide pellet was dried under vacuum. The oligonucleotides were further purified by reverse-phase HPLC using a Waters Breeze (2487 Dual λ Array Detector, 1525 Binary HPLC Pump) equipped with the column VP 250/32 C18 from Macherey Nagel. Oligonucleotides were purified using the following buffer system: buffer A: 100 mM NEt_3/HOAc, pH 7.0 in H_2O and buffer B: 100 mM NEt_3/HOAc in 80% (v/v) acetonitrile. A flow rate of 5 mL/min with a gradient of 0-25% of buffer B in 30 min was applied for the purifications, except where otherwise mentioned. Analytical RP-HPLC was performed on an analytical HPLC Waters Alliance (2695 Separation Module, 2996 Photodiode Array Detector) equipped with the column Nucleosil 120-2 C18 from Macherey Nagel using a flow of 0.5 mL/min, a gradient of 0-25% of buffer B in 30 min was applied, except where otherwise mentioned. Calculation of concentrations was assisted using the software OligoAnalyzer 3.0 (Integrated DNA Technologies: https://eu.idtdna.com/calc/analyzer). For strands containing Bi or Tri, the extinction coefficient of their corresponding canonical-only strand was employed without corrections. The structural integrity of the synthesized oligonucleotides was analyzed by MALDI-TOF mass measurement.

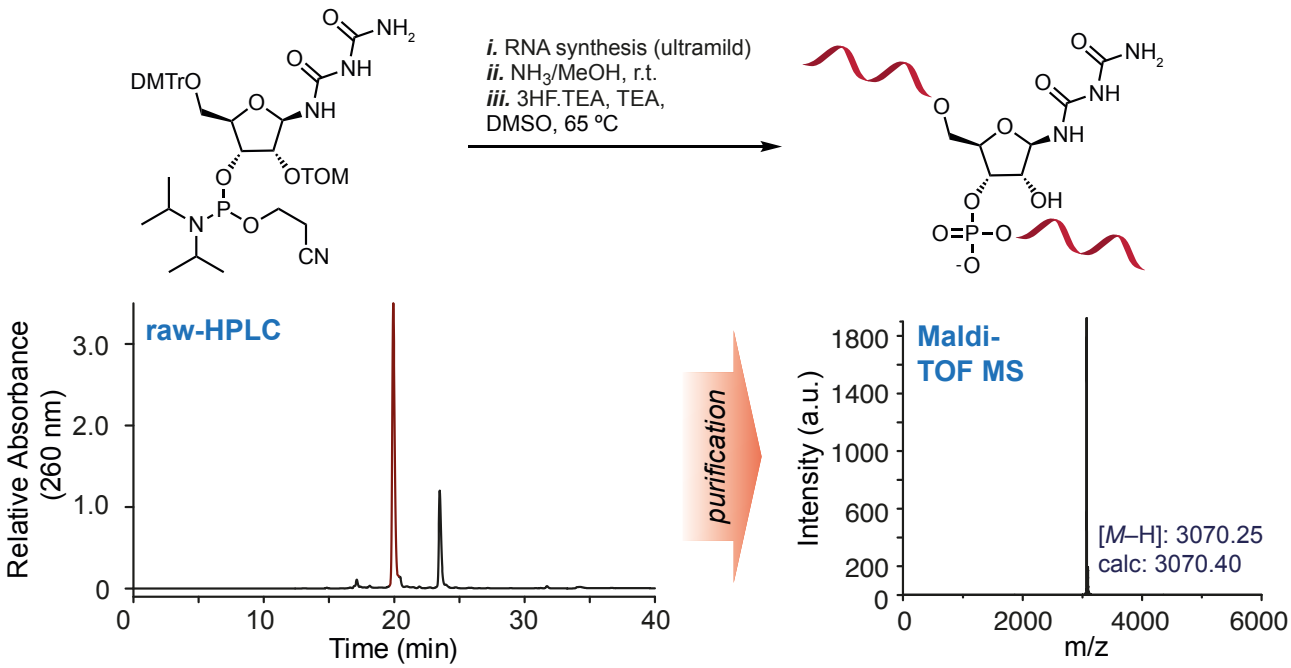


Figure S7. Preparatory HPLC and MALDI-TOF data for an example biuret nucleoside containing RNA strand (5'-CUUACXCUGA-3', X = Bi). The product peak is indicated in red.

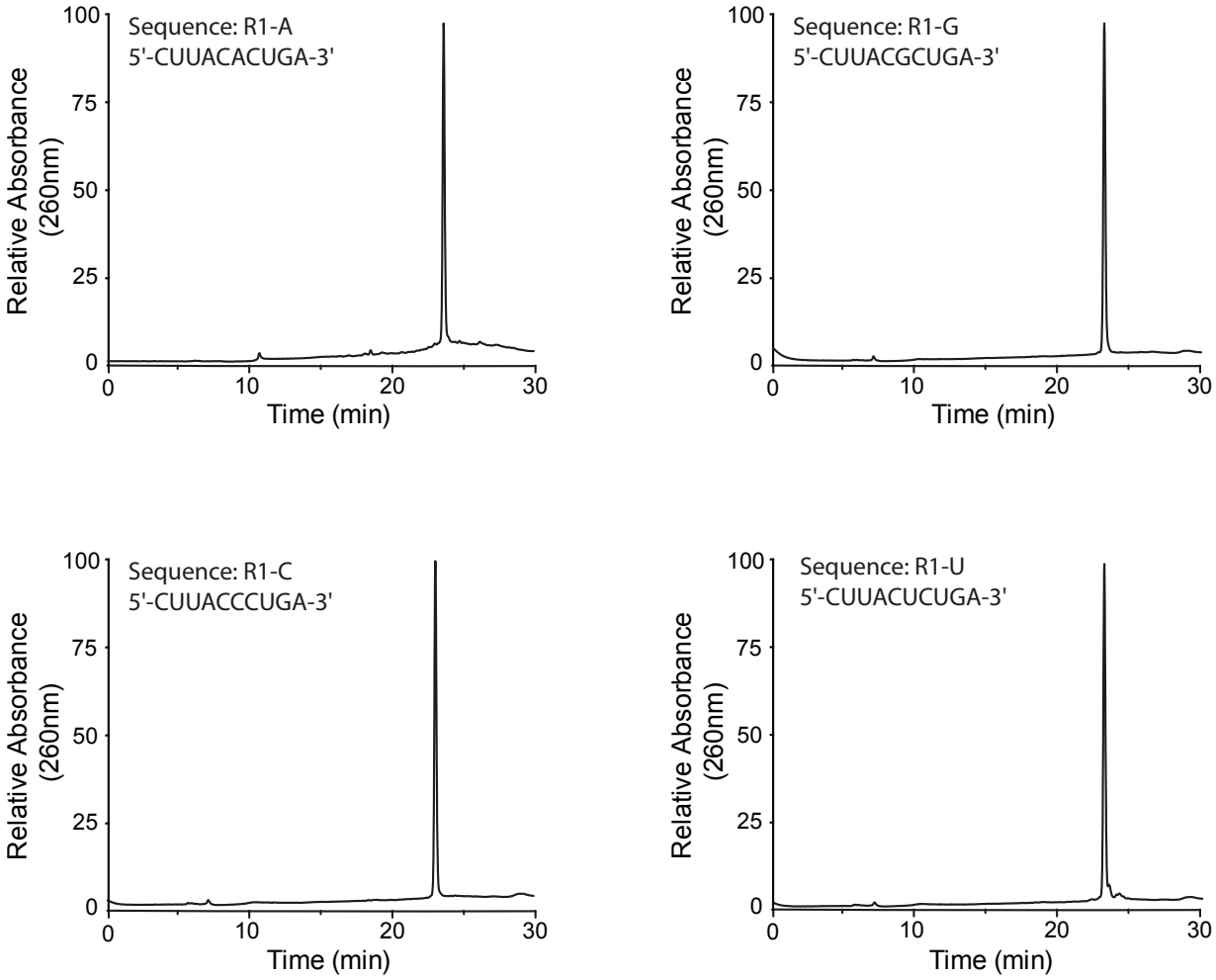


Figure S8. Analytical HPLC Traces of Purified Oligonucleotides

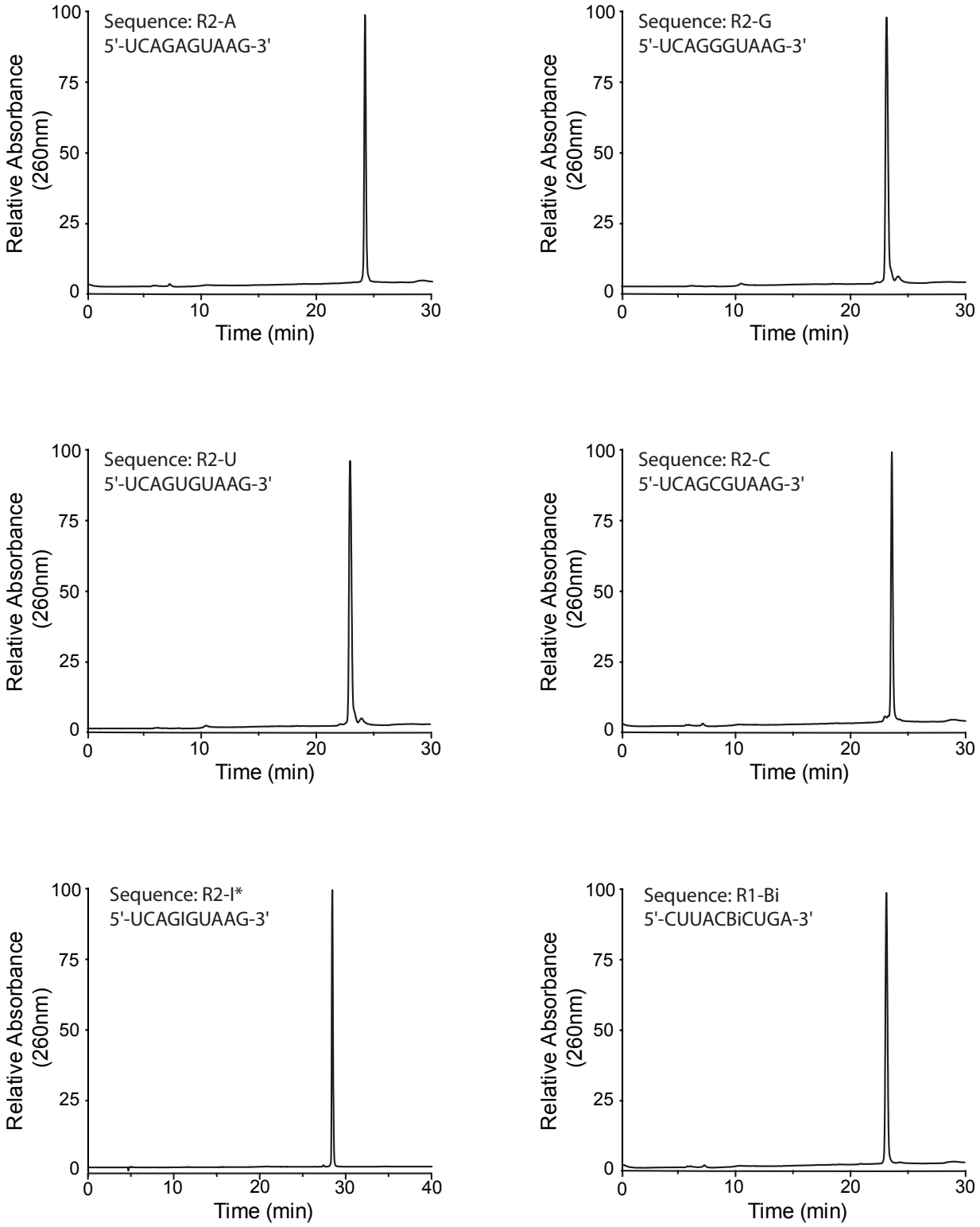


Figure S8. Analytical HPLC Traces of Purified Oligonucleotides (continued)

*A flow rate of 0.5 ml/min with a gradient of 0-45% of buffer B in 45 min was applied for this oligonucleotide.

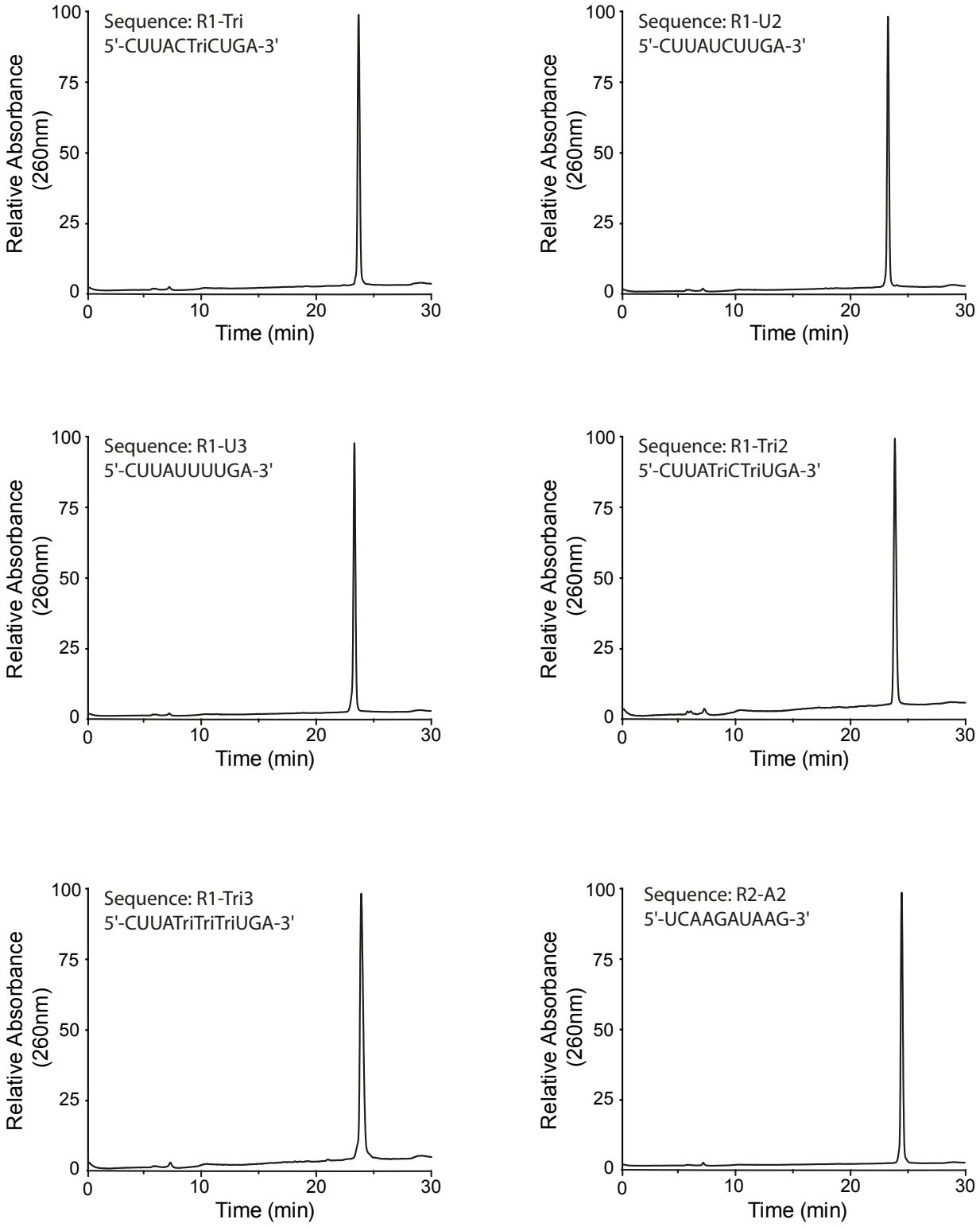


Figure S8. Analytical HPLC Traces of Purified Oligonucleotides (continued)

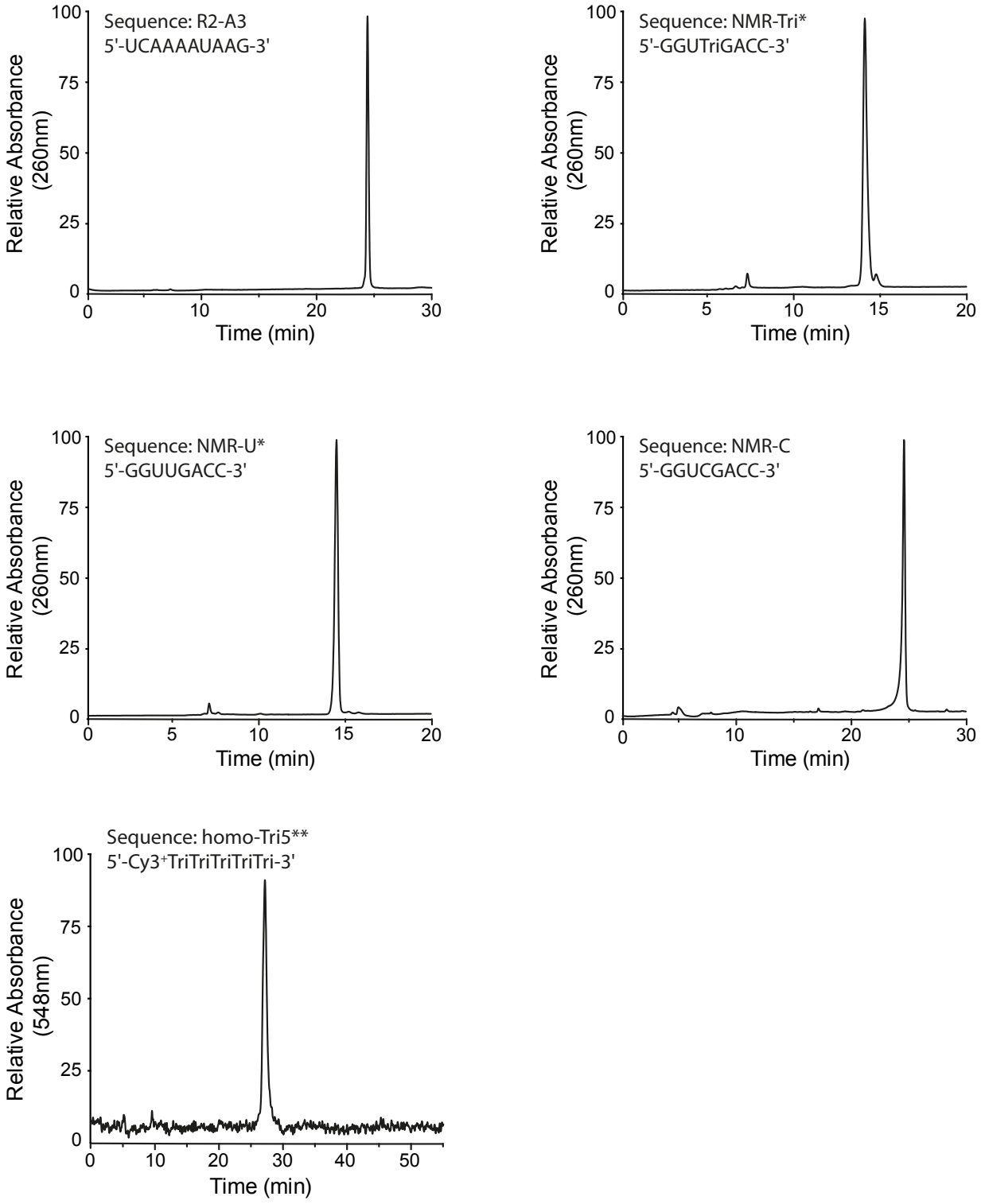


Figure S8. Analytical HPLC Traces of Purified Oligonucleotides (continued)

*A flow rate of 0.5 ml/min with a gradient of 10-15% of buffer B in 16 min was applied for these oligonucleotides.
** The crude HPLC trace is presented here as its appearance was identical after purification. A flow rate of 0.5 ml/min with a gradient of 0-90% of buffer B in 45 minutes was applied for this oligonucleotide.

Sequence name	Sequence (5’ → 3’)	Calculated [M-H]$^-$	Found (*m*/*z*)
R1-A	5'-CUUACACUGA-3'	3102.42	3102.37
R1-G	5'-CUUACGCUGA-3'	3118.42	3118.36
R1-C	5'-CUUACCCUGA-3'	3078.41	3078.67
R1-U	5'-CUUACUCUGA-3'	3079.39	3079.38
R2-A	5'-UCAGAGUAAG-3'	3205.47	3205.37
R2-G	5'-UCAGGGUAAG-3'	3221.47	3221.11
R2-C	5'-UCAGCGUAAG-3'	3181.46	3180.97
R2-U	5'-UCAGUGUAAG-3'	3182.44	3182.59
R2-I	5'-UCAGIGUAAG-3'	3206.46	3206.42
R1-Bi	5'-CUUACBiCUGA-3'	3070.40	3070.25
R1-Tri	5'-CUUACTriCUGA-3'	3113.41	3113.08
R1-U2	5'-CUUAUCUUGA-3'	3080.37	3080.64
R1-U3	5'-CUUAUUUUGA-3'	3081.35	3081.52
R1-Tri2	5'-CUUATriCTriUGA-3'	3148.41	3148.64
R1-Tri3	5'-CUUATriTriTriUGA-3'	3183.41	3183.26
R2-A2	5'-UCAAGAUAAG-3'	3189.47	3189.82
R2-A3	5'-UCAAAAUAAG-3'	3173.47	3173.23
NMR-Tri	5'-GGUTriGACC-3'	2557.38	2557.66
NMR-U	5'-GGUUGACC-3'	2523.36	2523.16
NMR-C	5'-GGUCGACC-3'	2522.38	2521.27
homo-Tri5	5’-Cy3$^+$TriTriTriTriTri-3’	2143.48 [M-2H]$^-$	2142.33

Table S1. List of synthesized oligonucleotides and their corresponding MALDI-TOF mass measurements

7. UV Melting Curve Measurements

The UV melting curves were measured on JASCO V-650 spectrometer using 10 mm QS cuvettes, purchased from Hellma Analytics. A solution (80 µL) of equimolar amounts of oligonucleotides (4 µM each) in the buffer solution containing 10 mM sodium phosphate buffer (pH 7.0) and 150 mM NaCl was heated at 50°C for 5 min and gradually cooled to 4°C prior to the measurement. Melting profiles were recorded at temperatures between 10 and 65°C with a ramping and scanning rate of 1°C/min at 260 nm. All samples were measured at least three times. T_m values from each measurement were calculated using the "fitting curve" method and presented as an average of three independent measurements. A full list of melting temperatures can be found in the main-text.

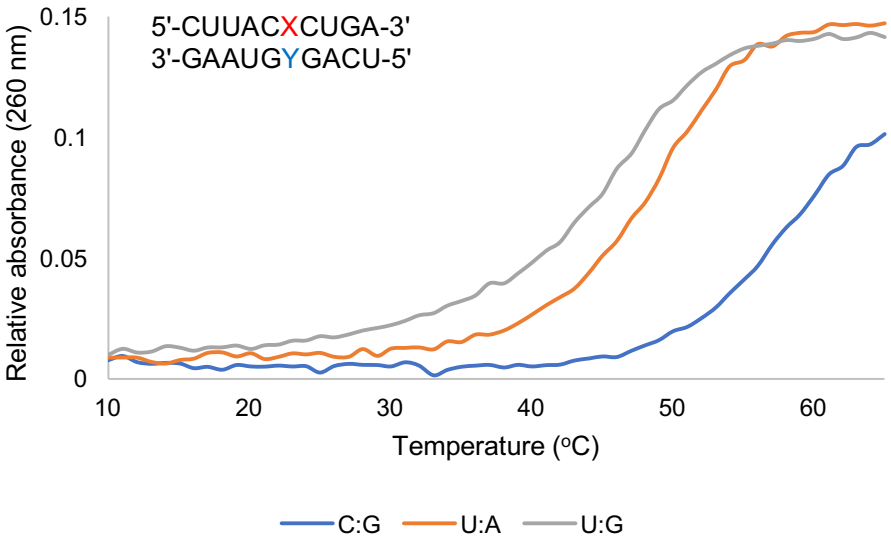


Figure S9. UV melting curves of dsRNA containing X:Y = C:G, U:A, U:G base pairs

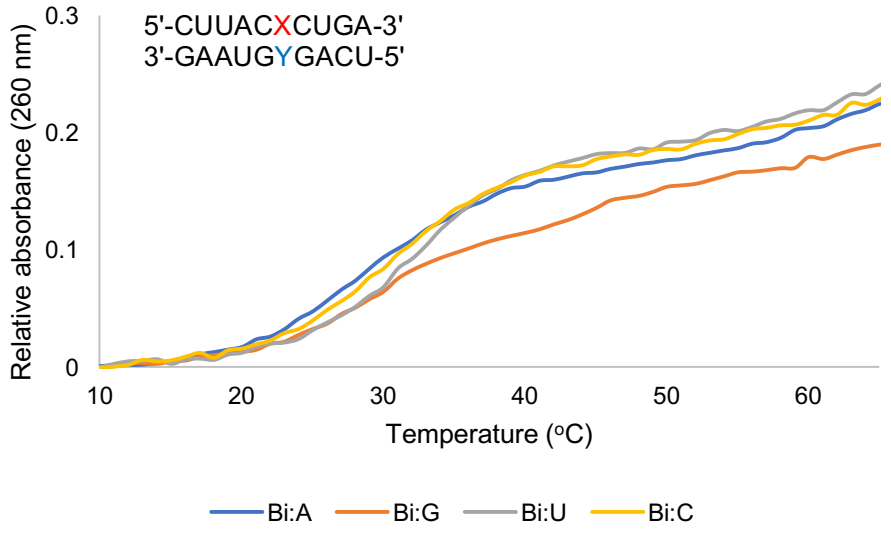


Figure S10. UV melting curves of dsRNA containing X =Bi and Y = A, G, U, C

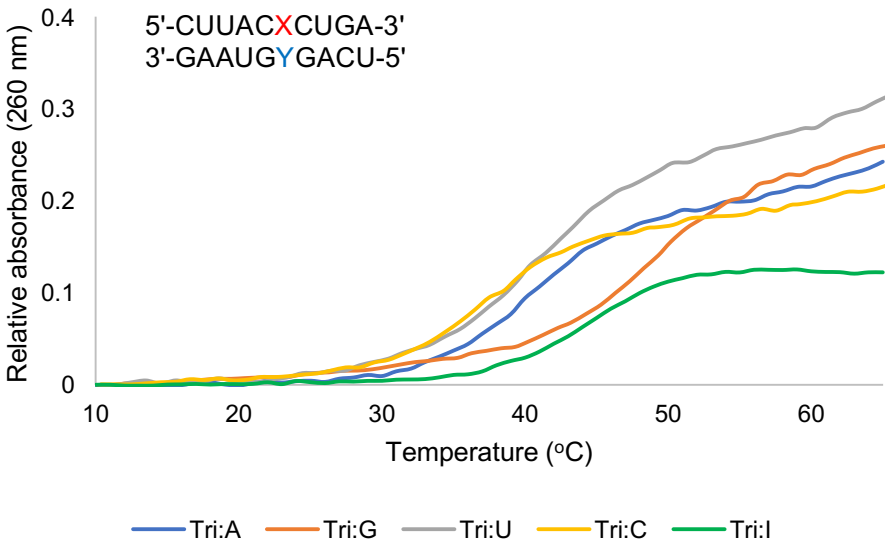


Figure S11. UV melting curves of dsRNA containing X =Tri and Y = A, G, U, C, I

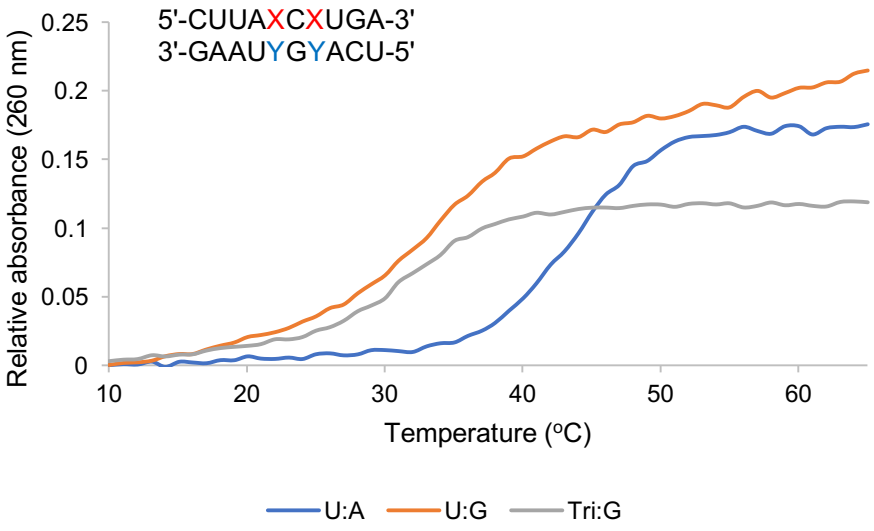


Figure S12. UV melting curves of dsRNA containing X:Y = U:A, U:G, Tri:G in two positions

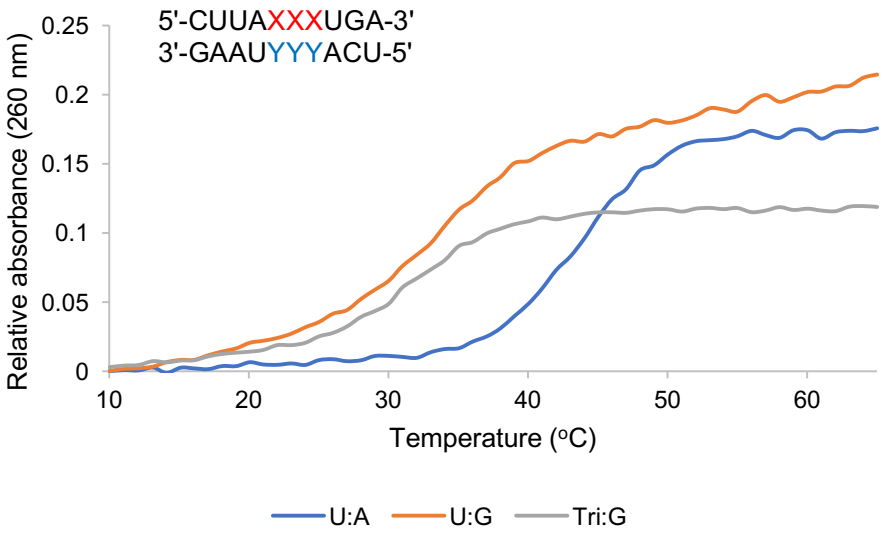


Figure S13. UV melting curves of dsRNA containing X:Y = U:A, U:G, Tri:G in three positions

8. CD Measurements

CD Spectra were measured on a JASCO CD Spectrophotometer using 10 mm QS cuvettes, purchased from Hellma Analytics. A solution (80 µL) of equimolar amounts of oligonucleotides (4 µM each) in the buffer solution containing 10 mM sodium phosphate buffer (pH 7.0) and 150 mM NaCl was heated at 50°C for 5 min and gradually cooled to 4°C prior to the measurement. The CD spectra were recorded at 5°C temperature increments between 10 and 65 °C, and at wavelengths between 400 and 210 nm. All samples were measured at least three times.

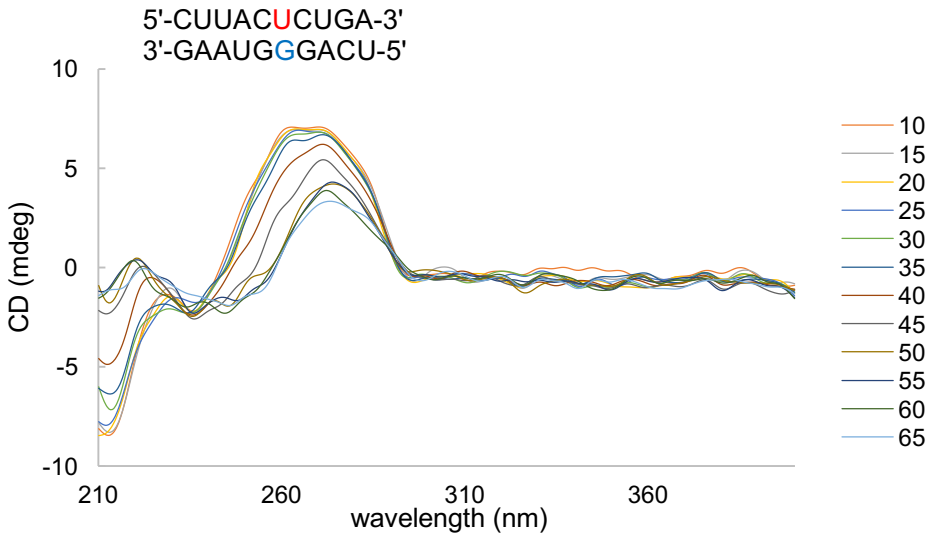


Figure S14. CD spectra of the dsRNA containing U:G base pair

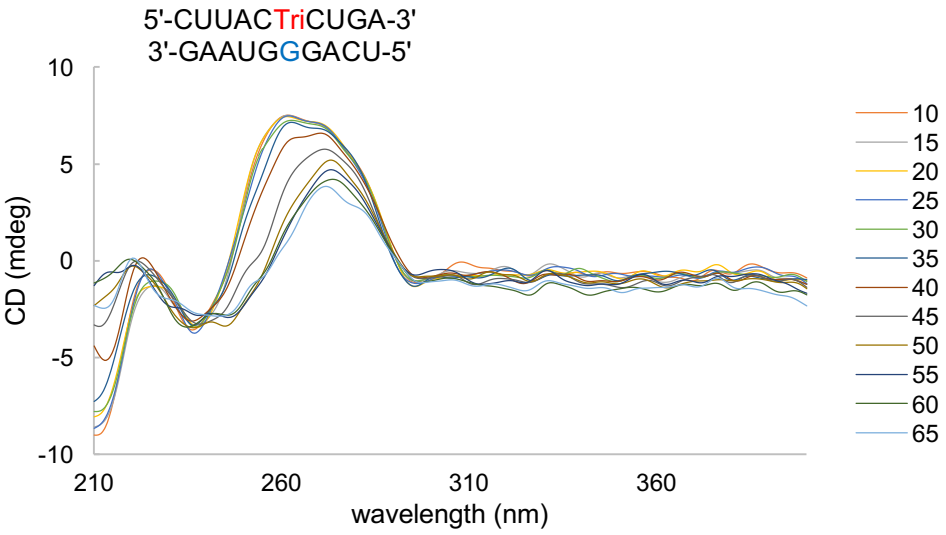


Figure S15. CD spectra of the dsRNA containing Tri:G base pair

9. High-resolution NMR Studies of dsRNA Oligonucleotides

The dsRNA samples (GGUXGACC, where X=C, U, or Triuret) were purified by HPLC and exchanged into NMR buffer (10 mM sodium phosphate, 150 mM NaCl, 0.05 mM, 10% D_2O), and filled into D_2O-matched 5mm Shigemi NMR tubes. The final sample concentrations were 0.8, 0.8 mM, and 0.86 mM for the X=C, X=U and X=Tri samples, respectively. After a series of NMR measurements, the samples were lyophilised and redissolved in D_2O, and a new series of NOESY spectra were recorded. Experiments were performed on a Bruker Avance III spectrometer operating on 800 MHz ^1H Larmor frequency equipped with a 5mm triple channel ^1H, ^{13}C, ^{15}N cryoprobe. Spectra were acquired and processed using Topspin 3.2 and further analysed by NMRFAM-Sparky [4]. The experiments were carried out at 298K. ^1H-^1H 2D NOESY spectra were measured with mixing time points at 40, 80, 120, 160, 200, 240, 280, and 320 ms; with 4098 x 256 complex points, 102 and 10.2 ms acquisition times in the direct and indirect dimension, with 1 s recycle delay, and 64 number of scans. Water suppression was achieved using excitation sculpting with gradients. The experimental time was ~6 hours for each 2D NOESY spectra. ^1H-^1H 2D TOCSY spectra were recorded with a mixing time of 30 ms using MLEV17 mixing scheme, with 4096 x 1024 complex points, 102 and 64 ms acquisition times in the direct and indirect dimension, with 1 s recycle delay, and 16 number of scans. Water suppression was achieved using excitation sculpting with gradients.

The resonance assignment of the NMR spectra followed common procedures. First, the fingerprint region (H8/H6 correlations with H5/H1') of the TOCSY/NOESY spectra were analysed. The peaks which appear both in the TOCSY and NOESY spectra correspond to the H5-H6 cross peaks of cytosine and uracil bases. The well-separated and intense cross peaks of CH5 base protons with CH41/H42 amine protons facilitated the discrimination between U and C bases. Second, the bases were sequentially connected based on the cross peaks between the H5/H8 of the i^{th} base and the H1' proton of the $(i\text{-}1)^{th}$ sugar. Third, the imino region was analysed (Fig. S8). The resonance of G1H1 was broadened beyond detection. The resonance for G2H1 was identified based on its cross peaks with C7H41, C7H42 of the other strand (that is C15H41, C15H42). In the spectrum of GGUTriGACC three extra "imino-like" peaks appeared at 9.68, 8.76, and 8.64 ppm which were assigned to be the amide protons of the triuret base. Their identities were confirmed by several inter- and intrastrand NOESY cross peaks. Finally, the sugar protons were assigned based on their sequential connections with the neighbouring sugar protons.

	Sugar protons						Base protons							
	H1'	H2'	H3'	H4'	H5'	H5''	H1	H2,H21, H22	H3	H41, H42	H5	H6	H71, H72	H8
G1	5.64	4.79	4.49	4.24	3.95	3.86								7.91
G2	5.89	4.51	4.59		4.12	4.12	13.42							7.52
U3	5.61	4.55	4.44		4.08	4.08			14.73		5.09	7.73		
Tri4	5.37	4.49	4.92	4.32	4.02	4.02	8.64		9.67		8.77		7.86, 7.03	
G5	5.61	4.72			4.16	4.16	10.49	6.01						7.70
A6	5.79	4.48	4.58	4.33	4.16	4.03		7.89						8.08
C7	5.39	4.17	4.35	4.49	4.04	4.04				8.41, 6.97	5.20	7.48		
C8	5.72	4.03	4.16	4.49	3.97	3.97				8.3, 6.88	5.49	7.48		7.66

Table S2. Assigned ^1H chemical shifts of the 5'-GGUTriGACC-3' dsRNA sample

	Sugar protons						Base protons						
	H1'	H2'	H3'	H4'	H5'	H5''	H1	H2,H21, H22	H3	H41, H42	H5	H6	H8
G1	5.73	4.90	4.53	4.32	4.01	3.90							7.99
G2	5.93	4.55	4.59		4.15	4.15	13.43						7.50
U3	5.61	4.66	4.43		4.12	4.12			14.25		5.13	7.71	
U4	5.70	4.48	4.56	4.49	4.14				11.87		5.86	7.97	
G5	5.66	4.72	4.45	4.54	4.17	4.17	10.08	6.20					7.76
A6	5.89	4.52	4.65		4.14	4.14		7.91					7.88
C7	5.42	4.15	4.37	4.52	4.08	4.03				8.37, 7.01	5.21	7.49	
C8	5.74	4.03	4.16		4.04	4.04				8.36, 6.87	5.46	7.65	

Table S3. Assigned ^{1}H chemical shifts of the 5'-GGUUGACC-3' dsRNA sample

	Sugar protons						Base protons						
	H1'	H2'	H3'	H4'	H5'	H5''	H1	H2,H21, H22	H3	H41, H42	H5	H6	H8
G1	5.638	4.827	4.486	4.239	3.93	3.853							7.912
G2	5.878	4.503	4.576		4.117		13.386						7.472
U3	5.575	4.507	4.108		4.106	4.106			14.386		5.103	7.806	
C4	5.59	4.459	4.12							8.34, 6.76	5.645	7.838	
G5	5.601	4.13	4.529	4.611			12.09						7.558
A6	5.926	4.512	4.642	4.365	4.134	4.028		7.739					7.931
C7	5.373	4.143	4.347	4.491	4.363	4.325				8.34, 6.95	5.211	7.467	
C8	5.712	4.028	4.143	4.487	3.974	3.86				8.32, 6.85	5.468	7.628	

Table S4. Assigned ^{1}H chemical shifts of the 5'-GGUCGACC-3' dsRNA sample

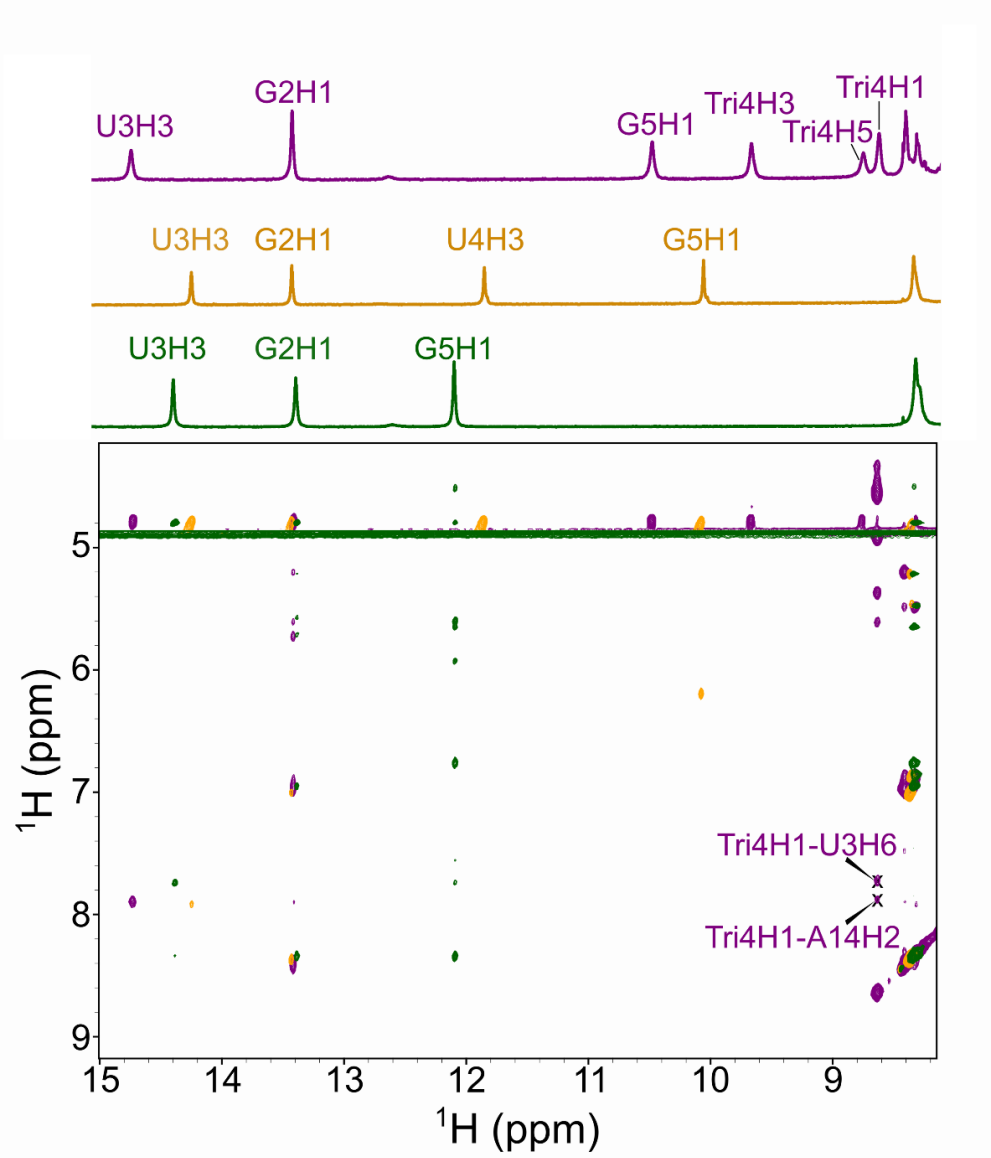


Figure S16. 1D and 2D imino proton spectra of 5'-GGUCGACC-3' (green), 5'-GGUUGACC-3' (yellow), and 5'-GGUtriGACC-3' (purple) at 298 K. Labels indicate the assigned imino and amid (triuret protons) resonanonces. In the 320 ms NOESY spectra of 5'-GGUTriGACC-3' cross-peaks indicate inter- and intrastrand NOE connections between the H1 amide proton of triuret and the H2 proton of A14 (same as A6), and between H1 amide proton of triuret and H6 proton of the neighbouring U3 base.

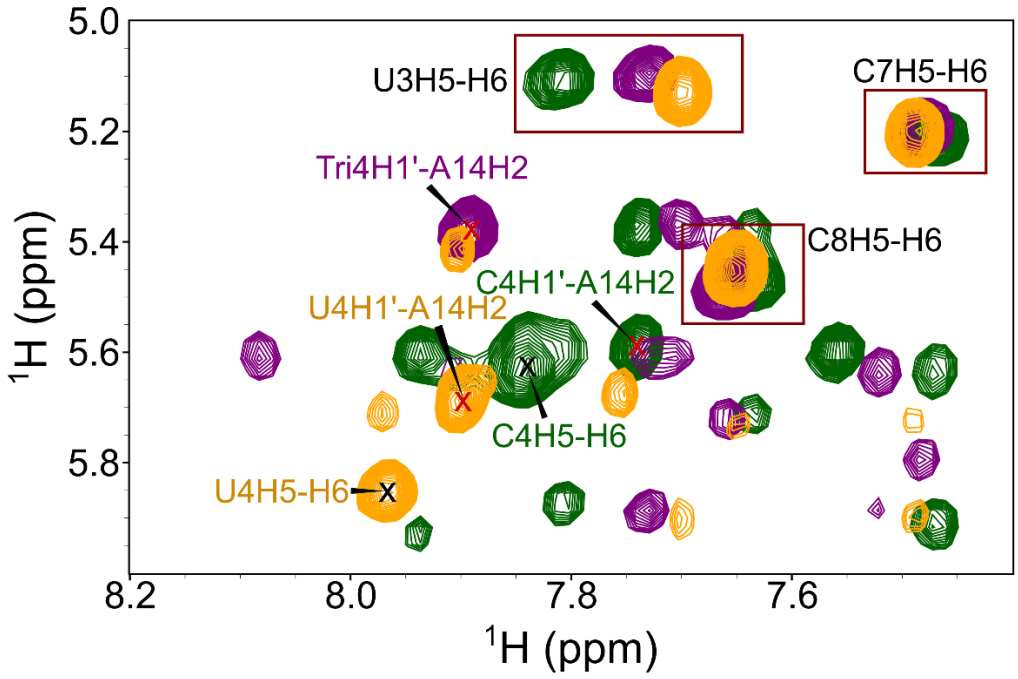


Figure S17. The fingerprint region of the 2D ^{1}H-^{1}H NOESY spectra of the three samples: 5'-GGUCGACC-3' (green), 5'-GGUUGACC-3' (yellow), and 5'-GGUXGACC-3' (purple) measured with a mixing time of 300 ms in D2O at 298 K at 800 MHz ^{1}H Larmor frequency. Boxed regions highlight the cross-peaks of the same spin pairs in the spectra. Red crosses show the inter-strand cross-peaks between the H1' sugar proton of the base at position 4 (cytosine, uracil or triuret) and the H2 base proton of adenine 14 (same as Ade6). The overall spectral similarity at the fingerprint region, as well as the strong X4H1'-A14H2 cross-peaks confirm that the overall fold of the three samples are comparable.

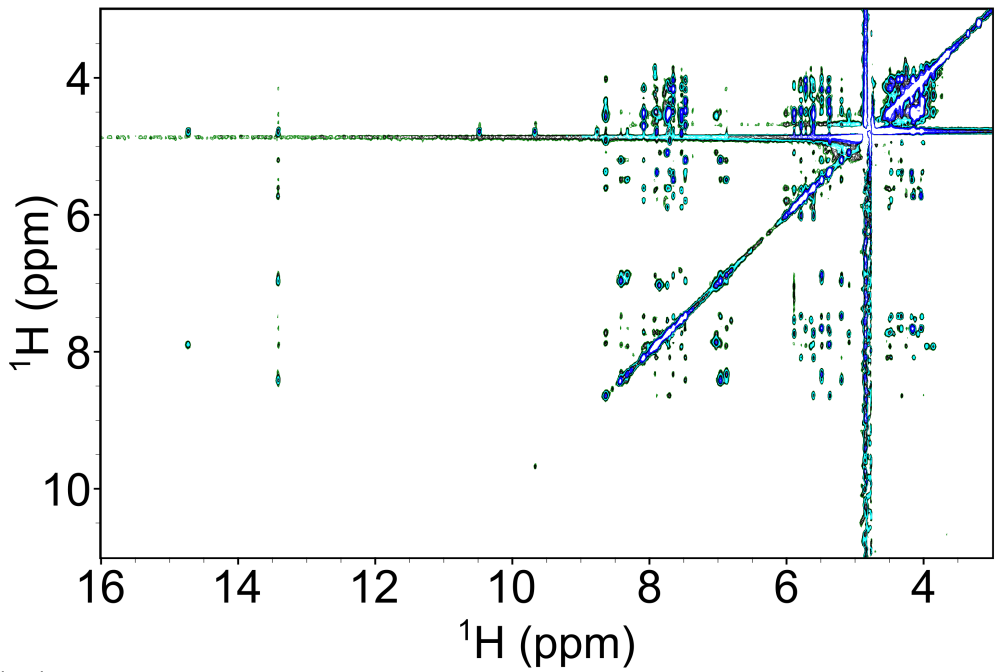


Figure S18. ^{1}H-^{1}H 2D NOESY spectrum of 5'-GGUXGACC-3', where X = Triuret recorded with 320 ms mixing time at 800 MHz ^{1}H Larmor frequency at 298 K.

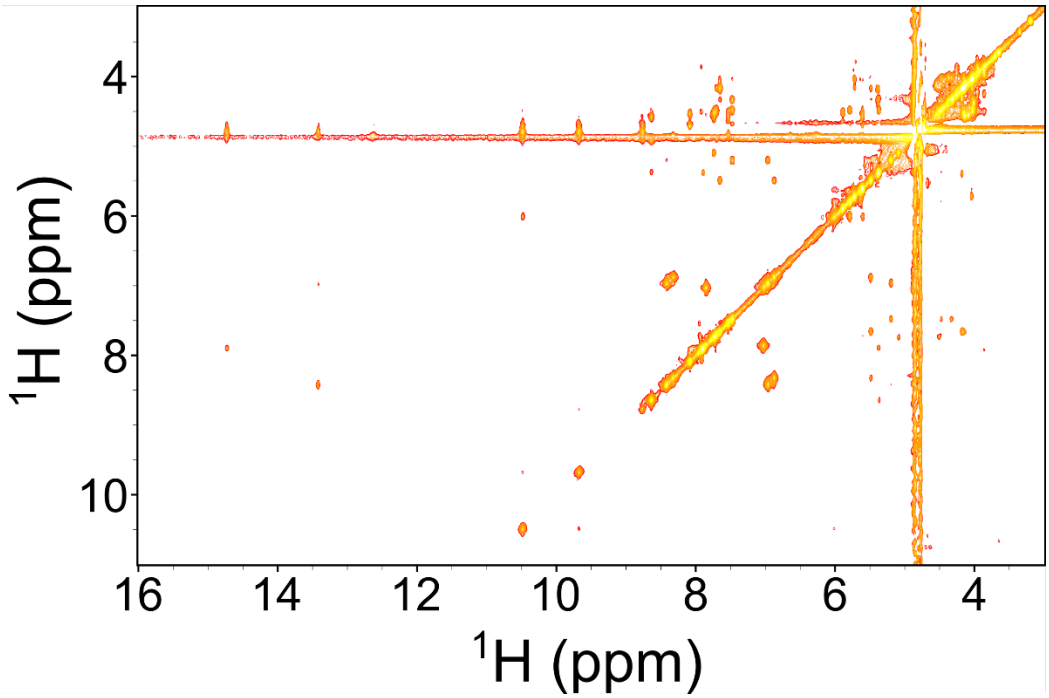
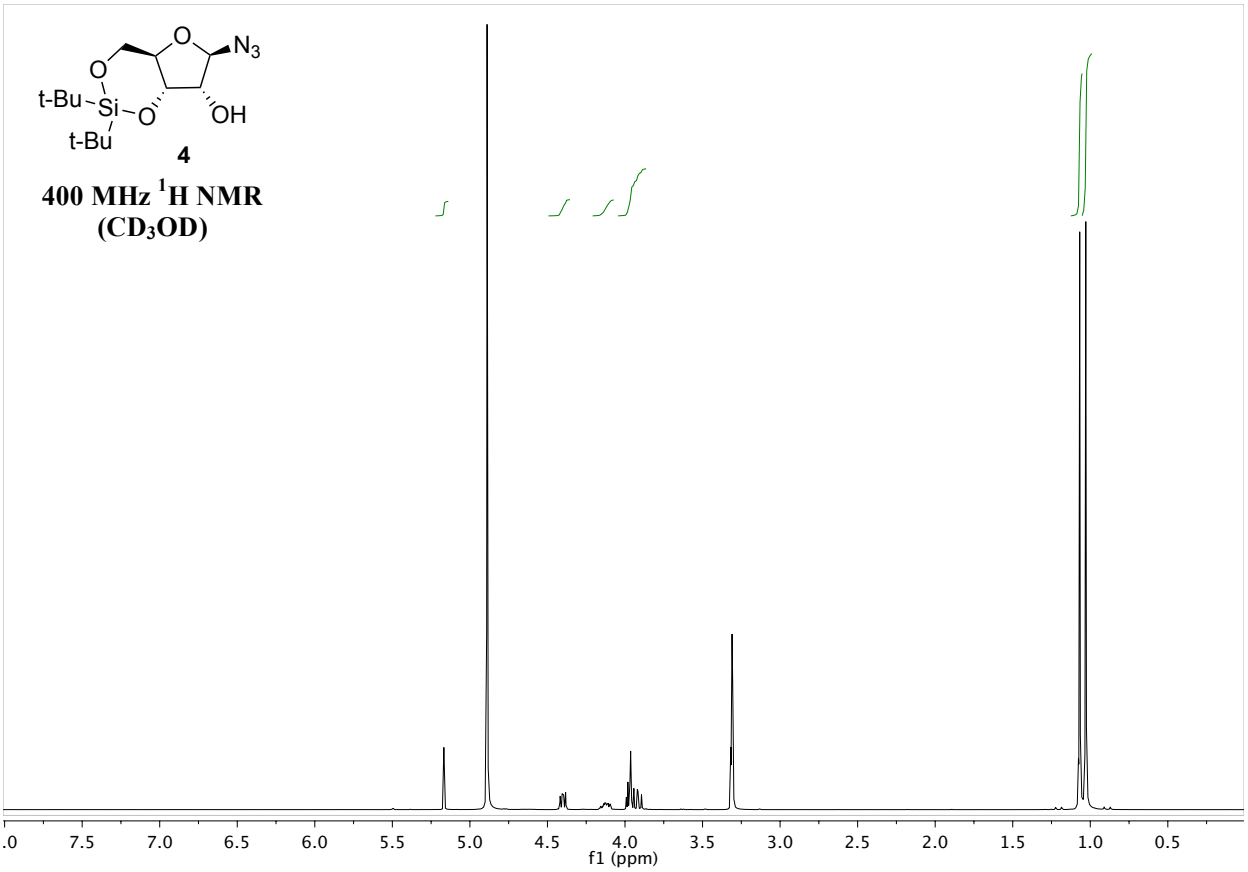
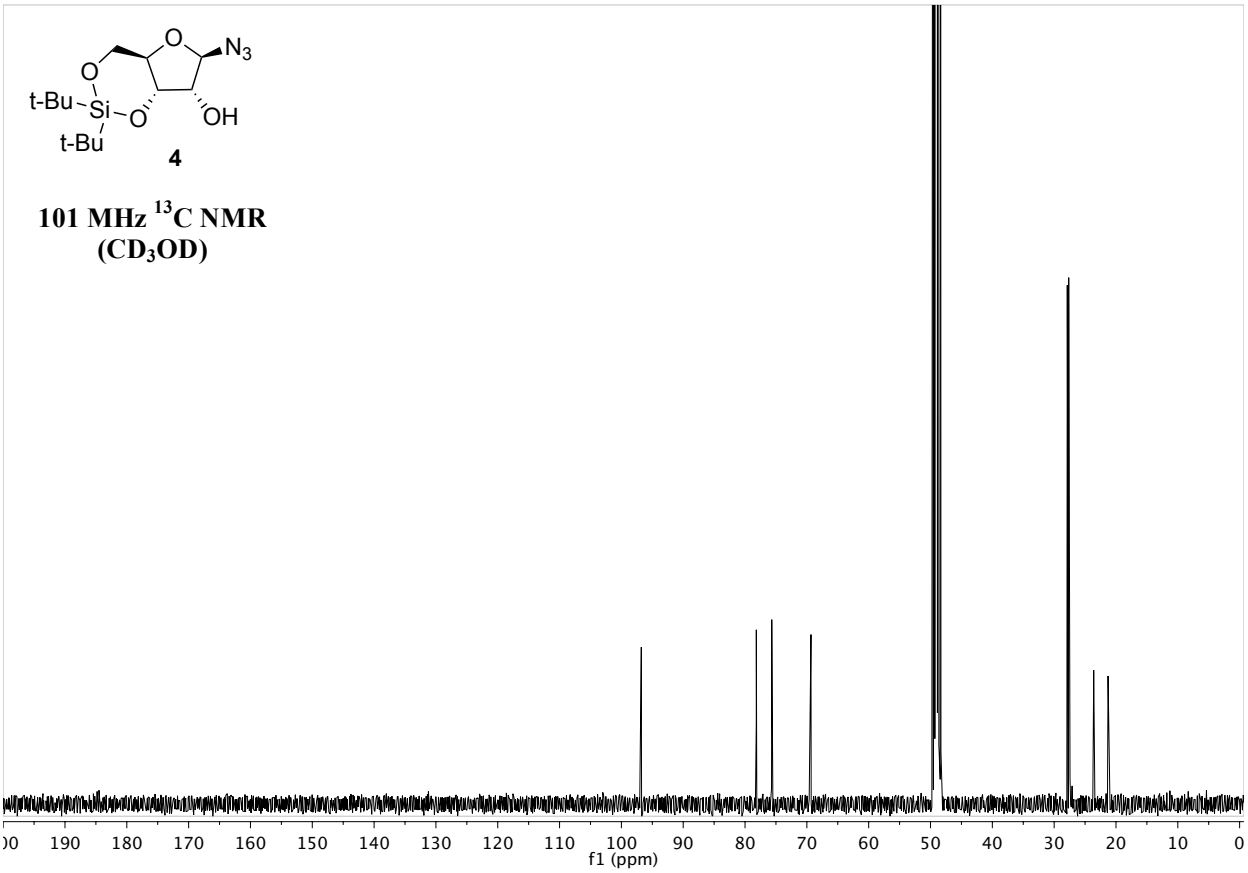
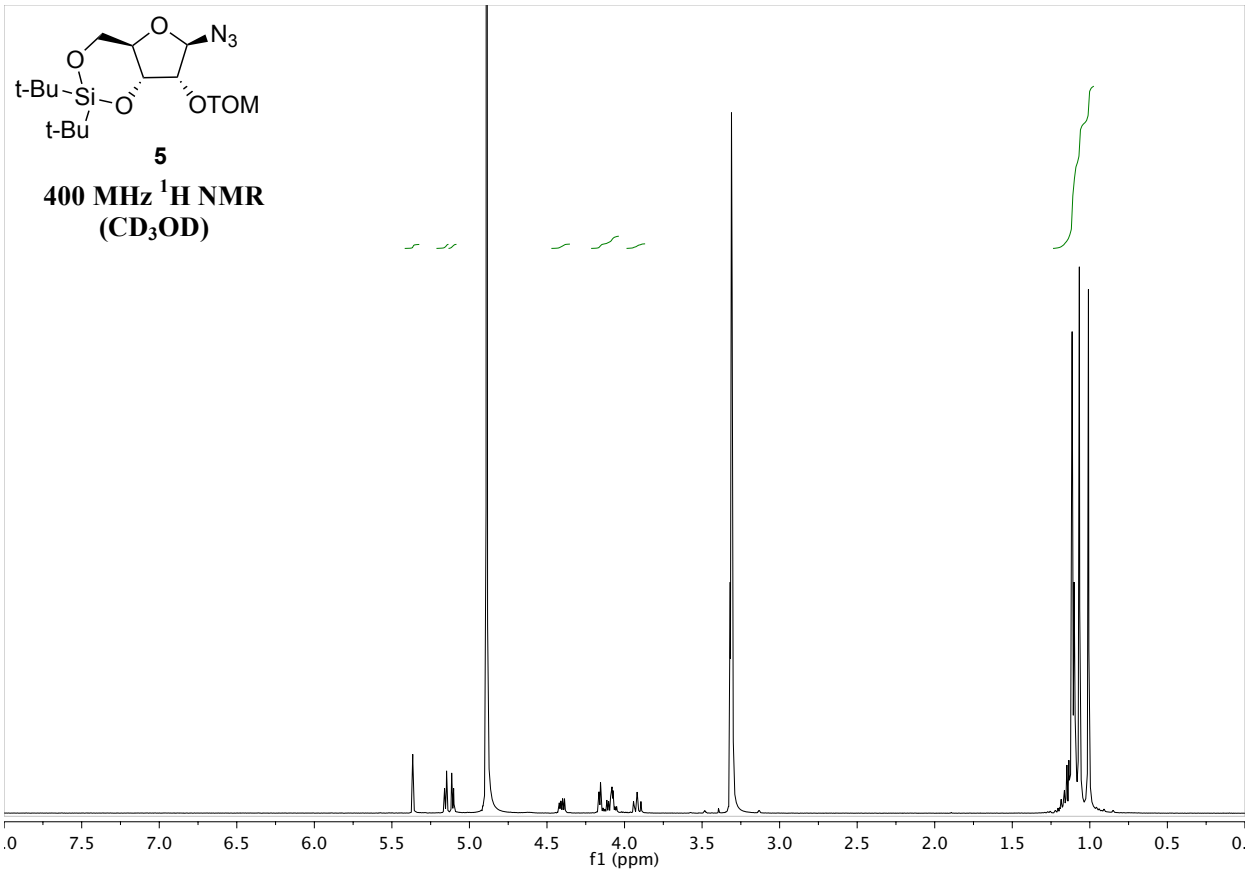


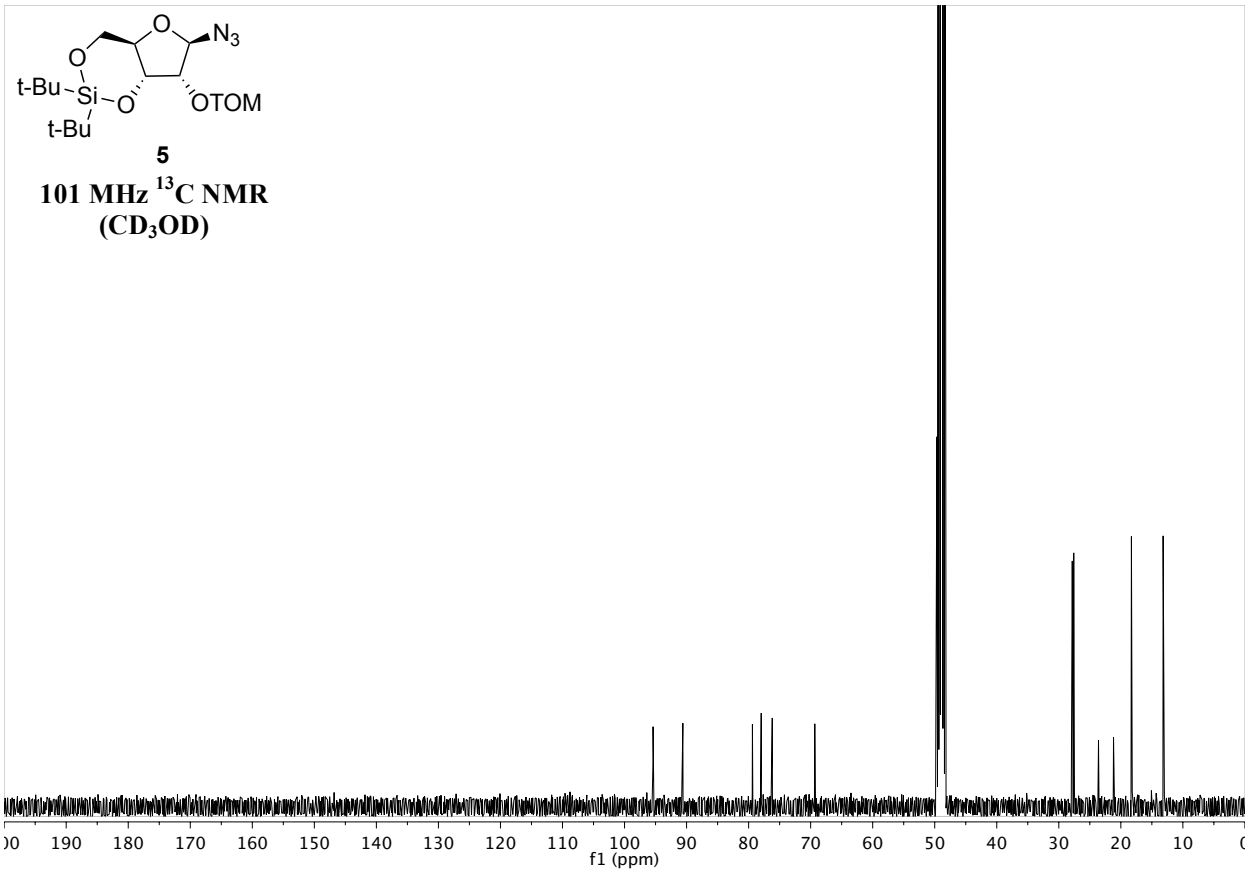
Figure S19. ^{1}H-^{1}H 2D NOESY spectrum of 5'-GGUXGACC-3', where X = Triuret recorded with 40 ms mixing time at 800 MHz ^{1}H Larmor frequency at 298 K.

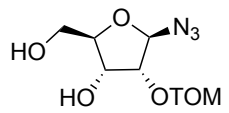
10. NMR Spectra of Synthesized Compounds





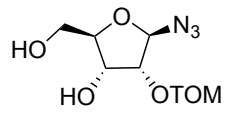

t-Bu
Si
O
t-Bu
O
N3
OTOM
5
400 MHz 1H NMR
(CD3OD)
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
f1 (ppm)


t-Bu
Si
O
t-Bu
O
N3
OTOM
5
101 MHz 13C NMR
(CD3OD)
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
f1 (ppm)



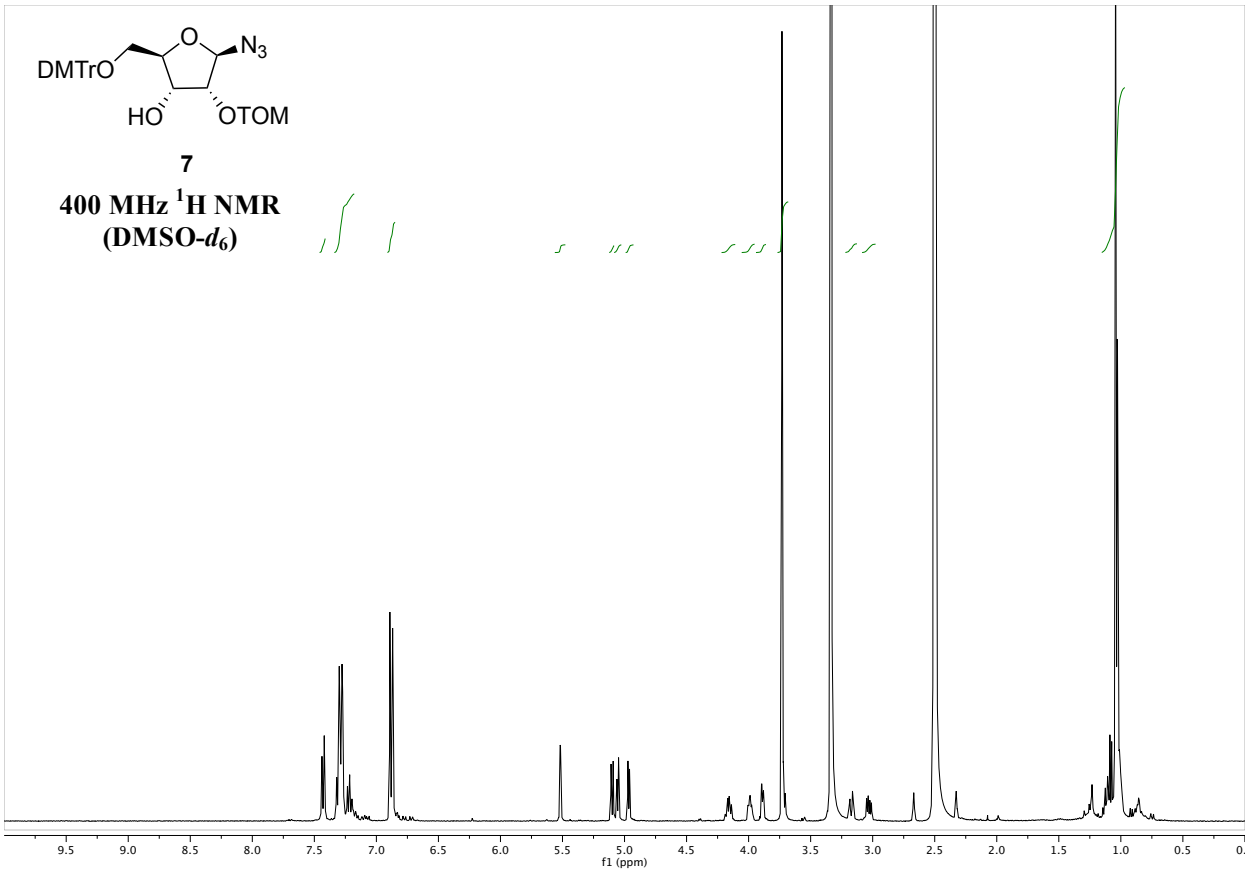
6

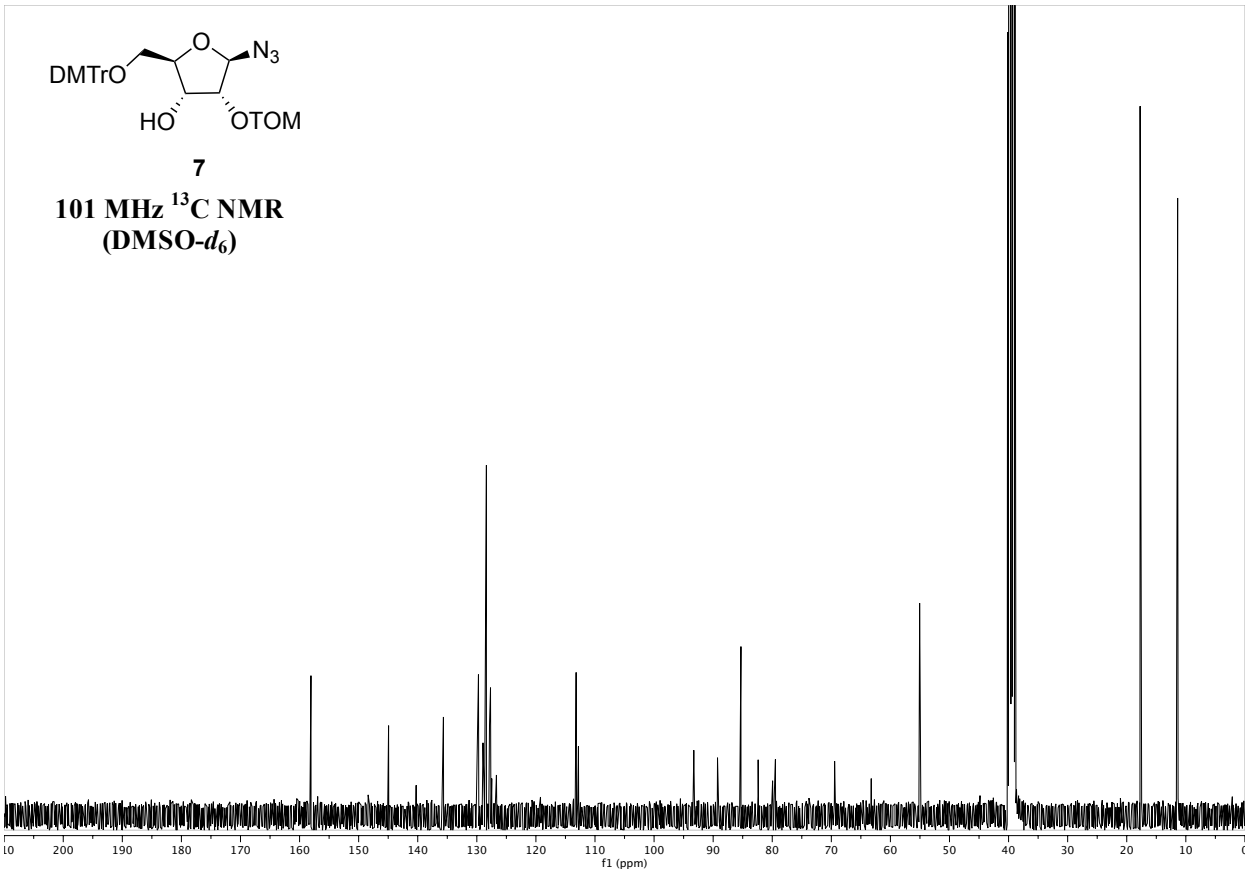
400 MHz ^{1}H NMR
(CD_3OD)

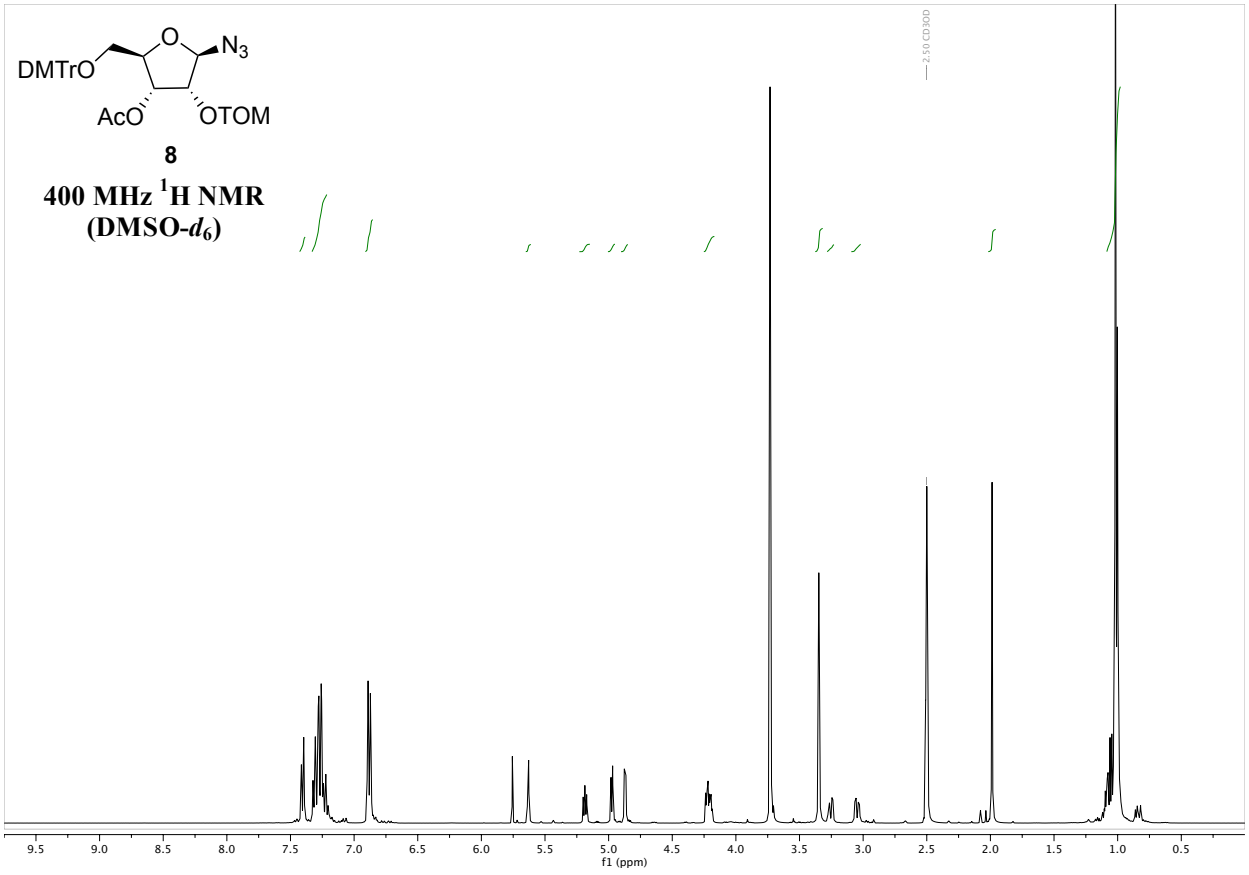


6

101 MHz ^{13}C NMR
(CD_3OD)

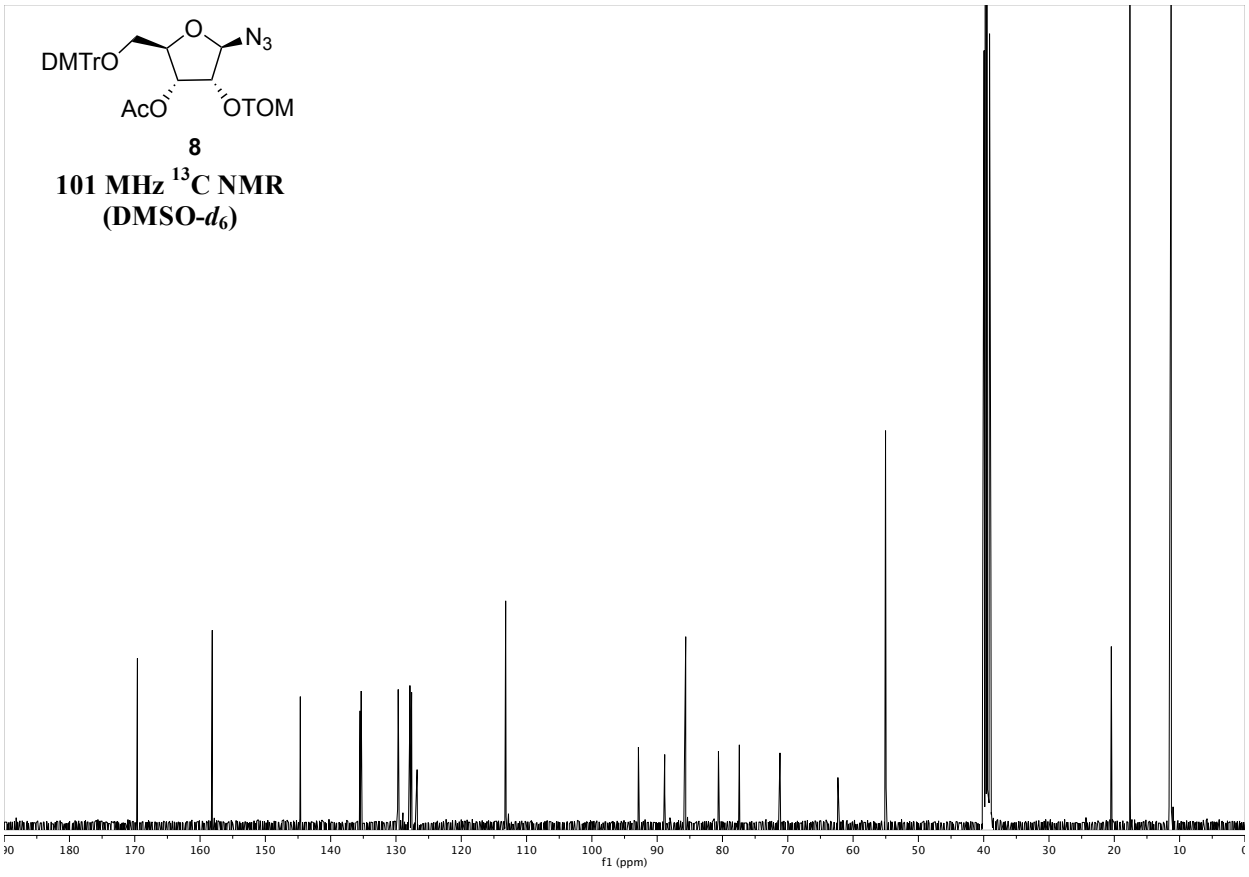

DMTrO
O
N3
HO
OTOM
7
400 MHz 1H NMR
(DMSO-d6)
f1 (ppm)


DMTrO
O
N3
HO
OTOM
7
101 MHz 13C NMR
(DMSO-d6)
f1 (ppm)



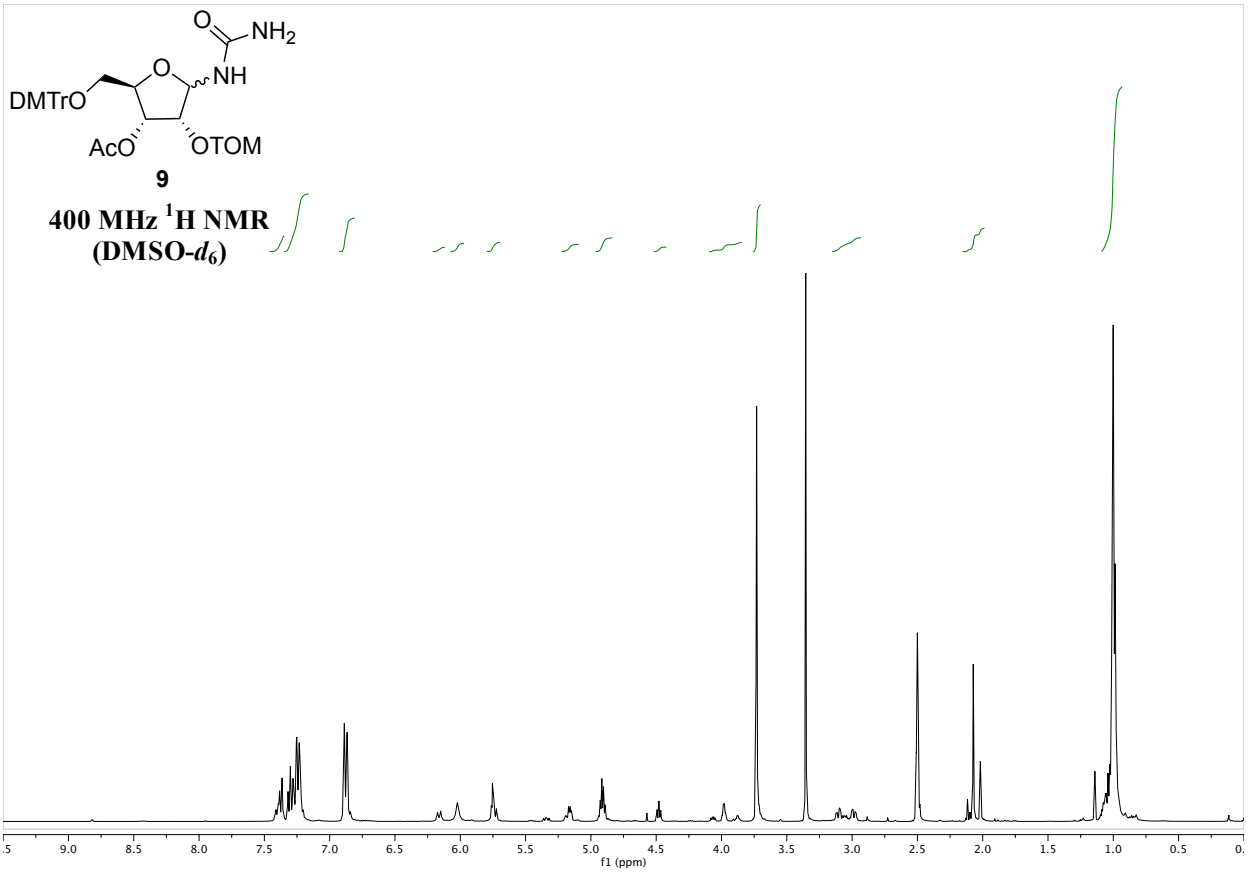
8

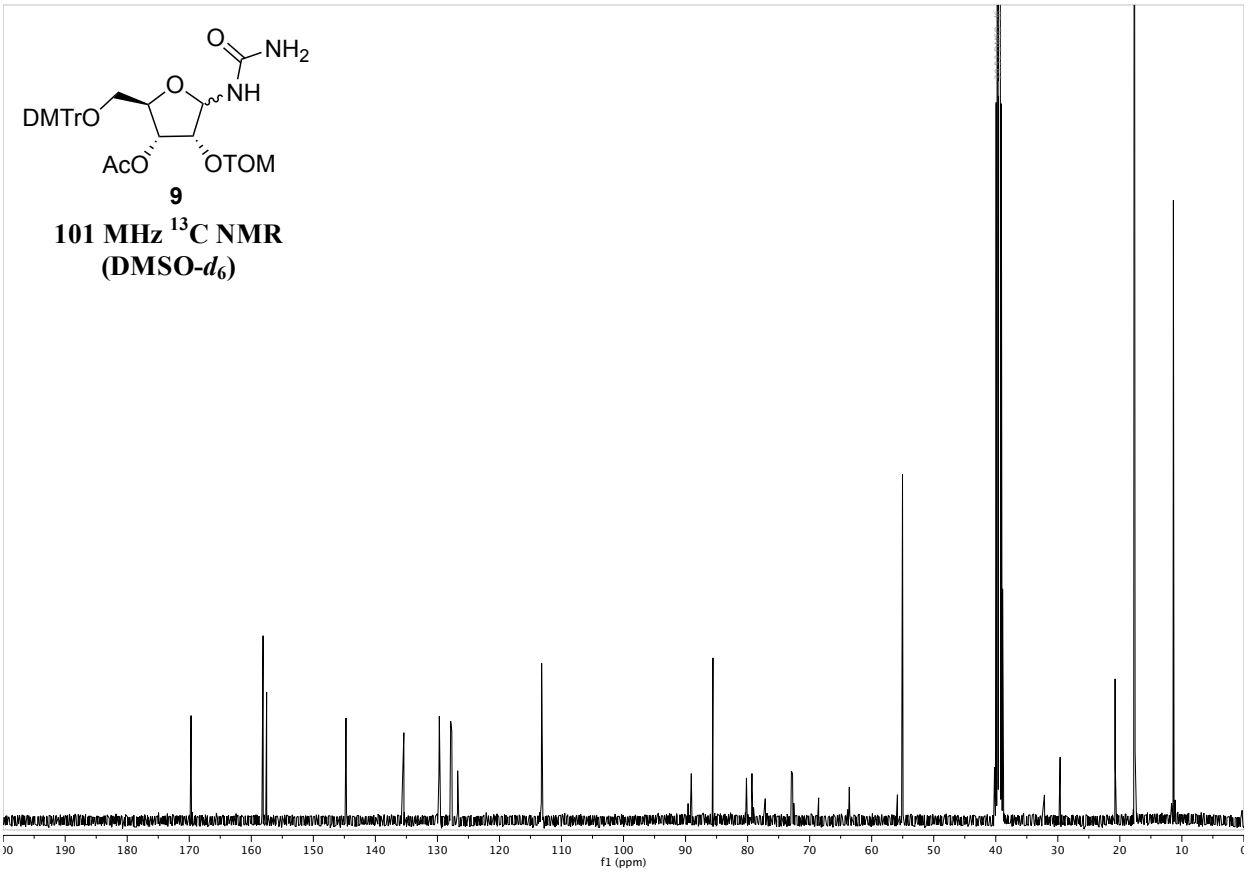
400 MHz ^{1}H NMR
(DMSO-d_6)

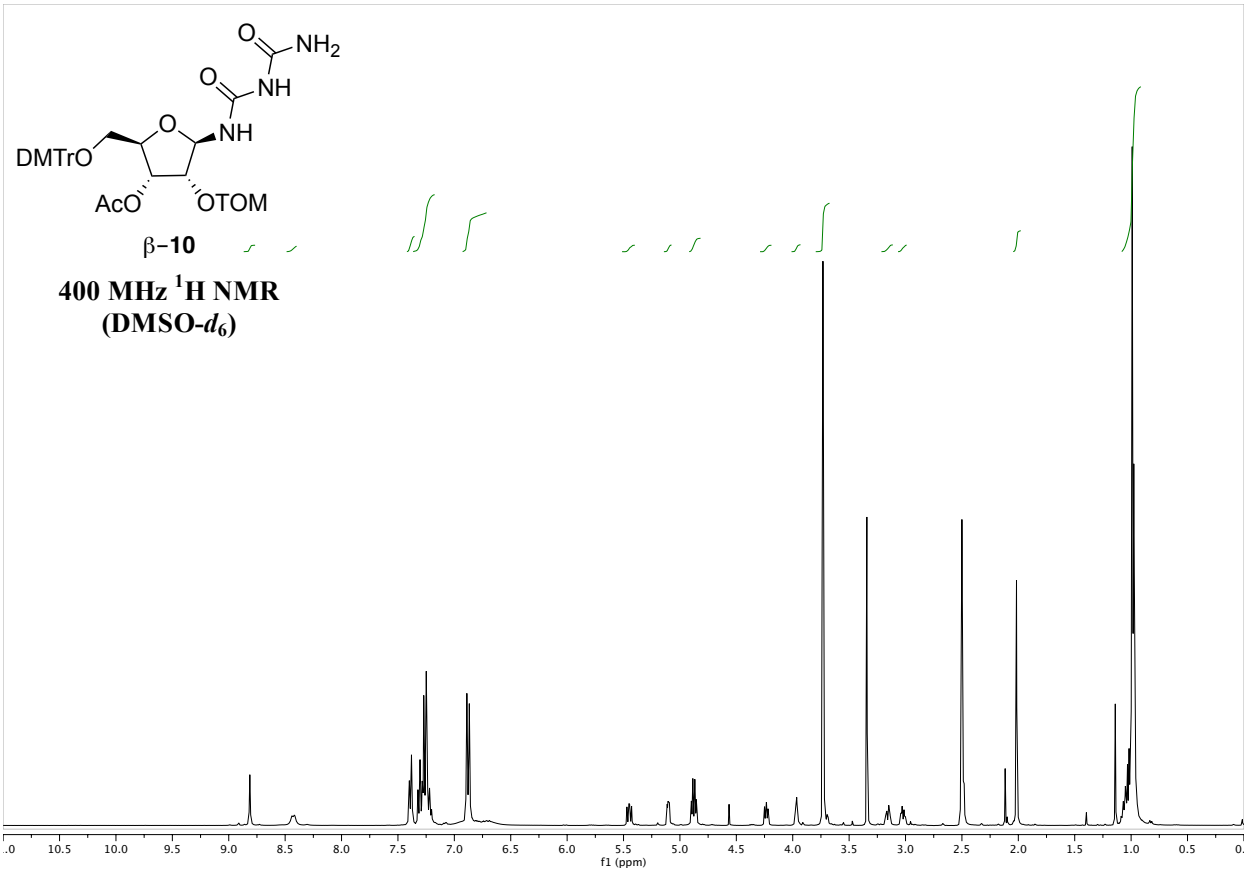


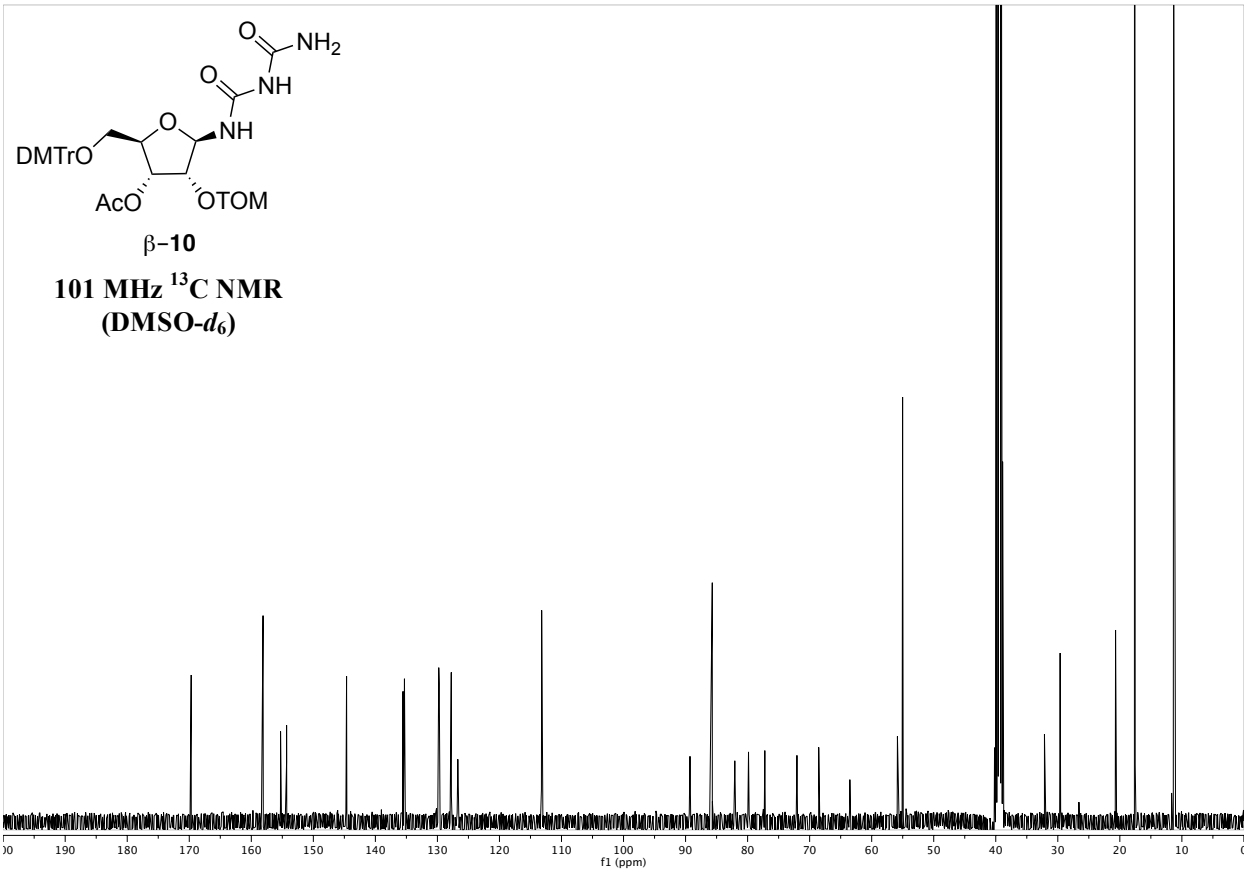
8

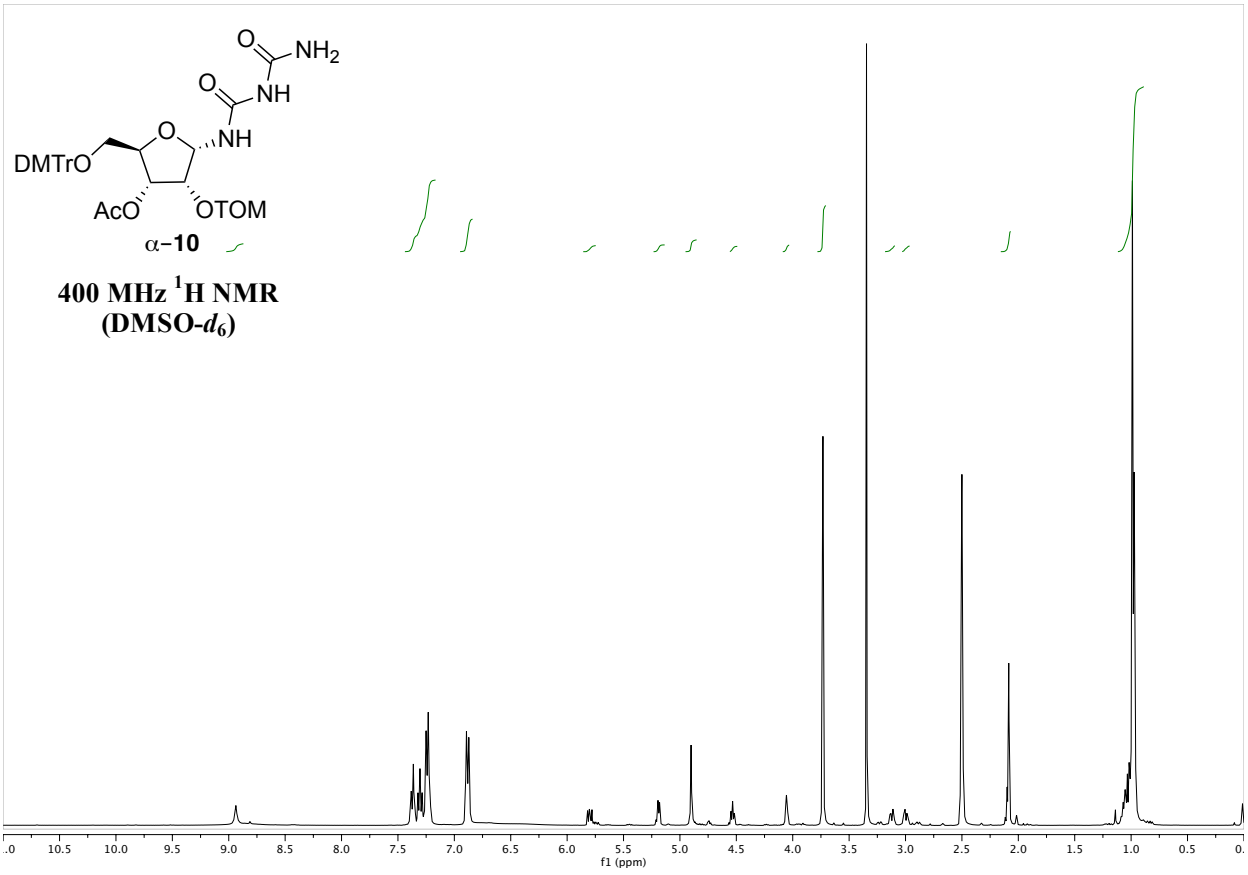
101 MHz ^{13}C NMR
(DMSO-d_6)

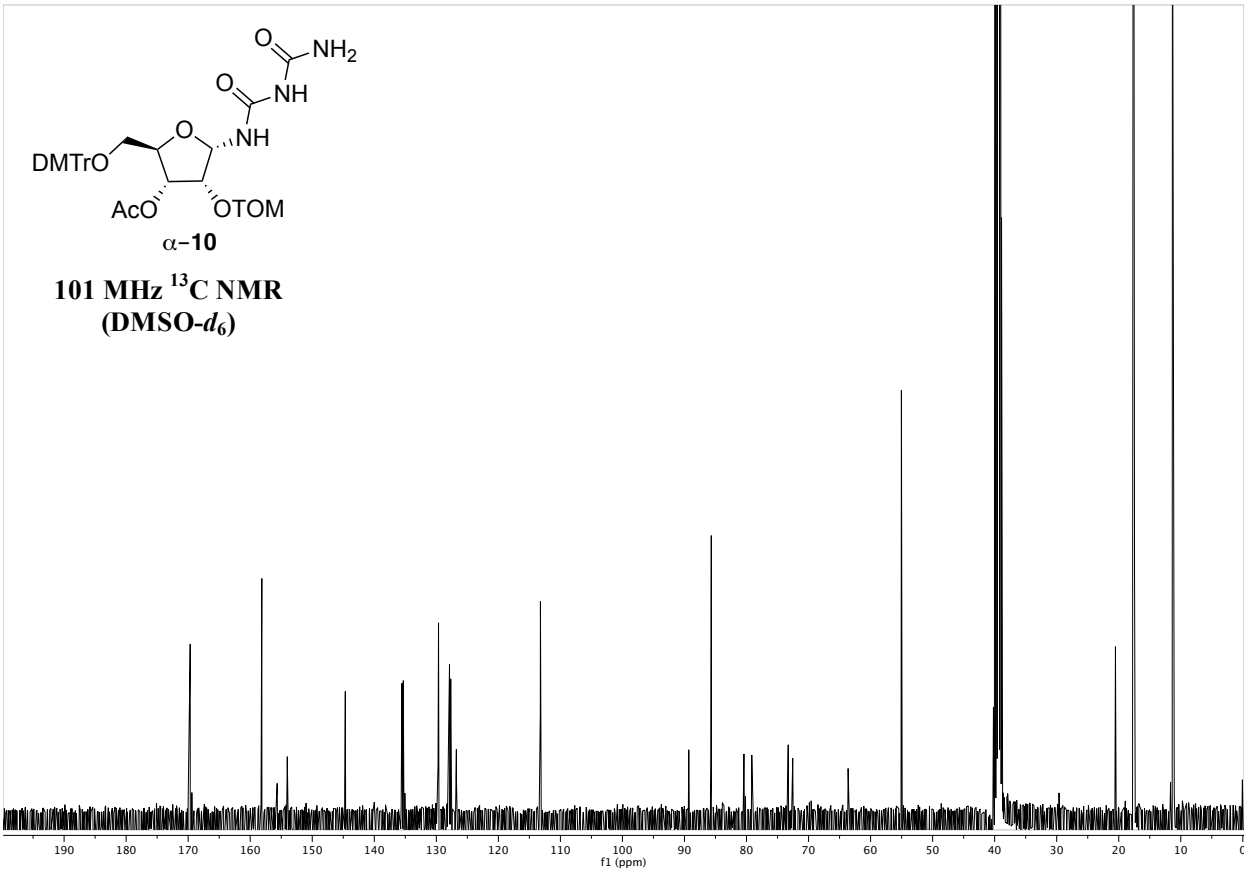

O
NH2
NH
O
DMTrO
AcO
OTOM
9
400 MHz 1H NMR
(DMSO-d6)
f1 (ppm)

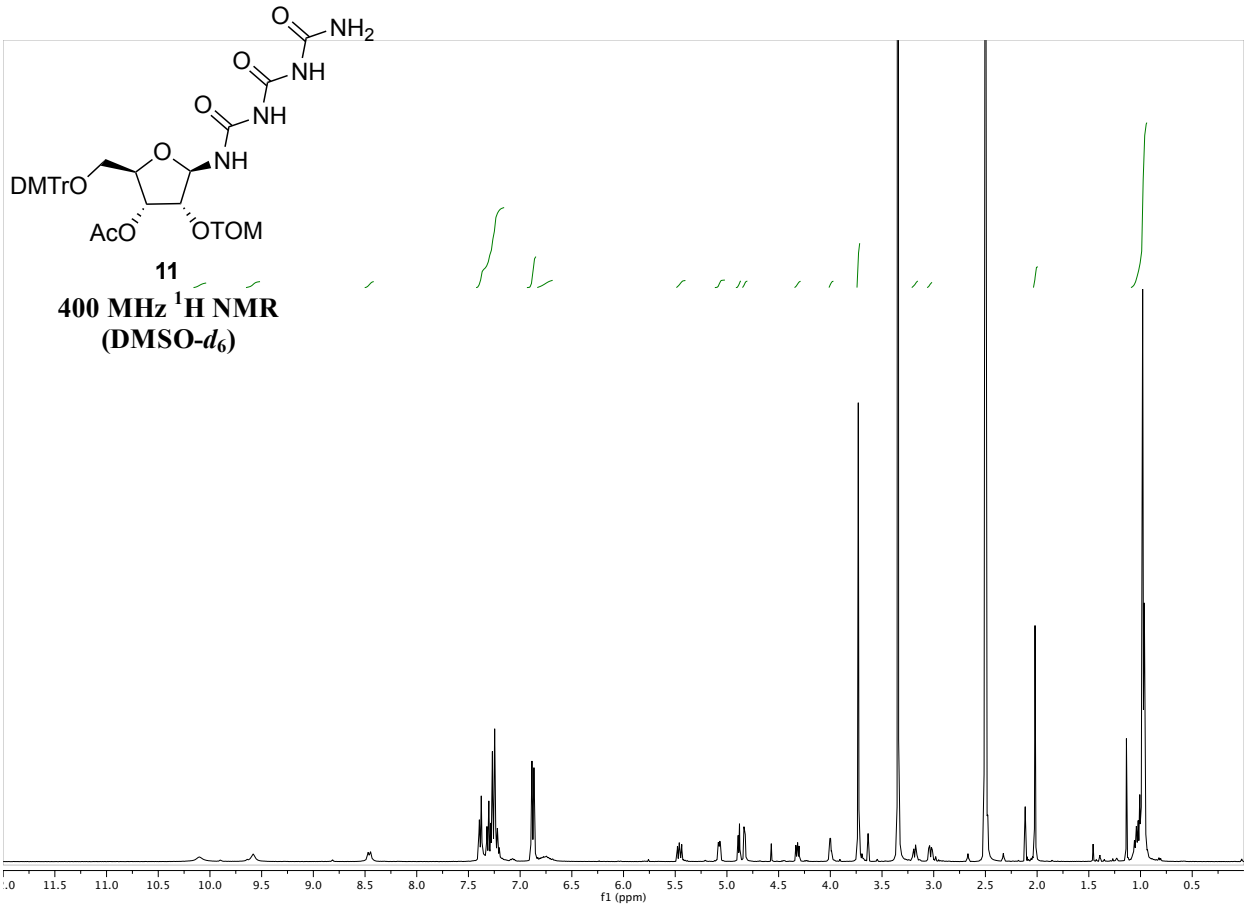

O
NH2
NH
O
DMTrO
AcO
OTOM
9
101 MHz 13C NMR
(DMSO-d6)
f1 (ppm)

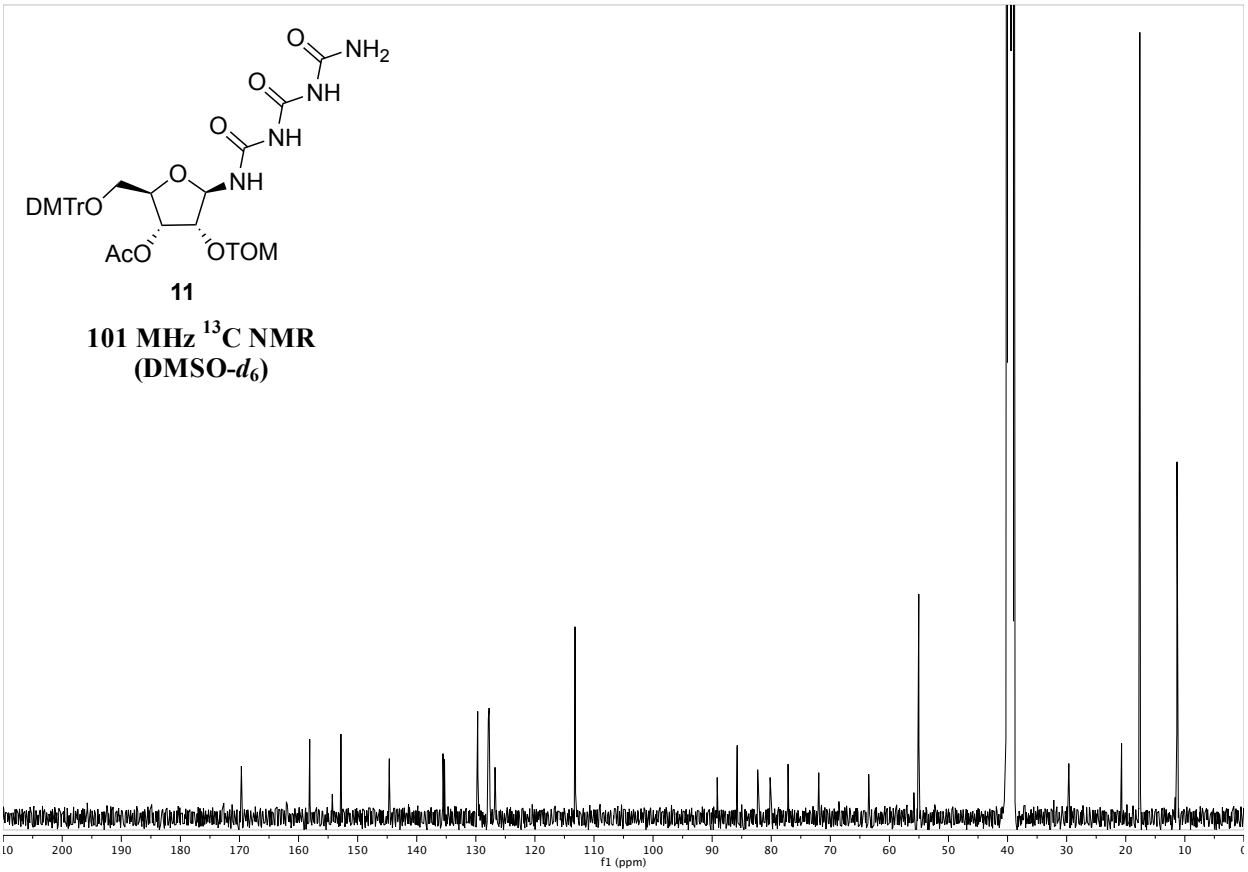

O
NH2
O
NH
NH
O
DMTrO
AcO
OTOM
β-10
400 MHz 1H NMR
(DMSO-d6)
f1 (ppm)

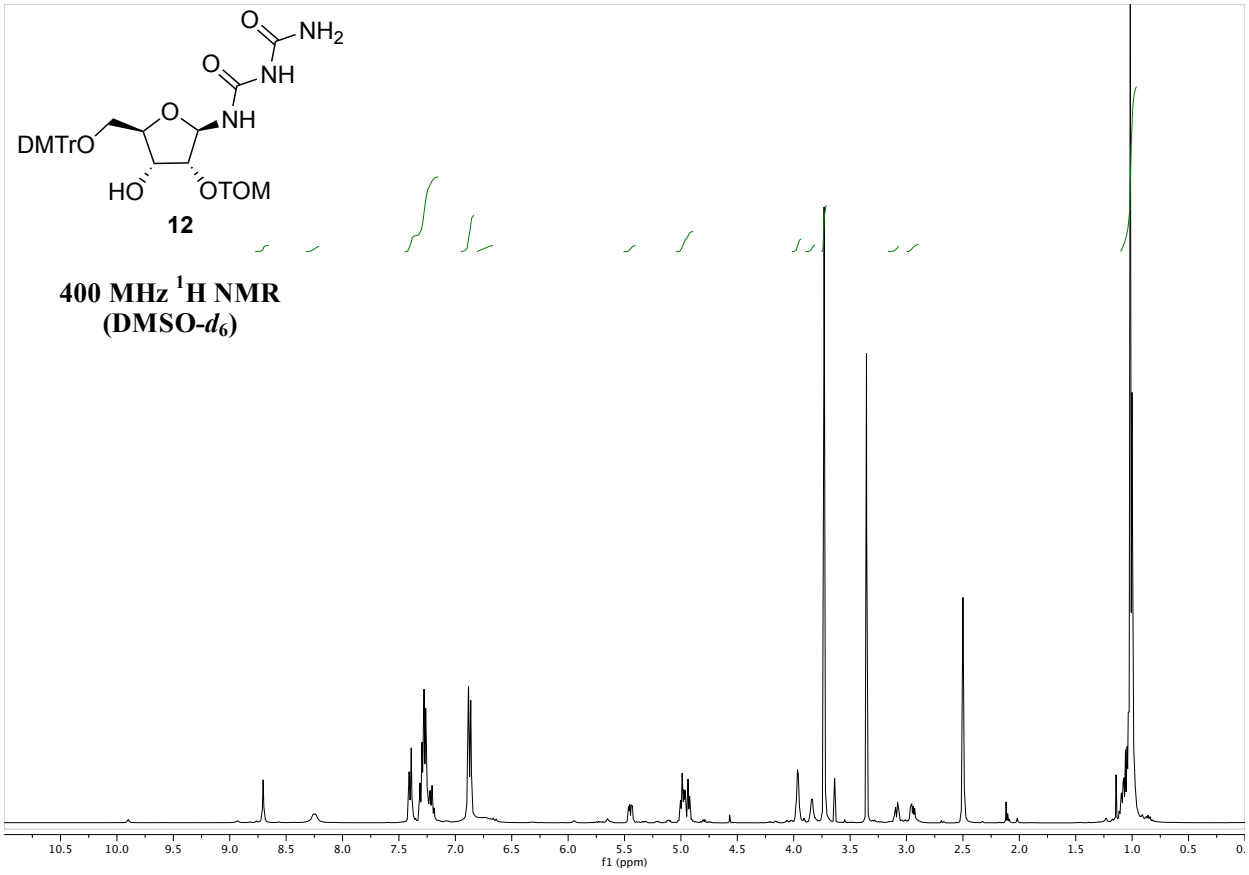

O
NH2
O
NH
NH
O
DMTrO
AcO
OTOM
β-10
101 MHz 13C NMR
(DMSO-d6)
f1 (ppm)

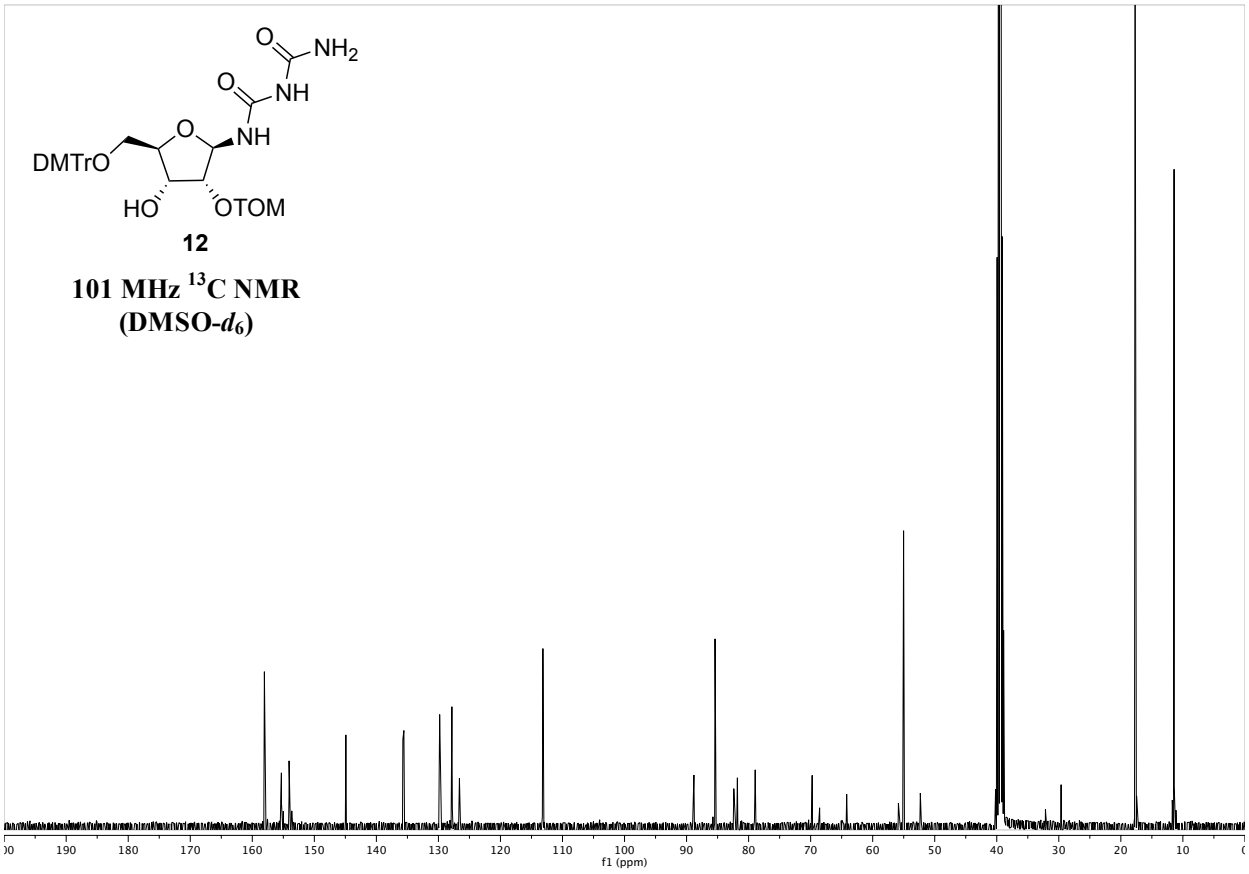

O
NH2
O
NH
O
NH
DMTrO
AcO
OTOM
α−10
400 MHz 1H NMR
(DMSO-d6)
f1 (ppm)

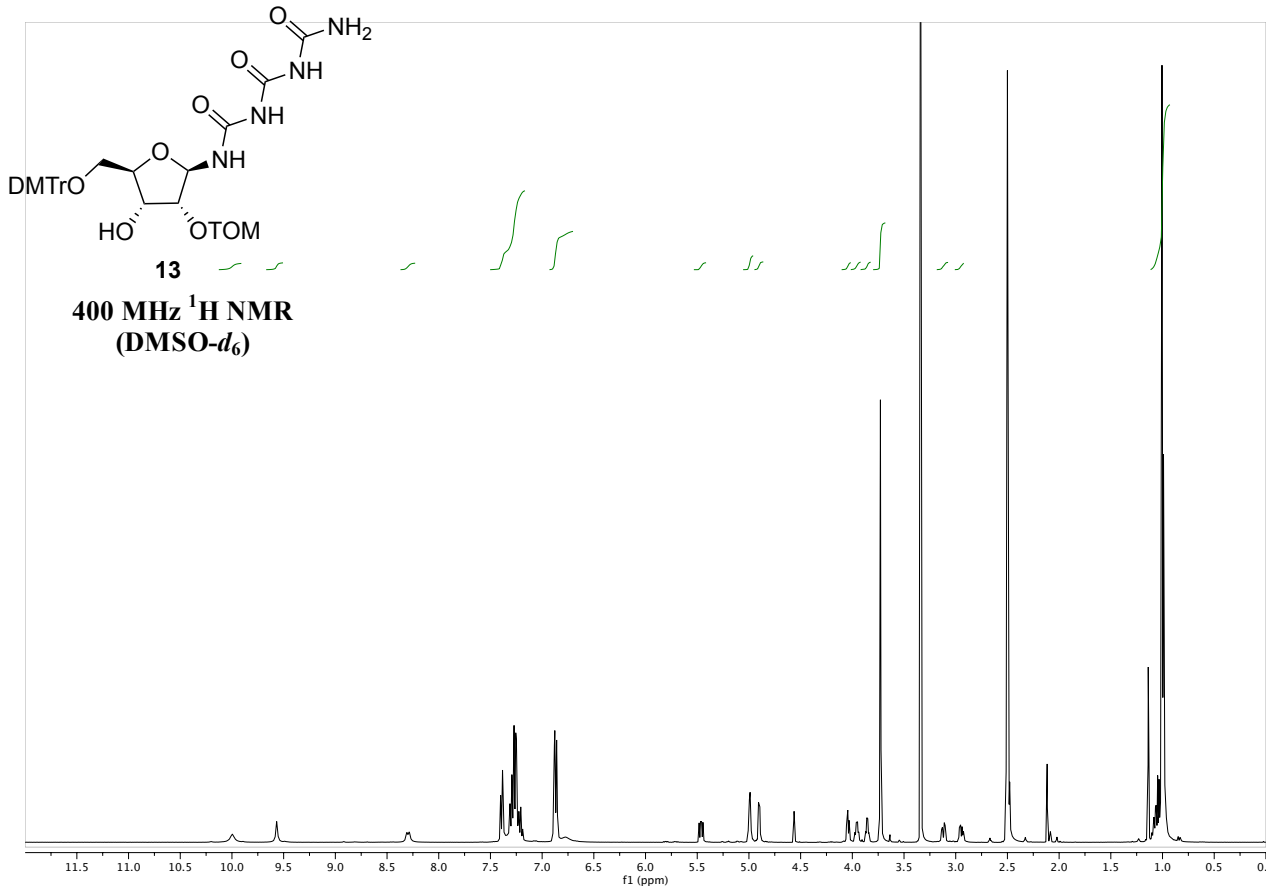

O
NH2
O
NH
O
NH
DMTrO
AcO
OTOM
α−10
101 MHz 13C NMR
(DMSO-d6)
f1 (ppm)


O
NH2
O
NH
O
NH
O
NH
DMTrO
AcO
OTOM
11
400 MHz 1H NMR
(DMSO-d6)
f1 (ppm)


O
NH2
O
NH
O
NH
O
NH
DMTrO
AcO
OTOM
11
101 MHz 13C NMR
(DMSO-d6)
f1 (ppm)

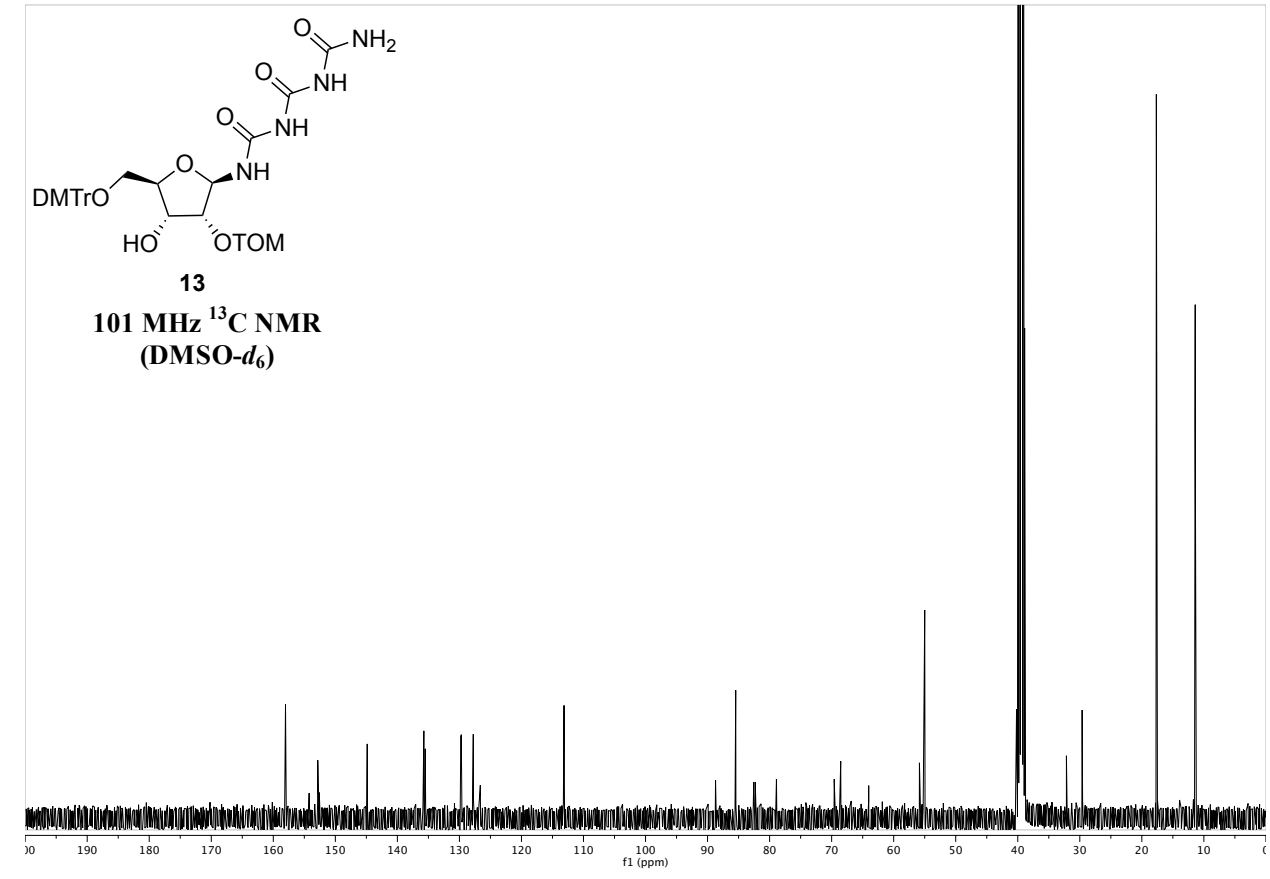

O
NH2
O
NH
O
NH
DMTrO
HO
OTOM
12
400 MHz 1H NMR
(DMSO-d6)
10.5 10.0 9.5 9.0 8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5
f1 (ppm)


O
NH2
O
NH
O
NH
DMTrO
HO
OTOM
12
101 MHz 13C NMR
(DMSO-d6)
190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10
f1 (ppm)



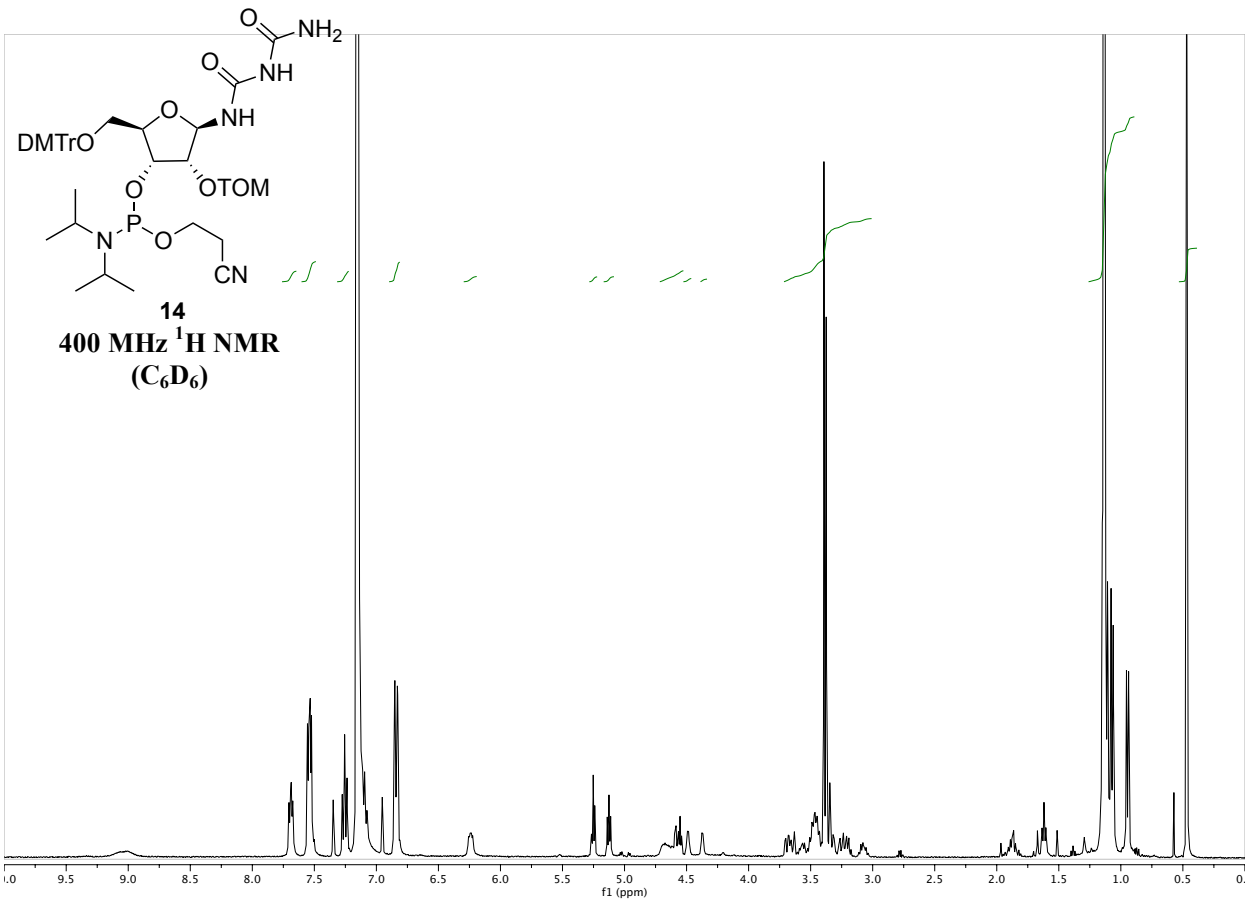
13

400 MHz ^{1}H NMR
(DMSO-d_6)

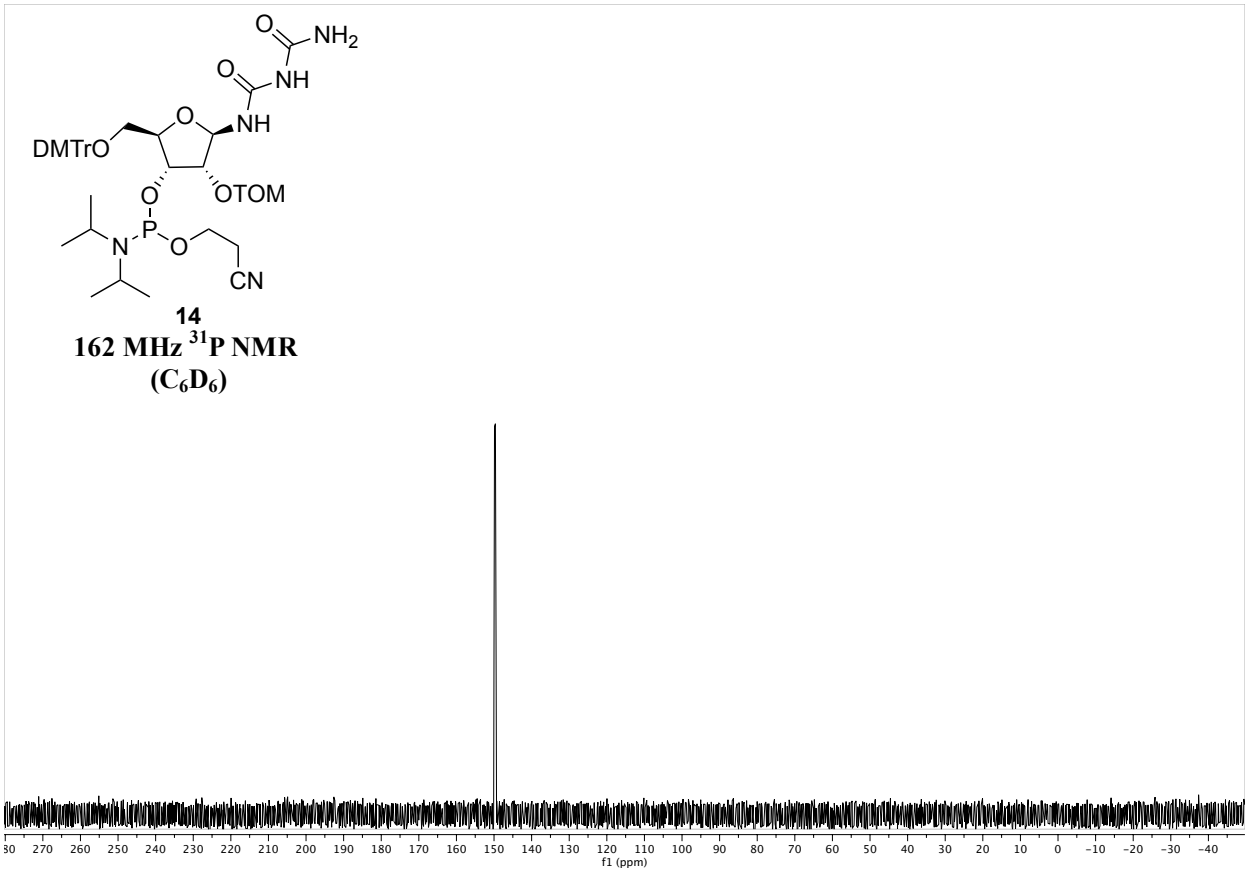


13

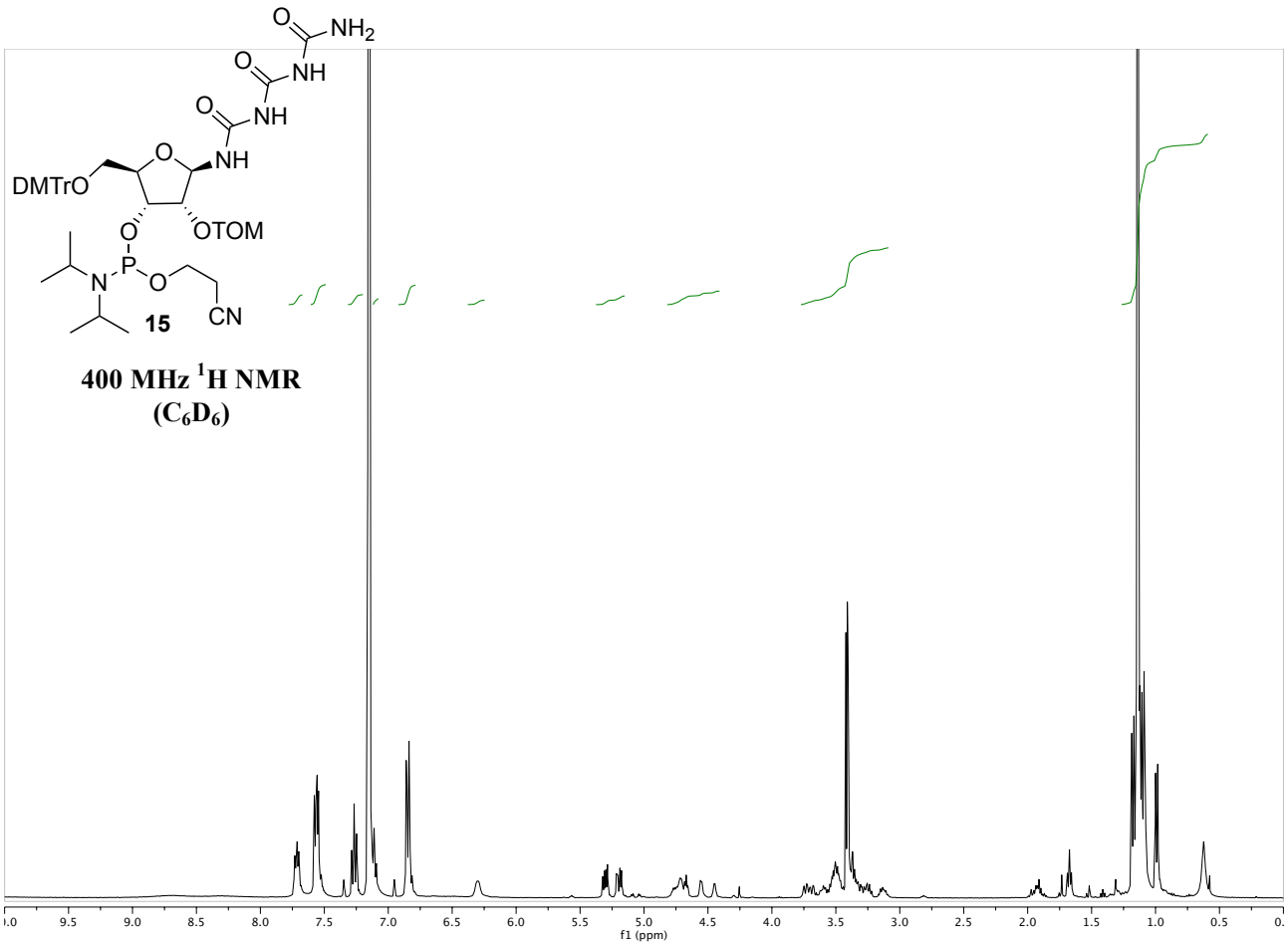
101 MHz ^{13}C NMR
(DMSO-d_6)

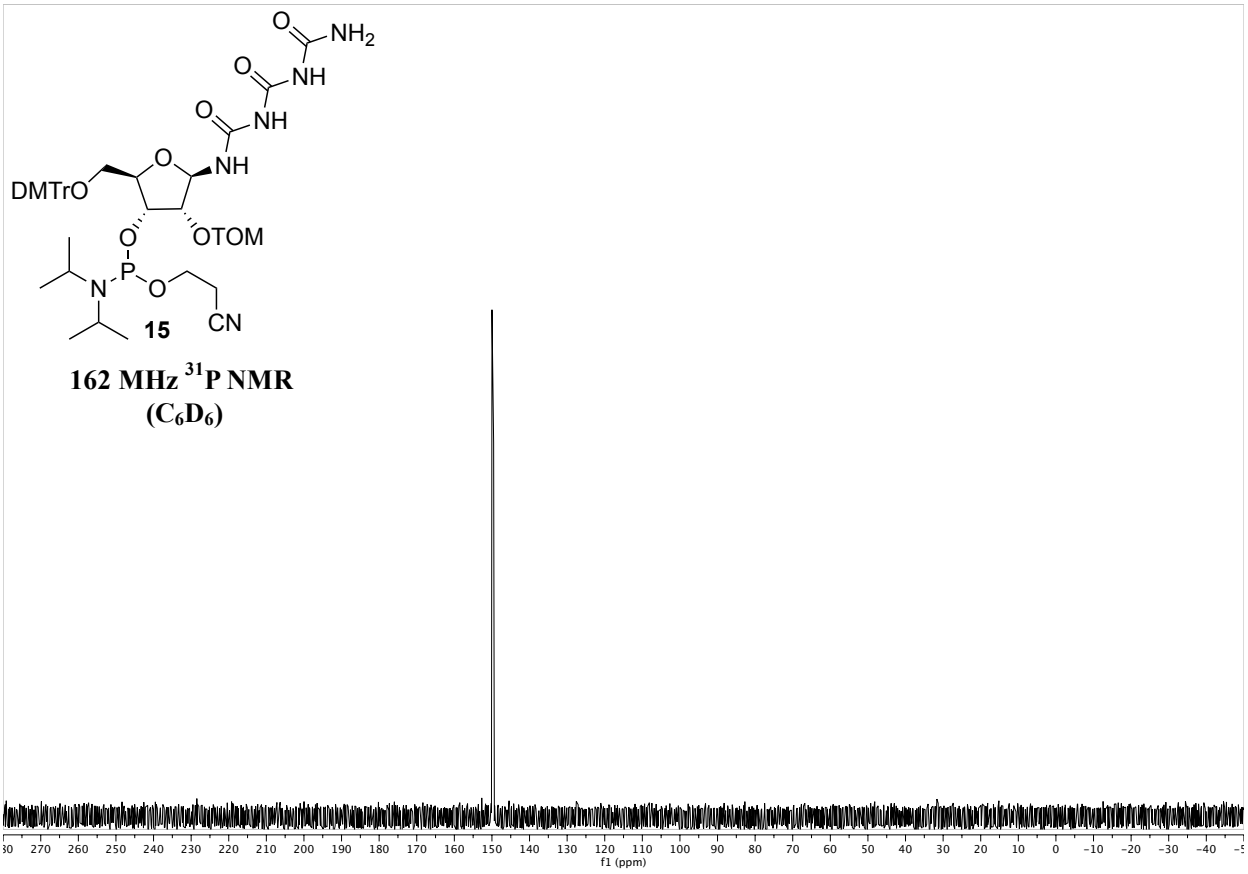


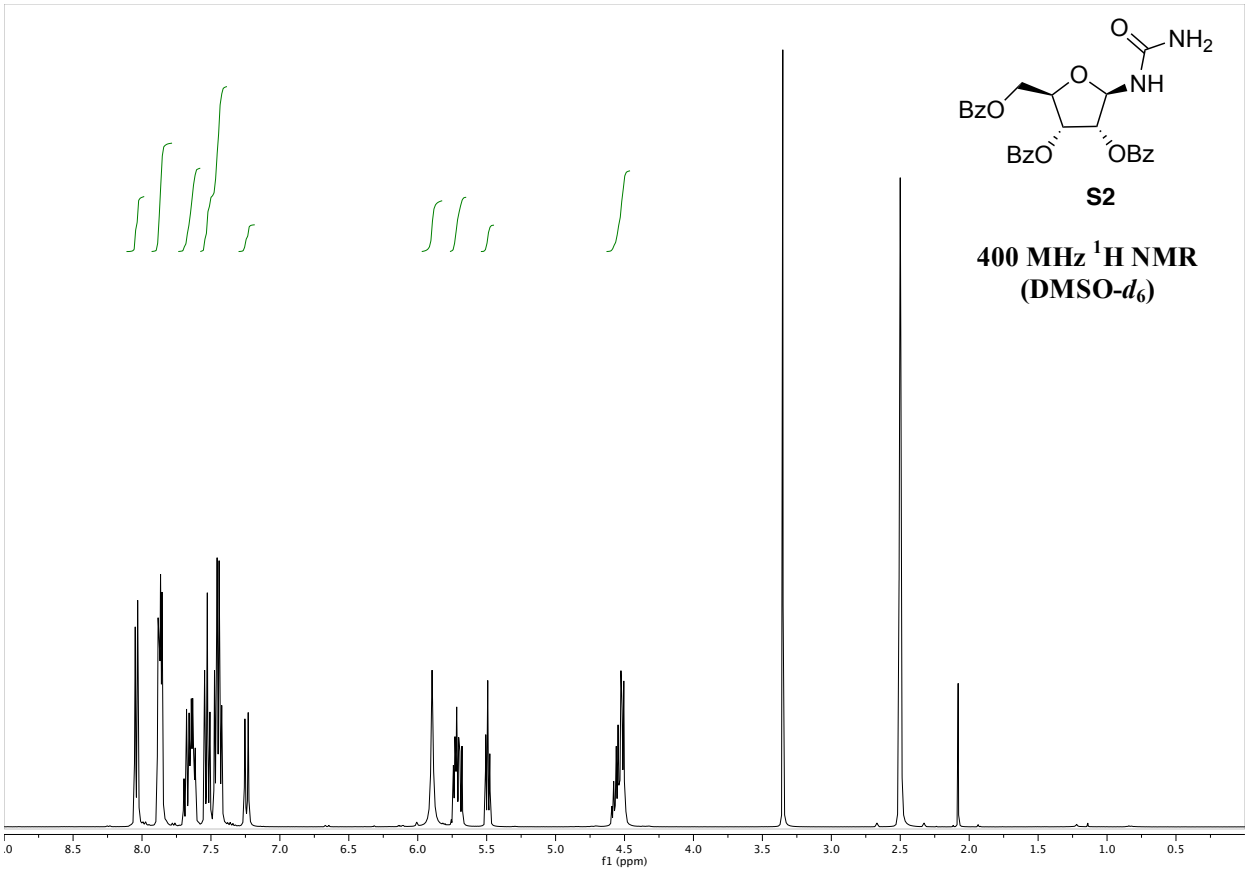
14
400 MHz ^{1}H NMR
(C_6D_6)



14
162 MHz ^{31}P NMR
(C_6D_6)

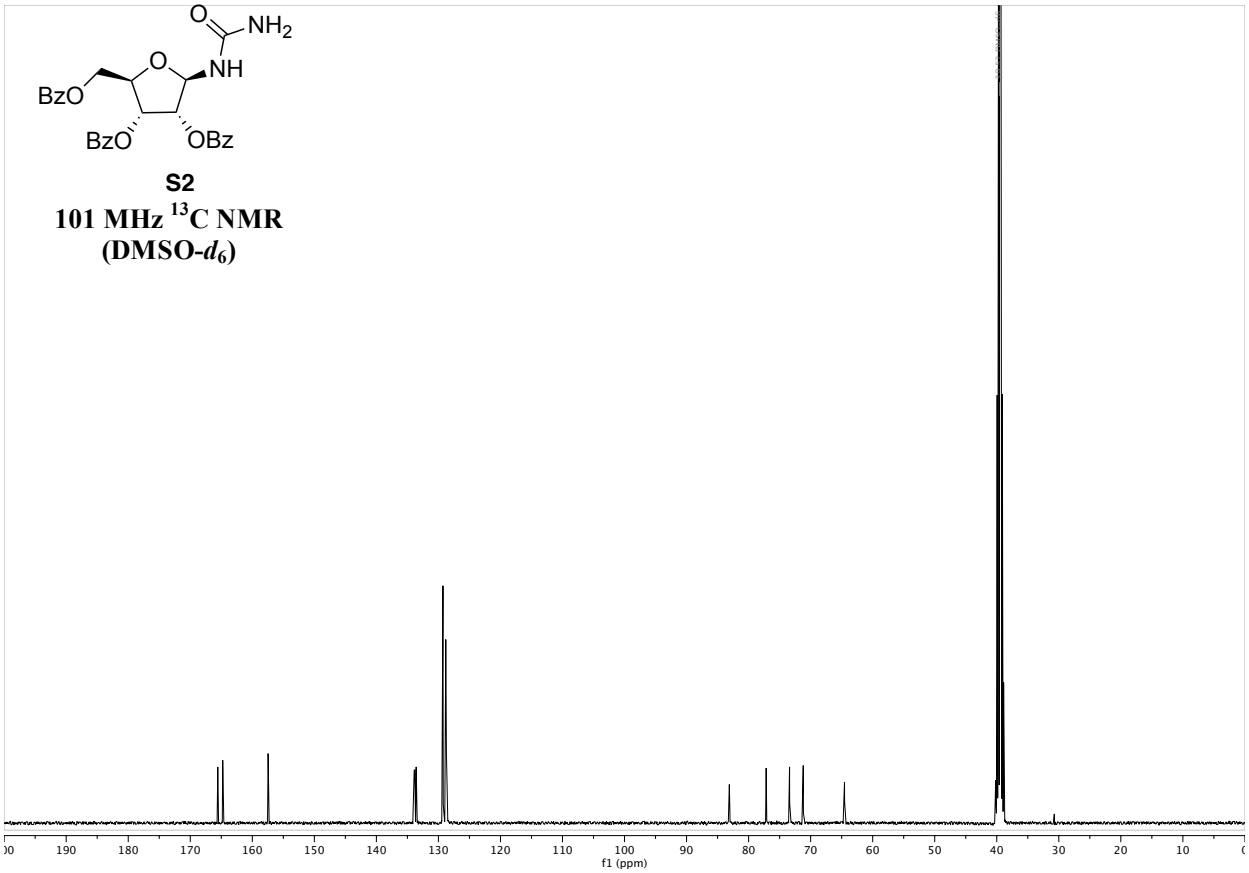

O
NH2
O
NH
O
NH
O
NH
DMTrO
O
OTOM
O
P
N
O
CN
15
400 MHz 1H NMR
(C6D6)
f1 (ppm)


O
NH2
O
NH
O
NH
O
NH
DMTrO
O
OTOM
O
P
N
O
CN
15
162 MHz 31P NMR
(C6D6)
f1 (ppm)



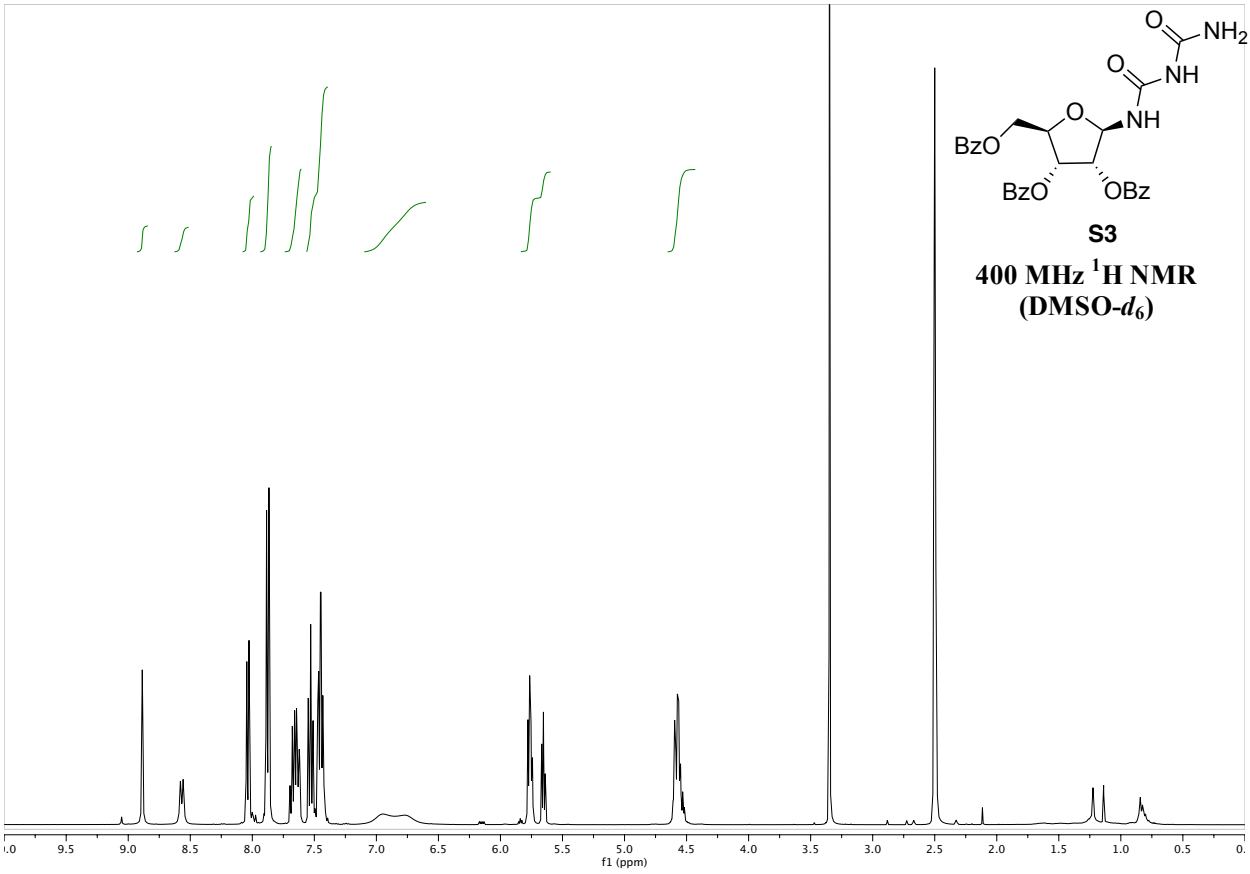
S2

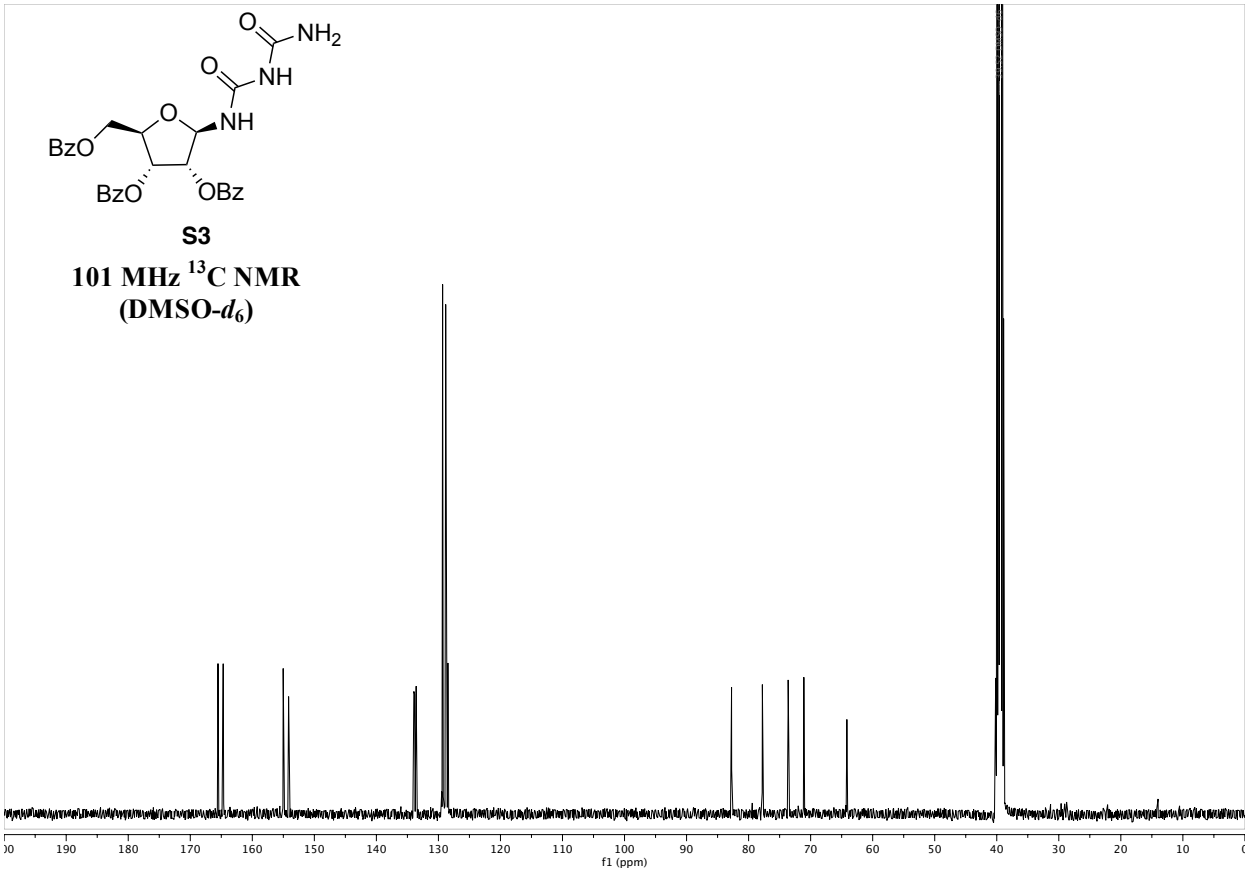
400 MHz ^{1}H NMR (DMSO-d_6)

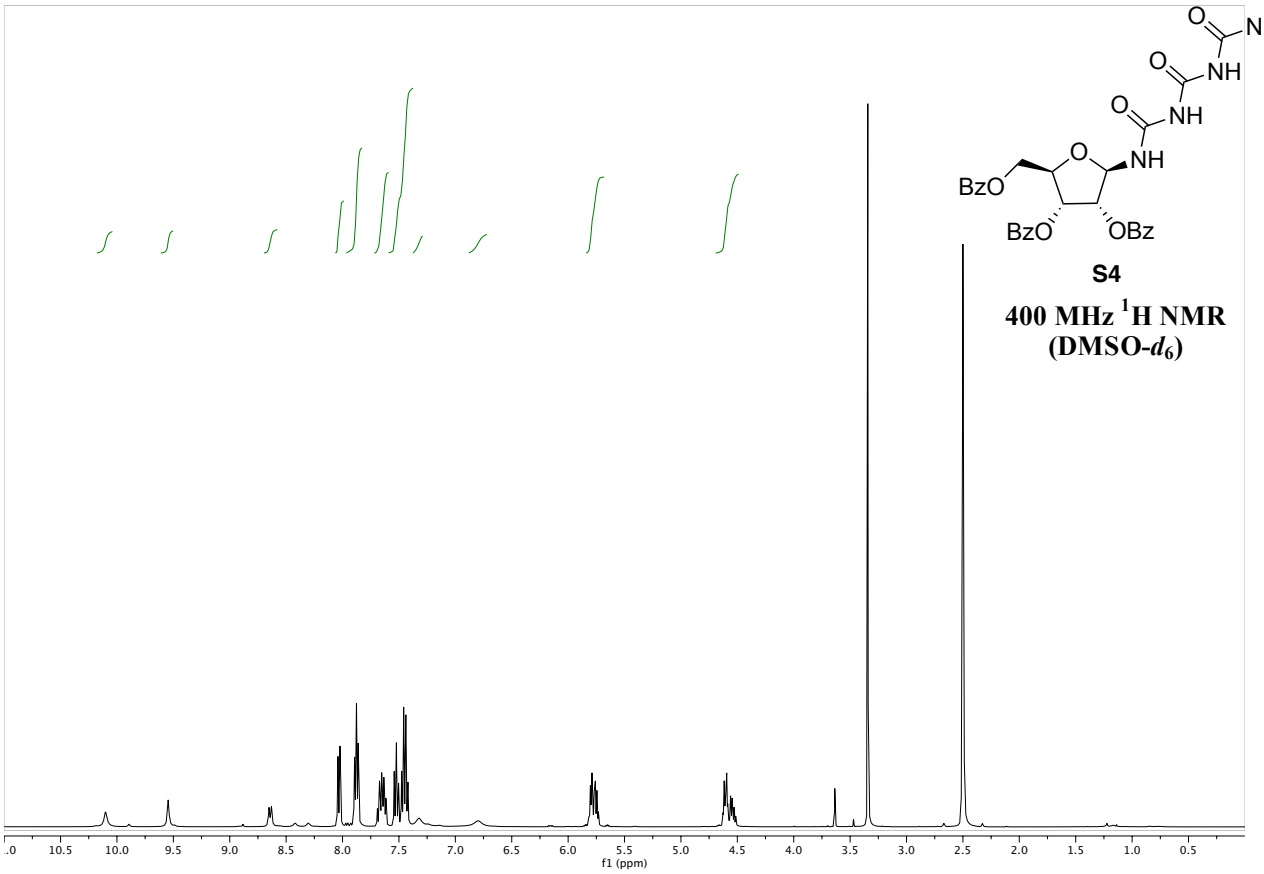


S2

101 MHz ^{13}C NMR (DMSO-d_6)

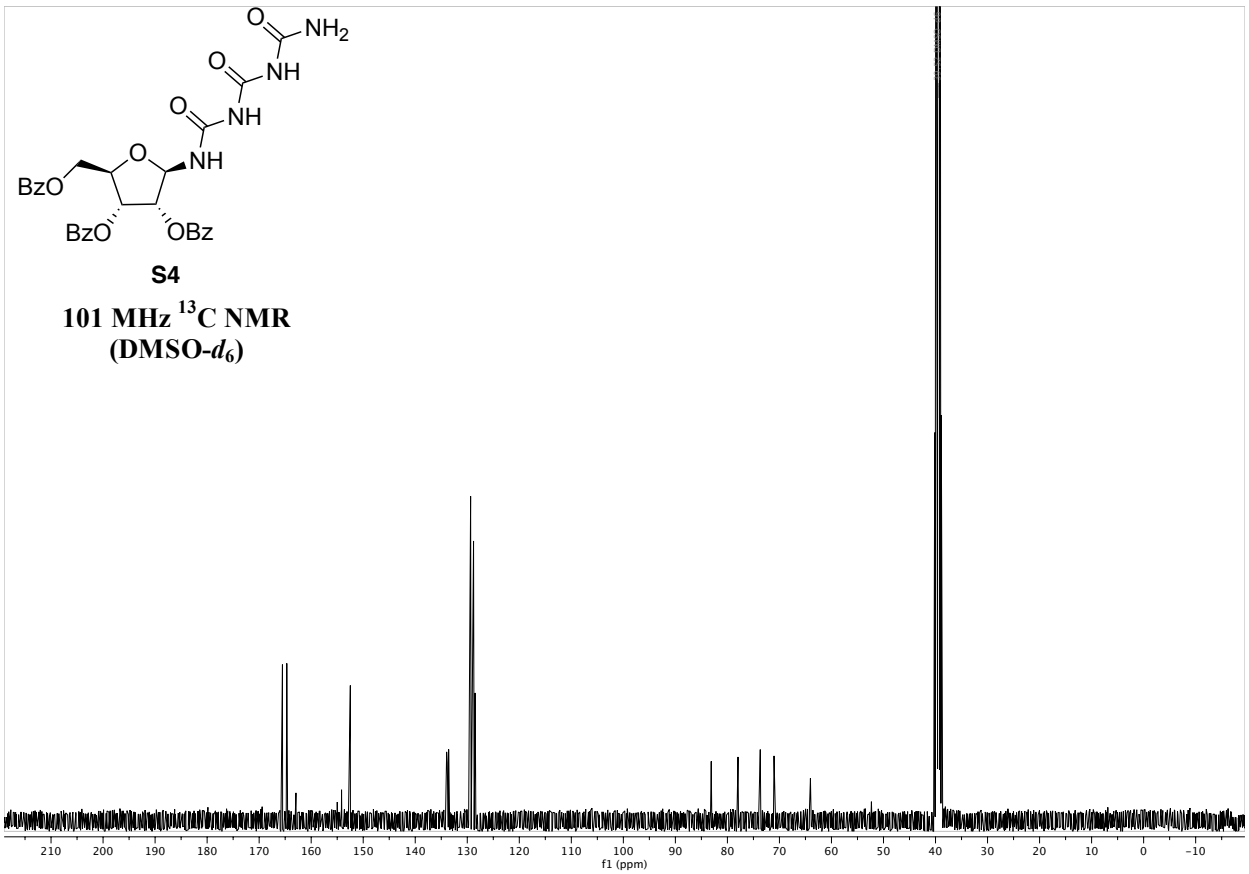

O
NH2
O
NH
O
NH
BzO
BzO
OBz
S3
400 MHz 1H NMR
(DMSO-d6)
f1 (ppm)


O
NH2
O
NH
O
NH
BzO
BzO
OBz
S3
101 MHz 13C NMR
(DMSO-d6)
f1 (ppm)



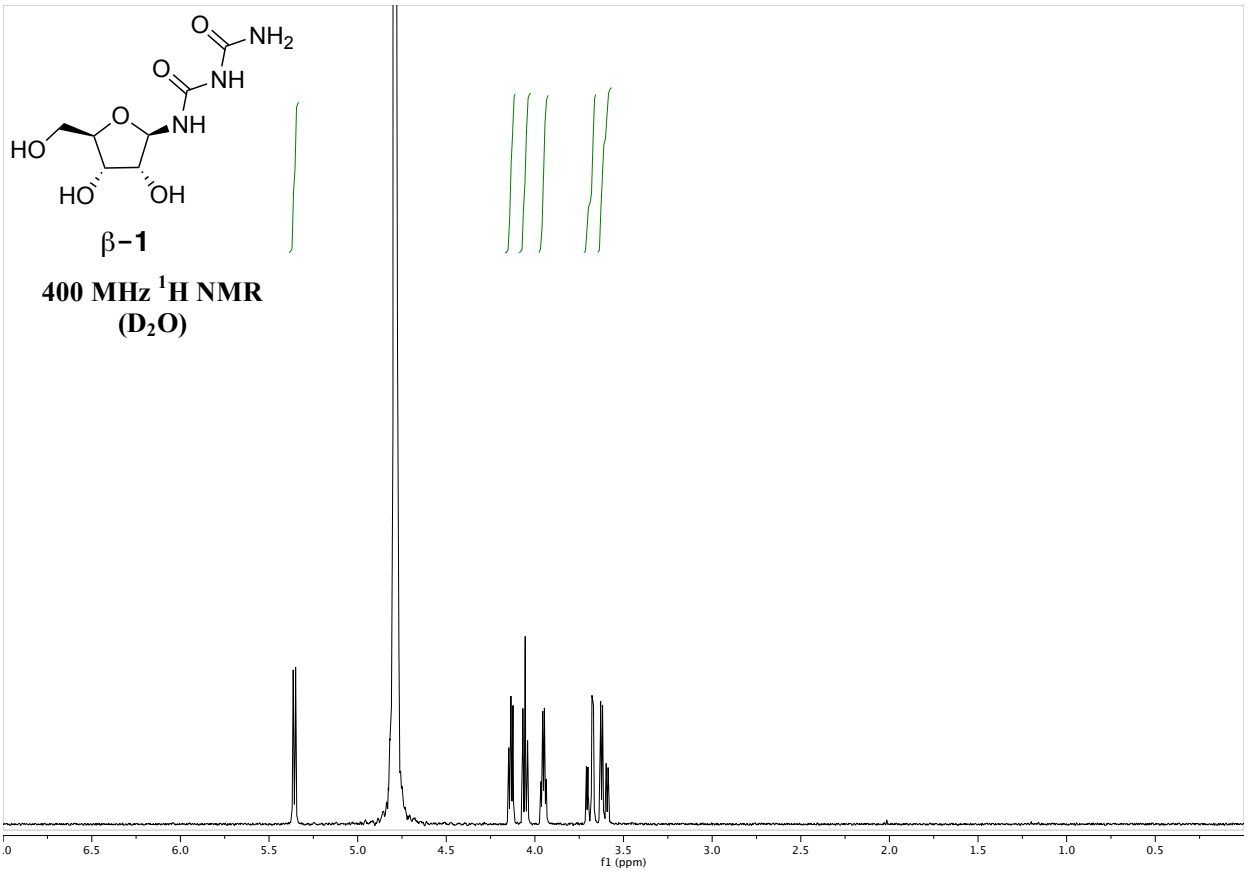
S4

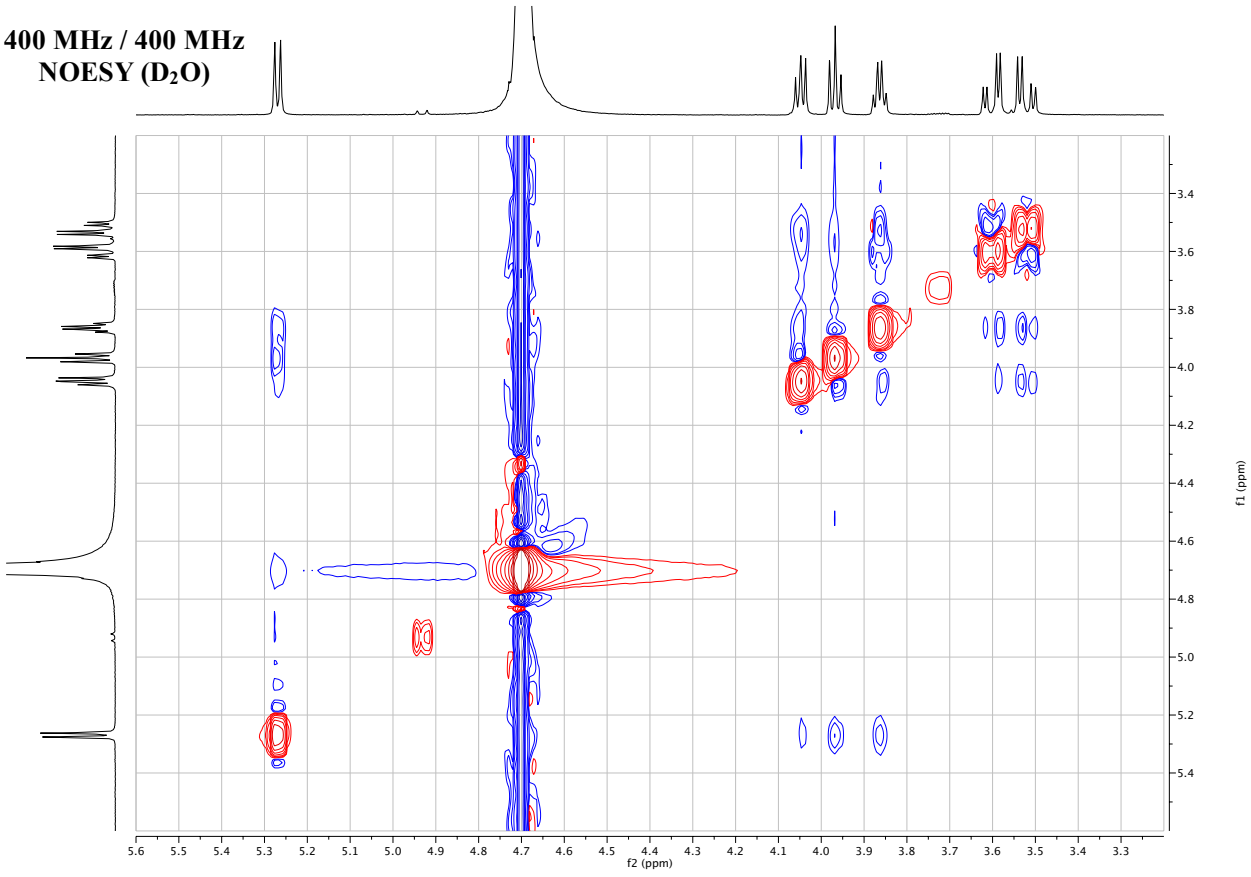
400 MHz ^{1}H NMR (DMSO-d_6)



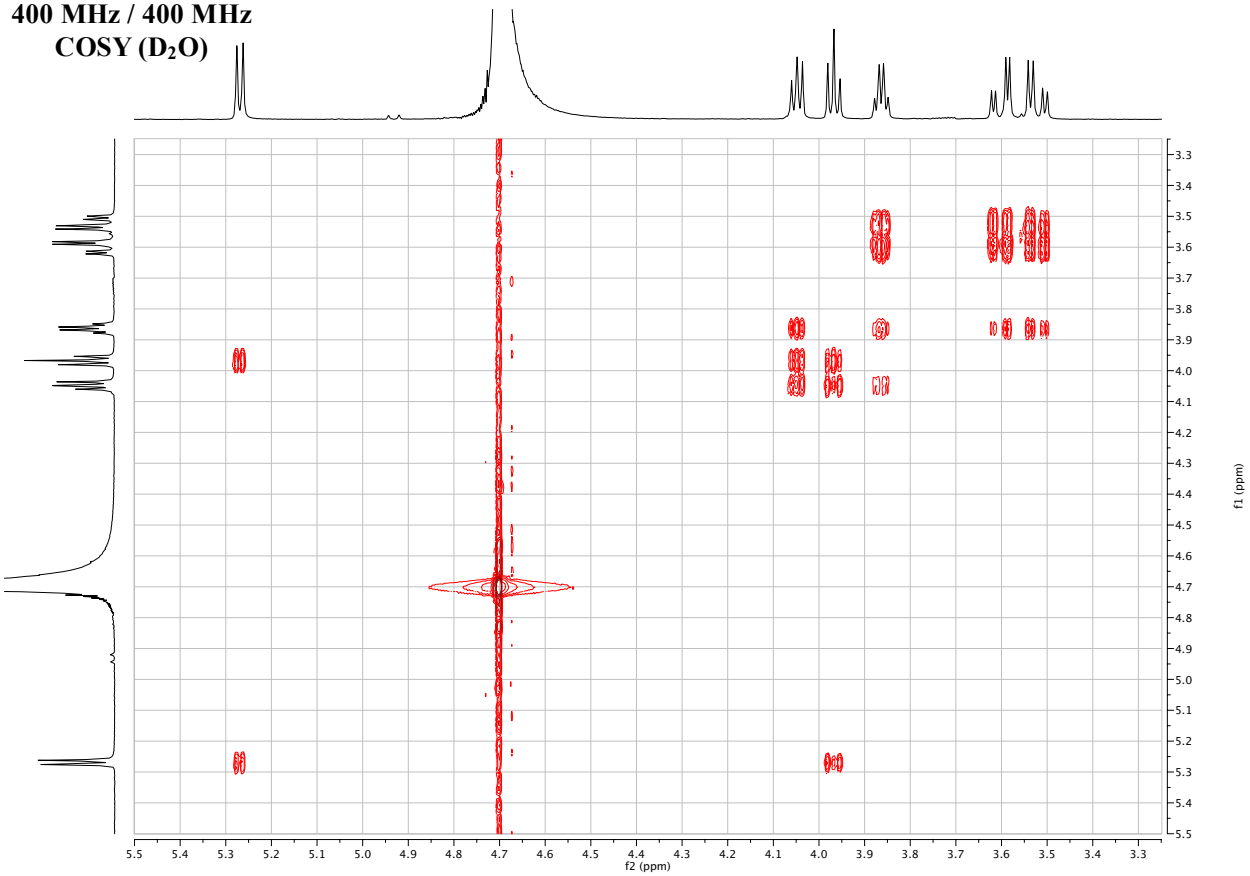
S4

101 MHz ^{13}C NMR (DMSO-d_6)

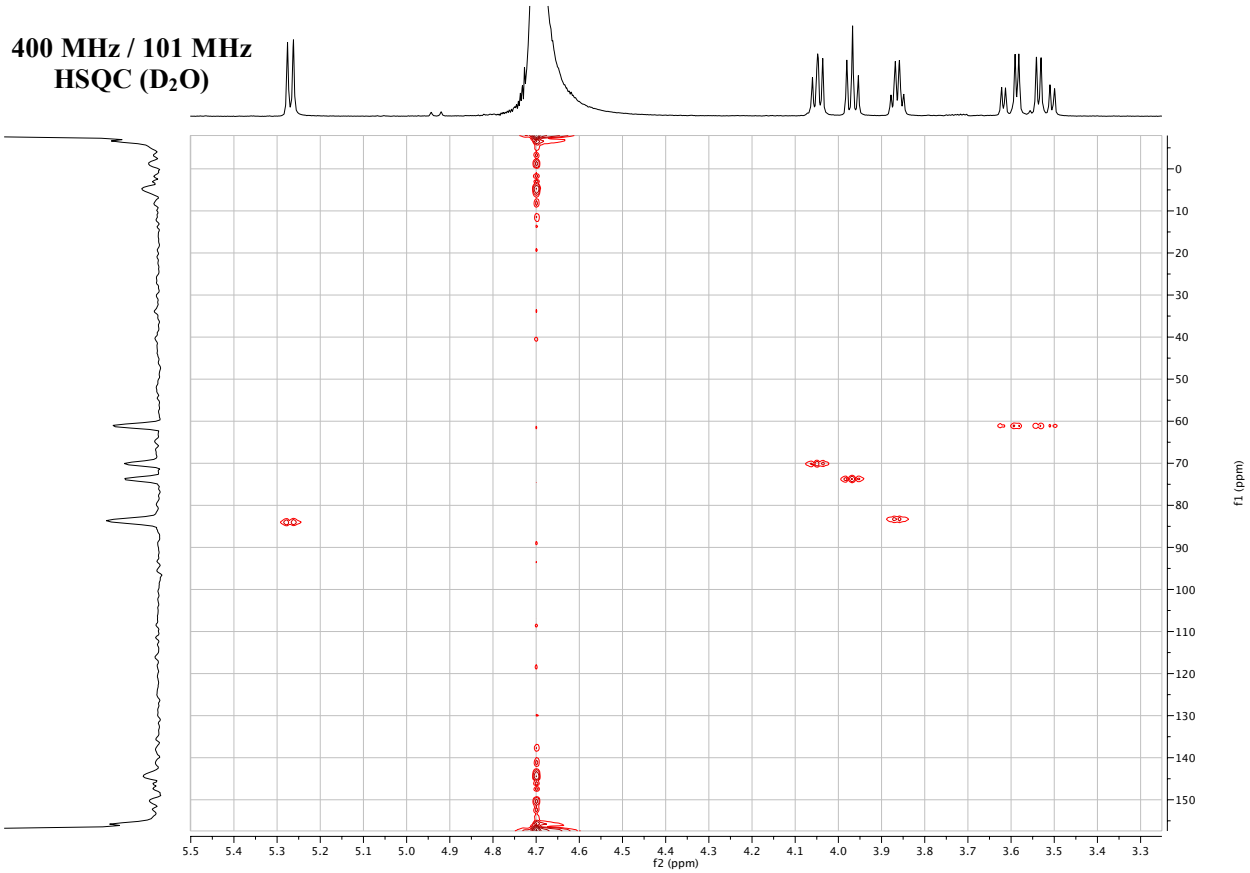

β−1
400 MHz ^{1}H NMR
(D_2O)
f1 (ppm)

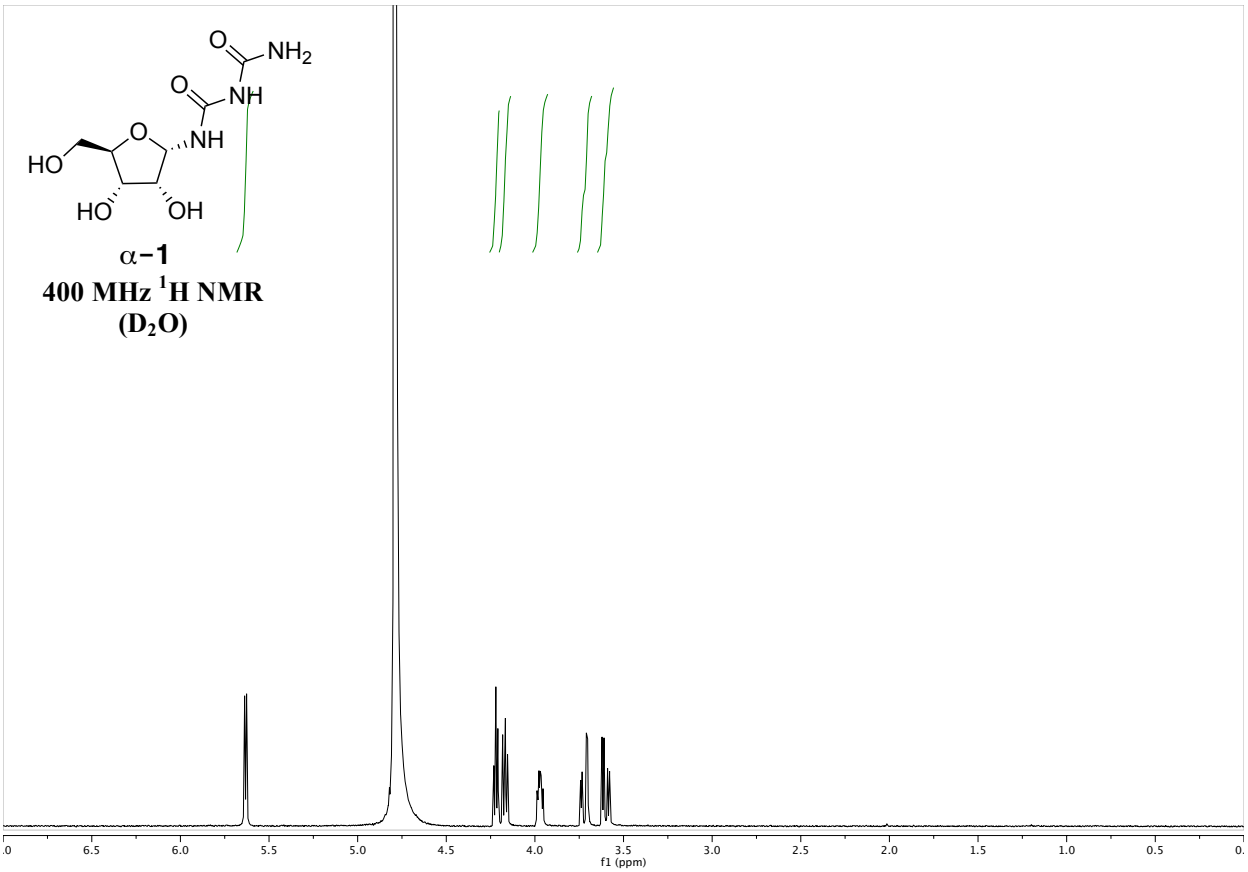

400 MHz / 400 MHz
NOESY (D_2O)
f2 (ppm)
f1 (ppm)

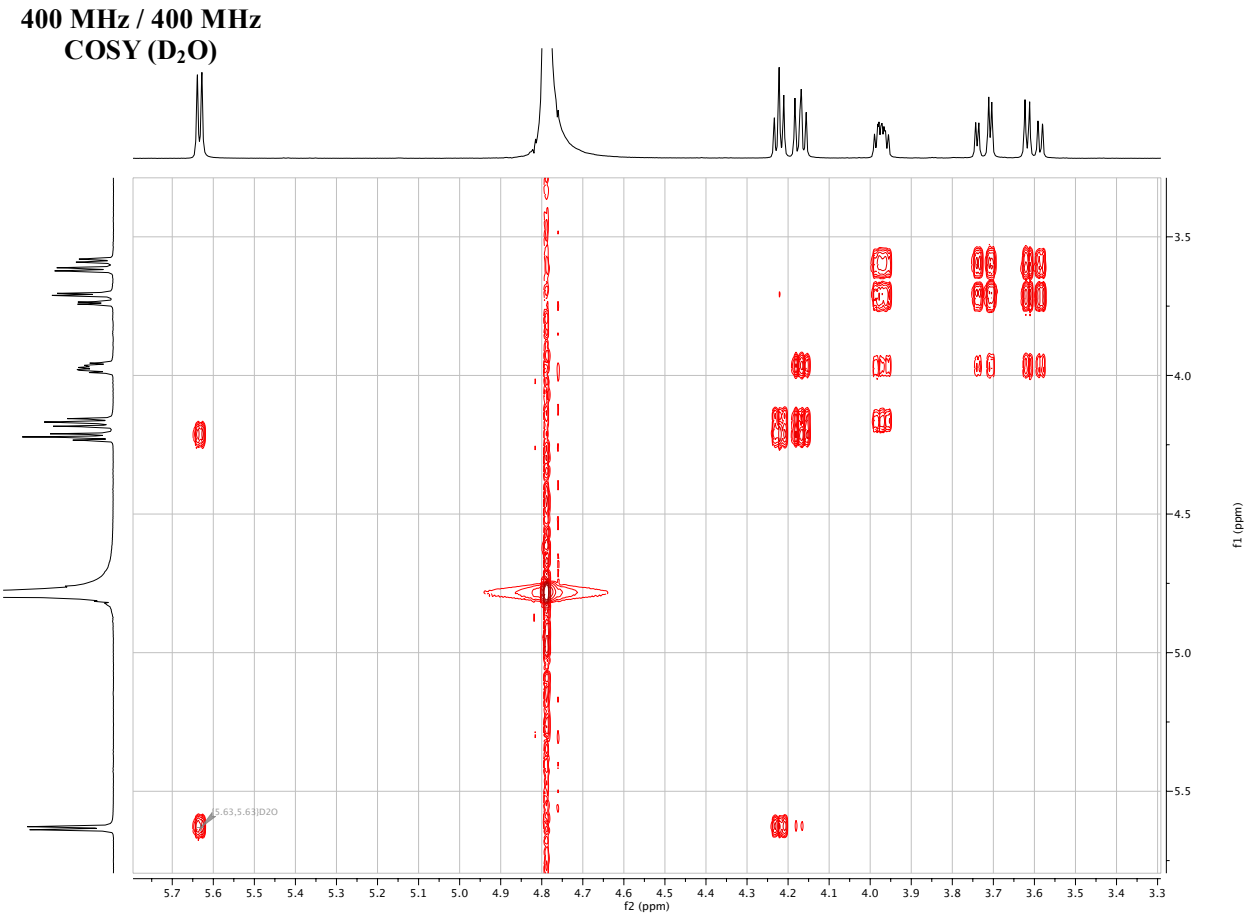
400 MHz / 400 MHz
COSY (D_2O)



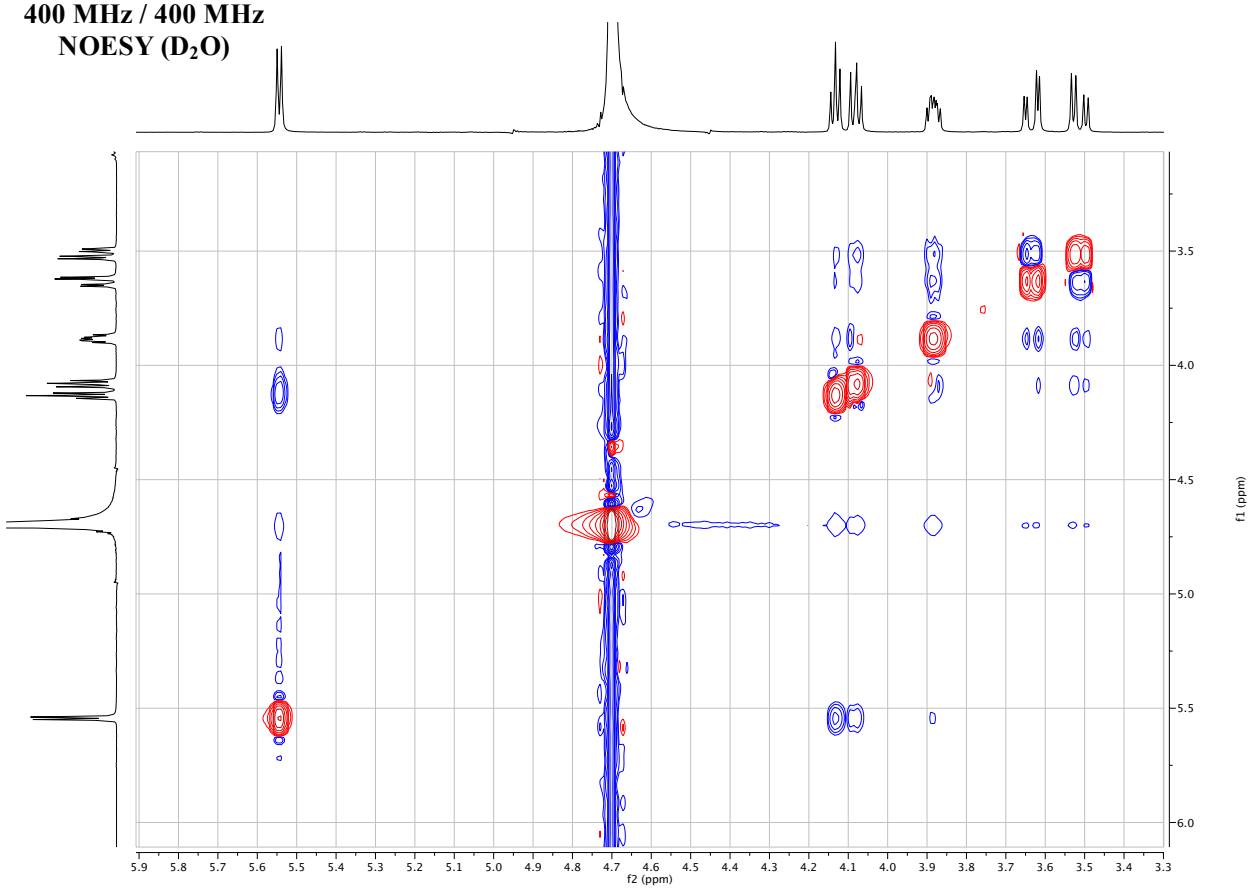
400 MHz / 101 MHz
HSQC (D_2O)



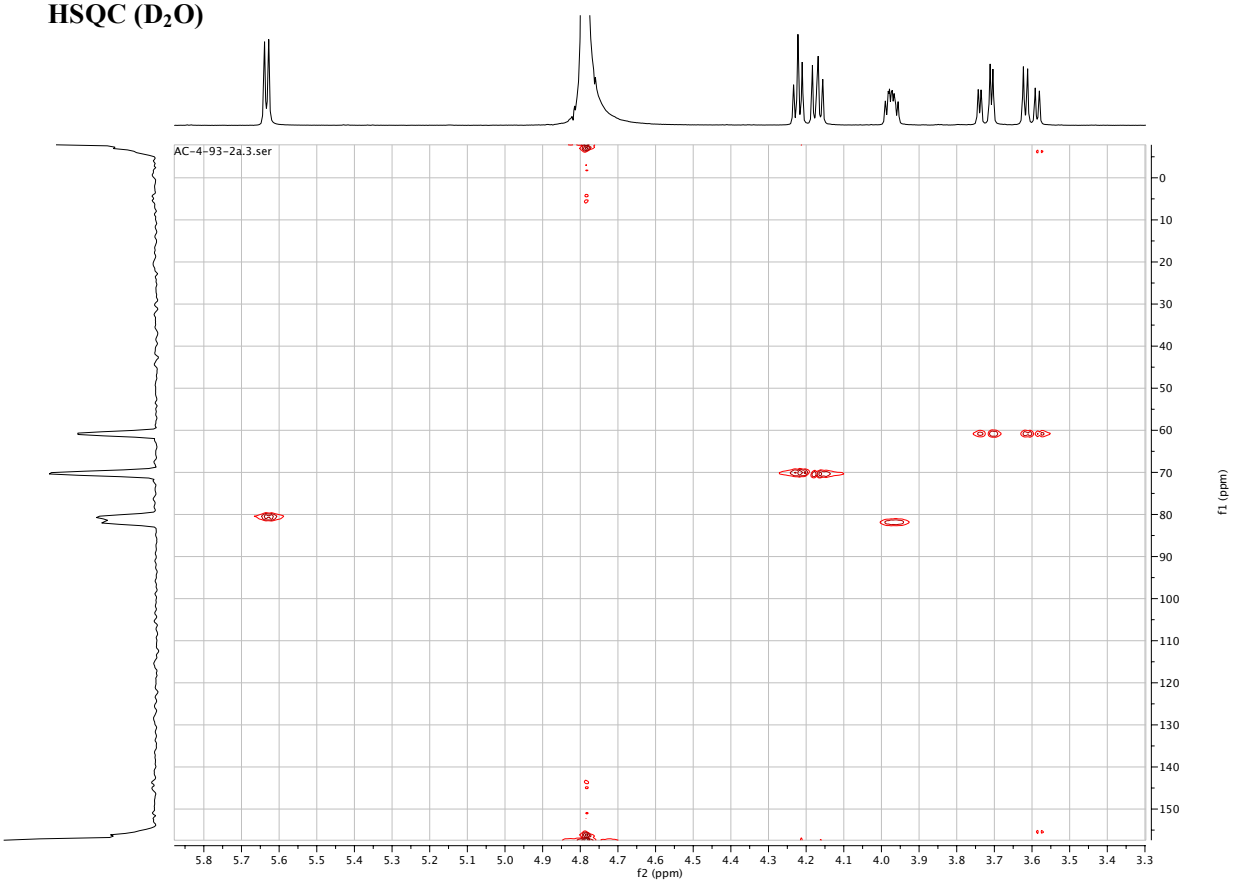

α−1
400 MHz ^{1}H NMR
(D_2O)
f1 (ppm)

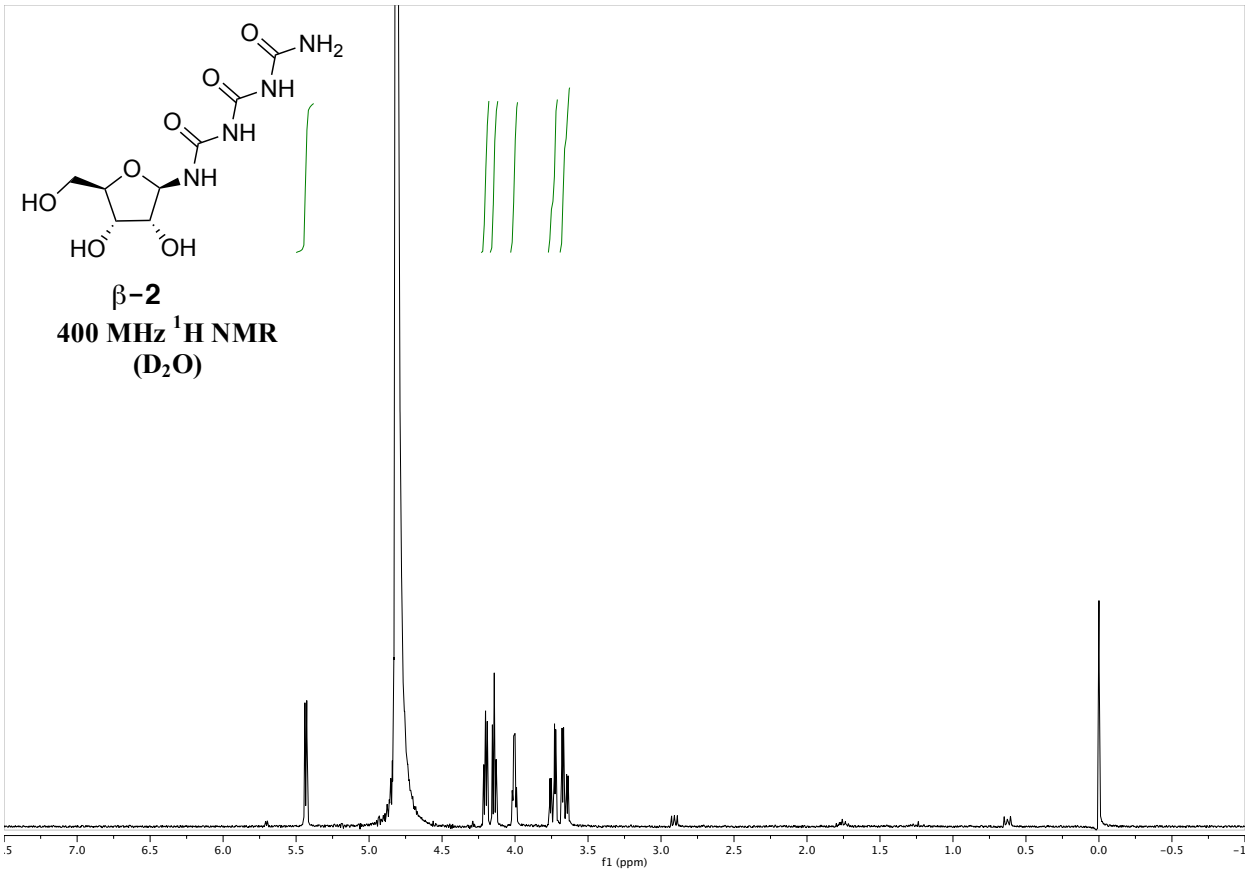

400 MHz / 400 MHz
COSY (D_2O)
f2 (ppm)
f1 (ppm)

400 MHz / 400 MHz
NOESY (D_2O)



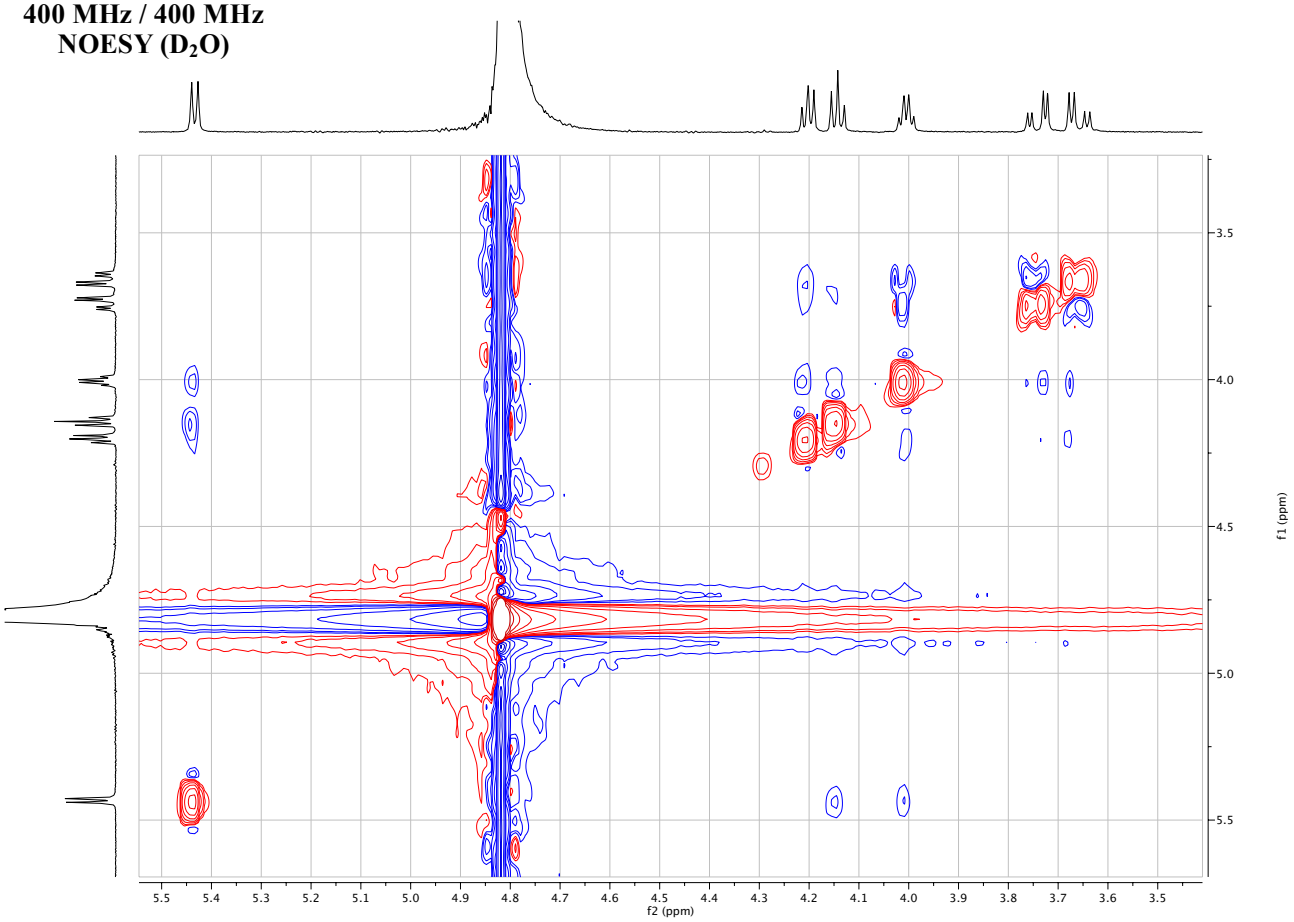
400 MHz / 101 MHz
HSQC (D_2O)



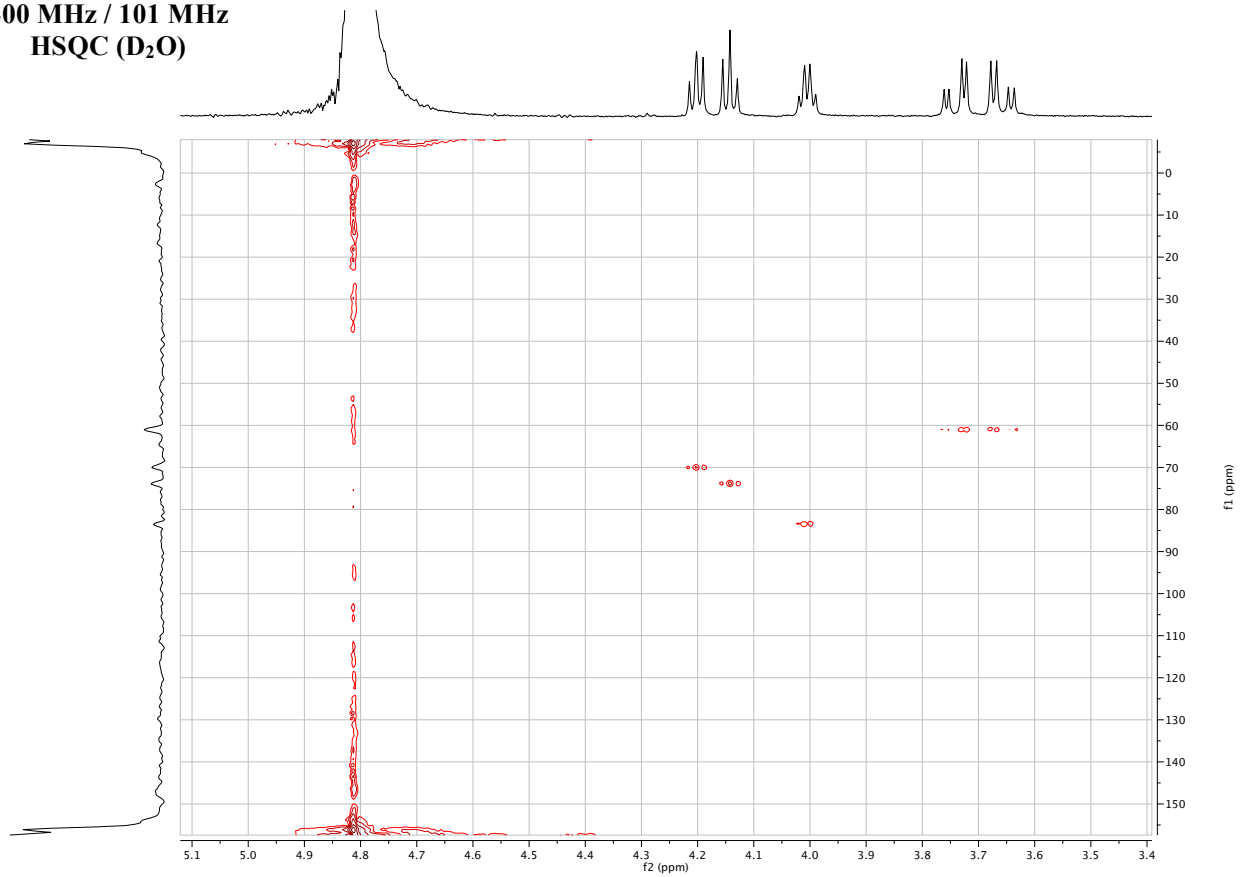


β−2
400 MHz ^{1}H NMR
(D_2O)

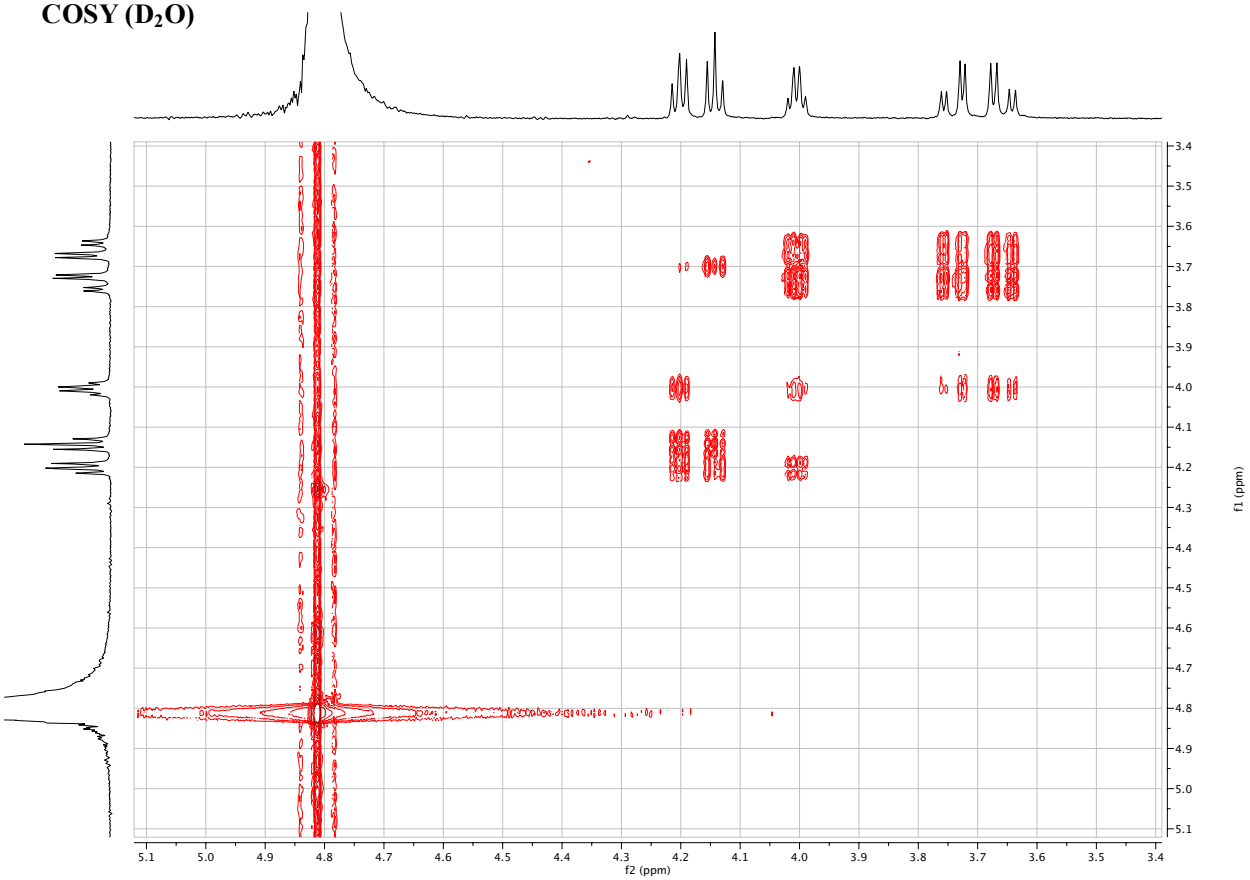
400 MHz / 400 MHz
NOESY (D_2O)

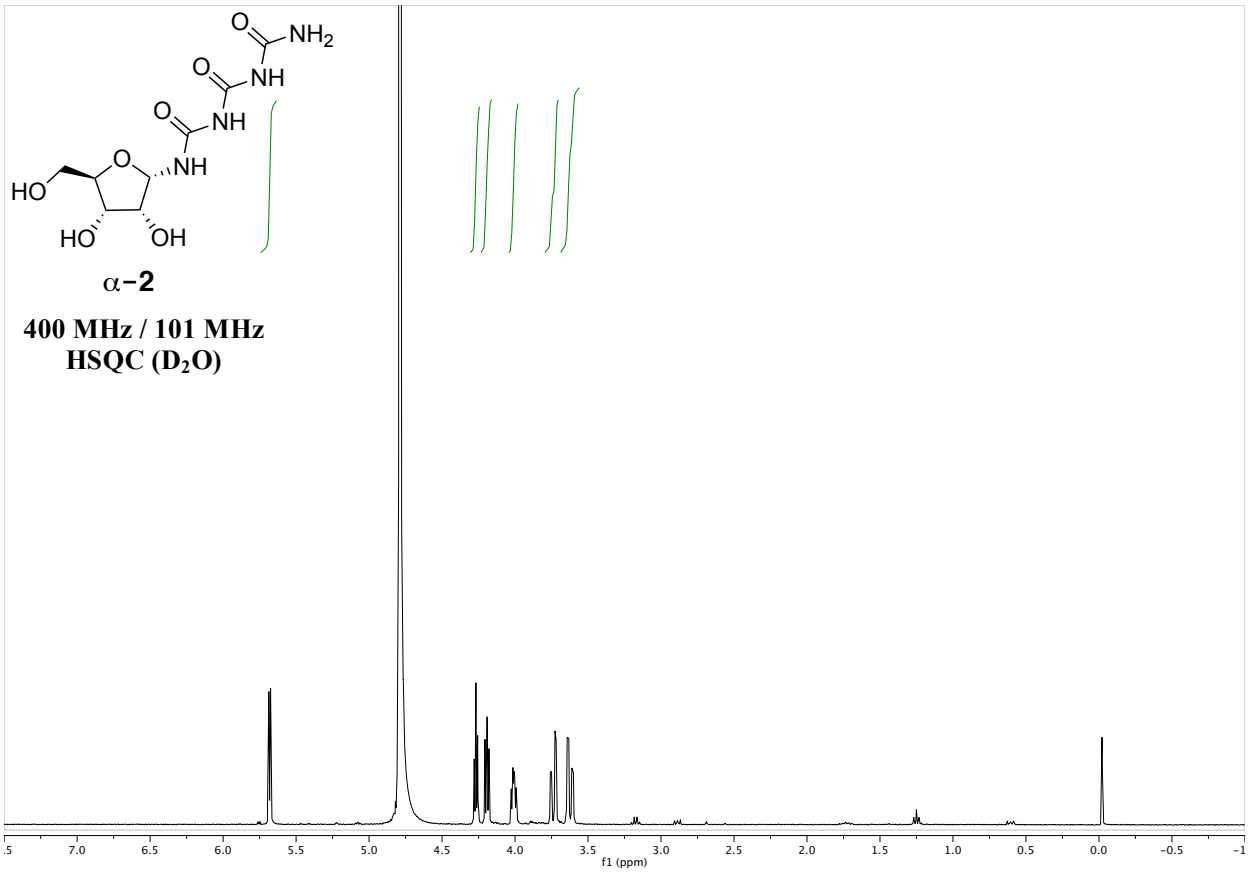


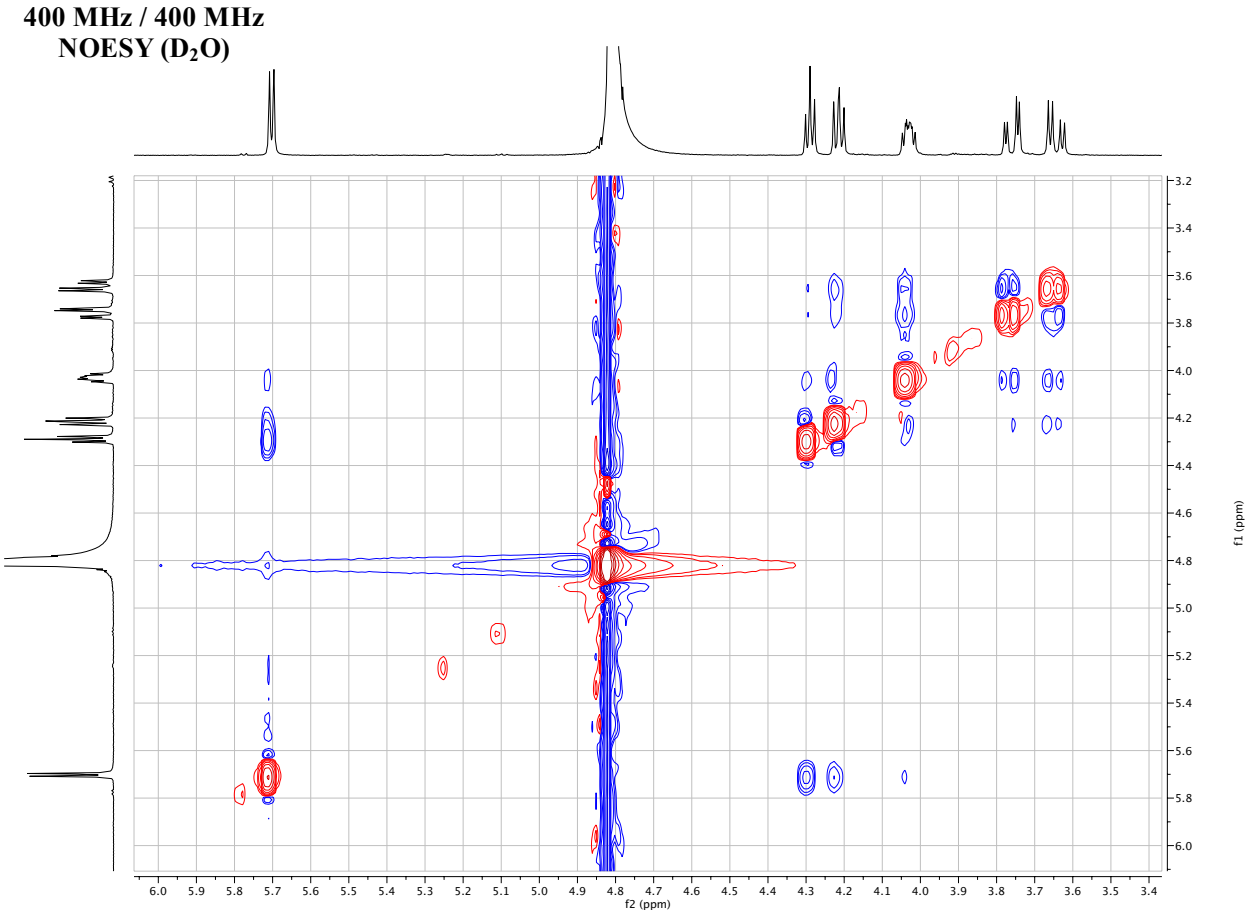
400 MHz / 101 MHz
HSQC (D_2O)



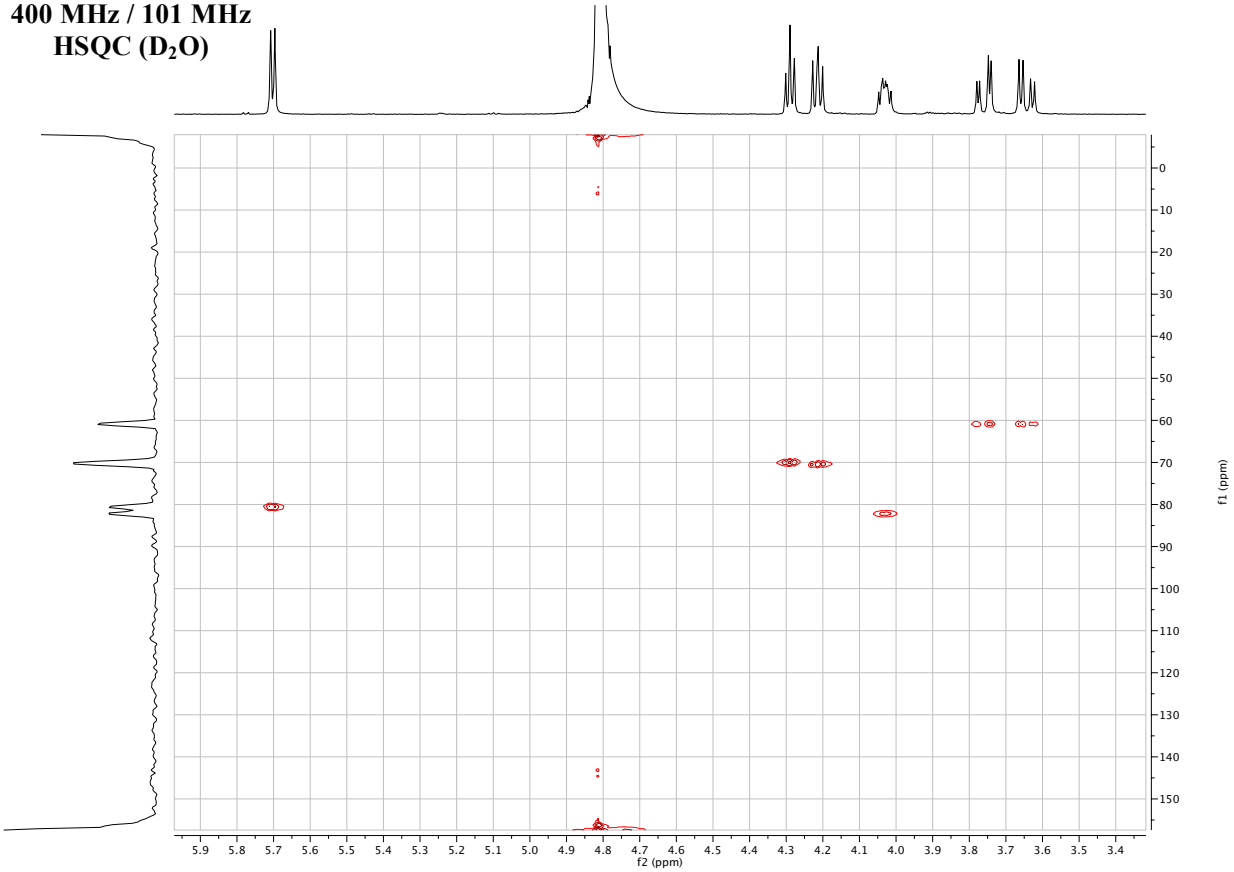
400 MHz / 400 MHz
COSY (D_2O)



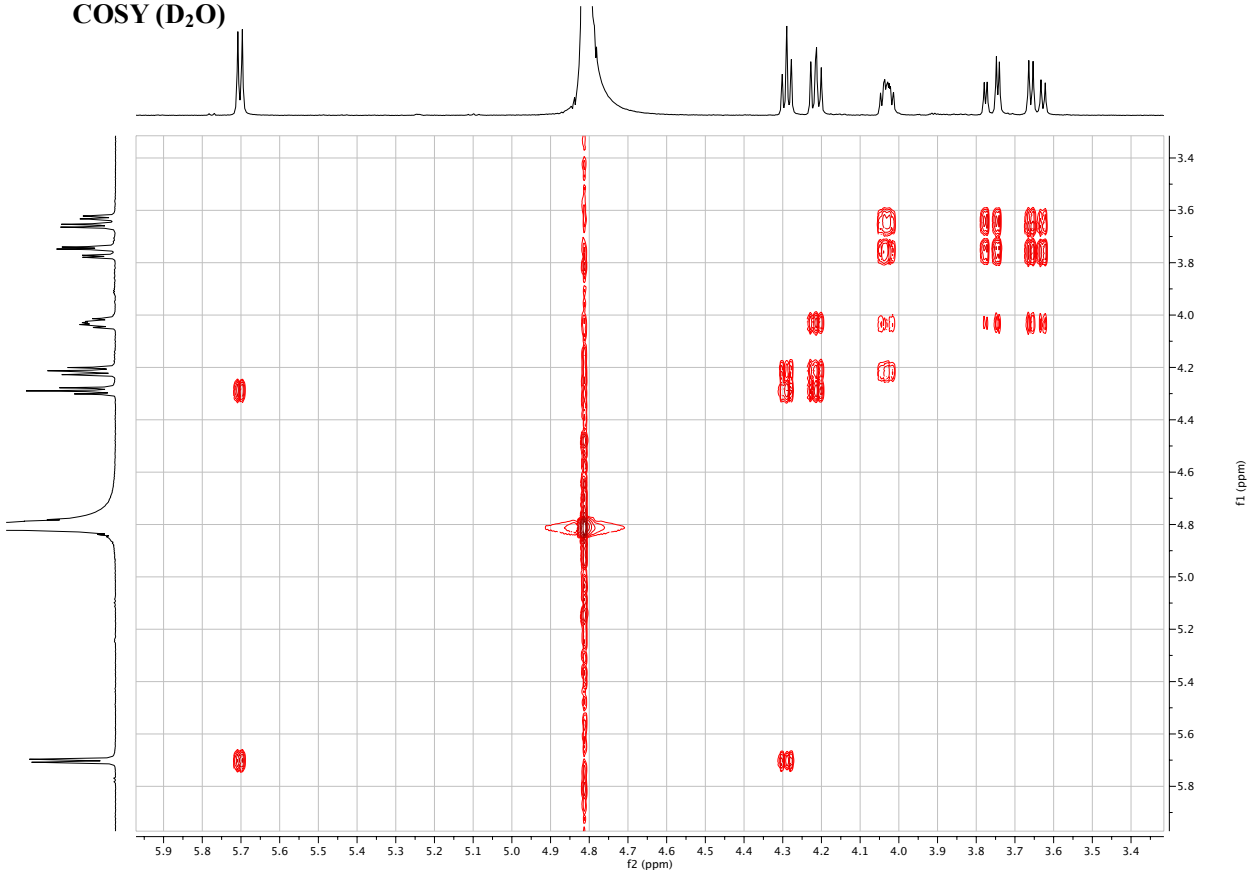

α−2
400 MHz / 101 MHz
HSQC (D2O)
f1 (ppm)


400 MHz / 400 MHz
NOESY (D2O)
f2 (ppm)
f1 (ppm)

400 MHz / 101 MHz
HSQC (D_2O)



400 MHz / 400 MHz
COSY (D_2O)



11. References

[1] M. W. Logue, B. H. Han, *Carbohydr. Res.* **1983**, *121*, 287-297.

[2] R. M. Pinto, R. I. Olariu, J. Lameiras, F. T. Martins, A. A. Dias, G. J. Langley, P. Rodrigues, C. D. Maycock, J. P. Santos, M. F. Duarte, M. T. Fernandez, M. L. Costa, *J. Mol. Struct.* **2010**, *980*, 163-171.

[3] E. Breitmaier, U. Hollstein, *Org. Magn. Reson.* **1976**, *8*, 573-575.

[4] W. Lee, M. Tonelli, J. L. Markley, *Bioinformatics* **2015**, *31*, 1325-1327.