Synthesis of biotin labelled cap analogue – incorporable into mRNA transcripts and promoting cap-dependent translation

Jacek Jemielity,* Maciej Lukaszewicz, Joanna Kowalska, Jakub Czarnecki, Joanna Zuberek, Edward Darzynkiewicz

Division of Biophysics, Institute of Experimental Physics, Faculty of Physics, University of Warsaw, Zwirki i Wigury 93,

02-089 Warsaw, Poland

**e-mail Address: jacekj@biogeo.uw.edu.pl*

Electronic Supplementary Information

List of content

I. Supplementary Tables **3**

II. Supplementary Figures **4**

III. Experimental procedures **9**

IV. RP HPLC profiles, negative ion ES-MS, ^{1}H and ^{31}P NMR spectra **20**

I. Supplementary Tables

Table S1.
Comparison of relative activity of luciferase translated from differently capped transcripts in the presence of various concentrations of streptavidin added along with mRNA to the RRL system. Each translation reaction mixture contained capped mRNA at ca. 9 nM concentration (the total reaction volume was 12.5 µL). The data are means from 3 ($m_2^{7,2'\text{-}O}$GpppG and $m^7G_{N\text{-}Biot}$pppG) or 2 (m^7GpppG) experiments.

		Relative luciferase activity				
	Streptavidin tetramer conc. [nM]	**0**	**1**	**2**	**8**	**16**
Cap analogue present at the 5'end of the transcript	m^7GpppG	1	0.99 ± 0.08	0.97 ± 0.06	1.05 ± 0.1	1.03 ± 0.12
	$m_2^{7,2'\text{-}O}$GpppG	1	1.06 ± 0.08	1.05 ± 0.02	1.09 ± 0.04	1.07 ± 0.07
	$m^7G_{N\text{-}Biot}$pppG	1	1.06147 ± 0.04	1.03 ± 0.01	1.03 ± 0.01	0.36 ± 0.02

Table S2.
Comparison of luciferase activity translated from differently capped transcripts in the presence of various concentrations of streptavidin added at two different time points from mRNA addition. Each translation reaction mixture contained capped mRNA at ca. 9 nM concentration (the total reaction volume was 12.5 µL). The data are from a single experiment.

			Luciferase activity (RLU x 10^7) after 60 min from mRNA addition				
	Streptavidin tetramer conc. [nM]		**0**	**1**	**2**	**8**	**16**
Cap analogue present at the 5'end of the transcript	$m_2^{7,2'\text{-}O}$GpppG	**0 min**	1.23	1.39	1.31	1.38	1.41
		15 min	1.34	1.33	1.32	1.26	1.32
	$m^7G_{N\text{-}Biot}$pppG	**0 min**	1.05	1.15	1.10	1.08	0.37
		15 min	9.84	1.02	0.92	0.93	0.56

II. Supplementary Figures

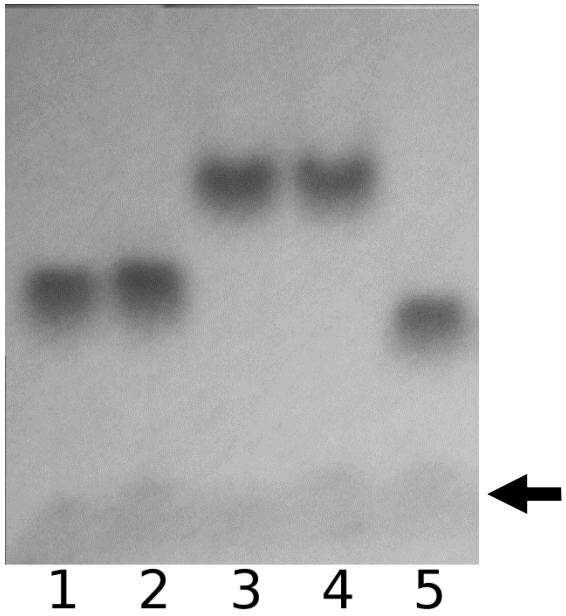


Figure S1.
Electrophoretic mobility of short (5 nt) transcripts capped *in vitro* with indicated dinucleotide cap analogues. Transcripts were obtained as described in Experimental procedures except that in the reaction shown in lane 3 GTP was omitted (in other reactions a 10:1 molar ratio of cap to GTP was used). Short transcripts were analyzed on 20% acrylamide gel (19:1) with 7M urea and visualized by UV shadowing.
Lane 1: $m_2^{7,2'\text{-}O}$GpppGCCCC
Lane 2: m^7GpppGCCCC
Lane 3: $m^7Gp_{N\text{-}Biot}$ppGCCCC (without GTP in reaction)
Lane 4: $m^7Gp_{N\text{-}Biot}$ppGCCCC
Lane 5: m^7Gp_NppGCCCC
The black arrow indicates expected position of a non-capped (GTP initiated) transcript

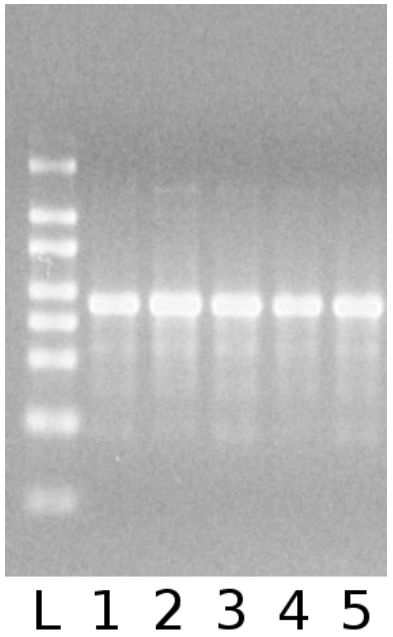


Figure S2.
Electrophoretic analysis of full-length luciferase mRNA transcripts capped *in vitro* with indicated dinucleotide cap analogues.
L: RNA size marker (RiboRuler RNA Ladder, High Range, Fermentas)
Lane 1: m^7GpppG-Luc-mRNA
Lane 2: $m_2^{7,2'\text{-}O}$GpppG-Luc-mRNA
Lane 3: m^7Gp_NppG-Luc-mRNA
Lane 4: $m^7Gp_{N\text{-}Biot}$ppG-Luc-mRNA
Lane 5: ApppG-Luc-mRNA

Figure S3.

Translational inhibition with m^{7}GTP of luciferase mRNA capped with m^{7}GpppG (▲), m$_2^{7,2'\text{-}O}$GpppG (□),m^{7}Gp$_{N\text{-}Biot}$ppG (●) and with ApppG (Δ) in rabbit reticulocyte lysate. Inhibition of cap-dependent translation in RRL (Flexi Rabbit Reticulocyte Lysate, Promega) was performed as described previously (Kowalska et al., 2009) except that in this experiment luciferase mRNA transcripts capped with different analogues were used. Briefly, the *in vitro* translation reactions were performed in 12.5 µL volume for 60 min at 30°C in conditions determined for cap dependent translation. The reaction mix was pre-incubated for 60 min at 30°C prior to addition of inhibitor and capped luciferase transcript and start of the translation reaction. Reactions were stopped by chilling on ice and the luciferase activity was measured in a luminometer (Glomax. Promega) using a Luciferase Assay System (Promega). Measured relative luminescence units (RLU) were plotted onto a graph against tested concentrations of m^{7}GTP.

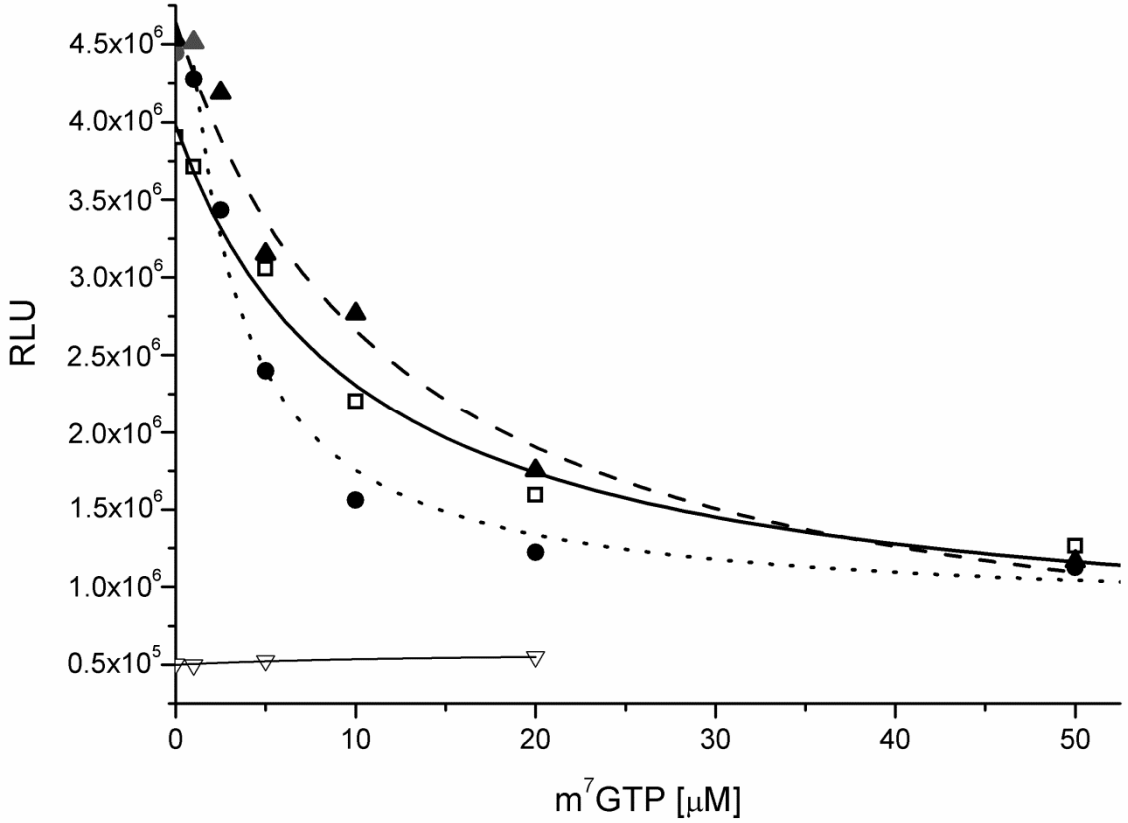


Figure S4.

Electrophoretic mobility of transcripts capped *in vitro* with m^{7}GpppG or m^{7}Gp$_{N\text{-}Biot}$ppG in the presence of streptavidin (RNA-EMSA). Capped RNA transcripts (1 pmol) and formed RNA-streptavidin complexes were analyzed using RNA High Sensitivity chips and an EXPERION automated electrophoresis system (BIORAD) as described in Experimental procedures. (L – RNA size marker; internal system lower alignment marker is indicated as pink triangle).

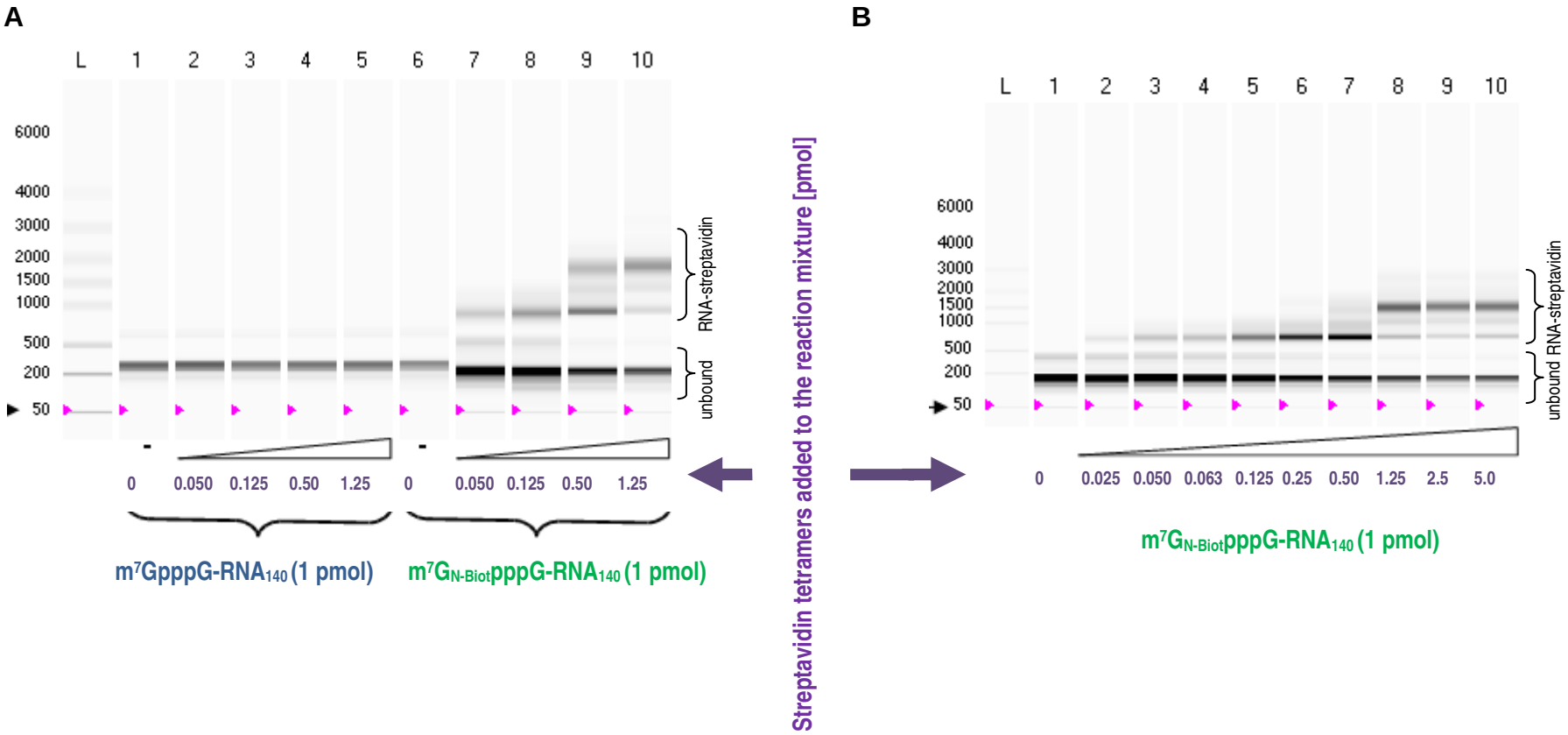


Figure S5.

Translation activity of luciferase mRNA transcripts capped *in vitro* with m^7GpppG (dense lines), $m_2^{7,2'\text{-}O}GpppG$ (spares lines), or $m^7Gp_{N\text{-}Biot}ppG$ (grid lines) observed in the presence of streptavidin in the translation system. Translation reactions were performed in conditions for cap-dependent translation as described earlier in the text (Figure S3), except that here, after the pre-incubation step, an appropriate dilution of streptavidin and capped luciferase transcript was added into the reaction mixture to start the translation reaction. After 60 min the reaction was stopped by putting on ice and luciferase activity was measured in a luminometer. The figure shows data normalized against observed luciferase activity for reactions without streptavidin [=1] (individually for each type of capped transcript). 2-3 experiments were performed and SD are shown). The streptavidin concentrations given are for the tetrameric form. Each translation reaction mixture contained capped mRNA at ca. 9 nM concentration (the total reaction volume was 12.5 µL). The streptavidin concentrations given are for the tetrameric form.

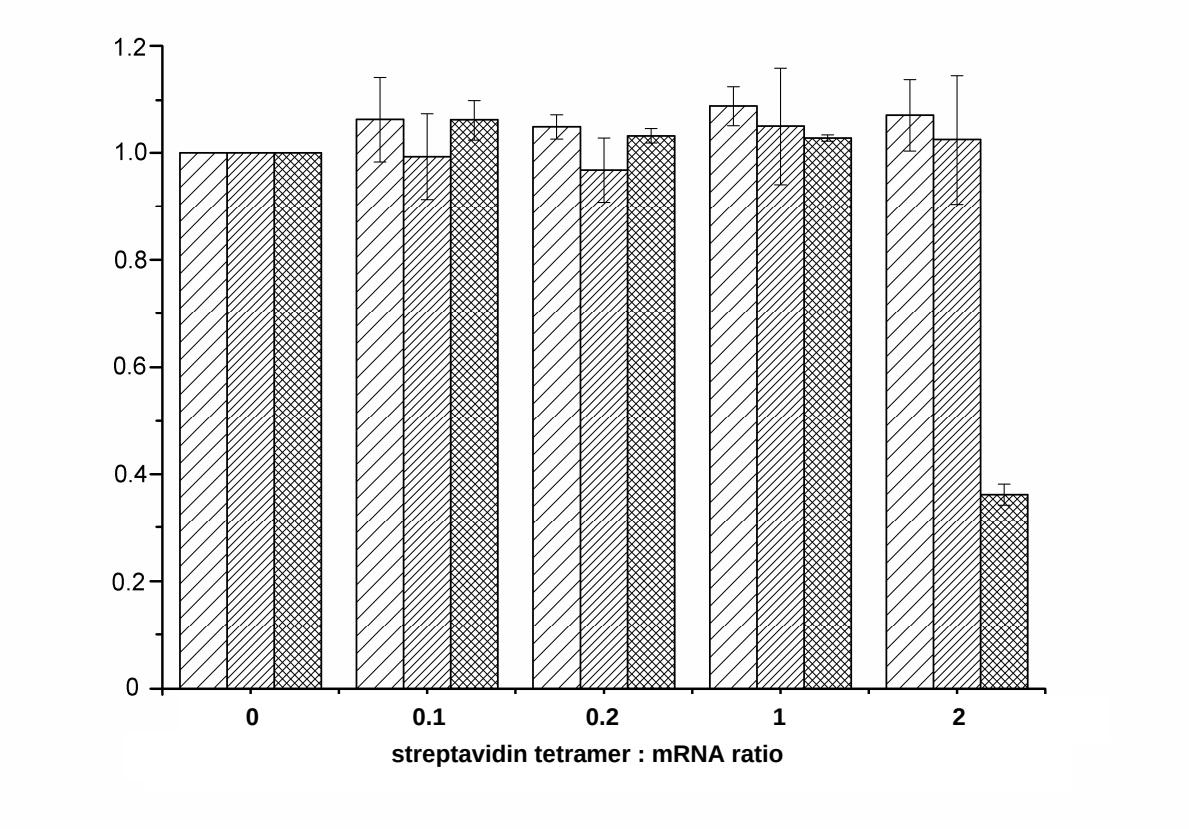
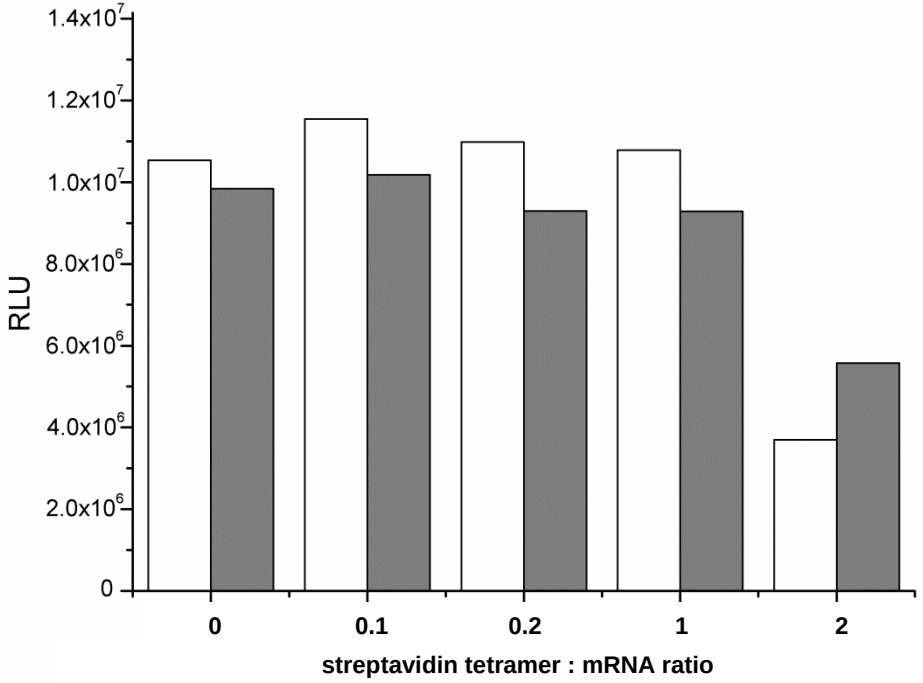


Figure S6.

Influence of streptavidin on translation of $m^7Gp_{N\text{-}Biot}ppG$-capped luciferase mRNA transcripts *in vitro* in RRL. Translation reactions were performed as described in Figure S5, except that streptavidin was added into the reaction either at the same time point as luciferase mRNA (white column bars) or 15 min after the translation reaction was started with mRNA transcript (grey column bars). Each translation reaction mixture contained capped mRNA at ca. 9 nM concentration (the total reaction volume was 12.5 µL). The streptavidin concentrations given are for the tetrameric form.



III. Experimental procedures

1. General information

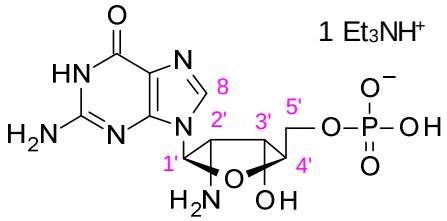
Solvents and chemical reagents were purchased from Sigma-Aldrich, unless otherwise stated. 2'-Amino-2'-deoxyguanosine was purchased from Jena Bioscience.

The synthesized nucleotides were purified by ion-exchange chromatography on a DEAE Sephadex A-25 (HCO_3^- form) column. The column was loaded with reaction mixture and washed thoroughly with deionized water (until the eluate did not precipitate with $AgNO_3$ solution). Then, nucleotides were eluted using a linear gradient of triethylammonium bicarbonate (TEAB) in deionized water. Collected fractions were analyzed spectrophotometrically at 260 nm and fractions containing the desired products, after being analyzed by RP HPLC, were poured together. After evaporation under reduced pressure with repeated additions of ethanol (to decompose TEAB), nucleotides were isolated as triethylammonium (TEA) salts. The final products (cap analogues **1** and **2** used for biological studies) were further purified by RP HPLC, in which case, after freeze-drying 3 times, they were isolated as ammonium salts. The reaction yields were calculated on the basis of optical density miliunit measurements (opt.mu = solution absorbance · solution volume in mL) of isolated products and appropriate starting materials. Optical unit measurements were performed in 0.1 M phosphate buffer (pH 6 for compound **4** and pH7 for all other nucleotides) at 260 nm.

Analytical HPLC was performed on Agilent Tech. Series 1200 using a Supelcosil LC-18-T HPLC column (4.6 x 250 mm, flow rate 1.3 mL/min) with a linear gradient 0–25 % of methanol in 0.05 M ammonium acetate buffer (pH 5.9) for 15 min and UV-detection at 254 nm. Semi-preparative HPLC was performed on the same apparatus equipped with a Discovery RP Amide C-16 HPLC column (25 cm x 21.2 mm, 5µm, flow rate 5.0 mL/min) with a linear gradient 0-15% of acetonitrile in 0.05 M ammonium acetate buffer (pH 5.9) and UV-detection at 260 nm. The structure and homogeneity of intermediates and final products was confirmed by re-chromatography on RP HPLC, low resolution mass spectrometry using negative electrospray ionization and ^{1}H NMR and ^{31}P NMR spectroscopy. Compounds **1** and **2** were also analyzed by high resolution mass spectrometry using positive electrospray ionization (HRMS ES(+)). Mass spectra were recorded on Thermo Scienific LTQ Orbitrap Velos (high resolution spectra) and AB Sciex API 3200 (low resolution spectra) spectrometers. NMR spectra were recorded at 25ºC on a Varian UNITY-plus spectrometer at 399.94 MHz (^{1}H NMR) and 161.90 MHz (^{31}P NMR). ^{1}H NMR chemical shifts were reported to sodium 3-trimethylsilyl-[2,2,3,3-D_4]-propionate (TSP) in D_2O as an internal standard. ^{31}P NMR chemical shifts were reported to 20 % phosphorus acid in D_2O as an external standard. The raw NMR data were processed using ACD/Labs 12.0 Software.

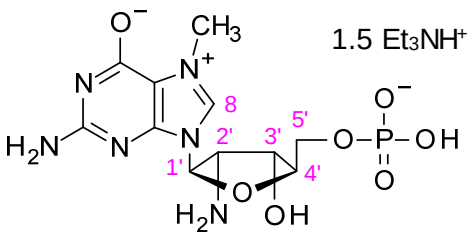
2. Chemical syntheses

2.1. 2'-amino-2'-deoxyguanosine 5'-monophosphate triethylammonium salt (3)



2'-amino-2'-deoxyguanosine (300 mg, 1.06 mmol) was suspended in trimethylphosphate (5 mL) and cooled on an ice-bath for 5 min. Then, $POCl_3$ (0.397 µL, 4.26 mmol) was added and the mixture was stirred at 0 °C. The reaction was monitored by analytical RP HPLC, quenched after 8 h by addition of water and adjusted to pH 6 with solid $NaHCO_3$. The mixture was separated on DEAE Sephadex using a 0-0.8 M gradient of TEAB yielding 10150 mOD_{260} (395 mg, 0.84 mmol, 79 %) of **1** as an triethylammonium (TEA) salt. Analytical RP HPLC R_t: 4.5 min, MS ESI(-) *m/z* 361.1 calc. $[M-H]^-$ 361.1; ^1H NMR (400 MHz, D_2O, 25 ℃) δ ppm: 8.11 (s, 1H, H8), 6.20 (d, *J*=7.7 Hz, 1H, H1'), 4.72 (dd, *J*=5.8, 1.5 Hz, 1H, H 3'), 4.52 (dd, *J*=7.7, 5.8 Hz, 1H, H2'), 4.45 (m, 1H, H4'), 4.08 (m, 2H, 2 x H 5'), 3.20 (q, *J*=7.3 Hz, 6H), 1.28 (t, *J*=7.3 Hz, 9H); ^{31}P NMR (D_2O, 162MHz, 25 ℃) δ ppm: 1.47 (bs, 1P)

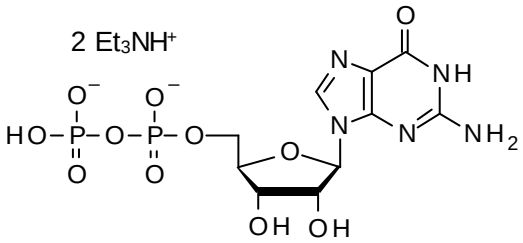
2.2. 2'-amino-2'-deoxy-7-methylguanosine 5'-monophosphate (4)



To a suspension of compound **3** (5075 mOD_{260}, 198 mg, 0.42 mmol) in DMSO (3 mL) methyl iodide (3.36 mmol, 210 µL) was added and the mixture was stirred at RT. The reaction was monitored by analytical RP HPLC. After 3h the reaction was quenched by addition of 30 mL of water and washed twice with 10 mL of diethyl ether. The aqueous layer was separated by DEAE Sephadex using a 0-0.8 M gradient of TEAB yielding 3115 mOD_{260} (130 mg, 0.27 mmol, 65 %) of **3** as a TEA salt. MS ESI(-)

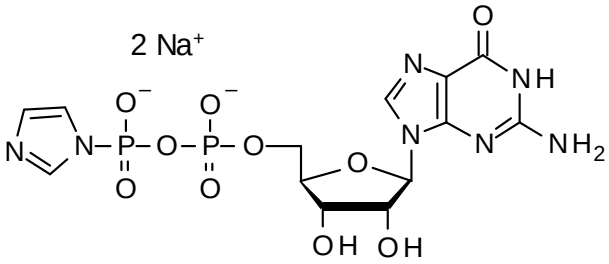
m/z 375.1 calc. $[M-H]^-$ 375.1; ^{1}H NMR (400 MHz, D_2O, 25 ℃) δ ppm: 6.25 (d, *J*=6.2 Hz, 1H, H1'), 4.66 (br. d, *J*=4.6 Hz, 1H, H 3'), 4.55 (m, 1H, H4'), 4.29 (dd, *J*=6.2, 4.6 Hz, 1H, H2'), 4.13 (s, 3H, N^7-CH_3), 4.00 - 4.17 (m, 2H, 2 x H5'), 3.21 (q, *J*=7.2 Hz, 9H), 1.28 (t, *J*=7.2 Hz, 12H); ^{31}P NMR (D_2O, 162MHz, 25 ℃) δ ppm: 2.81 (bs, 1P)

2.3. Guanosine 5'-diphosphate bis(triethylammonium) salt (5)



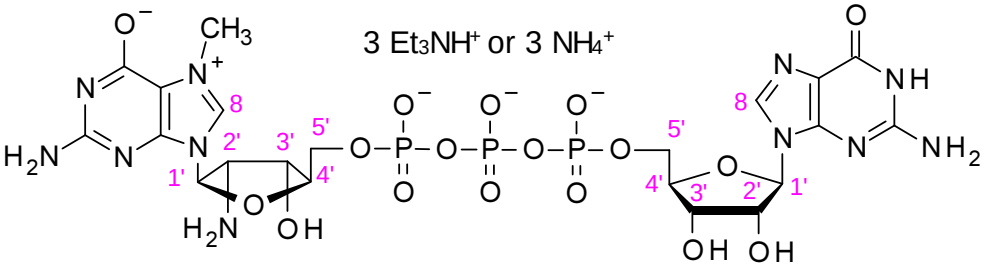
Compound **5** was synthesized from commercial GMP sodium salt according to the previously published procedure (Jemielity 2003). Briefly, GMP was converted into triethylammonium salt by passing through Dowex resin, evaporated to dryness and activated with imidazole employing a dithiodipyridine-triphenylphosphine system. GMP imidazolide was then coupled with phopshate triethylammonium salt (4 eq.) in DMF, in the presence of excess $ZnCl_2$ (8 eq.). The resultant GDP was purified by DEAE Sephadex and isolated as triethylammonium salt. Compound **3** may also be prepared from commercially available GDP trisodium salt by passing through Dowex resin in the triethylammonium form.

2.4. Guanosine 5'-diphosphate P-imidazolide disodium salt (6)



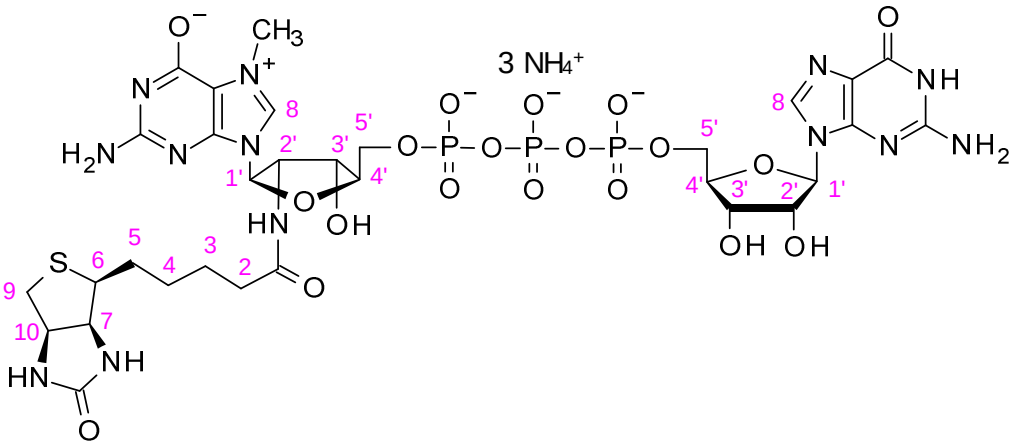
The title compound was synthesized from GDP triethylammonium salt *via* condensation with imidazole employing a dithiodipyridine-triphenylphosphine system according to the previously published procedure (Jemielity 2003).

2.5. P1-(2'-amino-2'-deoxy-*N7*-methylguanosin-5'-yl) P3-(guanosin-5'-yl) triphosphate m^7G_NpppG (1)



To a suspension of compounds **4** (1440 mOD_{260}, 60 mg, 0.13 mmol) and **6** (1570 mOD_{260}, 87 mg, 0.13 mmol) in DMF (1.5 mL) anhydrous $ZnCl_2$ (141 mg, 1.04 mmol) was added and the mixture was vigorously shaken until the reagents dissolved. The reaction progress was monitored by RP HPLC. The reaction was quenched after 6 h by addition of Na_2EDTA (390 mg, 1.04 mmol) solution in water (30 mL) and adjusted to pH 6 with solid $NaHCO_3$. Purification by DEAE Sephadex afforded 1720 mOD_{260} (84 mg, 0.076 mmol, 58 %) **5** as a triethylammonium salt as a white solid. In this form the compound was used in the biotinylation procedure (paragraph 2.6). For the purpose of the biological studies a part of the synthesized batch (ca. 10 mg) was additionally purified by semi-preparative RP HPLC using a linear gradient 0-15% of acetonitrile within 120 min (R_t = 35 min) and isolated as ammonium salt. HRMS ESI(+) *m/z* 802.1155 calc. $[M+H]^+$ 802.1107 ^{1}H NMR (400 MHz, D_2O, 25 ℃) δ ppm: 9.10 (s *slowly exchanging*, 1H, H8 m^7G), 8.09 (s, 1H, H8 G), 6.30 (d, *J*=6.4 Hz, 1H, H1' m^7G), 5.86 (d, *J*=5.8 Hz, 1H, H1' G), 4.86 (dd, *J*=5.5, 1.5 Hz, 1H, H3' m^7G), 4.70 (dd, *J*=5.8, 5.2 Hz, 1H, H2' G), 4.64 (br. s., 1H, H4' m^7G), 4.50 (dd, *J*=5.2, 3.5 Hz, 1H, H3' G), 4.48 (dd, *J*=6.4, 5.5 Hz, 1H, H2' m^7G), 4.37-4.21 (m, 5H, H4' G, 2 x H5' G, 2 x H5' m^7G), 4.08 (s, 3H, N^7-CH_3), ^{31}P NMR (D_2O, 162MHz, 25 ℃) δ ppm: -11.38 (d, *J*=18.6 Hz, 1P), -11.82 (d, *J*=18.6 Hz, 1P), -23.07 (t, *J*=18.6 Hz, 1P, P_β)

2.6. P1-(2'-aminobiotinyl-2'-deoxy-*N7*-methylguanosin-5'-yl) P3-(guanosin-5'-yl) triphosphate; $m^7G_{N\text{-}Biot}pppG$ (2)



To a solution of D-biotin (18 mg, 0.072 mmol) in 300 µL of DMSO were added TEA (10 µL, 0.099 mmol) and TSTU (22 mg, 0.073 mmol) and the mixture was shaken for 30 min. The resultant solution was added portion wise (10 additions over a period of 1h) to a solution of compound **1** (960 mOD_{260}, 0.048 mmol) in 0.5 ml of 500 mM aqueous borane buffer, pH 8.5. After each addition the pH was checked and, if necessary, carefully re-adjusted to 8.5 with 0.5 M NaOH. As RP HPLC analysis revealed complete conversion of **1** into **2**, the reaction mixture was diluted with 3 mL of water and neutralized with few drops of 50% acetic acid. The product was purified by semi-preparative RP HPLC using a linear gradient 0-15% of acetonitrile within 60 min (R_t = 32 min) yielding 580 mOD_{260} (29 mg, 0.026 mmol) of **2** (60%). HRMS ESI(+) *m/z* 1028.1944 calc. $[M+H]^+$ 1028.1883; ^1H NMR (400 MHz, D_2O, 25 ℃) *δ* ppm: 9.13 (s *slowly exchanging*, 1 H, H8 m^7G), 7.97 (s, 1 H, H8 G), 5.93 (d, *J* = 7.3 Hz, 1 H, H1' m^7G), 5.78 (d, *J* = 6.1 Hz, 1 H, H1' G), 4.74 (dd, *J* = 7.3, 5.6 Hz, 1 H, H2' m^7G), 4.66 (dd, *J* = 6.1, 5.6 Hz, 1 H, H2' G), 4.59 (dd, *J* = 7.9, 4.9 Hz, 1 H, H10 Biot), 4.52 (dd, *J* = 5.8, 1.4 Hz, 1 H, H3' m^7G), 4.47 (m overlapped with 4.46, 1 H, H4' m^7G), 4.46 (m overlapped with 4.47, H3' G), 4.36 - 4.33 (m, 4 H, H4' G; H5' H5'' G, H7 Biot), 4.32-4.20 (m, 2H, H5' H5'' m^7G), 4.11 (s, 3 H, N^7-CH_3), 3.21 (td, *J* = 4.9, 9.4 Hz, 1 H, H6

Biot), 2.96 (dd, *J* = 13.1, 4.9 Hz, 1 H, H9a Biot), 2.75 (d, *J* = 13.1 Hz, 1 H, H9b Biot), 2.26 (m, 2 H, H2 Biot), 1.64 (m, 1H, H5a Biot), 1.53 (m, 2 H, 2 x H3 Biot), 1.46 (m, 1H, H5b Biot), 1.25 (m, 1H, H4a Biot), 1.11 (m, 1H, H4b Biot); ^{31}P NMR (D_2O, 162MHz, 25 ℃) *δ* ppm: -11.53 (overlapped d, 2P, P_α and P_γ); -23.14 (t, *J* = 19.2, 1P, P_β)

3. Synthesis of short capped RNA transcripts by *in vitro* transcription reaction.

Short capped RNA transcripts were synthesized on the synthetic double stranded DNA template (28 bp in length, sequence: 5'-AGTTGTATTTAGGTGACACTATAGCCCC-3'), where the sequence of the DNA-dependent SP6 RNA polymerase promoter is adjacent to the GCCCC sequence that was transcribed into RNA (5 nt). The standard *in vitro* transcription reaction (in 25 µL final volume) contained: SP6 transcription buffer (Fermentas), 3.4 µM DNA template, 48 U of RiboLock ribonuclease inhibitor (Fermentas), 2 mM CTP, 0.1 mM GTP and 1 mM dinucleotide cap analogue (the molar ratio of cap analogue to GTP was 10:1). To obtain non-capped transcript (GCCCC) a dinucleotide cap analogue was omitted in the reaction and GTP concentration was increased to 1 mM. The reaction mix was incubated for 5 min at 37 ℃ before 30U of SP6 RNA polymerase was added (to final concentration 1.2 U/µl). The transcription reaction was continued for 60 min at 37 ℃. Synthesized short RNA transcripts (5 nt in length from reaction purely with GTP, and 6 nt transcripts from reactions where dinucleotide cap analogue was present) were separated on 20% denaturating polyacryamide gel (19:1) with 7 M Urea, in 1X TBE buffer (30 min at 15 mA). Directly after electrophoresis RNA transcripts were visualized by UV shadowing.

4. Translation efficiency of mRNA transcripts capped with dinucleotide cap analogues.

4.1. Transcription *in vitro*

Capped, polyadenylated luciferase mRNAs were synthesized *in vitro* on a dsDNA template (amplified by PCR reaction) that contained: an SP6 promoter sequence of DNA-dependent RNA polymerase, a 5'UTR sequence of rabbit β-globin RNA, the entire firefly luciferase ORF and a string of 31 adenosines (dsDNA template: SP6 promoter —5'UTRβ-globin—LUCIFERASE--$(A)_{31}$).

A typical *in vitro* transcription reaction mixture (40 µL final volume) contained: SP6 transcription buffer (Fermentas), 0.7 µg of DNA template, 40 U of RiboLock ribonuclease inhibitor (final concentration 1 U/µL, Fermentas), 0.5 mM ATP/CTP/UTP, 0.1 mM GTP and 0.5 mM dinucleotide cap analogue (the molar ratio of the cap analogue to GTP was 5:1). The reaction mixture was pre-incubated at 37 ℃ for 5 min before addition of 40 U of SP6 RNA polymerase (Fermentas) to a final concentration of 1 U/µL and the reaction was continued for 45 min at 37 ℃. After incubation, re action mixtures were treated with DNase RQ1 (Promega), in transcription buffer, for 20 min at 37 ℃ at a concentration of 1 U per 1 µg of template DNA. RNA transcripts were purified using NucAway Spin Columns (Ambion), the integrity of transcripts was checked on a non-denaturating 1% agarose gel and concentrations were determined spectrophotometrically. Purified RNA transcripts were stored at -80℃.

4.2. *In vitro* translation efficiency of capped, polyadenylated RNA transcripts in rabbit reticulocyte lysate system (RRL).

A translation reaction in RRL was performed in 10 µL volume for 60 min at 30 ℃, in conditions determi ned for cap-dependent

translation. A typical reaction mixture contained: 40% RRL lysate (Flexi Rabbit Reticulocyte Lysate System, micrococcal nuclease treated, Promega), a mixture of amino acids (0.01 mM), $MgCl_2$ (0.6 mM), potassium acetate (190 mM) and 5'-capped luciferase RNA transcript. Four concentrations of each analysed transcript were tested in an *in vitro* translation reaction: 0, 0.1, 0.2, and 0.4 ng/μL. The activity of synthesized luciferase was measured in a luminometer. Results were plotted onto a graph (where X axis was RNA concentration in ng/μL, and Y axis was relative light units RLU) and linear regression was used to fit a straight line to the obtained data points. The translation efficiency of RNA transcripts capped with different cap analogues was characterized as the slope constant value of the fitted line. Results were compared to the translation efficiency of *in vitro*-capped m^7GpppG-LUC-polyA RNA (translation efficiency normalized to 1).

5. Binding of streptavidin to biotinylated cap present on 5' end of RNA transcript. RNA-EMSA (RNA electrophoretic mobility shift assay).

5.1. Synthesis of 140 nt RNA transcripts capped with $m^7G_{N\text{-}Biot}$pppG (2) or standard m^7GpppG cap in *in vitro* transcription reaction.

Capped RNA transcripts (140 nucleotide in length) were synthesized on the dsDNA template that was a 157 bp PCR product that contained: an SP6 promoter sequence of DNA-dependent RNA polymerase, a 5'UTR sequence of rabbit β-globin RNA and the 97 base pairs of firefly luciferase ORF (dsDNA template: SP6 promoter —5'UTRβ-globin—LUCIFERASE (97 bp)). (The DNA template used here is a shortened version of the DNA template used to obtain full length luciferase RNA transcripts). Apart from the DNA

template concentration (0.5 µg of DNA template in 40 µL), the volume and composition of the *in vitro* transcription reaction mixture was the same as in the case of the transcription reaction for capped and polyadenylated luciferase RNA transcripts. The reaction mixture was pre-incubated at 37 ℃ for 5 min before addition of SP6 RNA polymerase (Fermentas) and the reaction was continued for 2 hours at 37 ℃. After incubation, reaction mi xtures were treated with DNase RQ1 (Promega), in transcription buffer, for 20 min at a 37 ℃ at a concentration of 1 U per 1 µg of template DNA. RNA transcripts were purified using NucAway Spin Columns (Ambion), the integrity of transcripts was checked on a 1.5 % agarose gel and concentrations were determined spectrophotometrically. Purified RNA transcripts were stored at -80℃.

5. 2. RNA-EMSA

RNA-EMSA experiments were performed on the synthesized *in vitro* 140 nt RNA transcripts capped with m^7GpppG or $m^7G_{N\text{-}Biot}$pppG **(2)** analogue. Before addition to the binding reaction mixture each capped RNA transcript was denaturated at 100℃ for 8 min and then cooled down on ice for at least 3 min. 85 ng (1.9 pmol)** of capped RNA transcript and increasing concentrations of streptavidin (Promega) (0.25 pmol, 0.5 pmol, 1 pmol, 4 pmol and 16 pmol) were used in binding reactions. The reaction mixture (9 µL final vol.) contained: 1 µL of RNA transcript, 1 µL of the proper dilution of streptavidin, 6 µL of the buffer (100mM Tris-Cl pH7.8 , 50mM NaCl, 10mM EDTA and 4U of RiboLock ribonuclease inhibitor) and 1 µL of H_2O. The binding reaction was incubated for 20 min on ice (4 ℃), and then 2 µL of ice-cooled sterile 80 % glycerol was added. Samples were loaded on 1.4 % agarose gel with ethidium bromide and electrophoresis was performed in 0.5X TBE buffer for 80 min at a constant voltage of 75 V (in a cold-room).

After electrophoresis gels were analysed under UV light.

** amounts of added RNA transcript were converted from nanograms (ng) to picomoles (pmol) using the estimated molecular weight of RNA molecules 140 nt in length that is ca. 48014 Da (g/mol).

*** amounts of streptavidin in moles were calculated for the tetrameric form of protein that has MW = 53.2 kDa. Dried pellets of purchased streptavidin were initially re-suspended in sterile H_2O to a concentration of 25 µM (to 1 mg, that is 0.0188 µmol of tetrameric streptavidin, was added 751.9 µL of H_2O; 25 µM solution gives 25 pmol/µL of streptavidin).

5. 3. RNA-EMSA on an EXPERION automated electrophoresis station (Biorad) using an Experion RNA HighSens Analysis Kit

RNA-EMSA experiments were performed on the synthesized *in vitro* 140 nt RNA transcripts capped with m^7GpppG or $m^7G_{N\text{-}Biot}$pppG **(2)** analogue. Before addition to the binding reaction mixture each capped RNA transcript was denaturated at 100℃ for 8 min and then cooled down on ice for at least 3 min. Usually 48 ng (1.06 pmol) of capped RNA transcript and increasing amounts of tetrameric streptavidin (Promega) (from 0.025 to 5 pmol) were used in binding reactions. The reaction mixture (3 µL final vol.) contained: 0.8 µL of RNA transcript, streptavidin (usually 0.5 to 1.5 µL of protein solution at proper concentration) and H_2O up to 3 µL. The binding reactions were incubated for 40 min on ice (4 ℃). During the incubation period the Exp erion RNA HighSens chip and the RNA Ladder were prepared according to manufacturer protocols. Then, 1 µL of each binding reaction was loaded into

separated sample wells on the chip and were run immediately on the Experion automated electrophoresis station. Data were collected using Experion software and presented as electrophoregrams or as virtual gels (that converted electrophoregram data into densitometric bands).

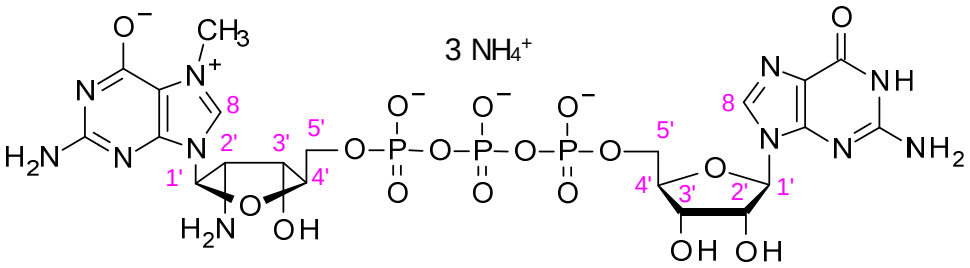
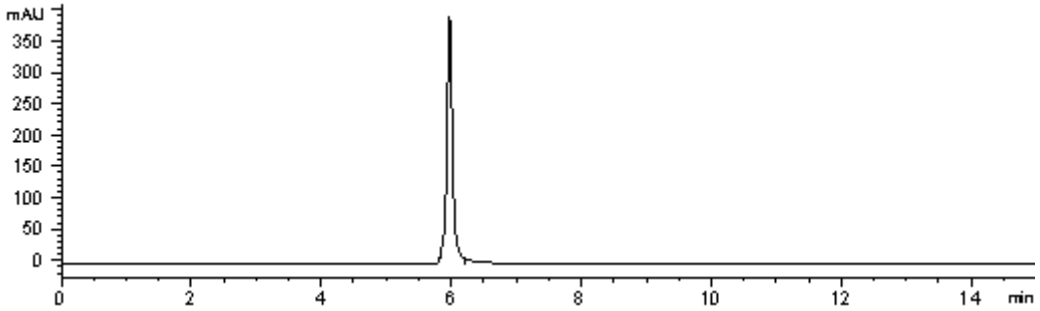
6. Supplementary references

J. Jemielity, T. Fowler, J. Zuberek, J. Stepinski, M. Lewdorowicz, A. Niedzwiecka, R. Stolarski, E. Darzynkiewicz and R.E. Rhoads, *RNA*, 2003, **9**, 1108-1122.

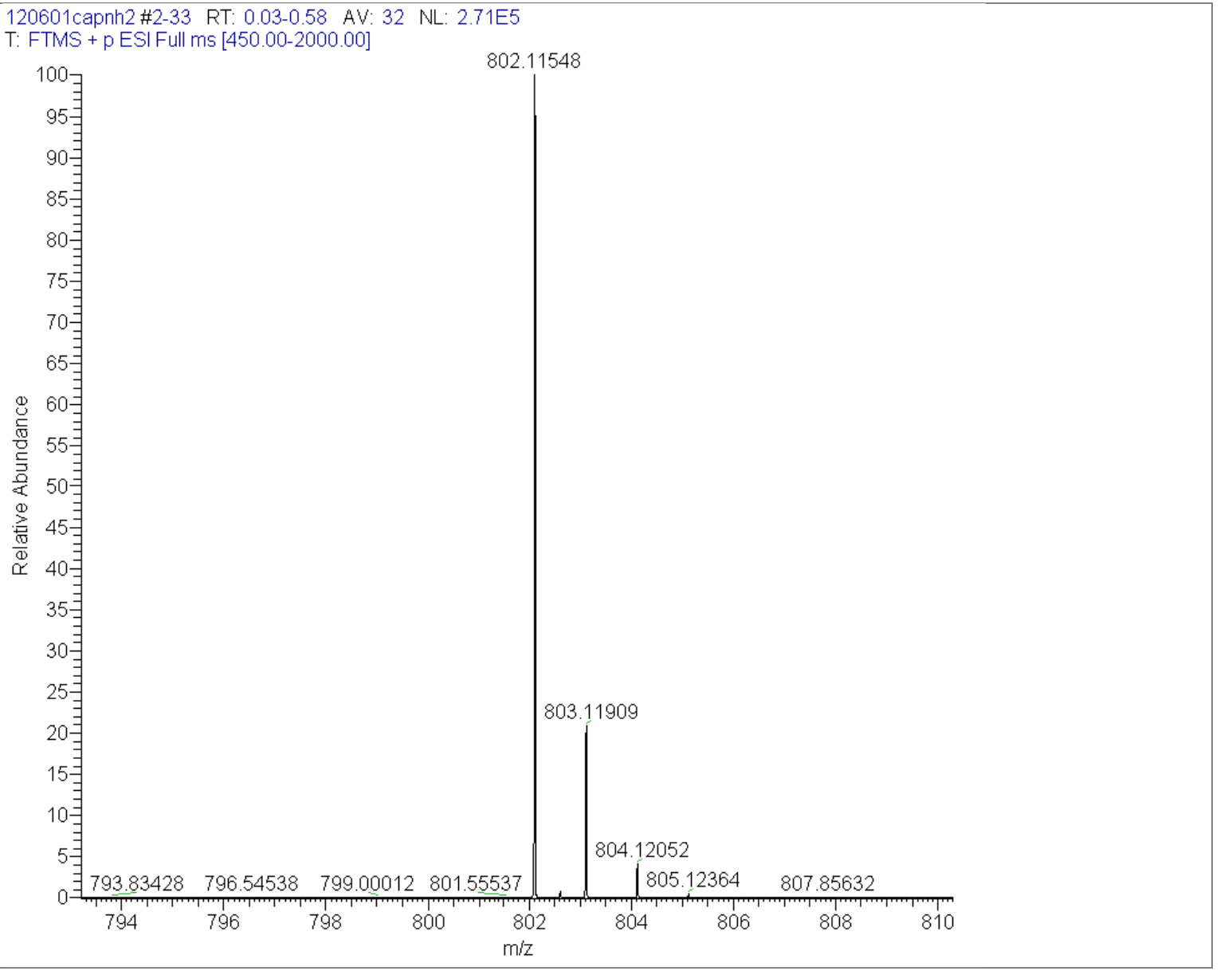
J. Kowalska, M. Lukaszewicz, J. Zuberek, M. Ziemniak, E. Darzynkiewicz and J. Jemielity, *Bioorg Med Chem Lett*, 2009, **19**, 1921-1925.

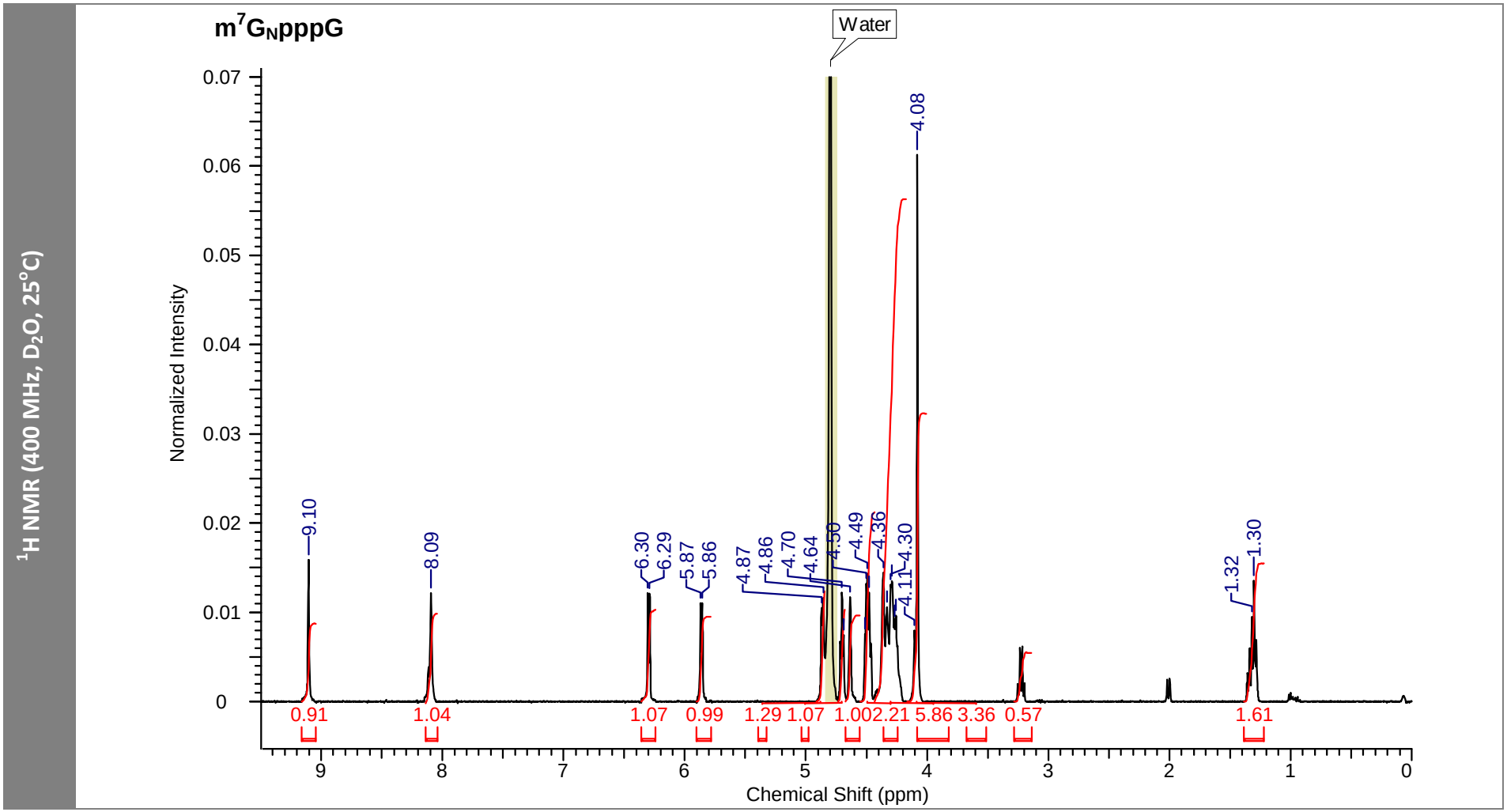
A. Niedzwiecka, J. Marcotrigiano, J. Stepinski, M. Jankowska-Anyszka; A. Wyslouch-Cieszynska; M. Dadlez; A. C. Gingras; P. Mak; E. Darzynkiewicz; N. Sonenberg; S. K. Burley; R. *J.* Stolarski *Mol. Biol.* 2002, **319,** 615-635.

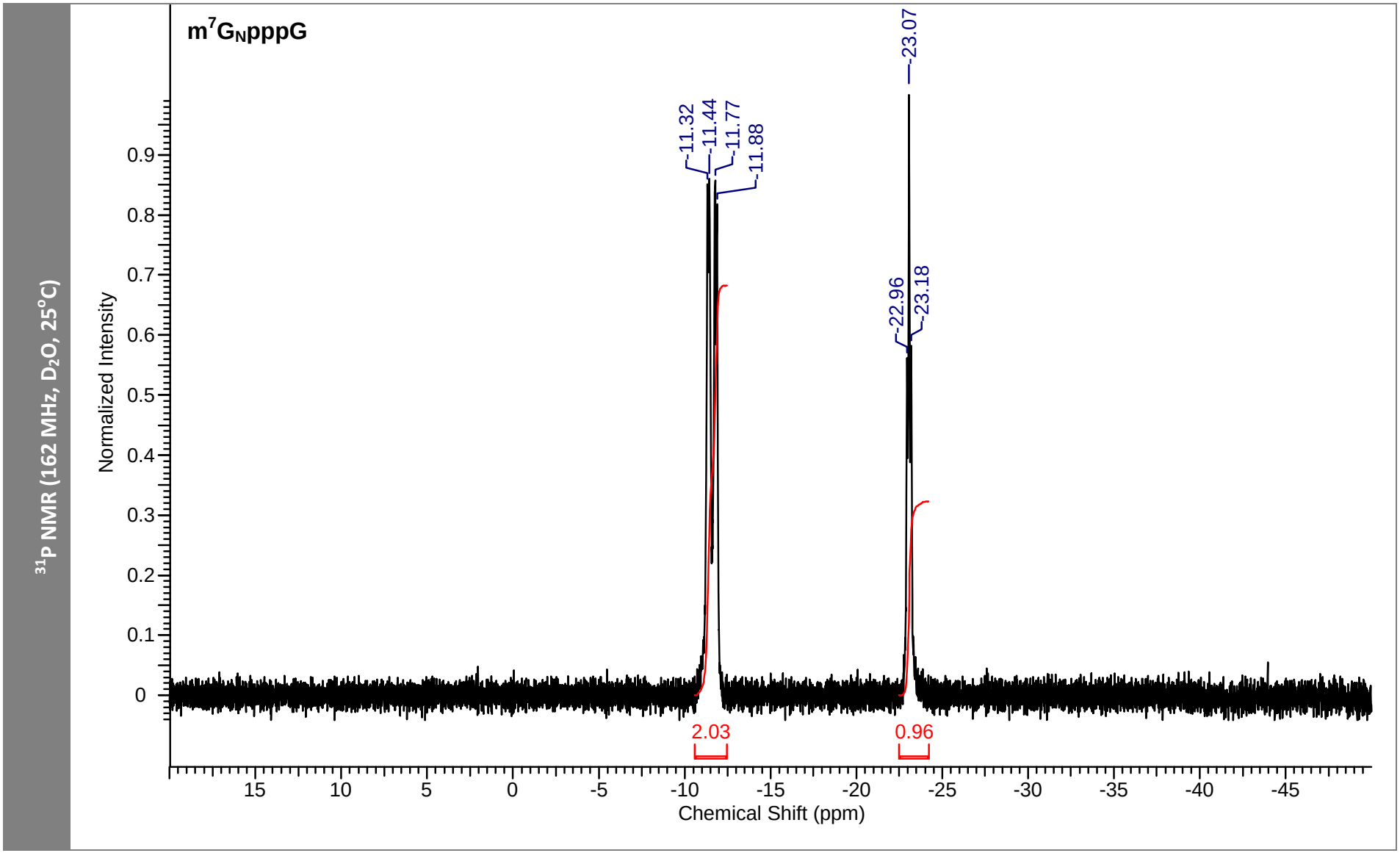
IV. RP HPLC profiles, negative ion ES-MS, ^{1}H and ^{31}P NMR spectra

	Compound 1 ($m7G_NpppG$)
Formula	
RP HPLC	

HRMS ES(+)

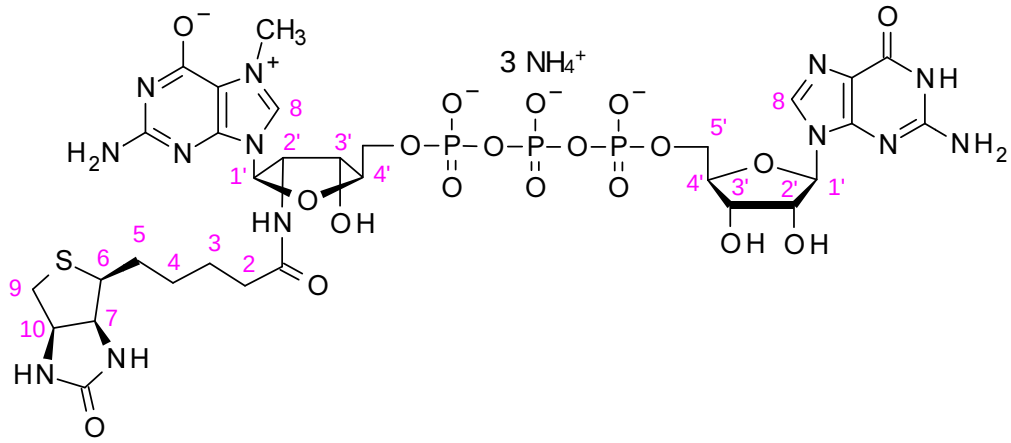

120601capnh2 #2-33 RT: 0.03-0.58 AV: 32 NL: 2.71E5
T: FTMS + p ESI Full ms [450.00-2000.00]
802.11548
803.11909
804.12052
805.12364
807.85632
793.83428
796.54538
799.00012
801.55537
Relative Abundance
m/z
100
95
90
85
80
75
70
65
60
55
50
45
40
35
30
25
20
15
10
5
0
794
796
798
800
802
804
806
808
810


m7GNpppG
Water
1H NMR (400 MHz, D2O, 25°C)
Normalized Intensity
Chemical Shift (ppm)
9.10
8.09
6.30
6.29
5.87
5.86
4.87
4.86
4.70
4.64
4.50
4.49
4.36
4.30
4.11
4.08
1.32
1.30
0.91
1.04
1.07
0.99
1.29
1.07
1.00
2.21
5.86
3.36
0.57
1.61

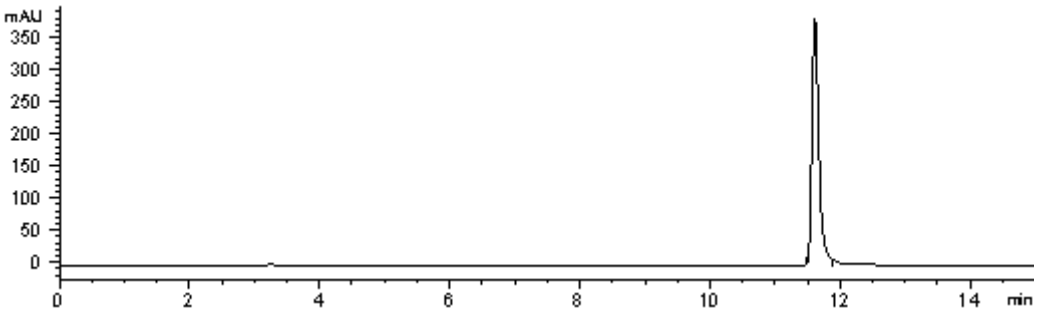

m7GNpppG
31P NMR (162 MHz, D2O, 25°C)
Normalized Intensity
Chemical Shift (ppm)
11.32
11.44
11.77
11.88
23.07
22.96
23.18
2.03
0.96

Compound 2 ($m^7Gp_{N\text{-}Biot}ppG$)

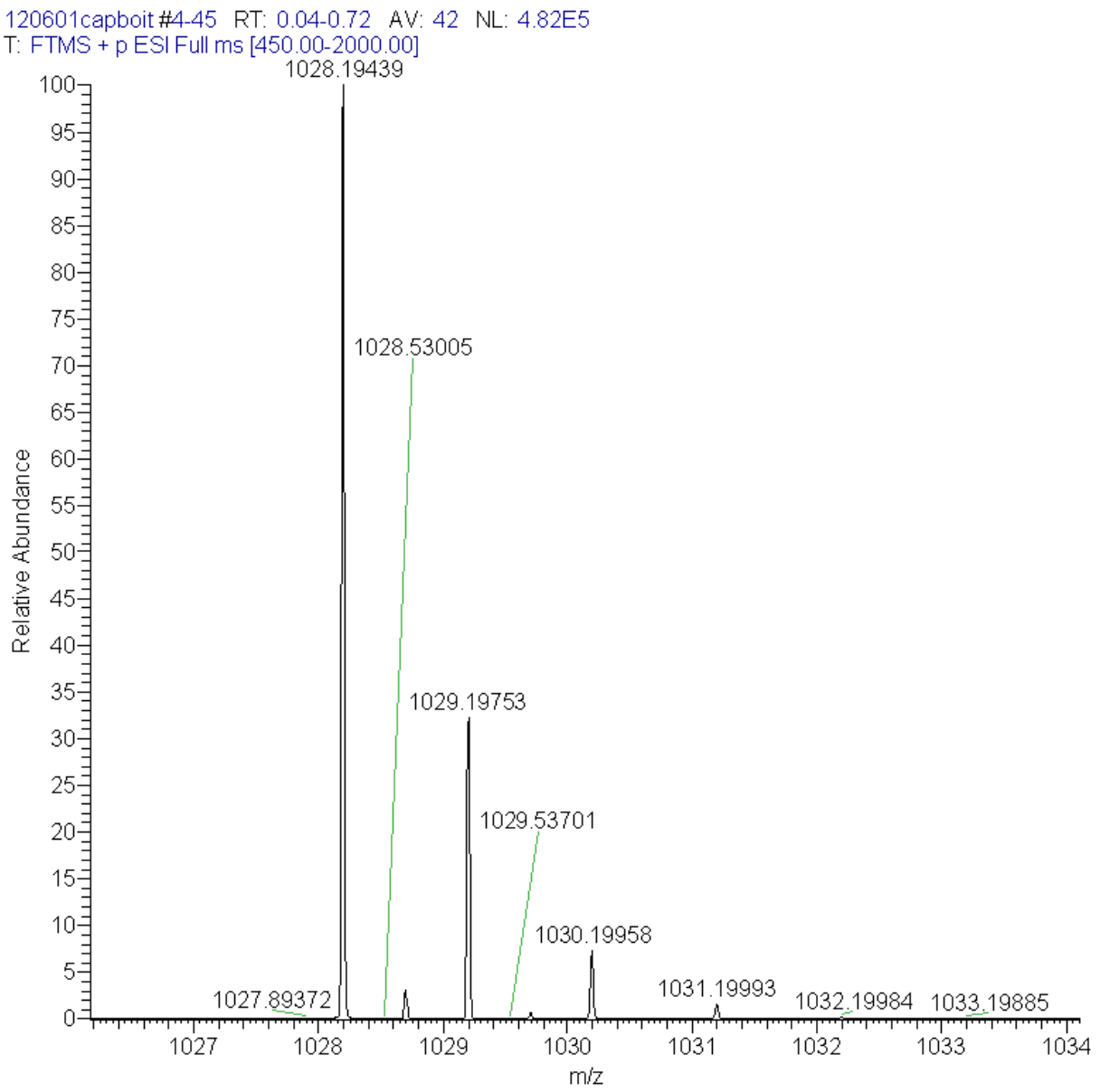
Formula

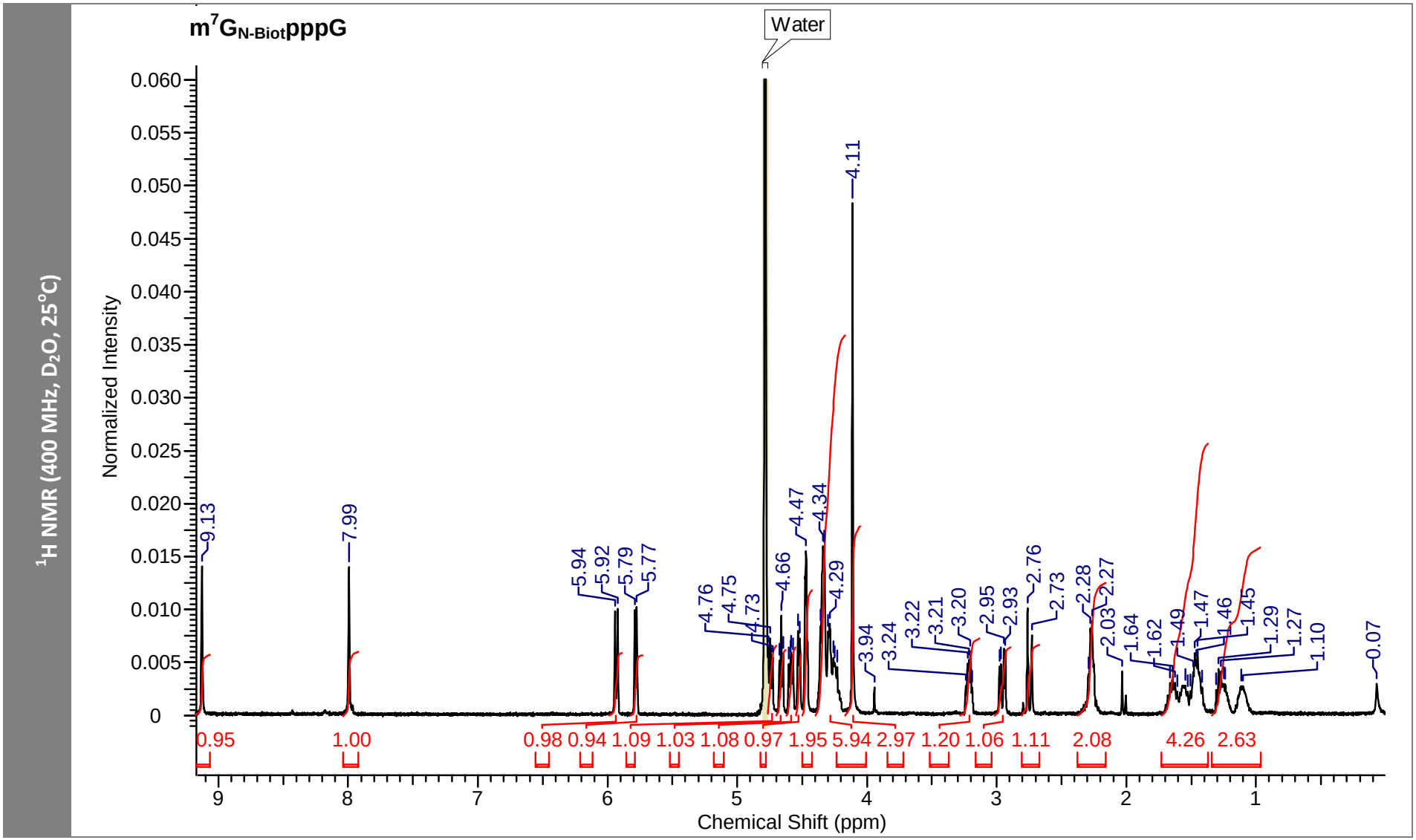


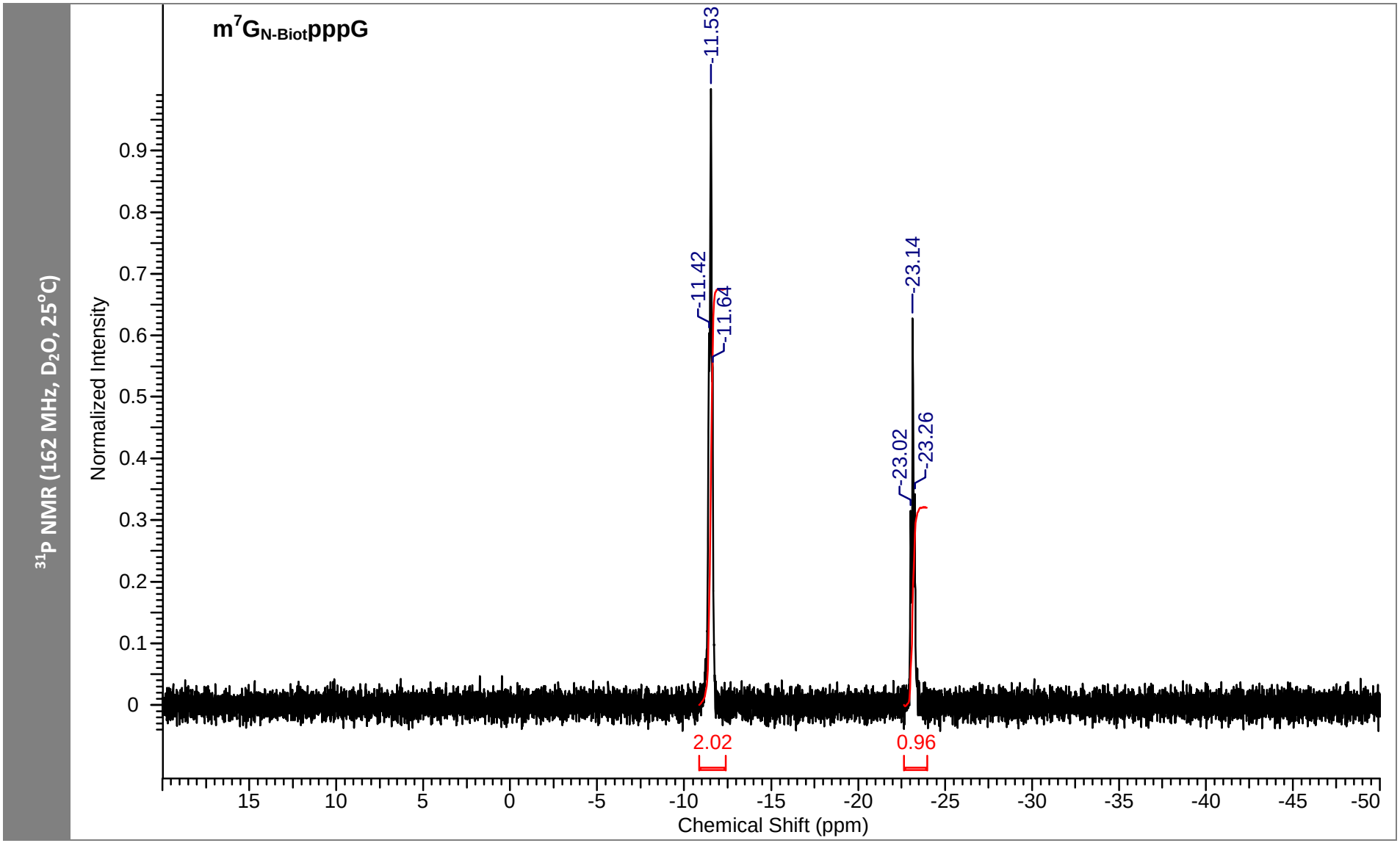
RP HPLC



HRMS ES(+)

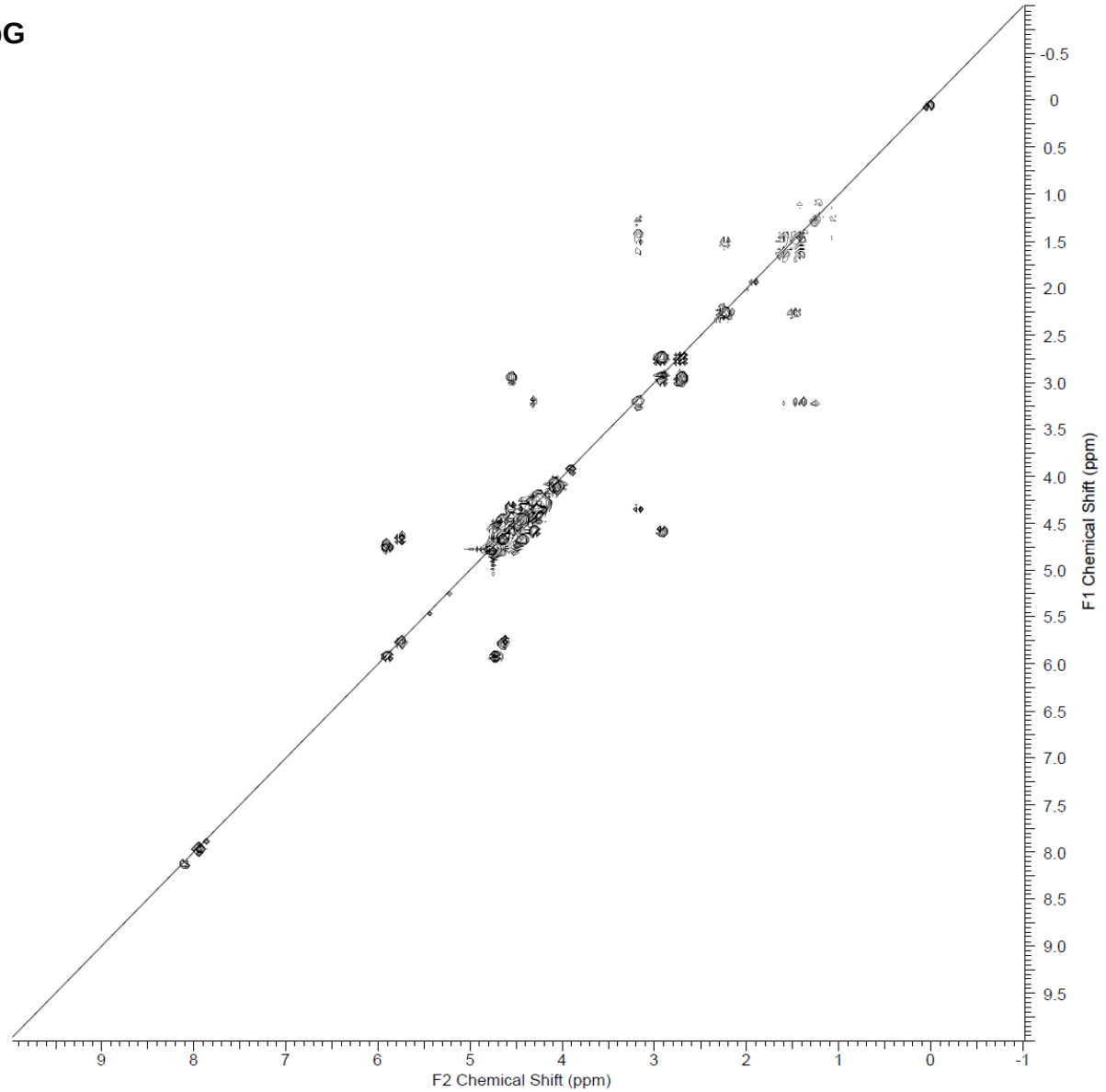


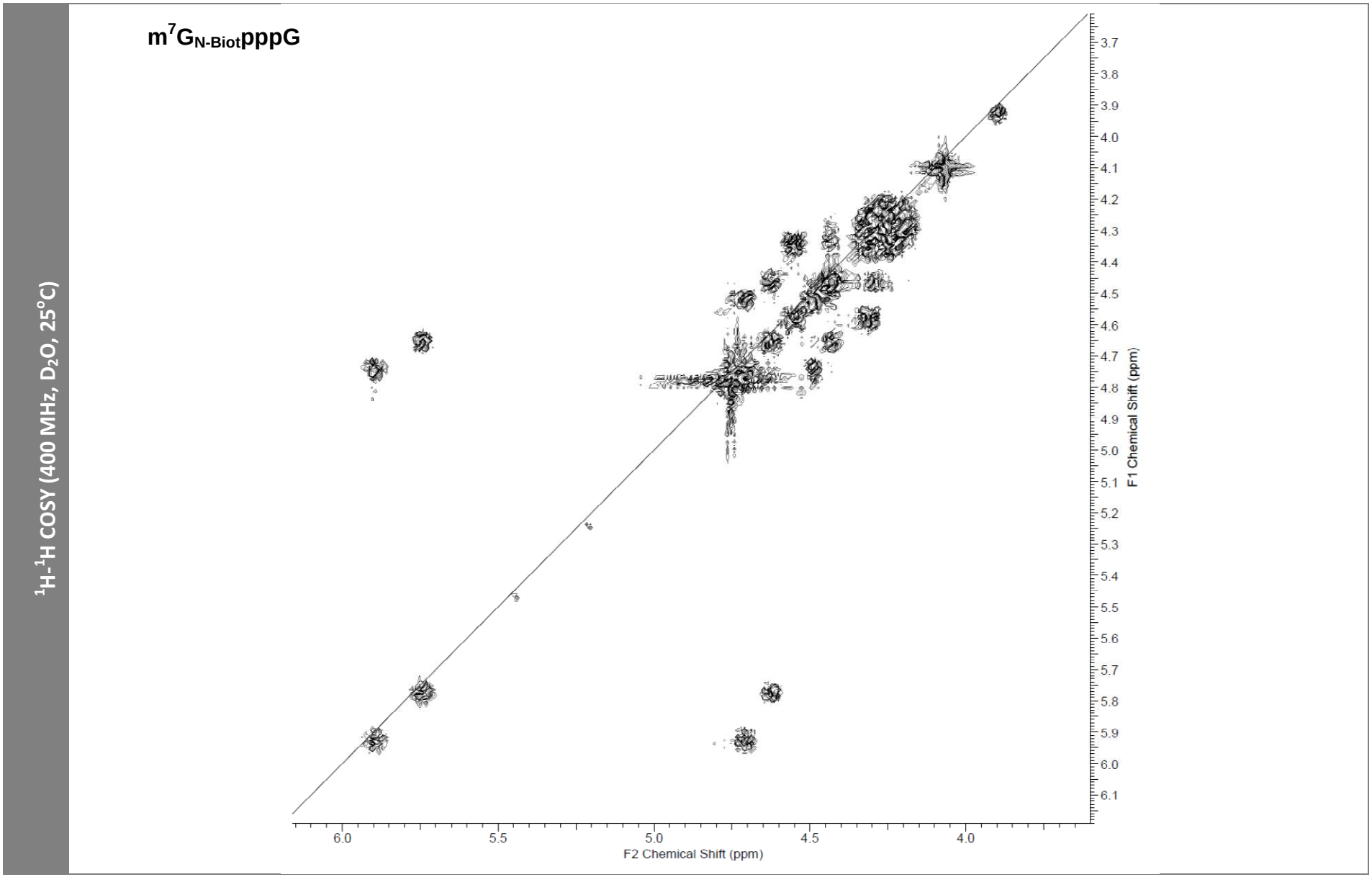

m7GN-BiotpppG
Water
1H NMR (400 MHz, D2O, 25°C)
Normalized Intensity
Chemical Shift (ppm)
9.13
7.99
5.94
5.92
5.79
5.77
4.76
4.75
4.73
4.66
4.47
4.34
4.29
4.11
3.94
3.24
3.22
3.21
3.20
2.95
2.93
2.76
2.73
2.28
2.27
2.03
1.64
1.62
1.49
1.47
1.46
1.45
1.29
1.27
1.10
0.07
0.95
1.00
0.98 0.94 1.09 1.03 1.08 0.97 1.95 5.94 2.97 1.20 1.06 1.11 2.08 4.26 2.63

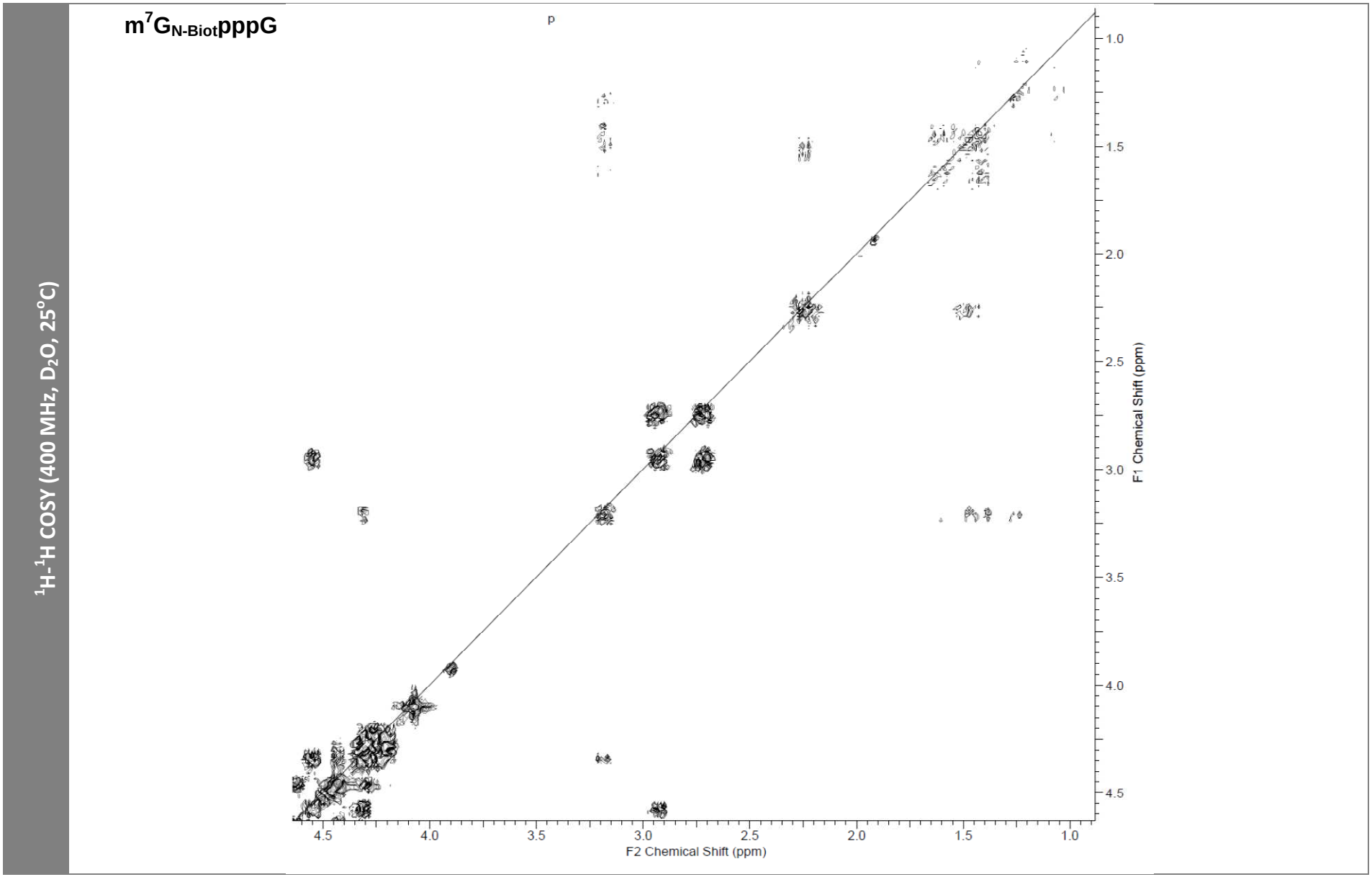


$m^7G_{N\text{-}Biot}pppG$

^1H-^1H COSY (400 MHz, D_2O, 25°C)




m7GN-BiotpppG
1H-1H COSY (400 MHz, D2O, 25°C)
F2 Chemical Shift (ppm)
F1 Chemical Shift (ppm)
6.0
5.5
5.0
4.5
4.0
3.7
3.8
3.9
4.0
4.1
4.2
4.3
4.4
4.5
4.6
4.7
4.8
4.9
5.0
5.1
5.2
5.3
5.4
5.5
5.6
5.7
5.8
5.9
6.0
6.1


m7GN-BiotpppG
p
1H-1H COSY (400 MHz, D2O, 25°C)
F1 Chemical Shift (ppm)
F2 Chemical Shift (ppm)
1.0
1.5
2.0
2.5
3.0
3.5
4.0
4.5