

The identity and methylation status of the first transcribed nucleotide in eukaryotic mRNA 5' cap modulates protein expression in living cells

Pawel J. Sikorski^{1,*} Marcin Warminski,² Dorota Kubacka,² Tomasz Ratajczak,¹ Dominika Nowis,^{1,3} Joanna Kowalska,^{2,*} Jacek Jemielity^{1,*}

¹ Centre of New Technologies, University of Warsaw, Banacha 2C, 02-097 Warsaw, Poland

² Division of Biophysics, Institute of Experimental Physics, Faculty of Physics, University of Warsaw, Pasteura 5, 02-093 Warsaw, Poland

³ Department of Genomic Medicine, Medical University of Warsaw, Nielubowicza 5, 02-097 Warsaw, Poland

* To whom correspondence should be addressed. Tel: +48 22 55 43774 ; Fax: +48 22 55 43771;
Email: p.sikorski@cent.uw.edu.pl, jkowalska@fuw.edu.pl, j.jemielity@cent.uw.edu.pl

SUPPLEMENTARY DATA

Table S1 List of used primers for RT-qPCR analysis.

human		
GAPDH	ACCCACTCCTCCACCTTTGAC	TGTTGCTGTAGCCAAATTCGTT
IFIT1	GATCAGCCATATTTCAATTTGAATC	GAAAATTCTCTTCAGCTTTTCTGTG
IFIT2	AAGAGGAAGATTTCTGAAGAGTGC	TCTCCAAGGAATTCTTATTGTTCTC
IFIT3	GAAGGAACTGGGCCGCCTGCTAAG	GCCCTGGCCCATTTCCTCACTACC
IFIT5	CGCTGAAGGAGGCCAGTATAG	CTGAAAGCGGCCATAGTGGTA
IL-6	AGACAGCCACTCACCTCTTCAG	TTCTGCCAGTGCCTCTTTGCTG
INFB1	TCTCCTGTTGTGCTTCTCCAC	GGCAGTATTCAAGCCTCCCAT
MDA5	GAGTCAAAGCCCACCATCTGA	CAGACCTTCTTCTGCCACTGT
RIG-I	ATGTGCTCCTACAGGTTGTGG	AACTGGGATCTGATTCGCAA
mouse		
GAPDH	AGGTCGGTGTGAACGGATTTG	TGTAGACCATGTAGTTGAGGTCA
Ifit1	ACATTGAAGAAGCCCTCAGCA	TCTACGCGATGTTTCCTACGG
Ifit1b12	GGGTTAGGAGGGACAAAGCAA	TCCTCTCTCTTTCCCCTGTAT
Ifit2	ACAGCAGACAGTTACACAGCA	TAGCTGTCGCAGATTGCTCTC
Ifit3	GCTCAGGCTTACGTTGACAAGG	CTTTAGGCGTGTCCATCCTTCC
IL-6	CTTCTTGGGACTGATGCTGGT	GGTCTGTTGGGAGTGGTATCC
INFB	GTCCGAGCAGAGATCTTCAGG	CCACCACTCATTCTGAGGCAT
MDA5	ATCTGCTTATCGCTACGACGG	TCGTGACAAGGCCATAACGAA
PKR	CACTCTGAACCCTCTGCCATT	CTGAGAGAAGCCACTCAAGGG
RIG-I	TGGAGTTGATGAGCCAATGCT	CACCAGCTTGAAACCAACCAG
TRL3	GTGAGATAACAACGTAGCTGACTG	TCCTGCATCCAAGATAGCAAGT
TRL7	GTTCTATGGAGAGCCGGTGATA	ATTCTTTAGATTTGGCGGCATA
TRL8	AATTTGCCTCAGAGCCTCCAA	ATCCAGCAAGTGAAGGTGAGG

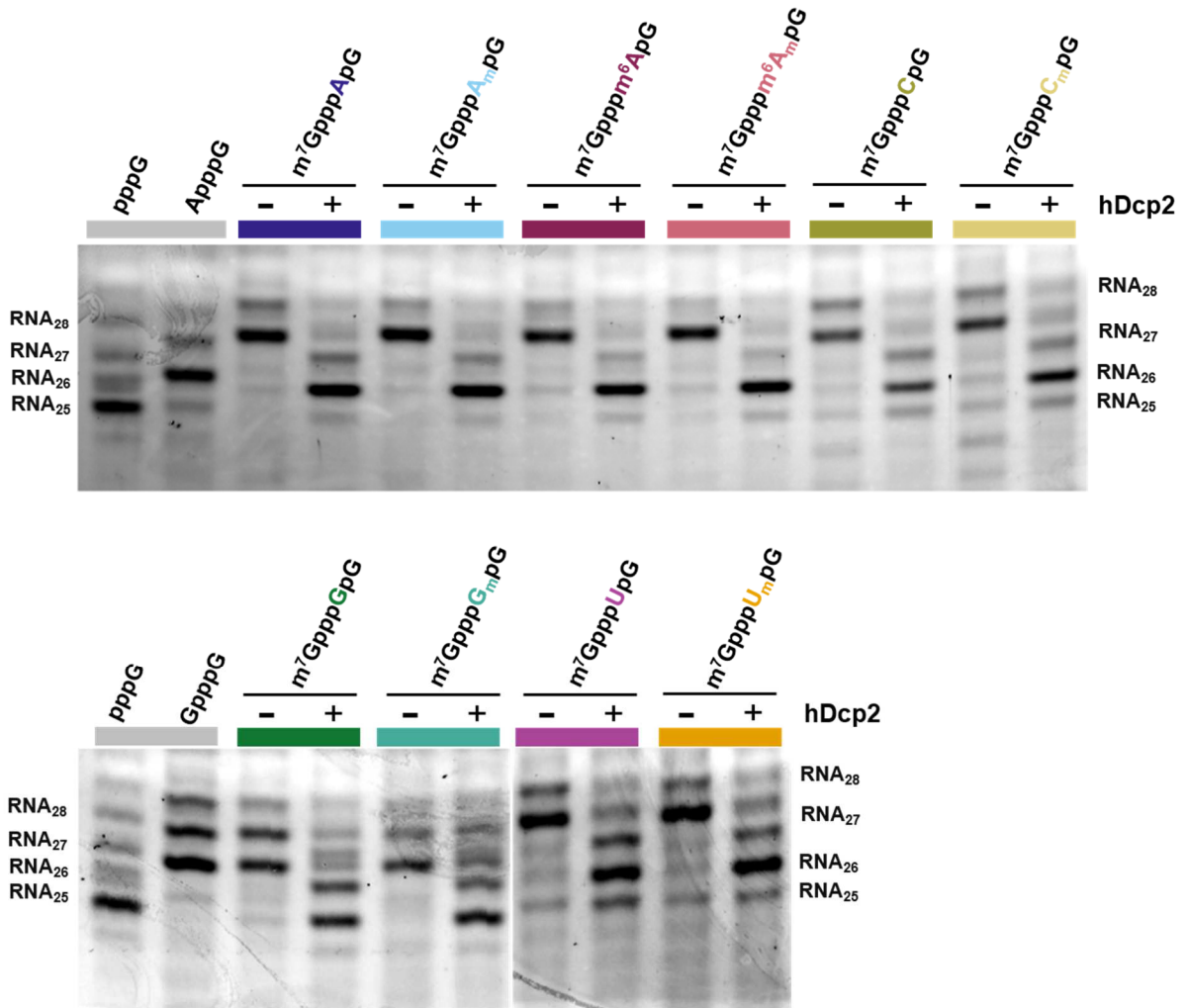


Figure S1. All trinucleotide cap analogs are incorporated in correct orientation. Short 25-nt transcripts were produced by IVT, followed by 3' end trimming by DNAzyme 10-23 and removal of uncapped RNAs by 5'-polyphosphates and Xrn1 treatment. Purified RNA (30 ng each) was subjected to exhaustive treatment with hDcp2 (100 nM) for 60 min. Aliquots taken at time 0 (-) and after 60 min of incubation (+) were resolved by PAGE.

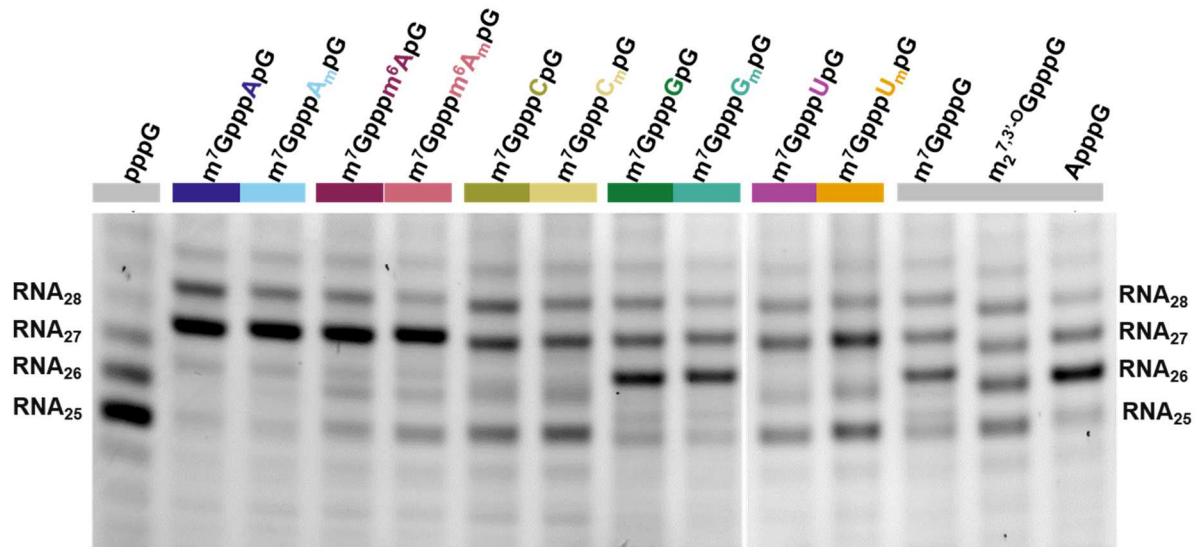


Figure S2. Analysis of short RNAs obtained by in vitro transcription using T7 RNA polymerase in the presence of different cap analogs (second replicate). IVT RNAs were obtained as described in Figure 2 (main manuscript).

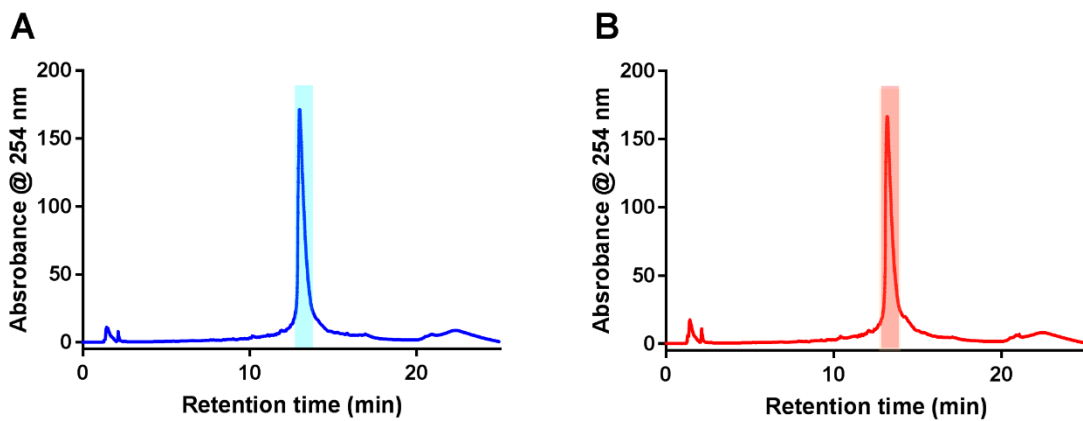


Figure S3. Example RP-HPLC profiles from the purification of *Gaussia* mRNAs capped with $m^7GpppApG$ (A) and m^7GpppA_mPG (B). The fraction collected as purified mRNA is marked by the coloured rectangle.

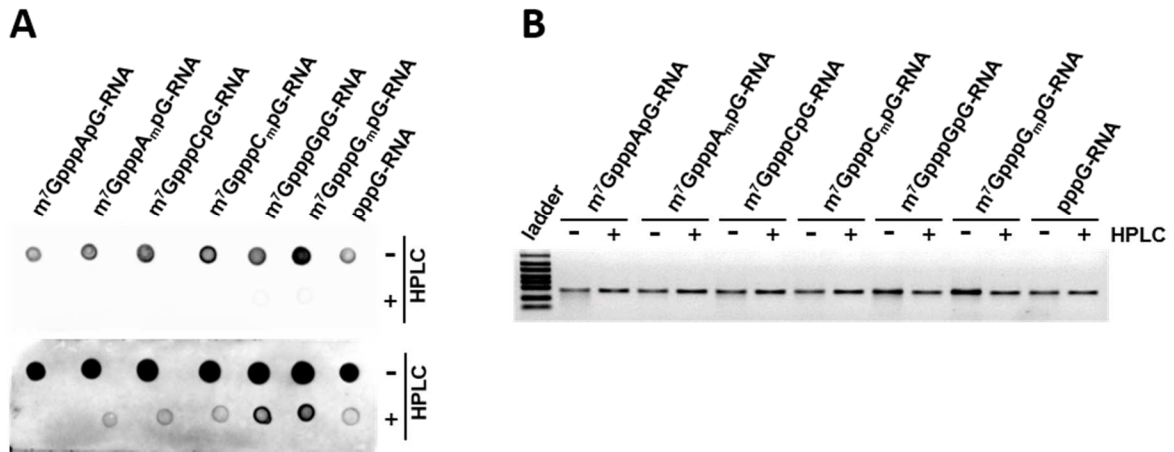


Figure S4. HPLC purification of in vitro transcribed mRNA removes dsRNA contaminants. (A) 25 ng of *Gaussia* mRNA capped with selected analogues with and without HPLC purification were blotted and analysed with J2 dsRNA-specific (@dsRNA) antibodies. Upper and lower panel present short and long exposition time, respectively. (B) As a control of the amount of RNA analysed in (A) also 25 ng of each mRNA was run on 1.2% TBE agarose gel.

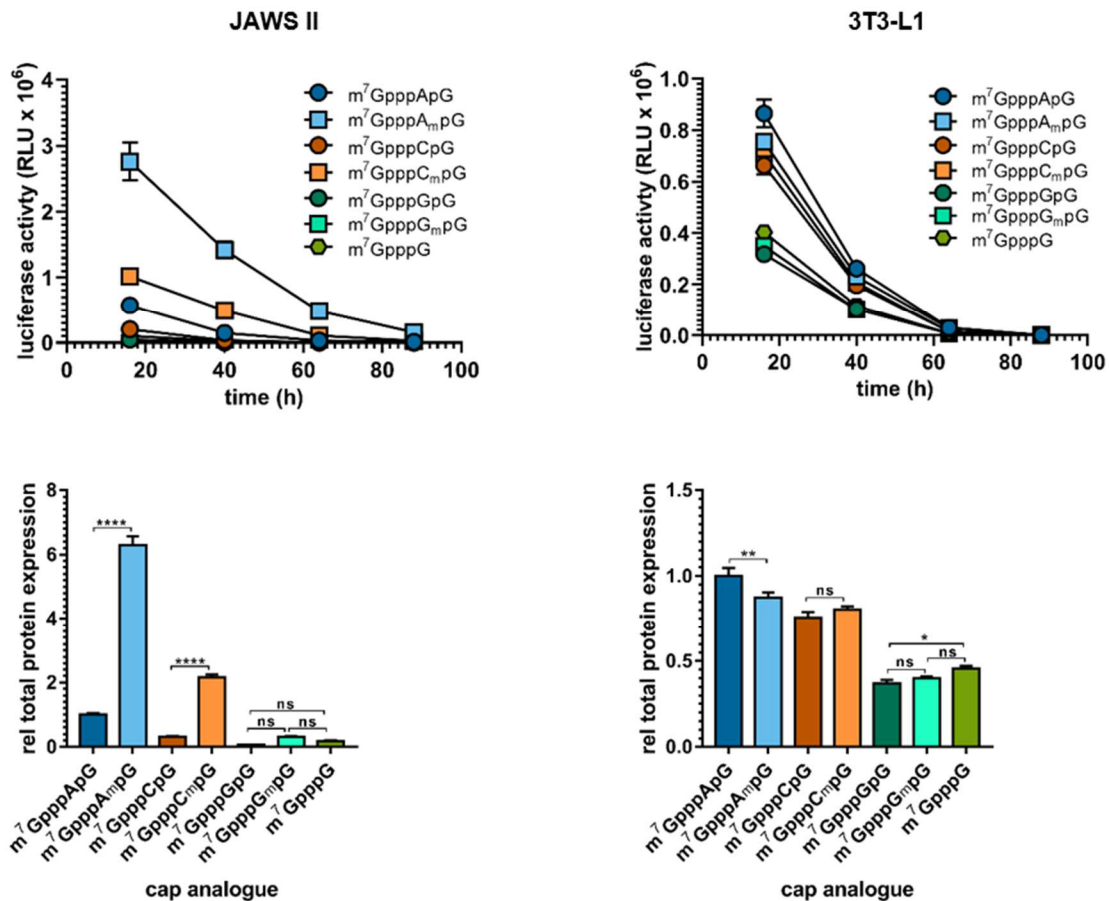


Figure S5. The influence of G at the first transcribed nucleotide site on translational properties of HPLC-purified IVT mRNA. (TOP) Time course of Gaussia luciferase activity in the supernatant of JAWS II and 3T3-L1 cells starting 16 hours after transfection with IVT mRNAs bearing different nucleotides at TSS at their 5' ends and continuing for 3 days. Data points present mean values \pm SD from single independent biological replicate, which consists of 3 independent transfection reactions. (BOTTOM) Total protein expression (cumulative luminescence) produced over 4 days by JAWS II and 3T3-L1 cells transfected with capped mRNAs. Bars represent mean value normalized to m⁷GpppApG-RNA \pm SD. Statistical significance: n.s.: not significant, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$ (one-way ANOVA with Turkey's multiple comparisons test).

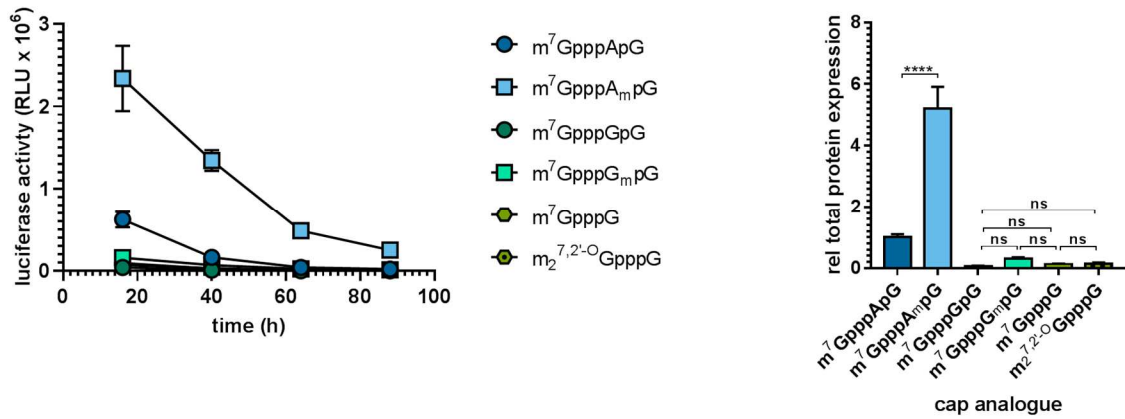


Figure S6. The influence of ARCA modification on translational properties of HPLC-purified IVT mRNA. (LEFT) Time course of *Gaussia* luciferase activity in the supernatant of JAWS II cells starting 16 hours after transfection with IVT mRNAs bearing different nucleotides at TSS at their 5' ends and continuing for 3 days. Data points present mean values \pm SD from single independent biological replicate, which consists of 3 independent transfection reactions. (RIGHT) Total protein expression (cumulative luminescence) produced over 4 days by JAWS II cells transfected with capped mRNAs. Bars represent mean value normalized to m⁷GpppApG-RNA \pm SD. Statistical significance: n.s.: not significant, **** p < 0.0001 (one-way ANOVA with Turkey's multiple comparisons test).

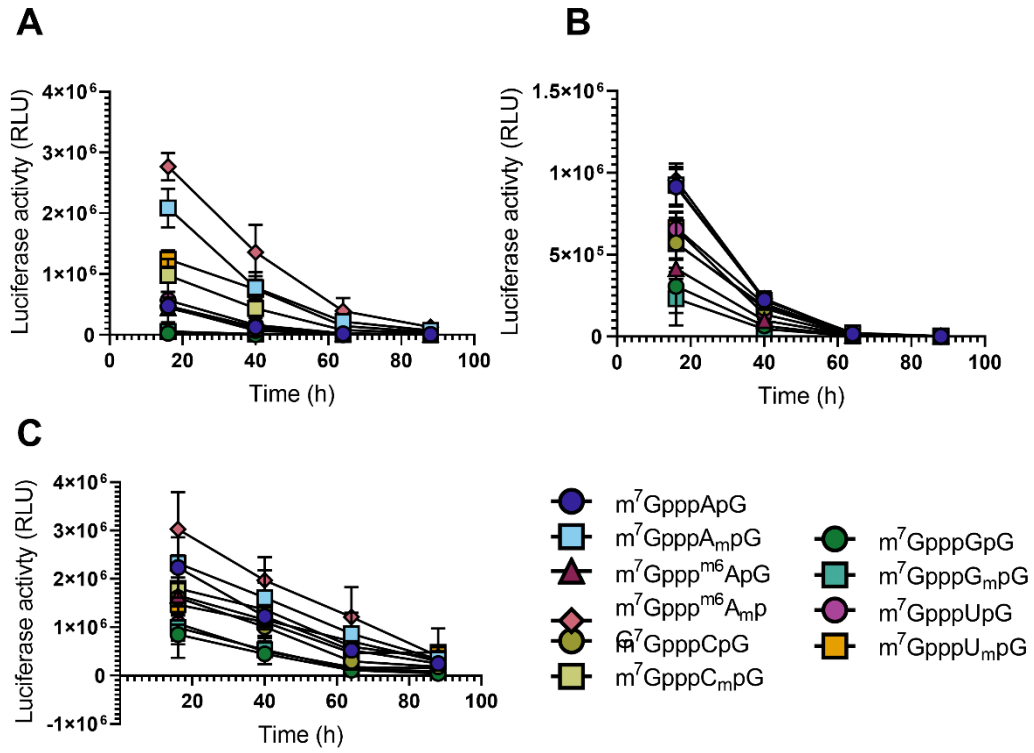


Figure S7. *Gaussia* luciferase activity in the supernatant of 3T3-L1, HeLa, and JAWS II cells measured after 16, 40, 64, and 88 h from transfection with IVT mRNAs bearing various trinucleotides at their 5' ends.

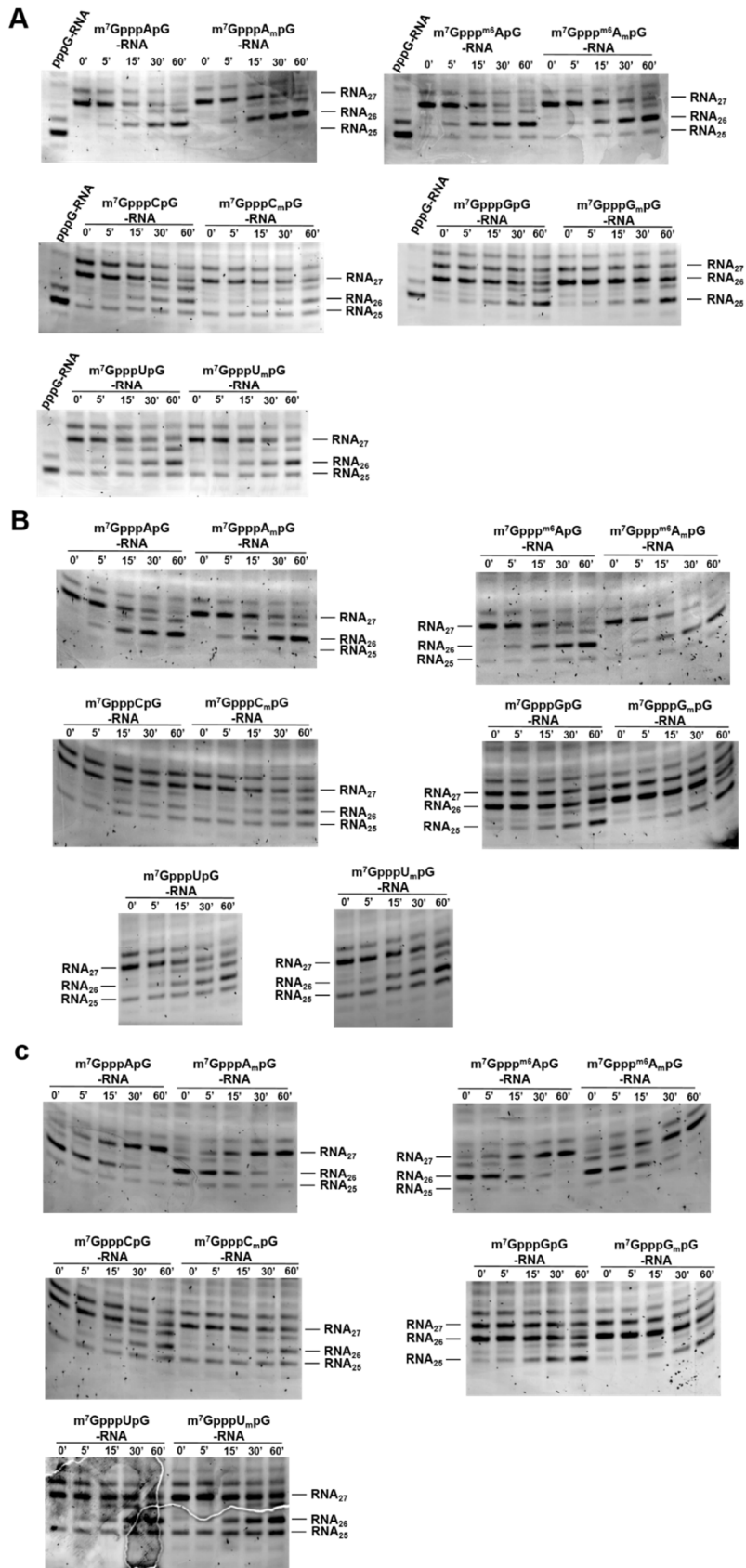


Figure S8. hDcp2 susceptibility assay for differently capped short RNAs. A-C represent PAGE analyses of three independent replicates.

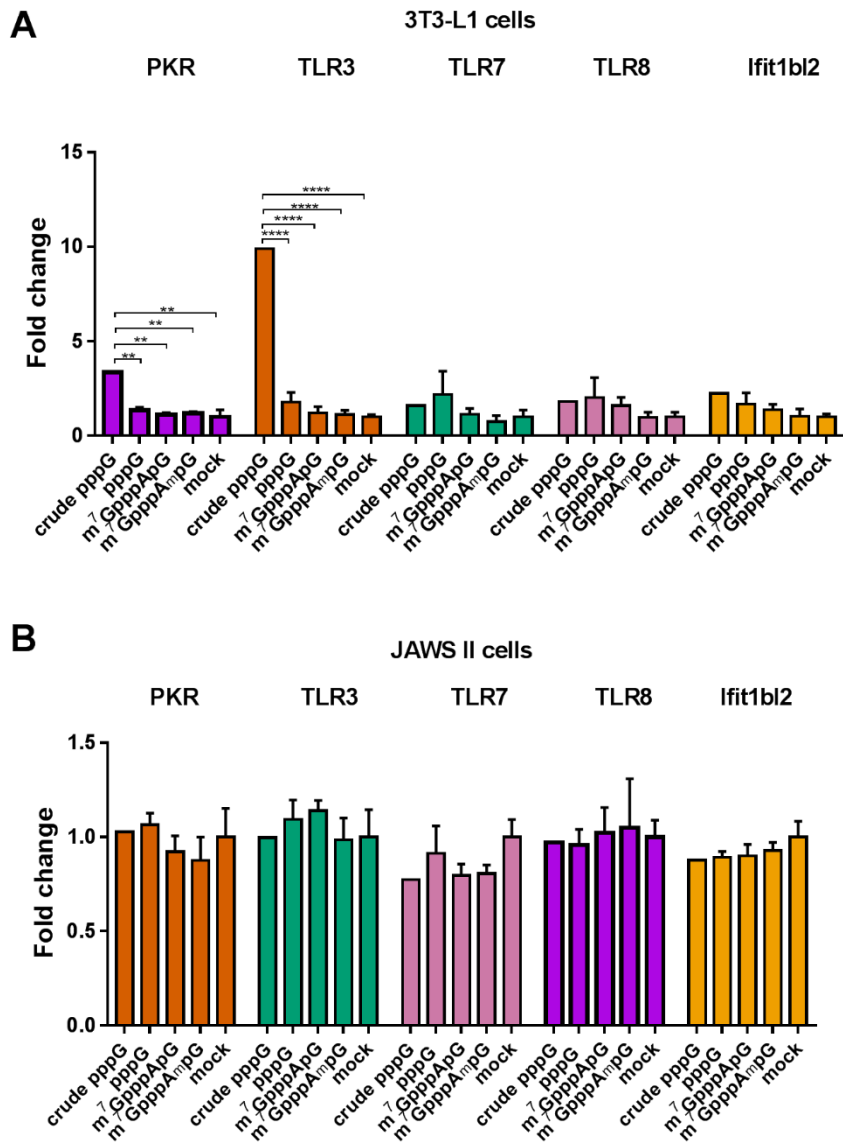


Figure S9. Changes in gene expression upon cell transfection with HPLC-purified capped IVT mRNA. Cells were transfected with 25 ng of HPLC-purified mRNA (A-B) for 5 h. mRNA expression analysis pre- and post-transfection for the indicated genes was carried out using RT-qPCR. Bars represent mean value of mRNA level change (fold change) \pm SEM from 3 independent biological replicates, each independent biological replicate consisting of single transfection. The data was obtained with mRNAs generated in two independent in vitro transcription reactions. Statistical significance: ** $p < 0.01$, **** $p < 0.0001$ (one-way ANOVA with Turkey's multiple comparisons test). Only statistically significant differences were marked on the graph. Before averaging the results, the data sets in each independent replicate were normalized to unpurified 5'-triphosphorylated RNA (pppG-RNA); then, the average values were normalized to mock samples to give the final normalized fold change values.

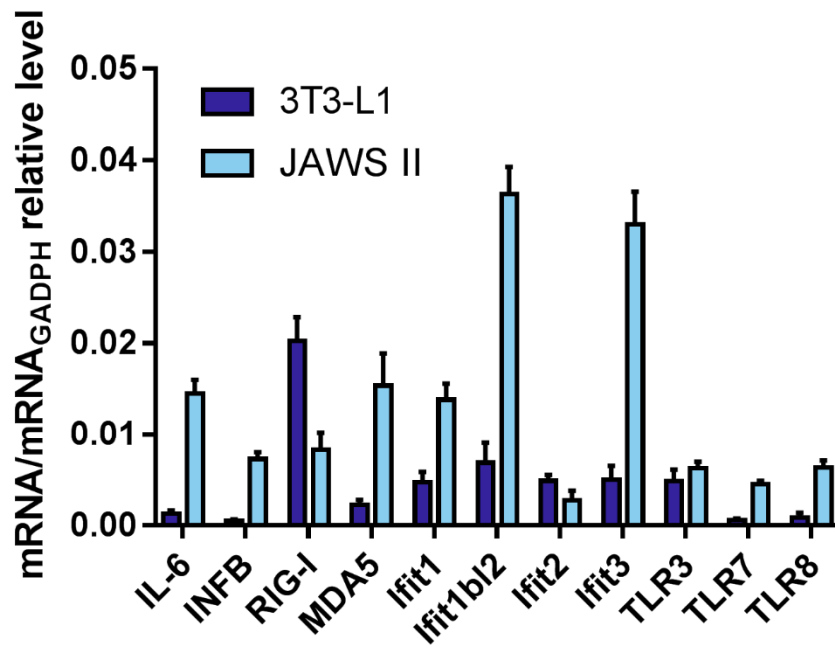
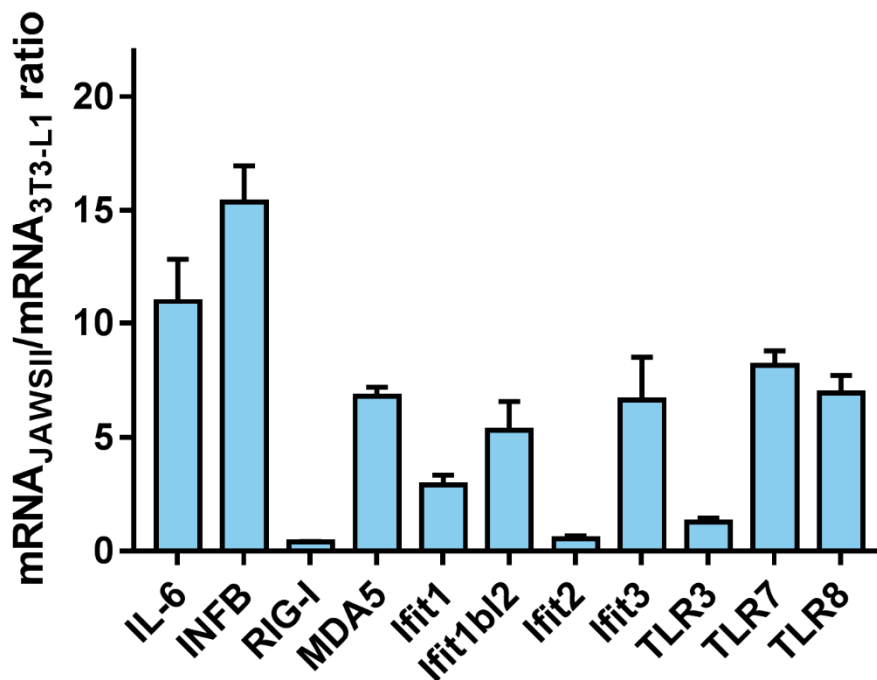
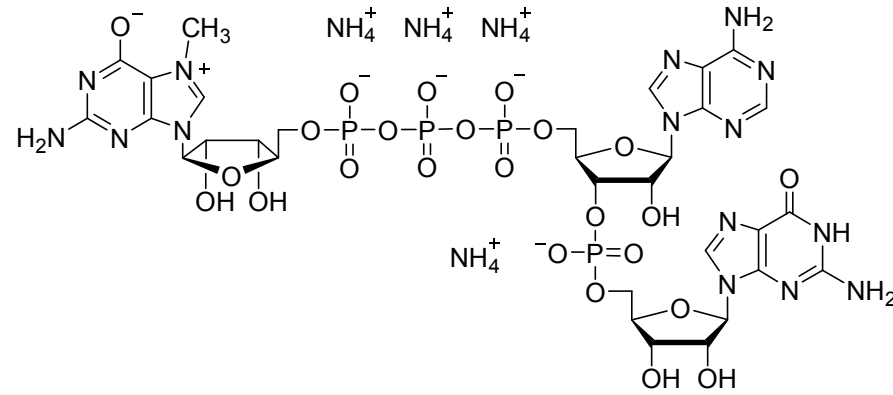
A**B**

Figure S10. Baseline expression of chosen genes in 3T3-L1 and JAWS II cells assessed by RT-qPCR. (A) Bars represent mean value of gene expression relative to GAPDH mRNA \pm SEM from 3 to 8 independent biological repetitions. (B) Bars represents gene expression (relative to GAPDH) levels ratio in JAWS II vs. 3T3-L1 cells \pm SEM.

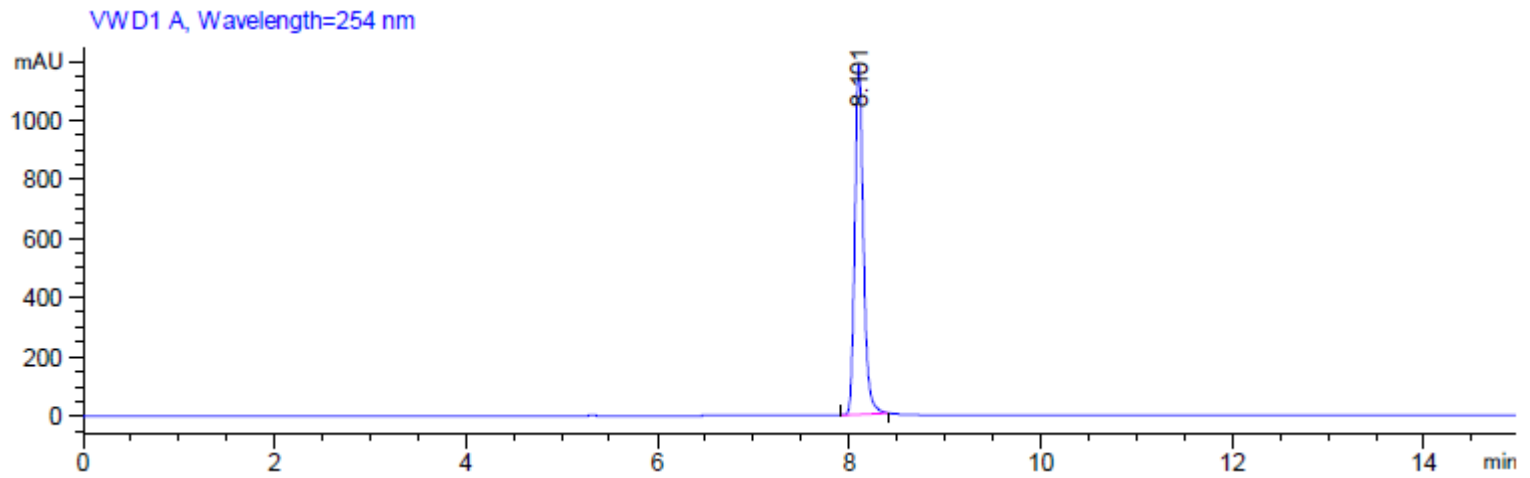
Figure S11. Cytokine production assay. JAWS II cells were transfected for 24 h with 25 ng of mRNA and concentration of secreted cytokines was measured on flow cytometer. (A, B) Comparison of immunogenic potential of (A) HPLC-purified and (B) crude mRNA bearing different cap analogues. (C-E) Direct comparison of immunogenic potential of mRNA with cap 0 (C), cap 1 (D) and with triphosphate group (E) before and after HPLC purification. Bars represents mean value \pm SEM from 2 independent biological replicates, each independent biological replicate consisting of a single transfection normalized to mock treated cells (each biological replicate was measured in duplicate). * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$ (one-way ANOVA with Turkey's multiple comparisons test and on (C - E) t-test was applied). Only statistically significant differences were marked on the graph.

m⁷GpppApG

Chemical structure

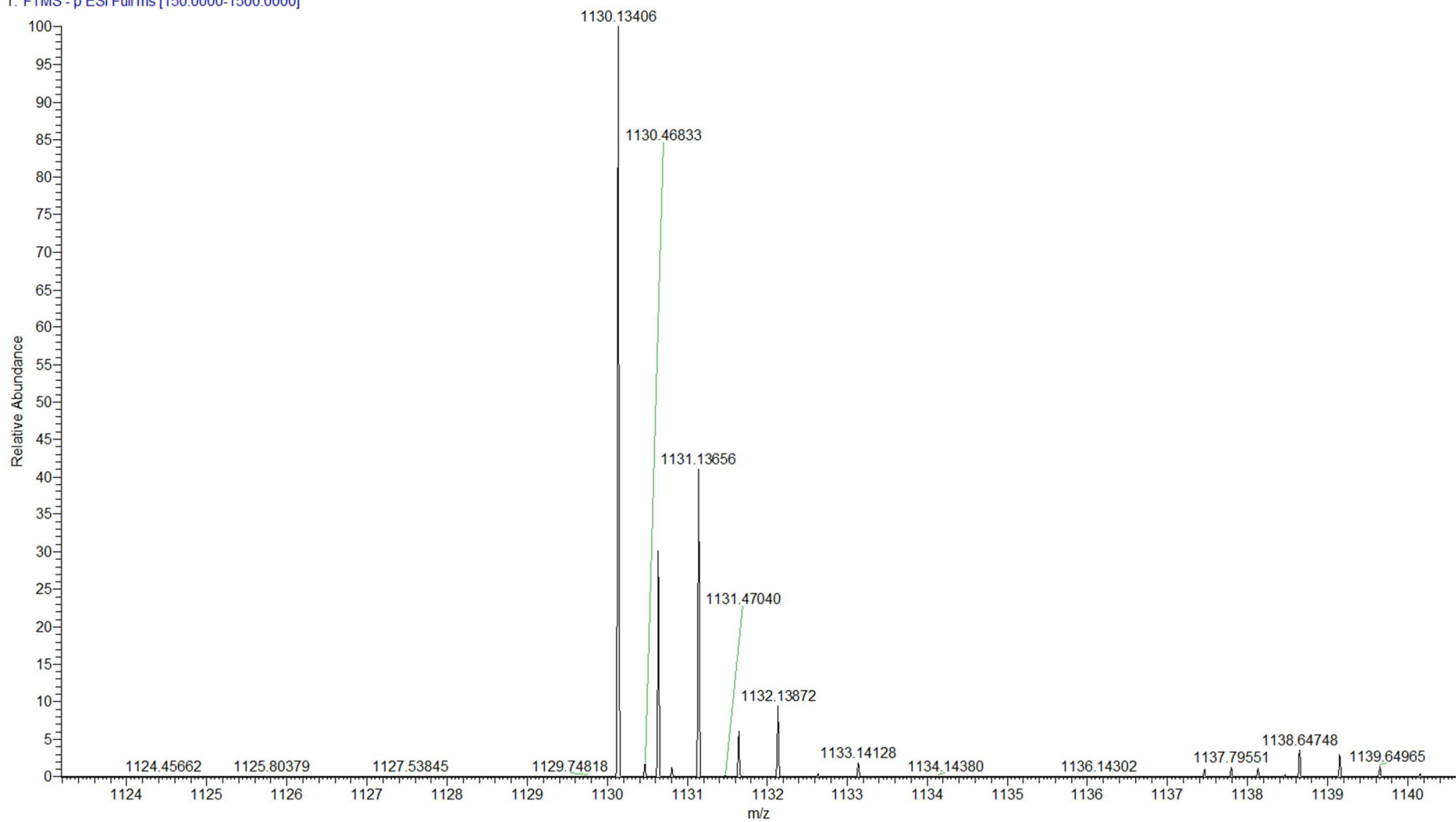


RP HPLC



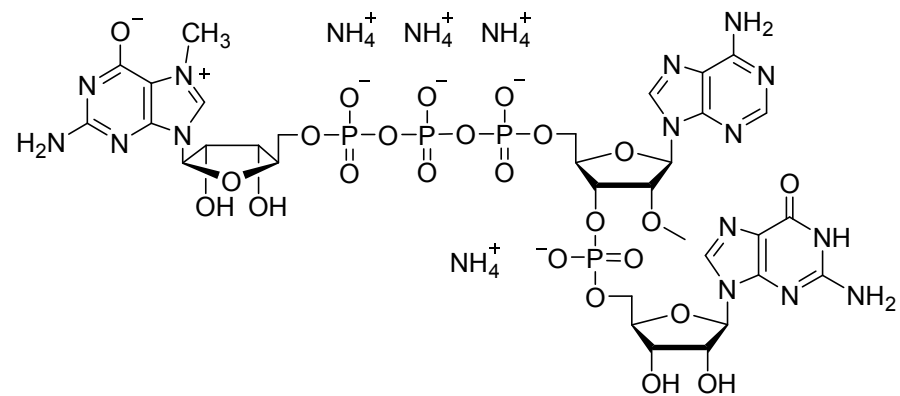
MS (-) ESI
(Calc. [M-H]⁻ C₃₁H₄₀N₁₅O₂₄P₄⁻ 1130.13266)

171213_TP_007#66-131 RT: 0.58-1.14 AV: 66 NL: 7.05E5
T: FTMS - p ESI Full ms [150.0000-1500.0000]

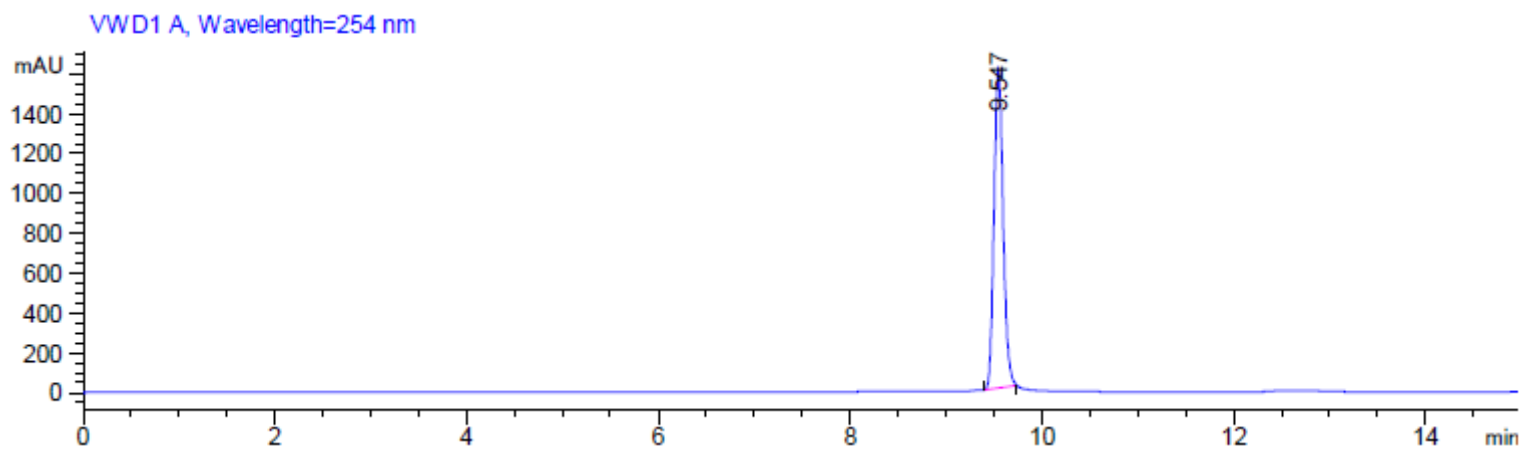


m⁷GpppA_{mp}G

Chemical structure

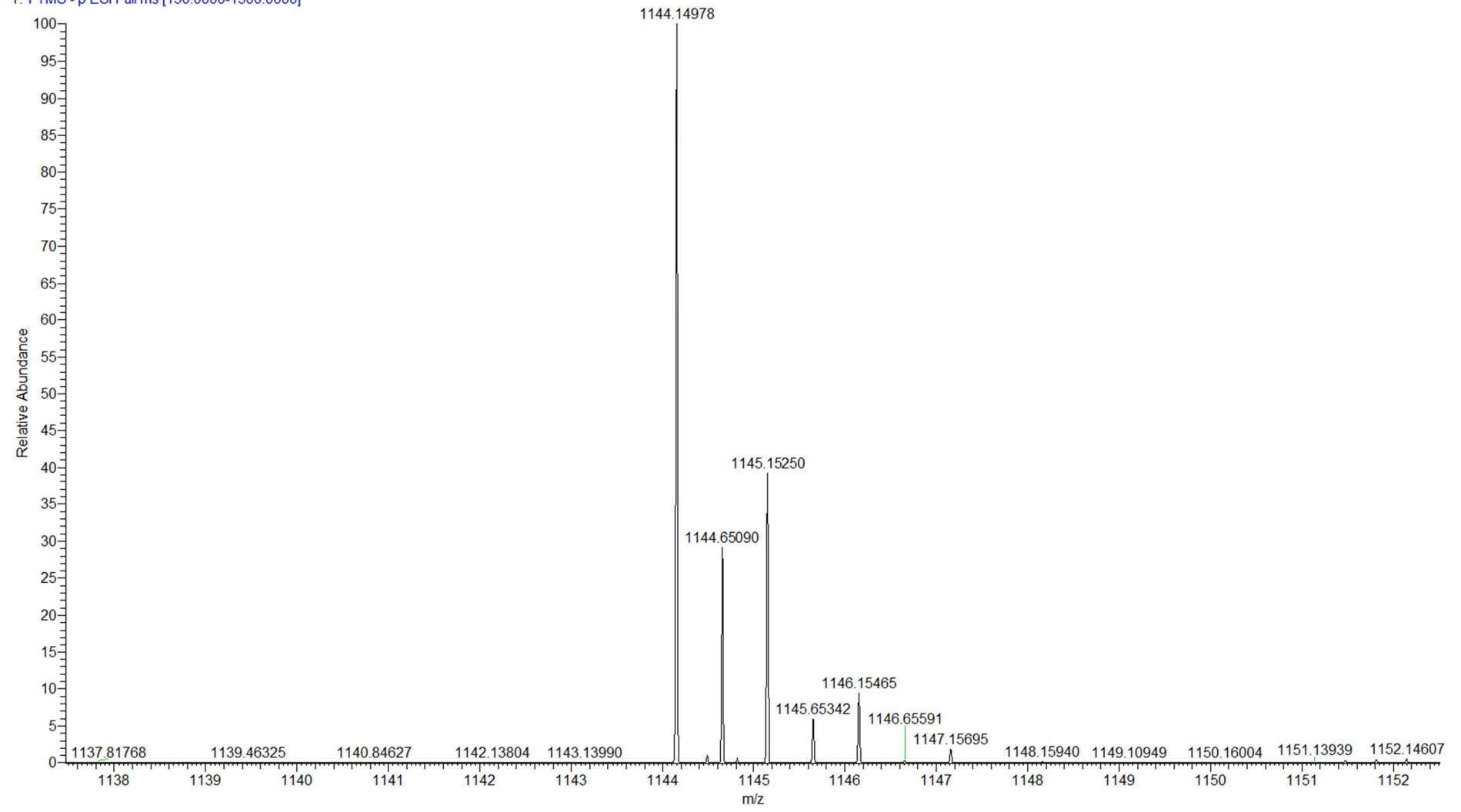


RP HPLC



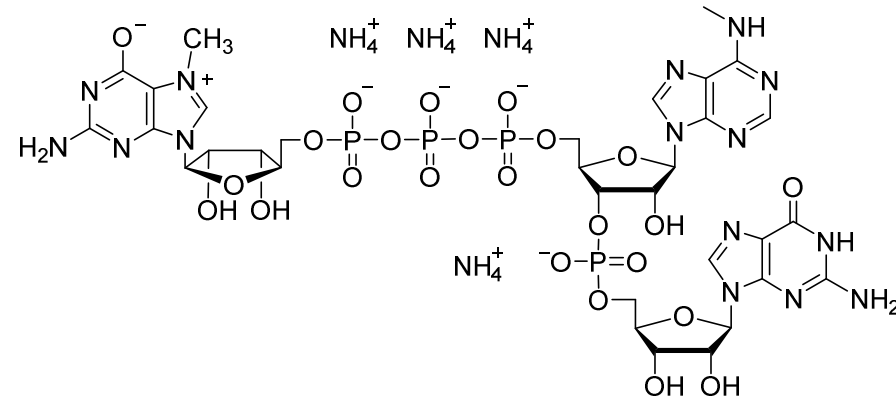
MS (-) ESI
(Calc. [M-H]⁻ C₃₂H₄₂N₁₅O₂₄P₄⁻ 1144.14831)

171213_TP_012#108-186 RT: 0.94-1.62 AV: 79 NL: 6.82E5
T: FTMS - p ESI Full ms [150.0000-1500.0000]

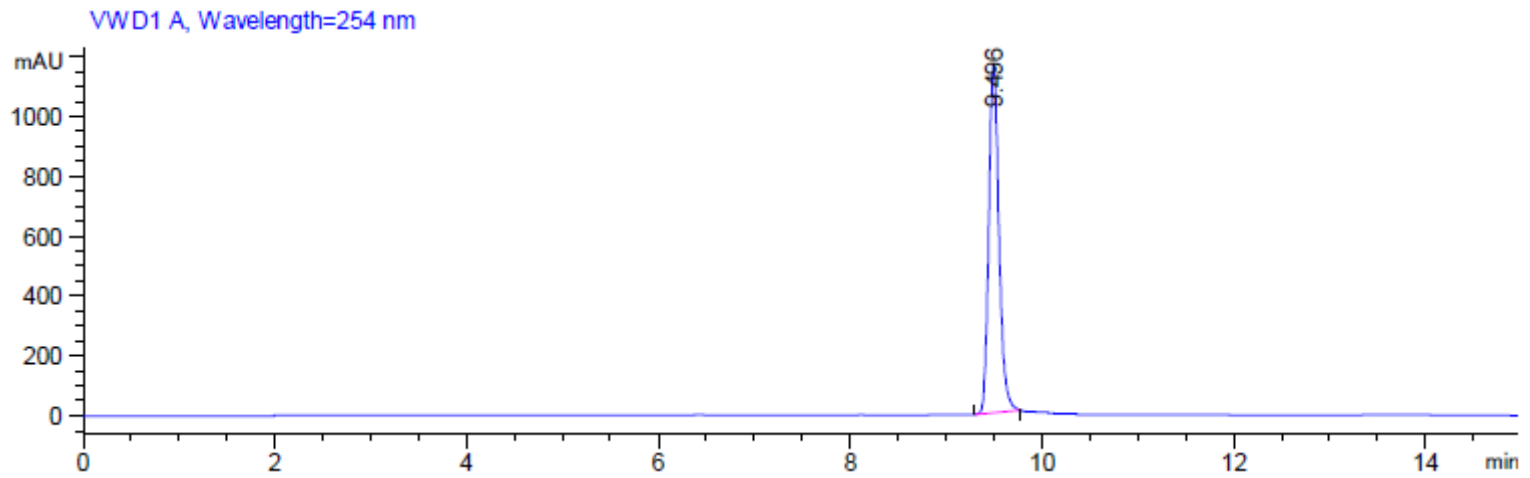


m⁷Gppp^m6ApG

Chemical structure

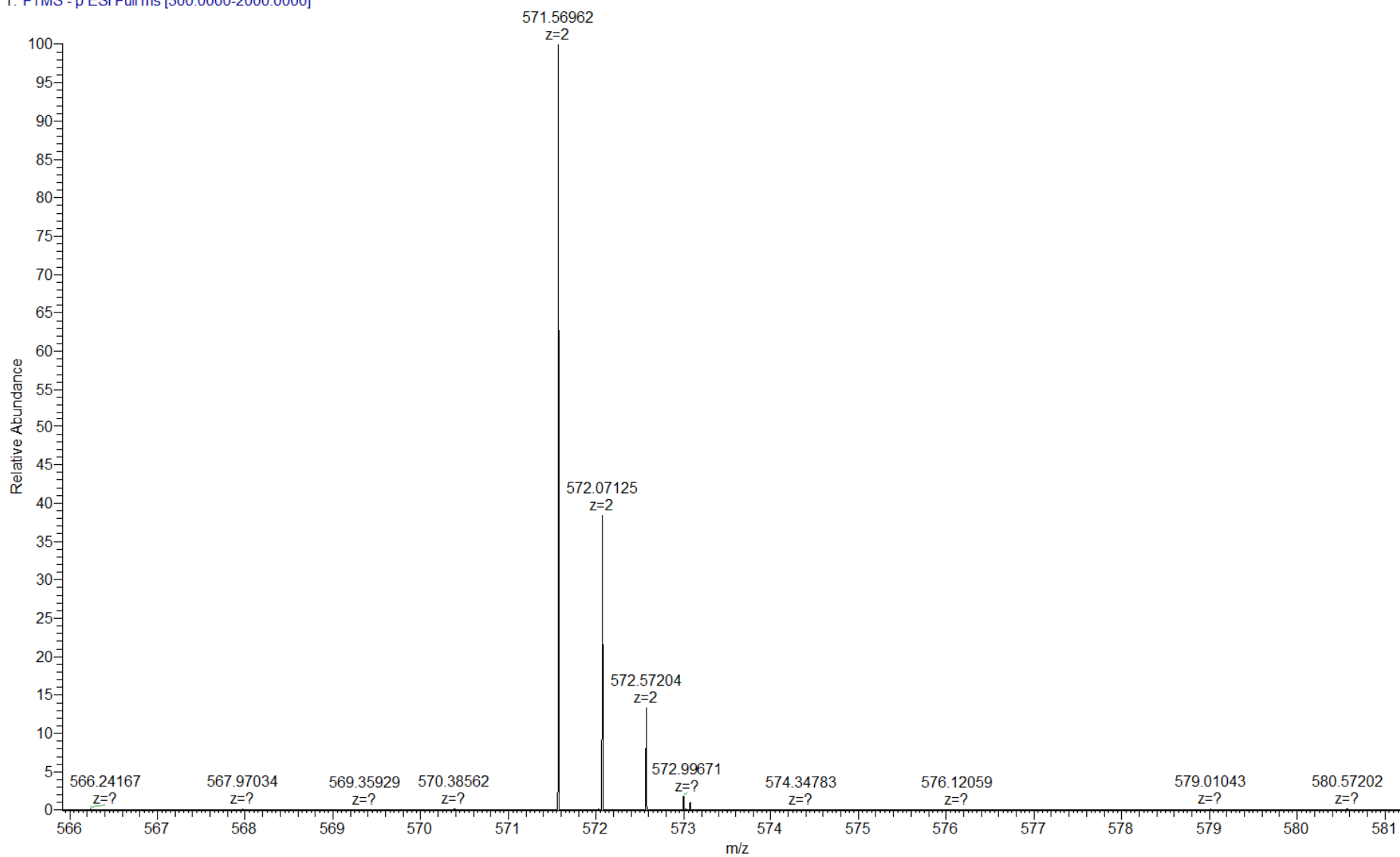


RP HPLC



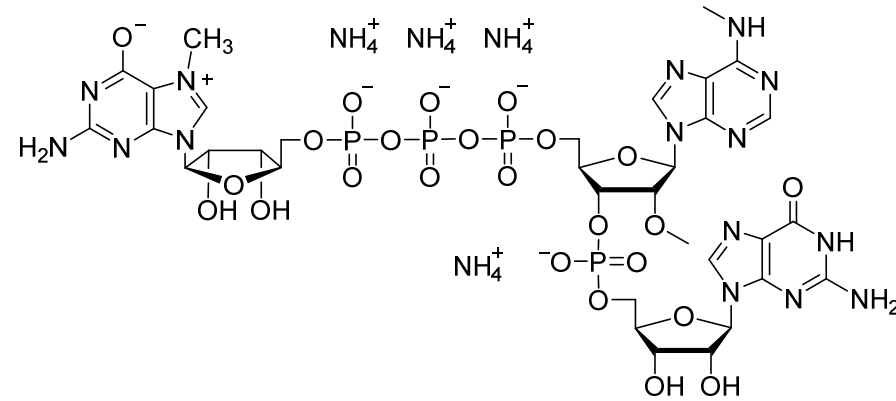
170711_TP002 #31-49 RT: 0.40-0.63 AV: 19 NL: 1.00E4
T: FTMS - p ESI Full ms [300.0000-2000.0000]

MS (-) ESI
(Calc. $[M-2H]^{2-}$ $C_{32}H_{41}N_{15}O_{24}P_4^{2-}$ 571.57052)

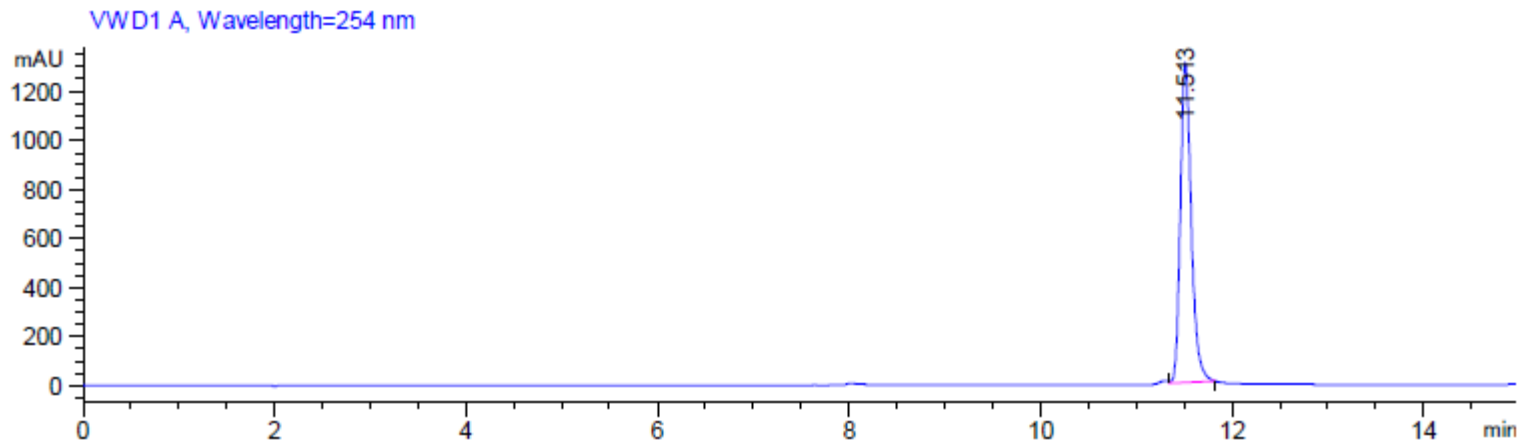


m⁷Gppp^{m6}AmpG

Chemical structure

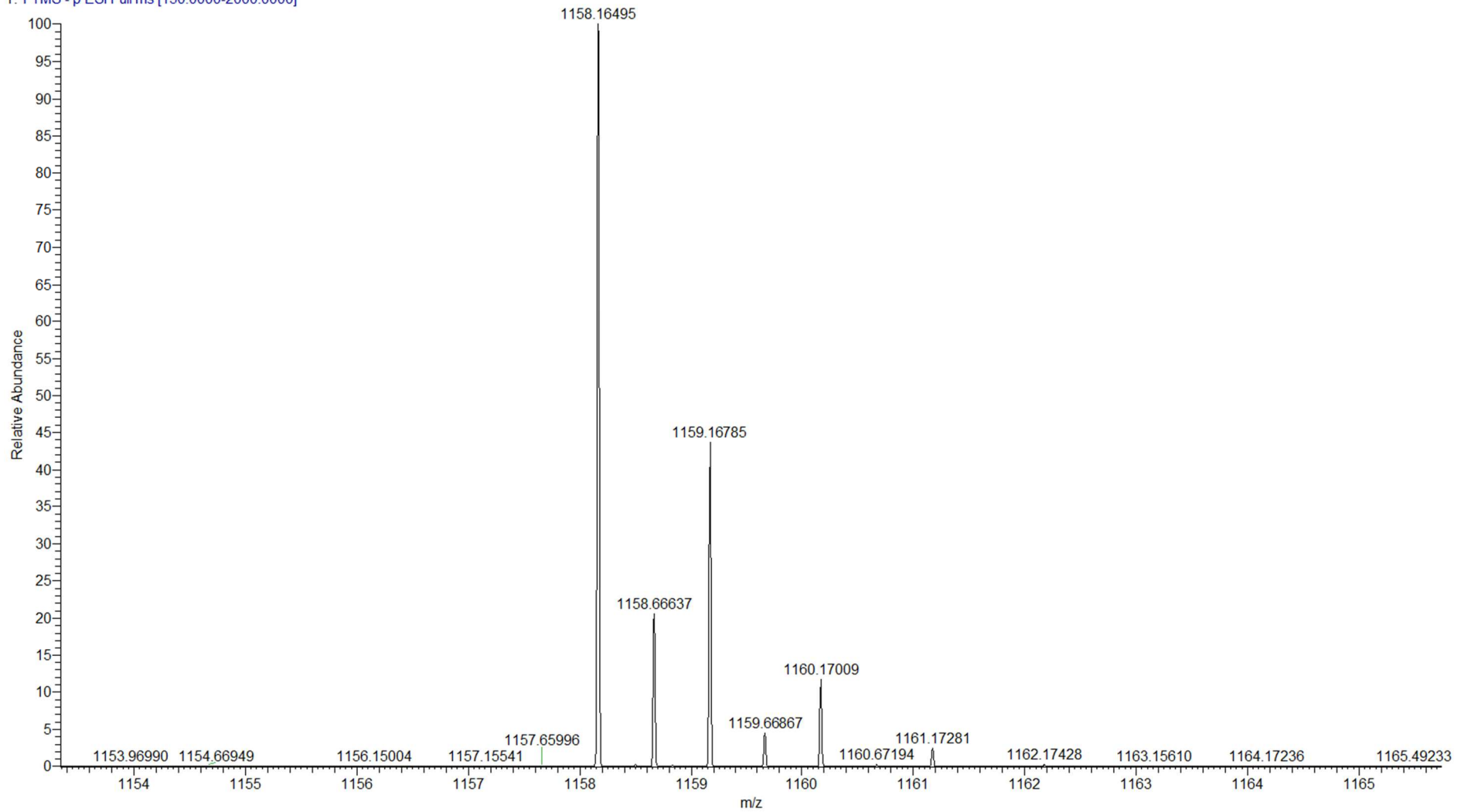


RP HPLC



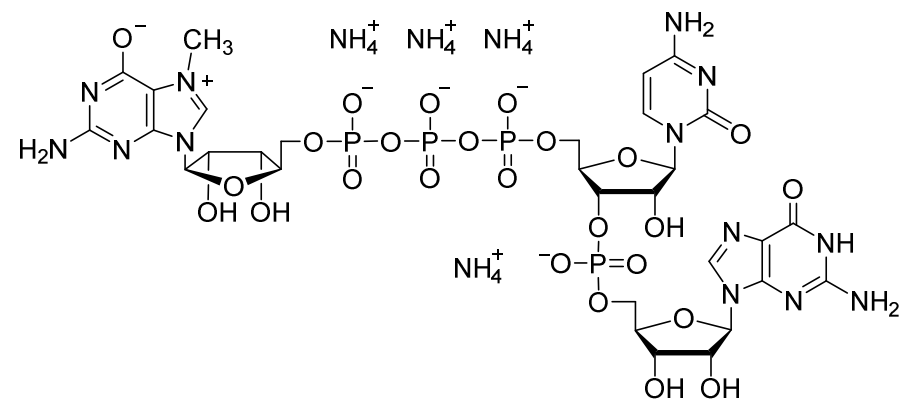
MS (-) ESI
(Calc. [M-H]⁻ C₃₃H₄₄N₁₅O₂₄P₄⁻ 1158.16396)

180517_MW_112 #65-256 RT: 0.65-2.59 AV: 192 NL: 9.14E5
T: FTMS - p ESI Full ms [150.0000-2000.0000]

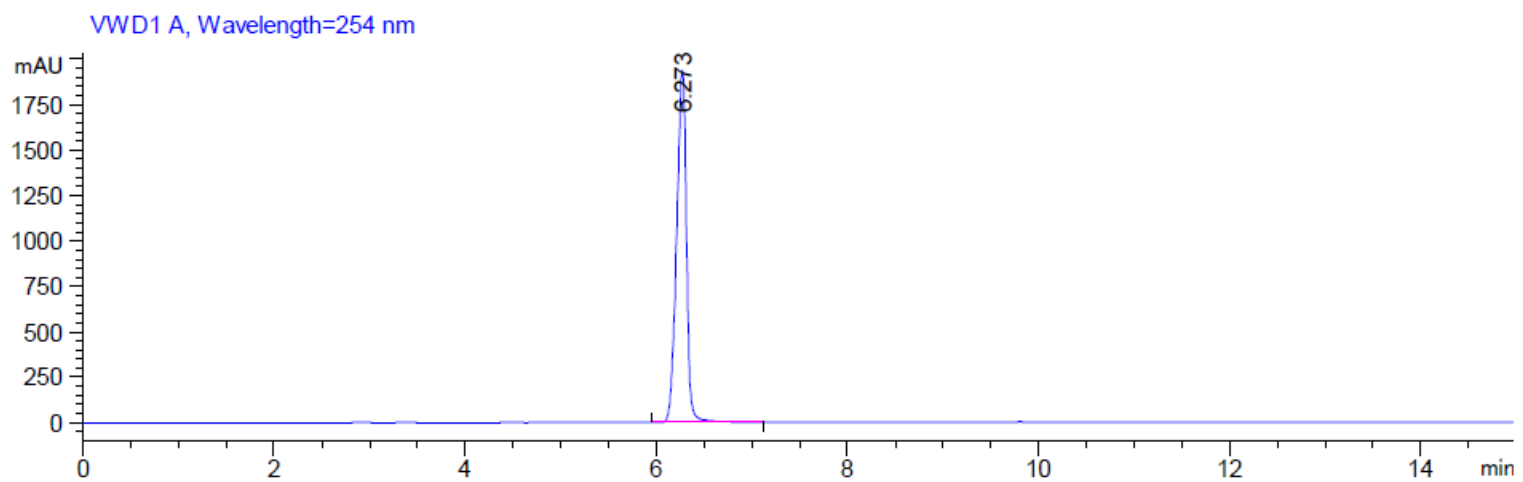


m⁷GpppCpG

Chemical structure

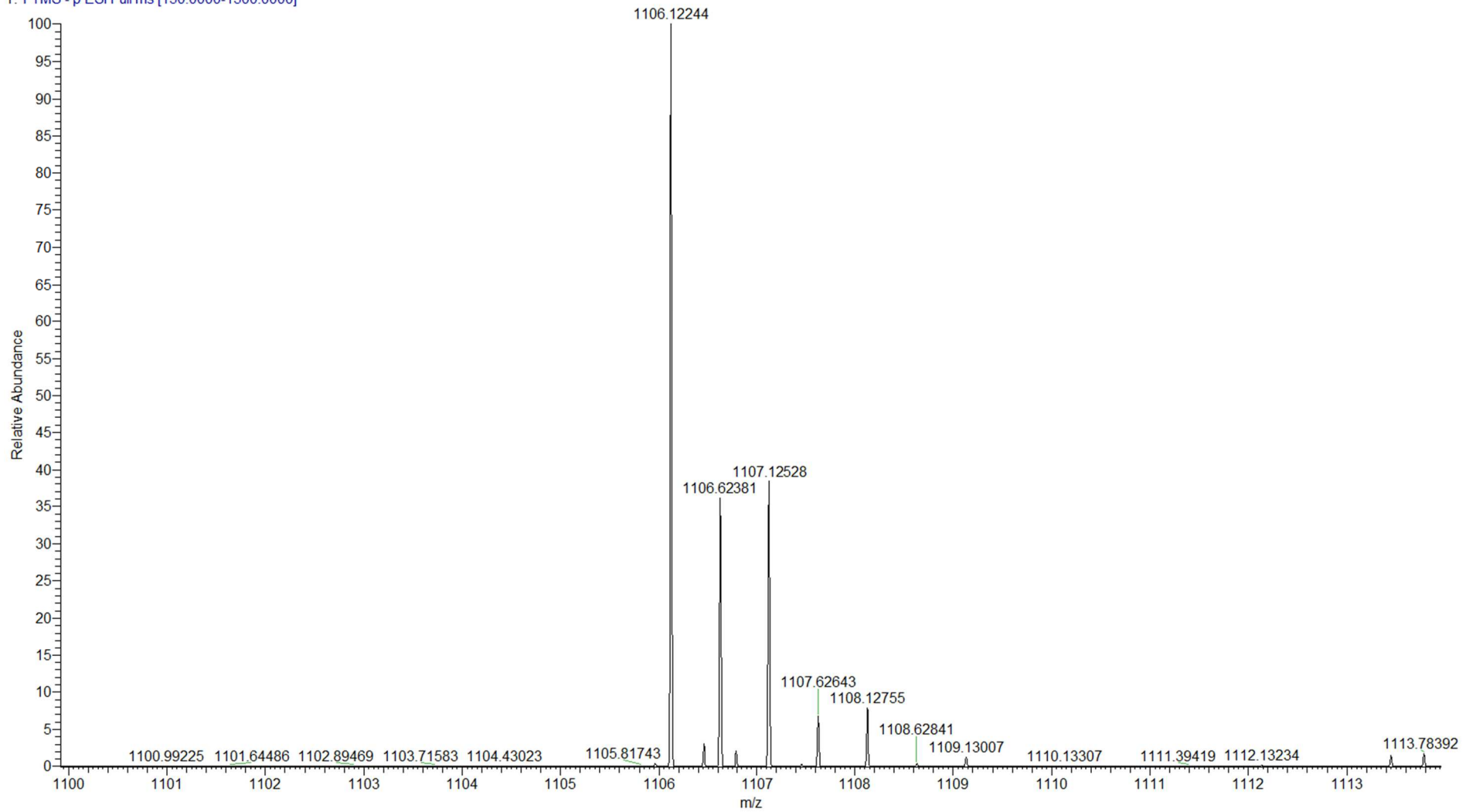


RP HPLC



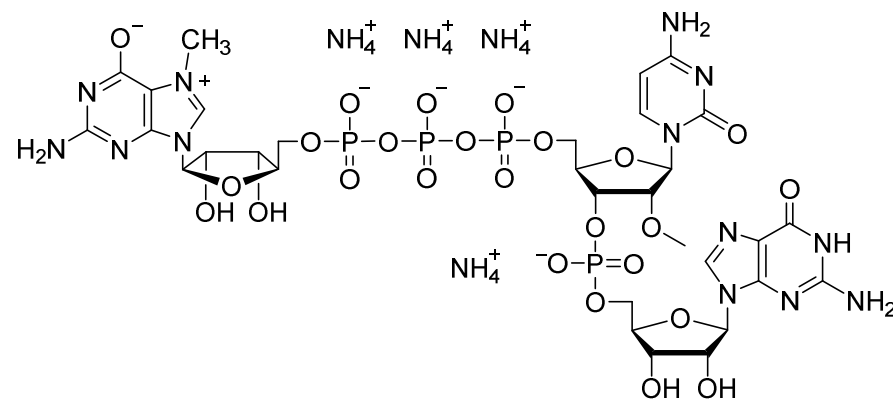
MS (-) ESI
(Calc. [M-H]⁻ C₃₀H₄₀N₁₃O₂₅P₄⁻ 1106.12142)

171213_TP_010#19-88 RT: 0.17-0.77 AV: 70 NL: 6.04E5
T: FTMS - p ESI Full ms [150.0000-1500.0000]

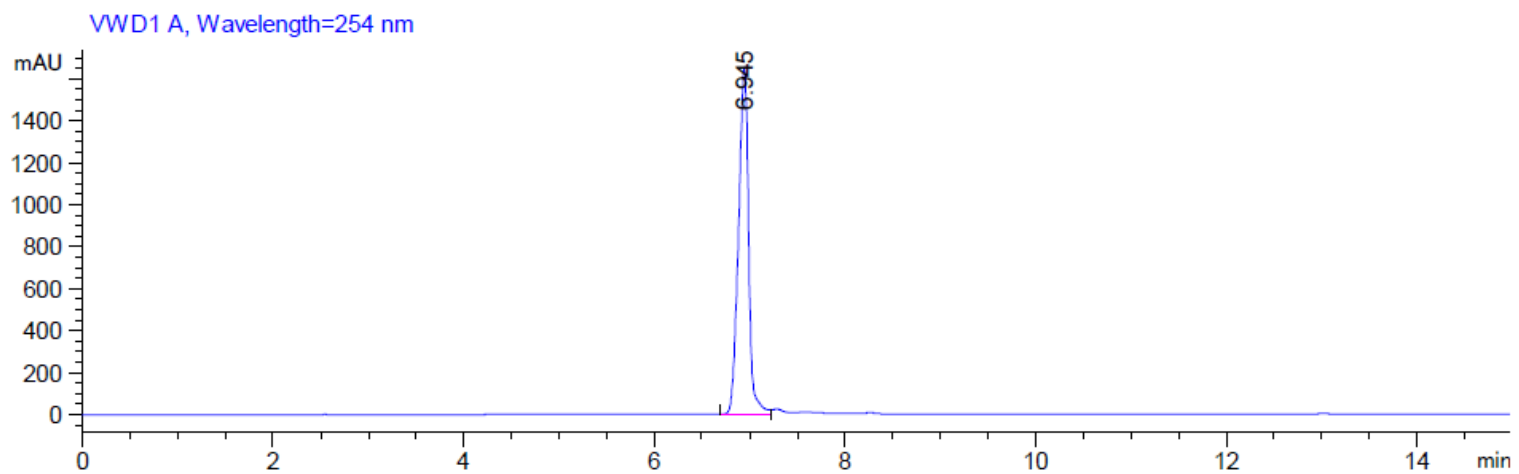


m⁷GpppCmpG

Chemical structure

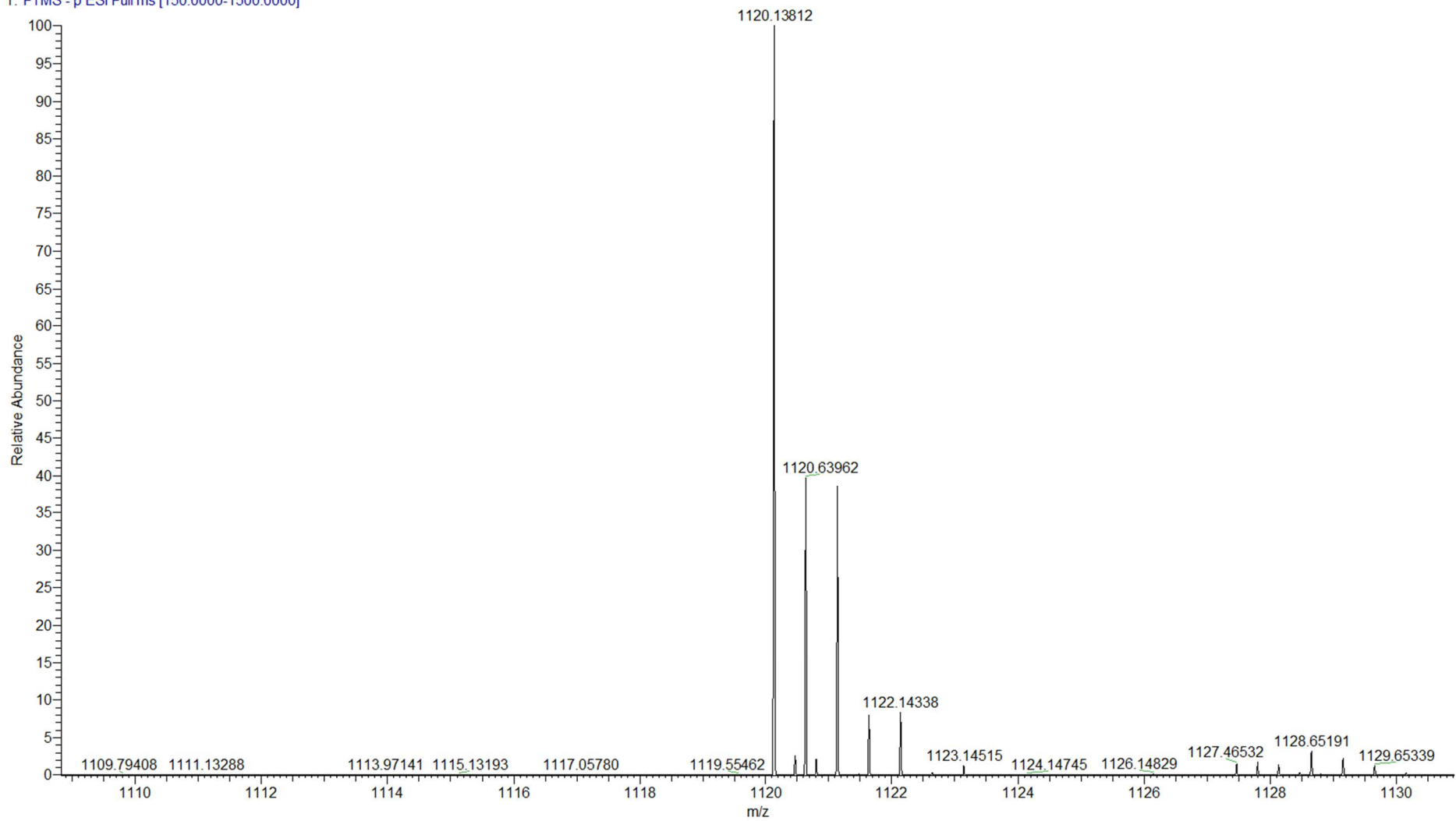


RP HPLC



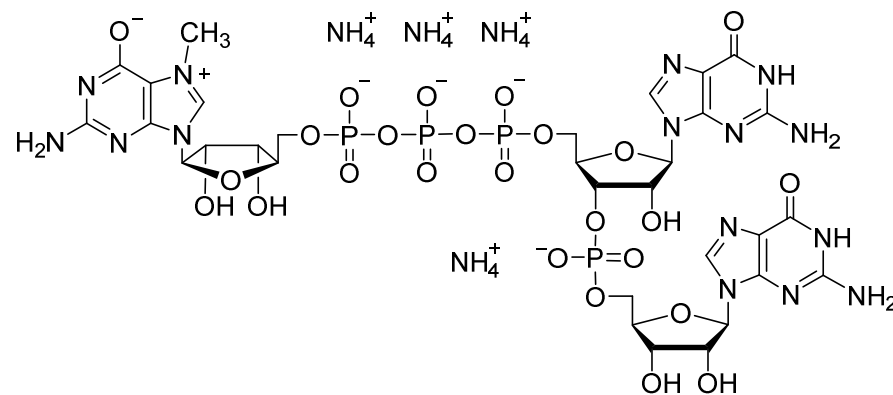
MS (-) ESI
(Calc. [M-H]⁻ C₃₁H₄₂N₁₃O₂₅P₄⁻ 1120.13707)

171213_TP_018 #10-50 RT: 0.09-0.44 AV: 41 NL: 3.88E5
T: FTMS - p ESI Full ms [150.0000-1500.0000]

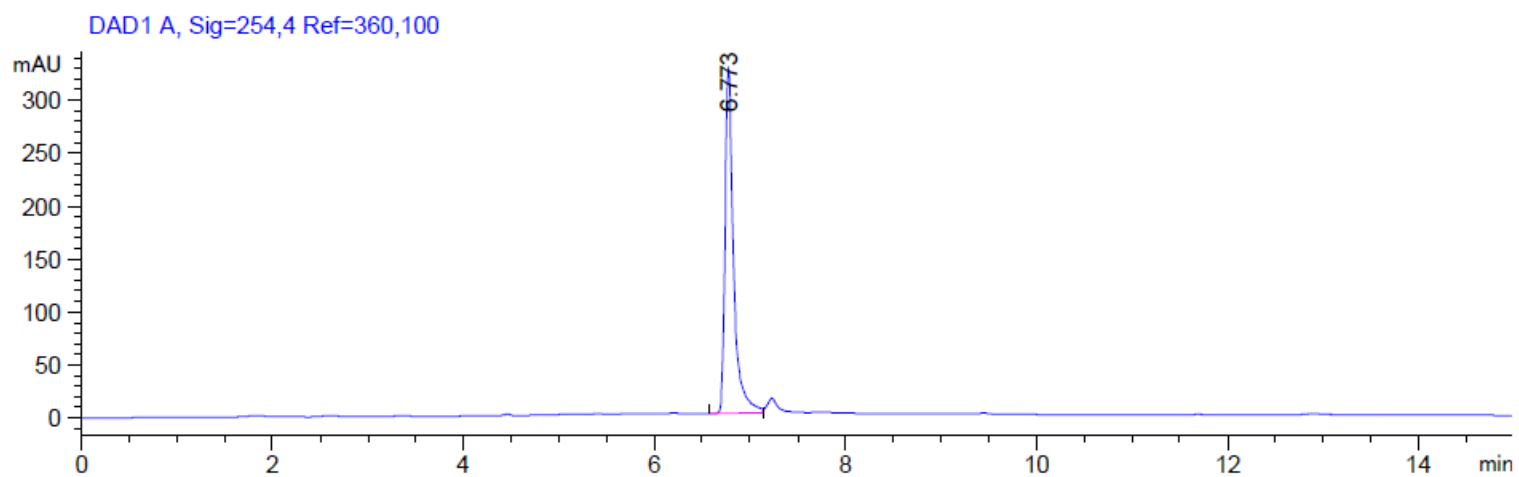


m⁷GpppGpG

Chemical structure

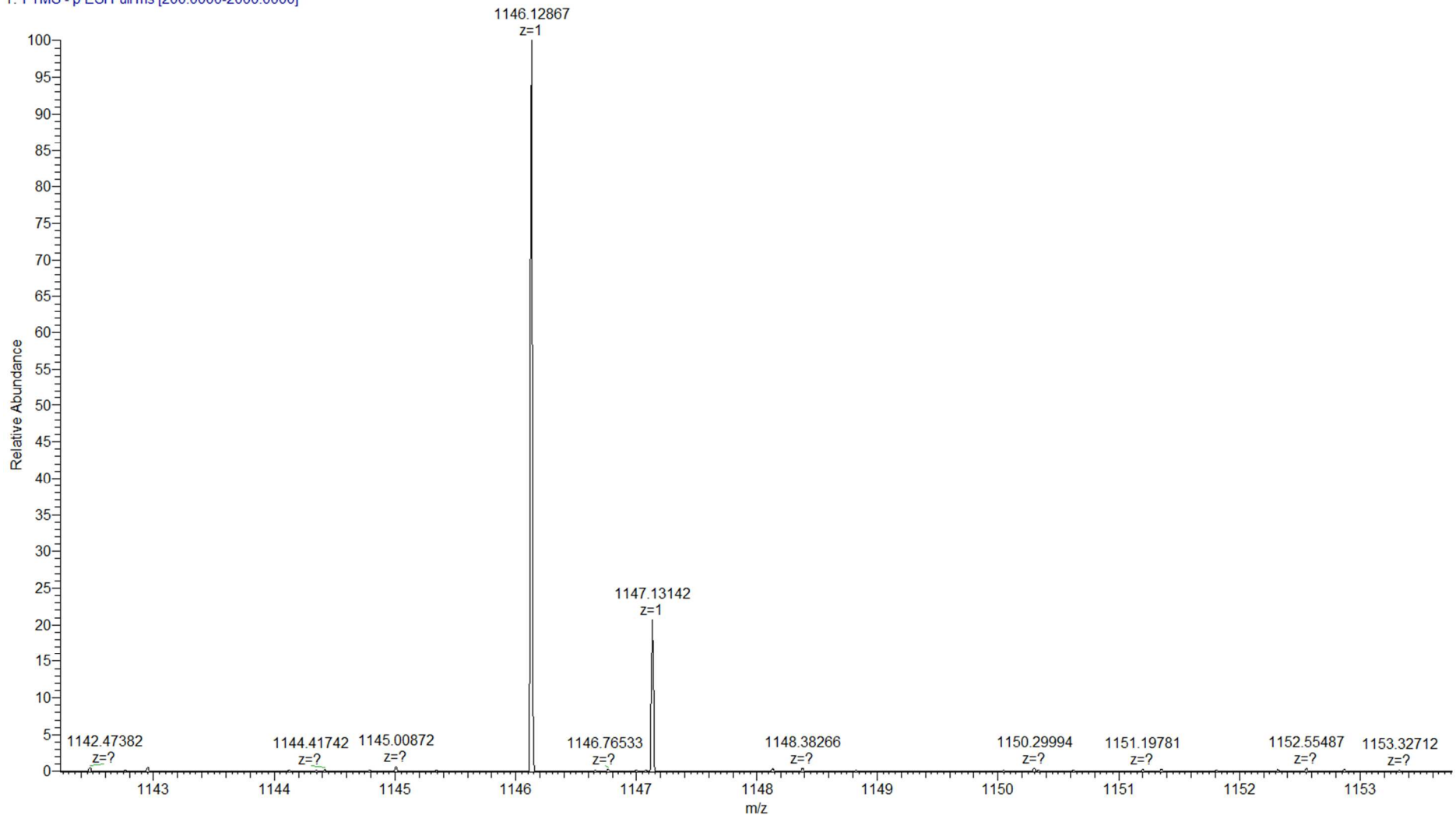


RP HPLC



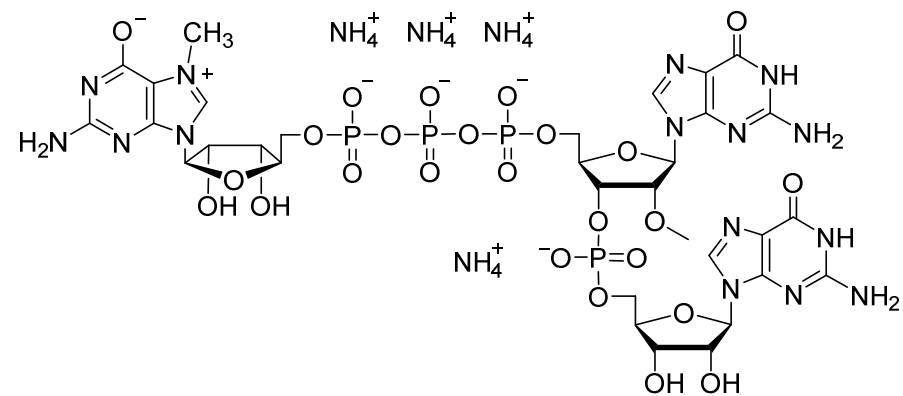
MS (-) ESI
(Calc. [M-H]⁻ C₃₁H₄₀N₁₅O₂₅P₄⁻ 1146.12757)

90128_TR_105 #5-112 RT: 0.05-1.15 AV: 108 NL: 4.60E3
T: FTMS - p ESI Full ms [200.0000-2000.0000]

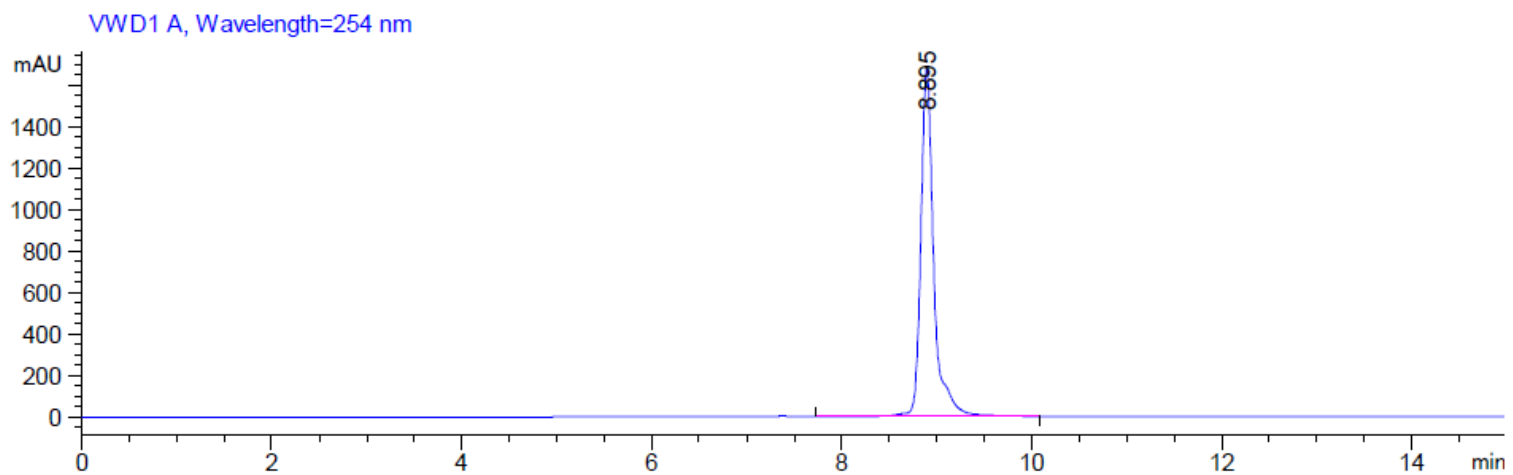


m⁷GpppG_{mp}G

Chemical structure

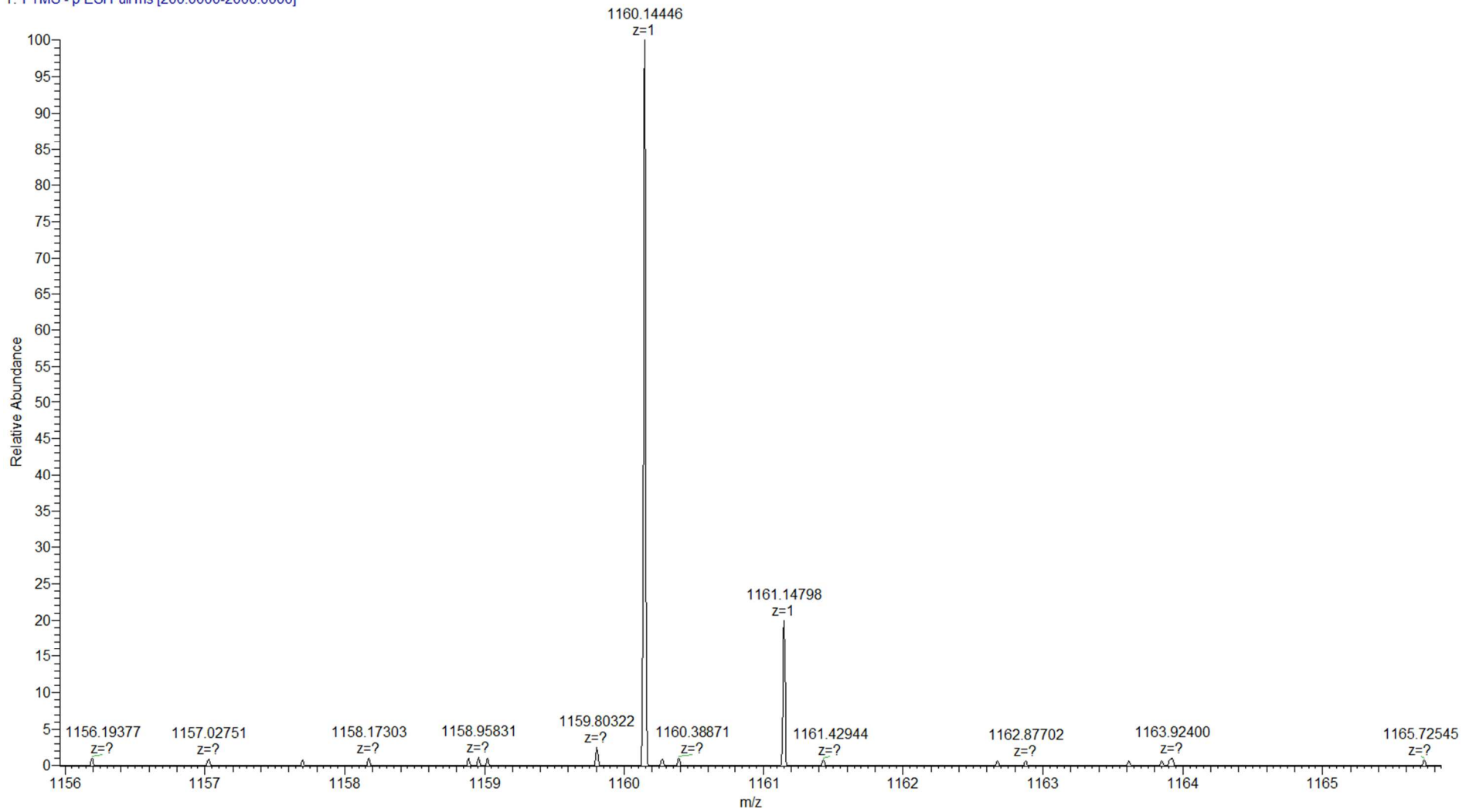


RP HPLC



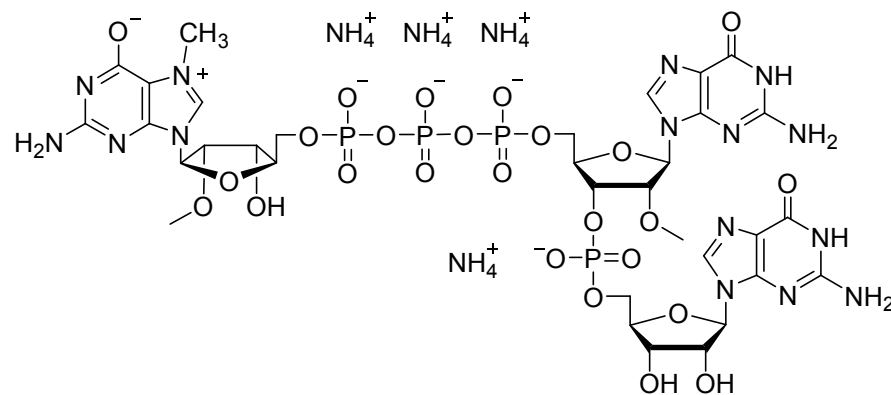
MS (-) ESI
(Calc. [M-H]⁻ C₃₂H₄₂N₁₅O₂₅P₄⁻ 1160.14322)

90128_TR_104 #36-64 RT: 0.38-0.66 AV: 29 NL: 6.31E3
T: FTMS - p ESI Full ms [200.0000-2000.0000]

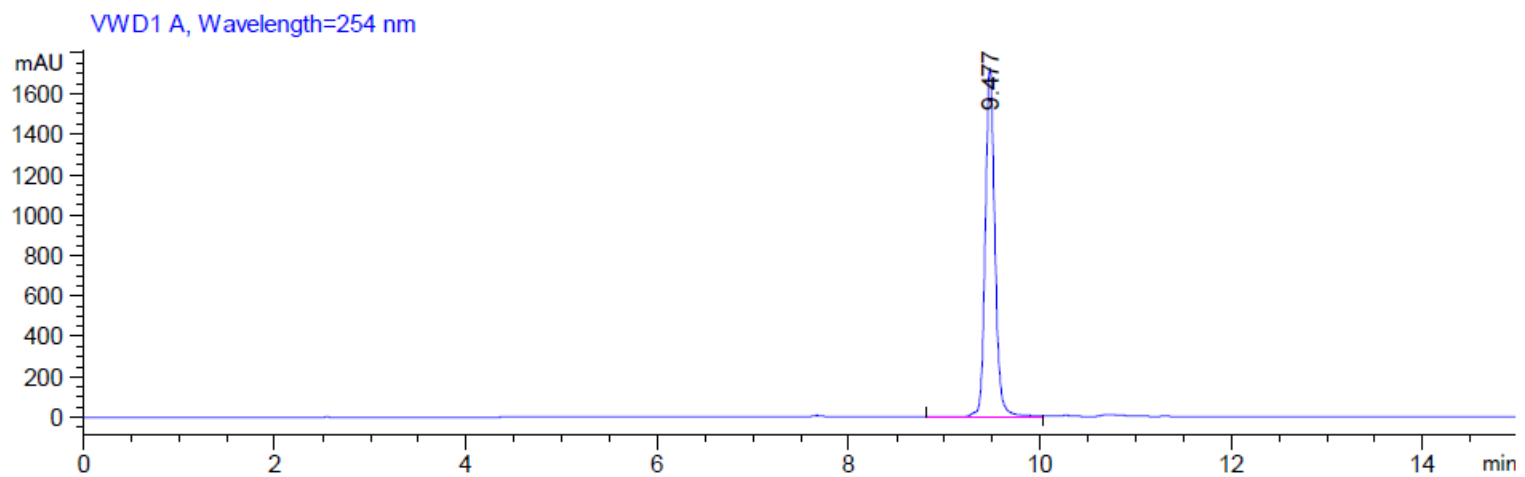


$m_2^{7,2-O}$ GpppG_mpG

Chemical structure

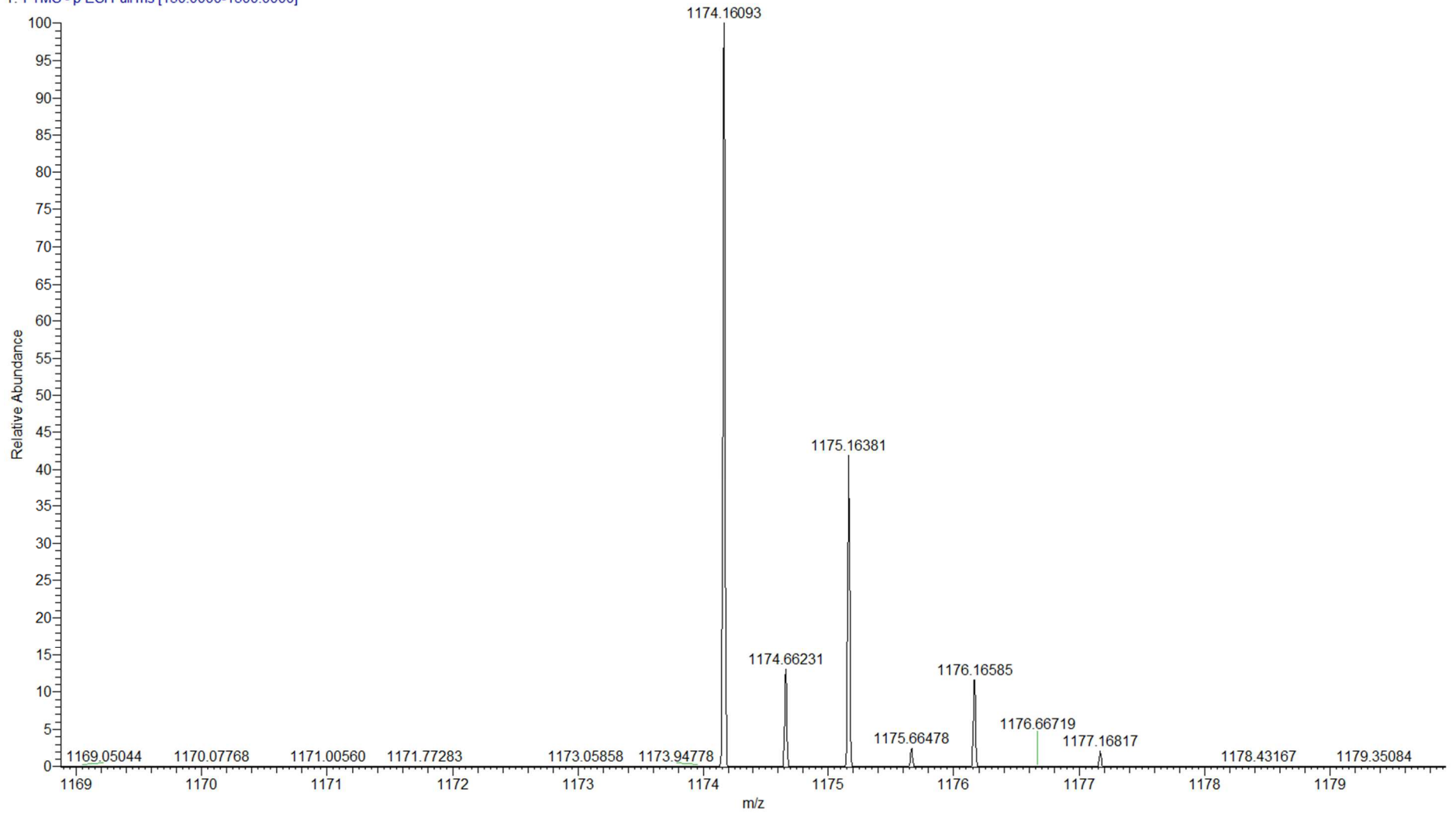


RP HPLC



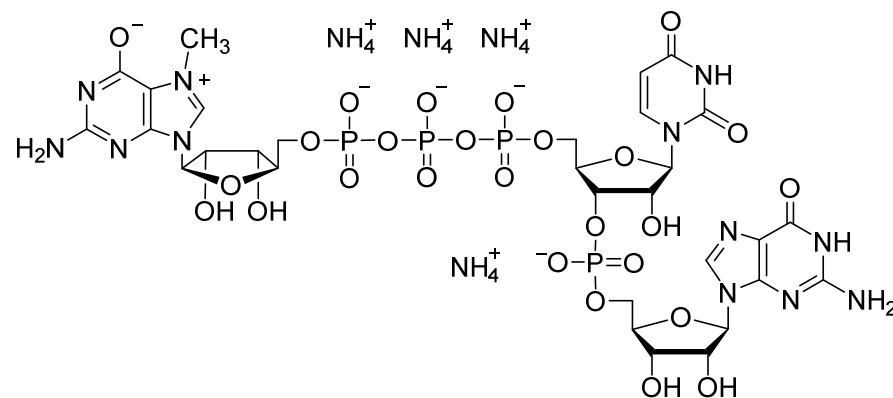
MS (-) ESI
(Calc. [M-H]⁻ C₃₃H₄₄N₁₅O₂₅P₄⁻ 1174.15887)

171213_TP_022#39-119 RT: 0.34-1.04 AV: 81 NL: 1.50E5
T: FTMS - p ESI Full ms [150.0000-1500.0000]

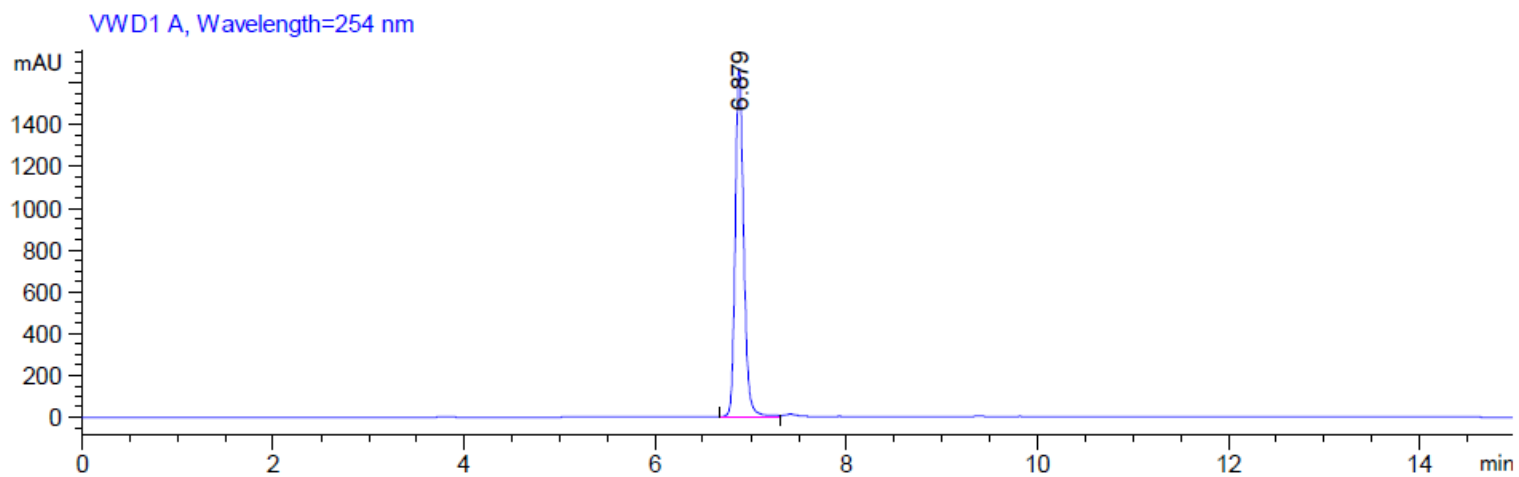


m⁷GpppUpG

Chemical structure

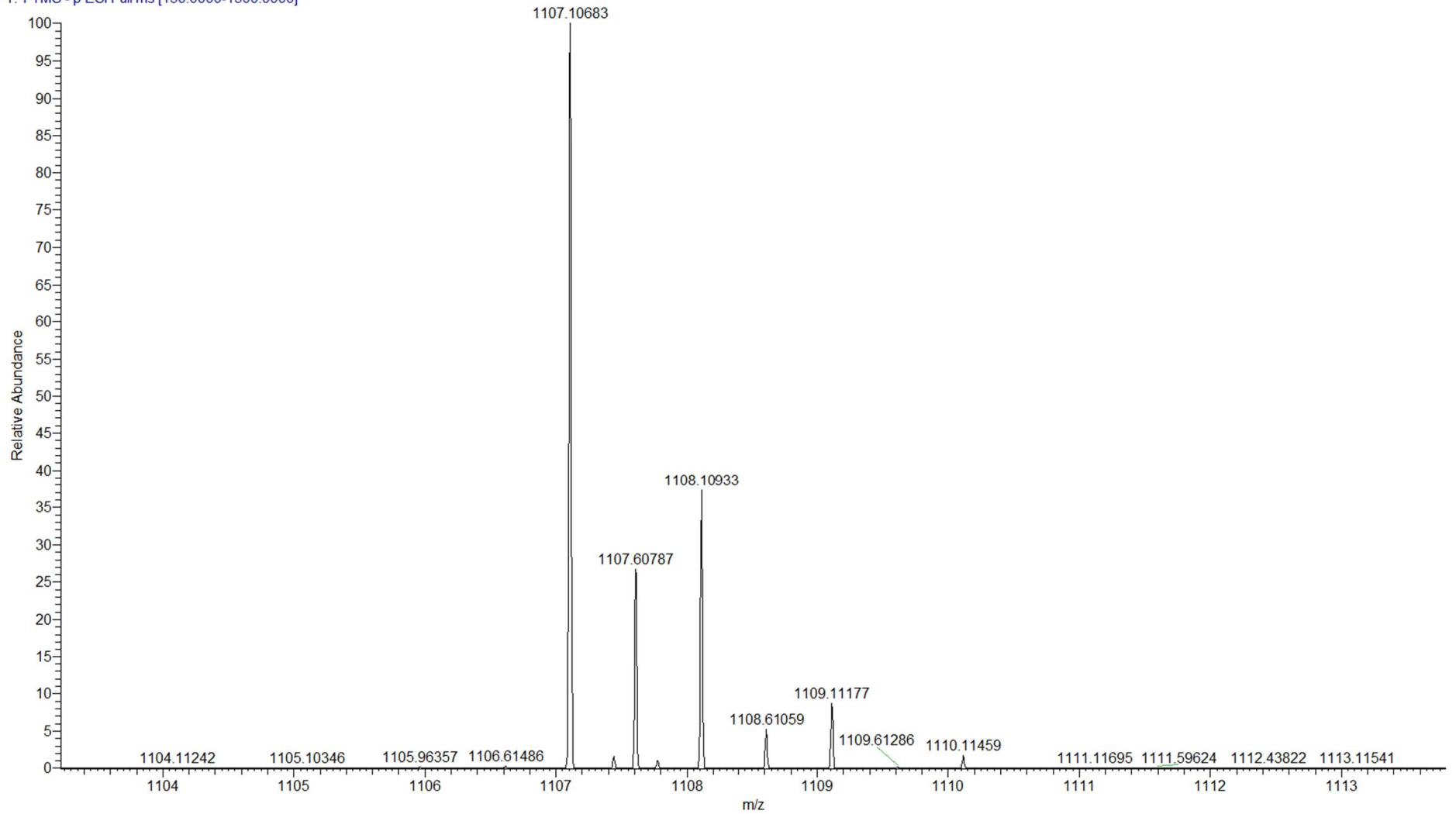


RP HPLC



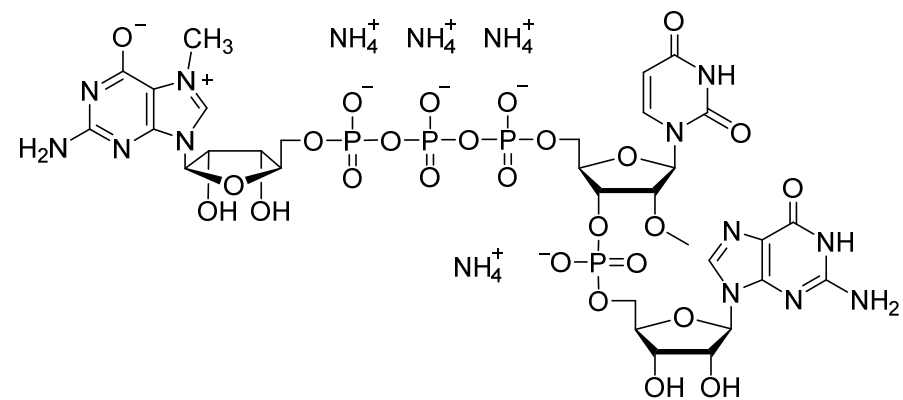
MS (-) ESI
(Calc. [M-H]⁻ C₃₀H₃₉N₁₂O₂₆P₄⁻ 1107.10544)

171213_TP_011#44-119 RT: 0.38-1.04 AV: 76 NL: 8.74E5
T: FTMS - p ESI Full ms [150.0000-1500.0000]

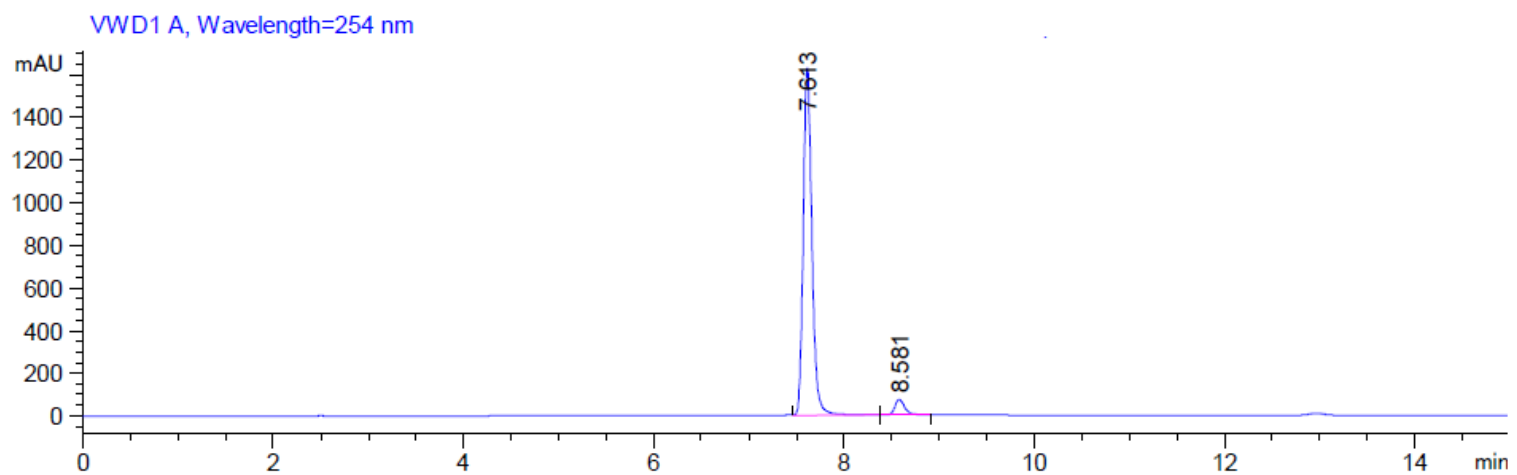


m⁷GpppU_{mp}G

Chemical structure

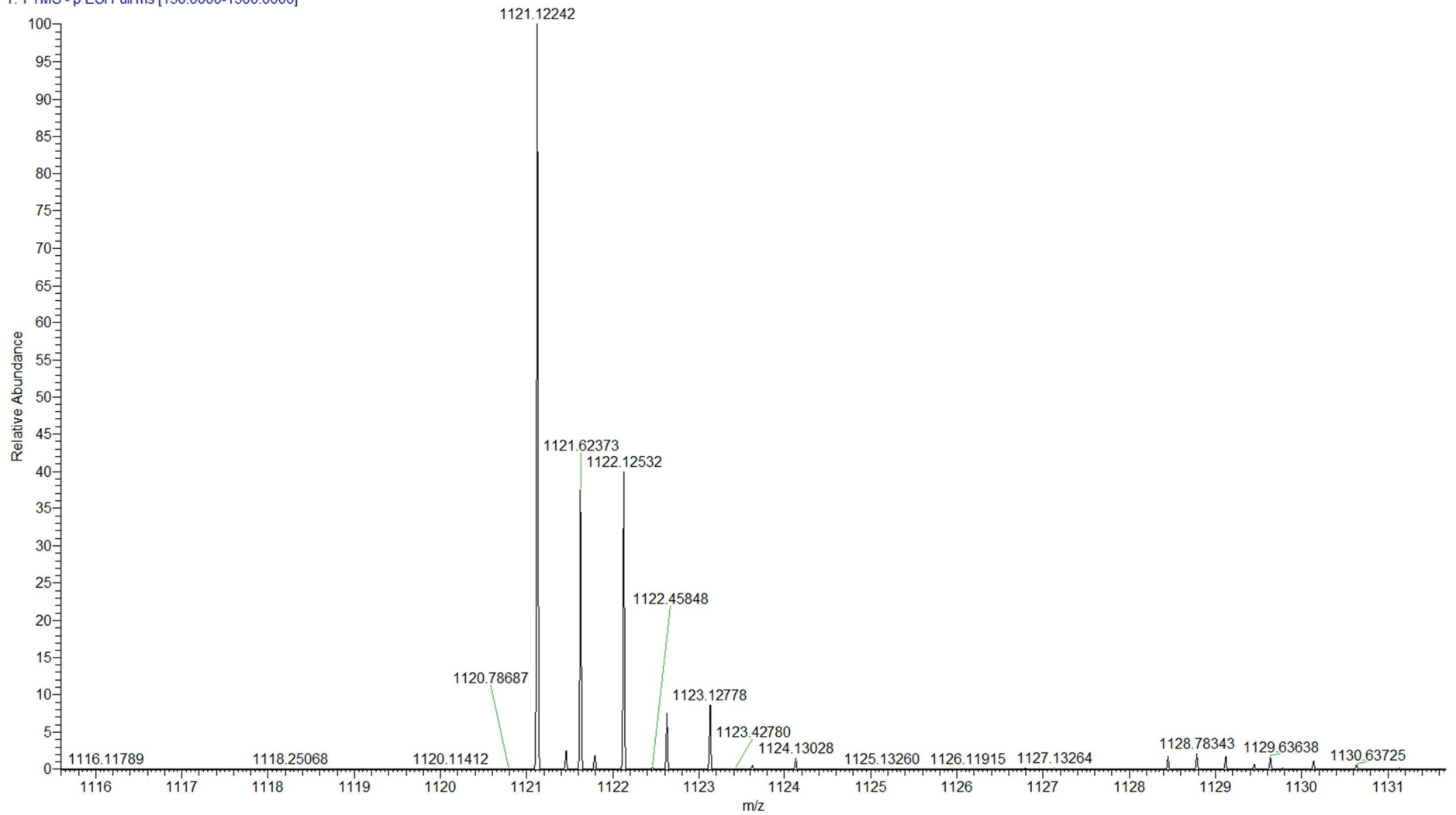


RP HPLC

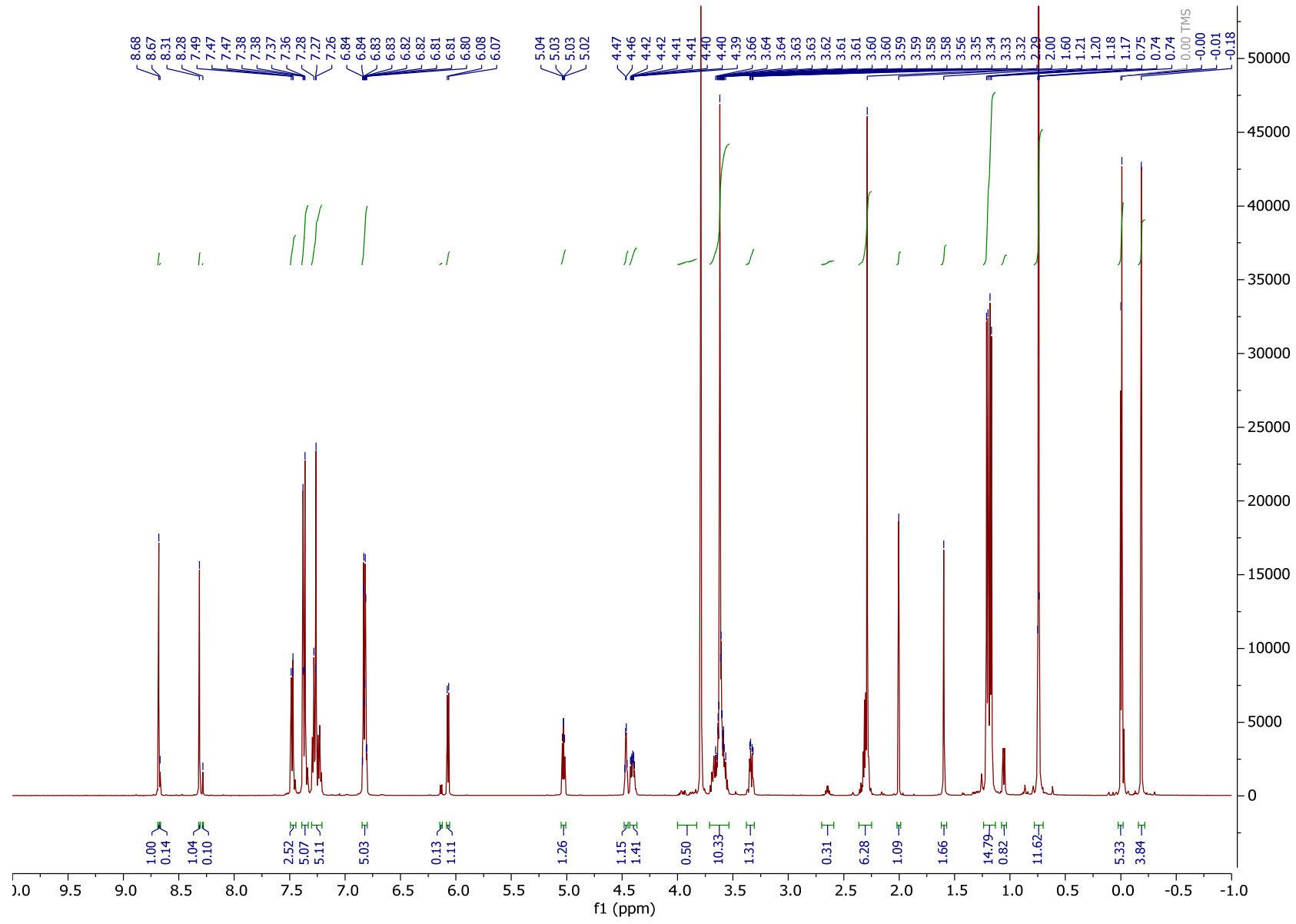


MS (-) ESI
(Calc. [M-H]⁻ C₃₁H₄₁N₁₂O₂₆P₄⁻ 1121.12109)

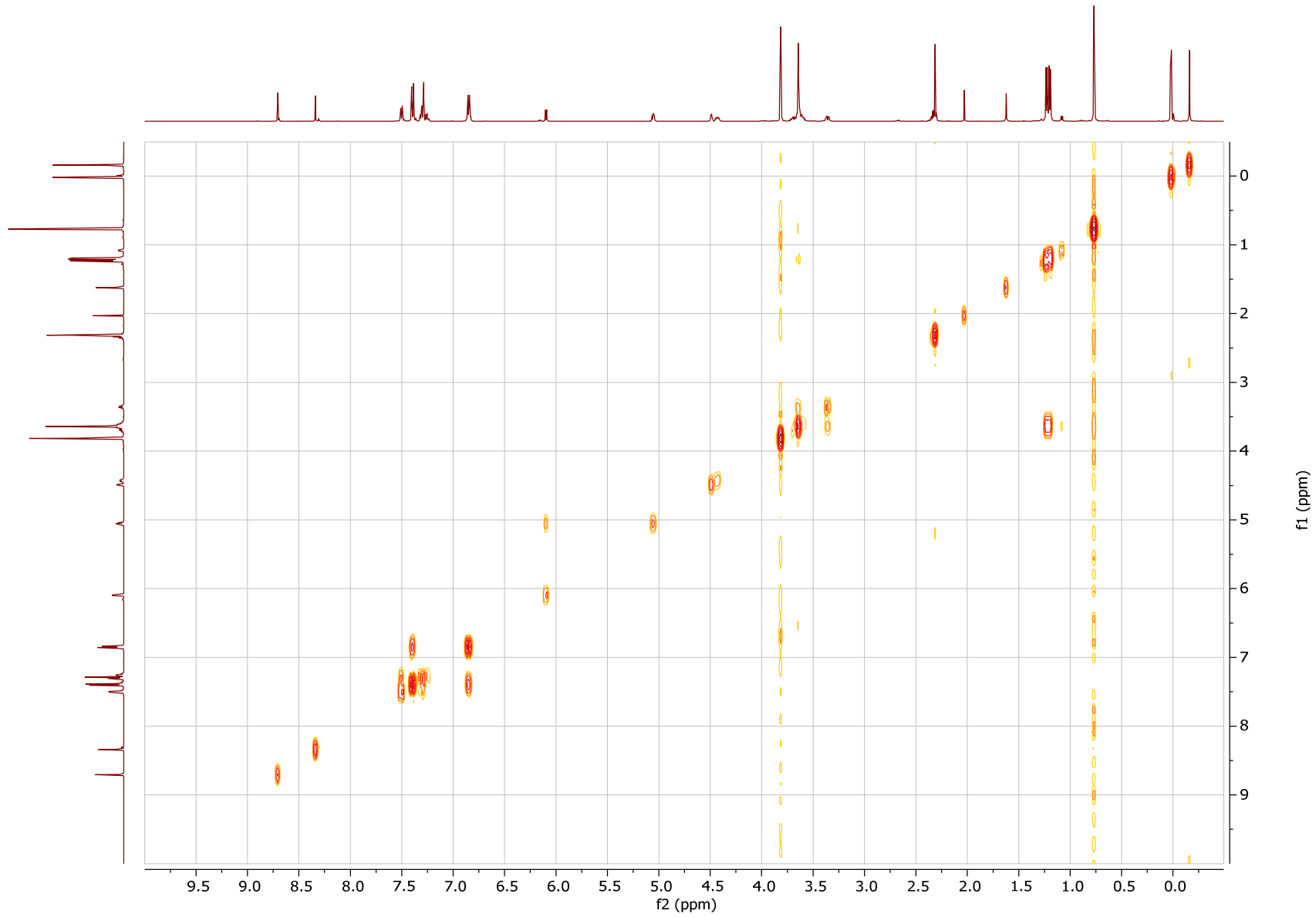
171213_TP_020#18-74 RT: 0.16-0.65 AV: 57 NL: 8.84E5
T: FTMS - p ESI Full ms [150.0000-1500.0000]



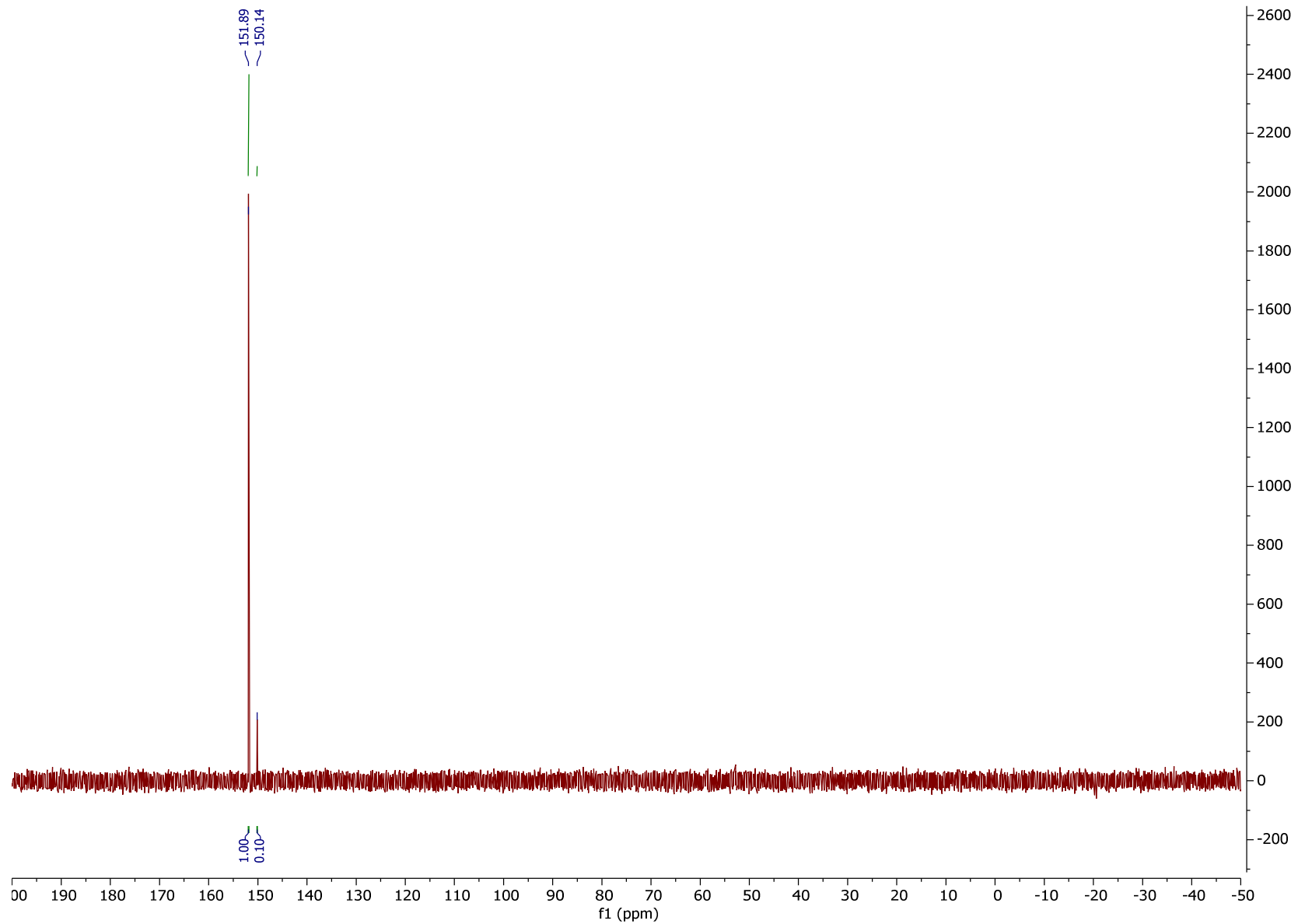
¹H NMR



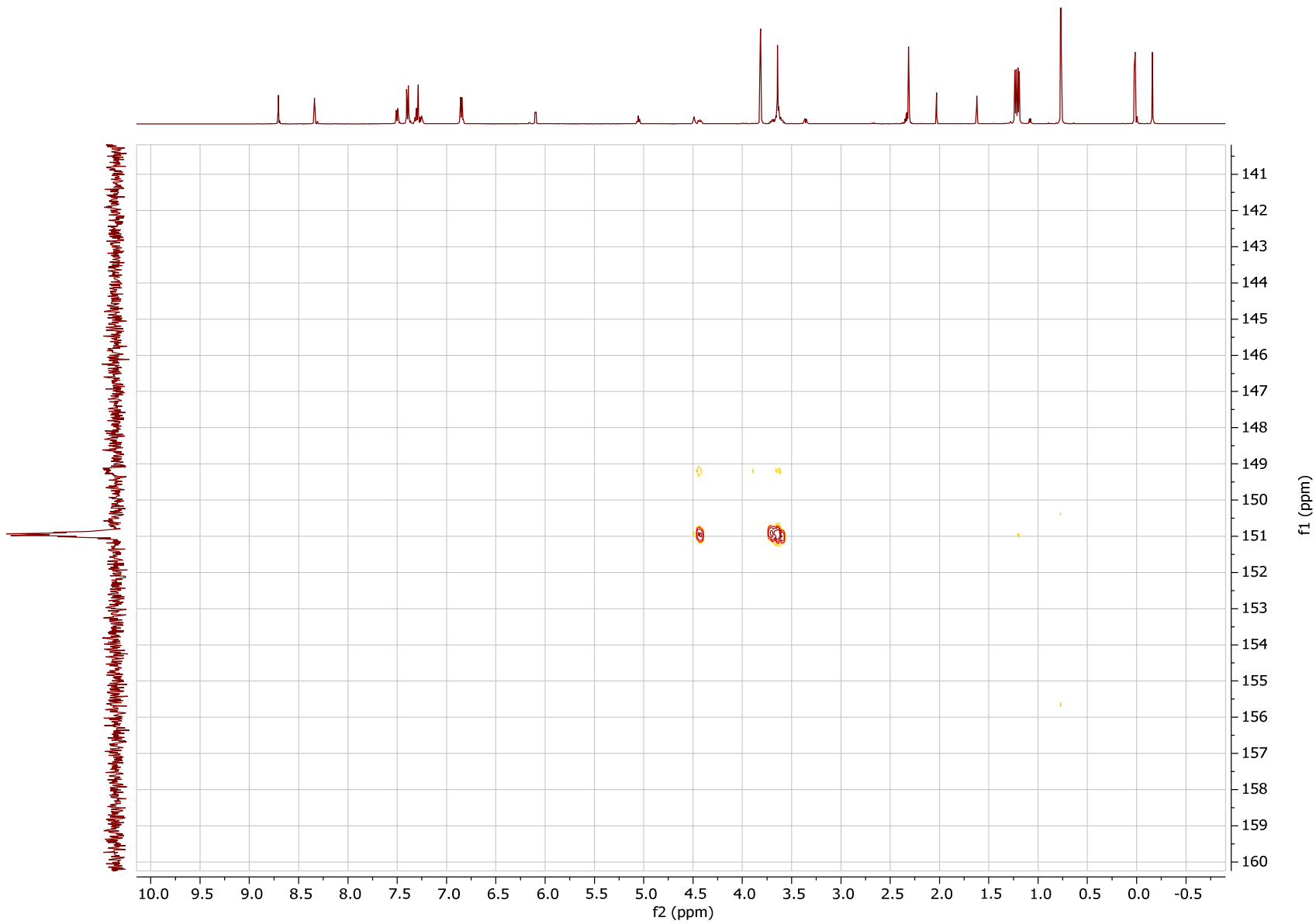
COSY NMR



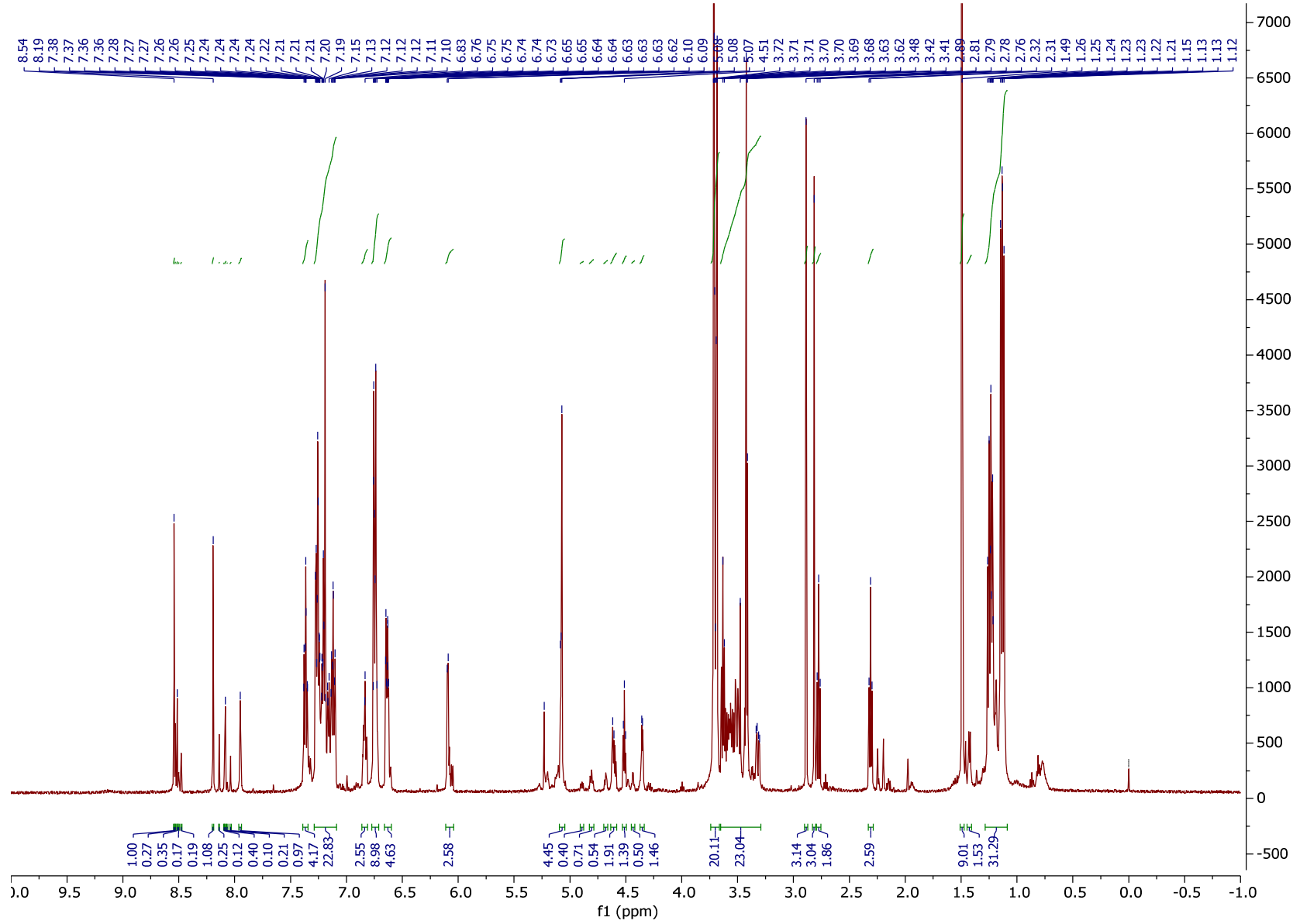
³¹P NMR



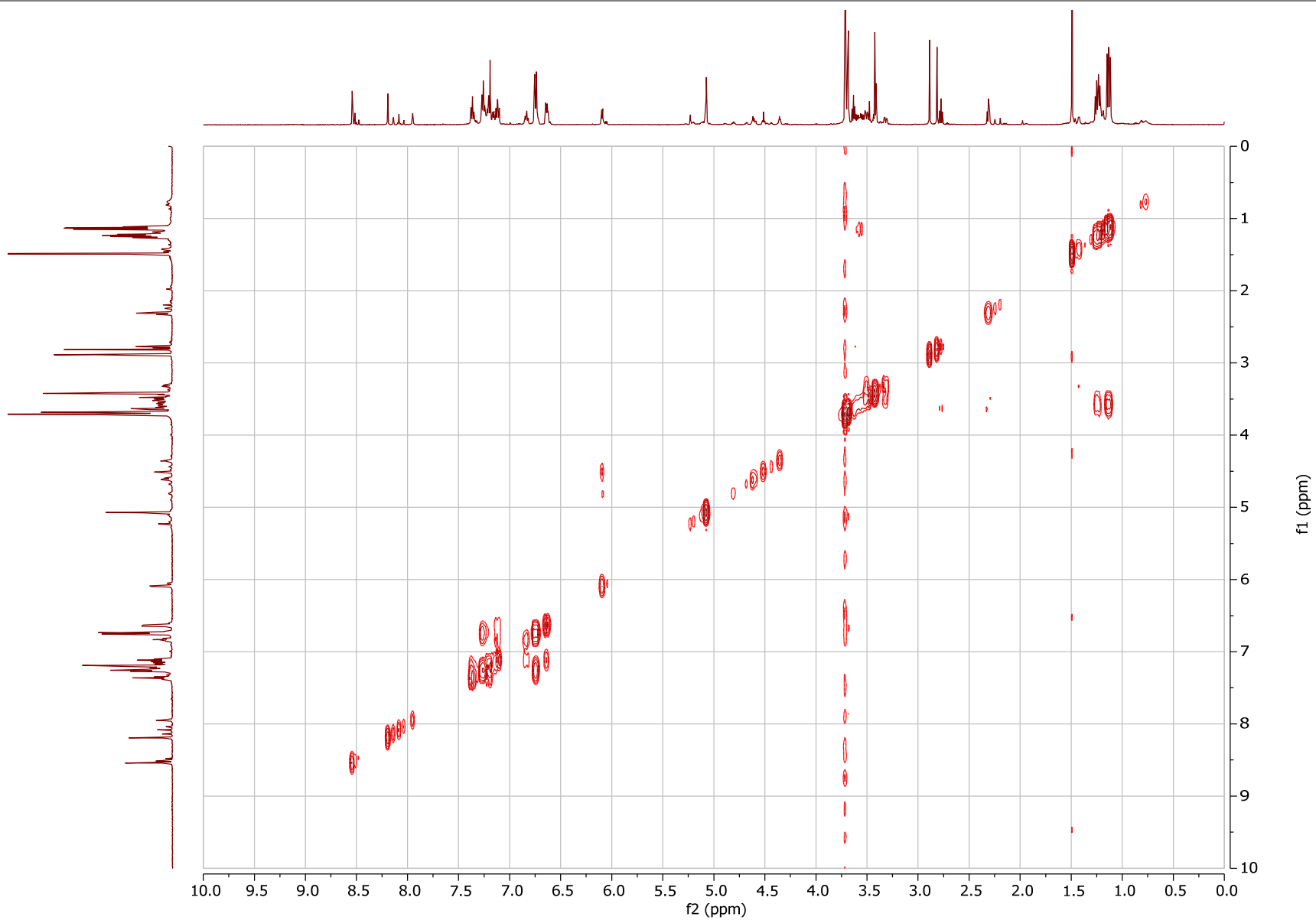
^1H - ^{31}P HMBC



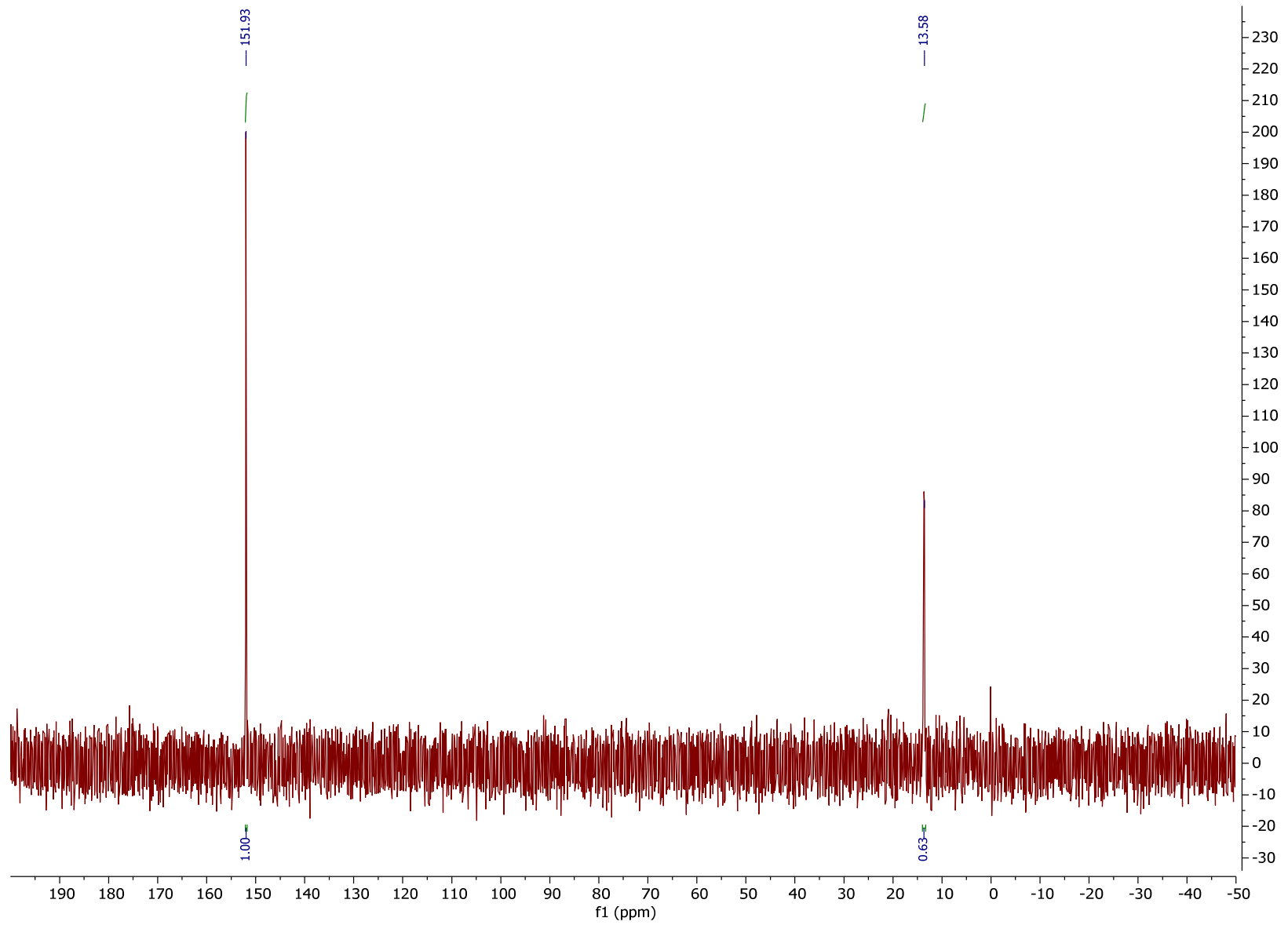
¹H NMR



COSY NMR

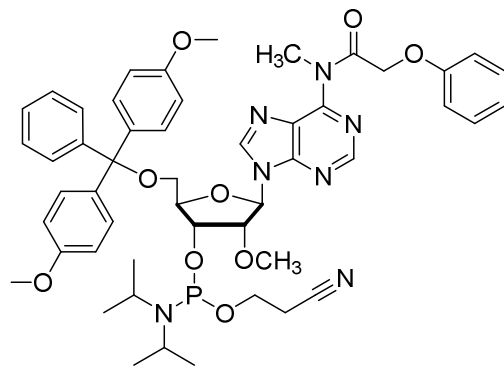


³¹P NMR

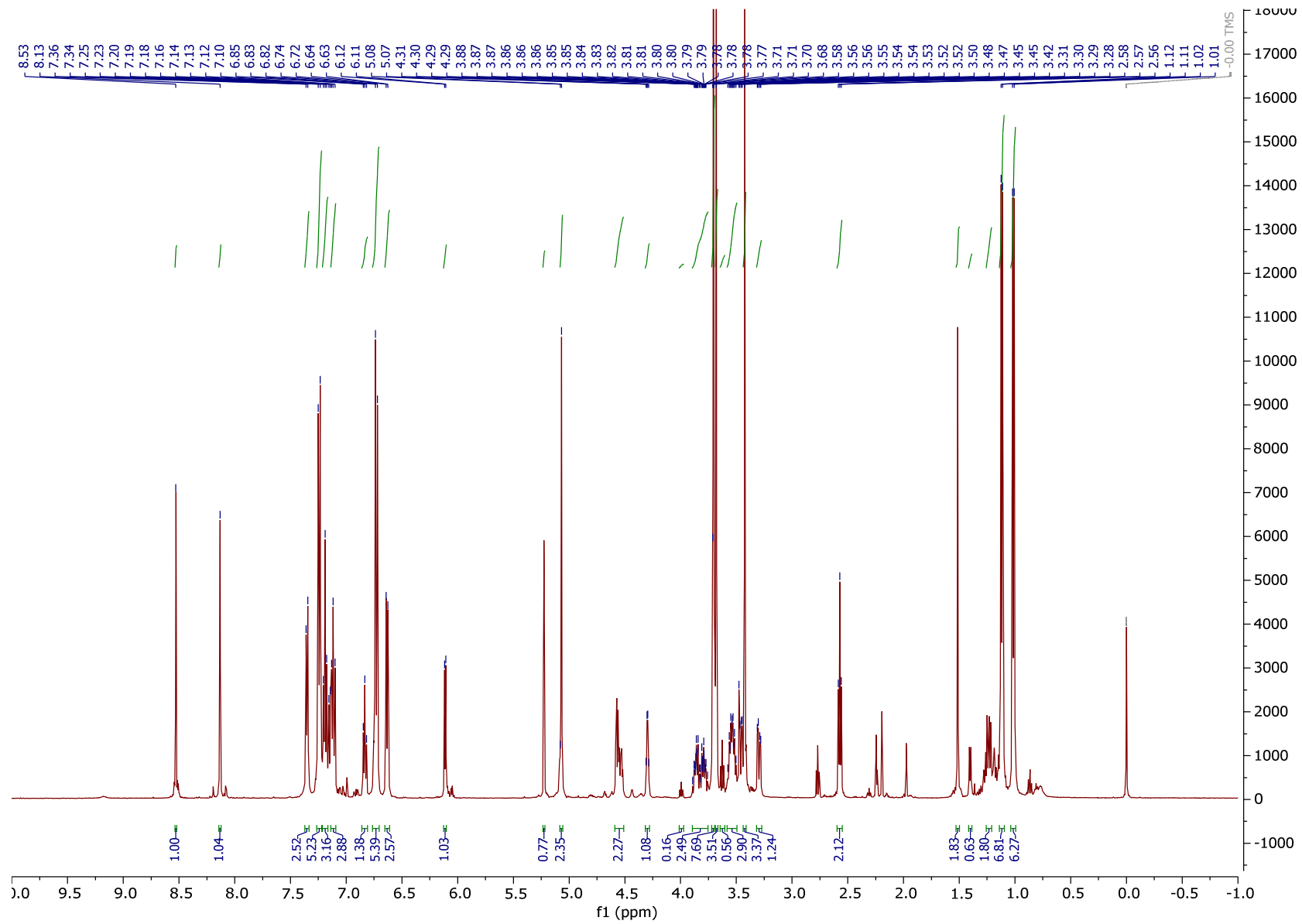


m⁶A_m phosphoramidite (fr. 2)

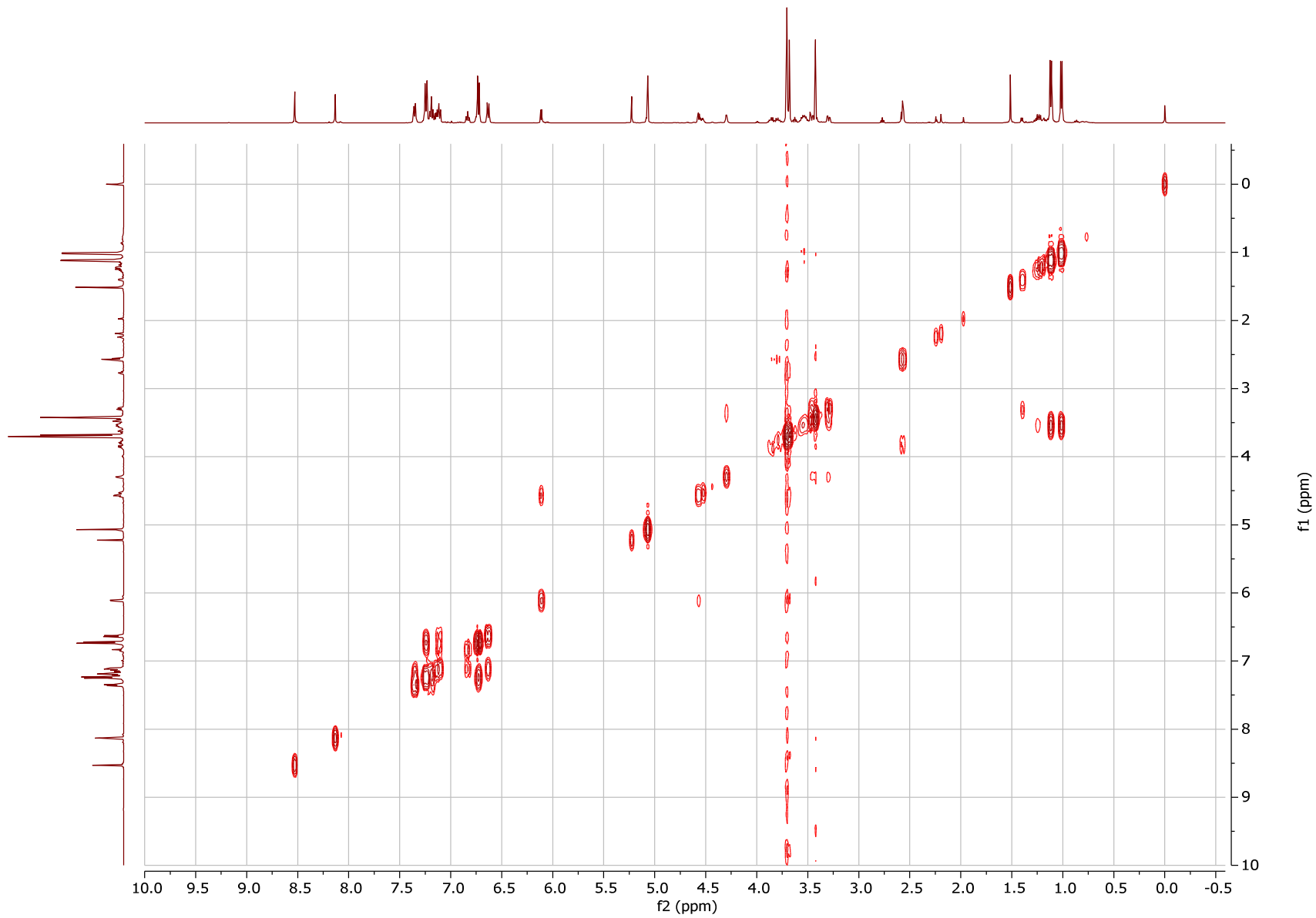
Chemical structure



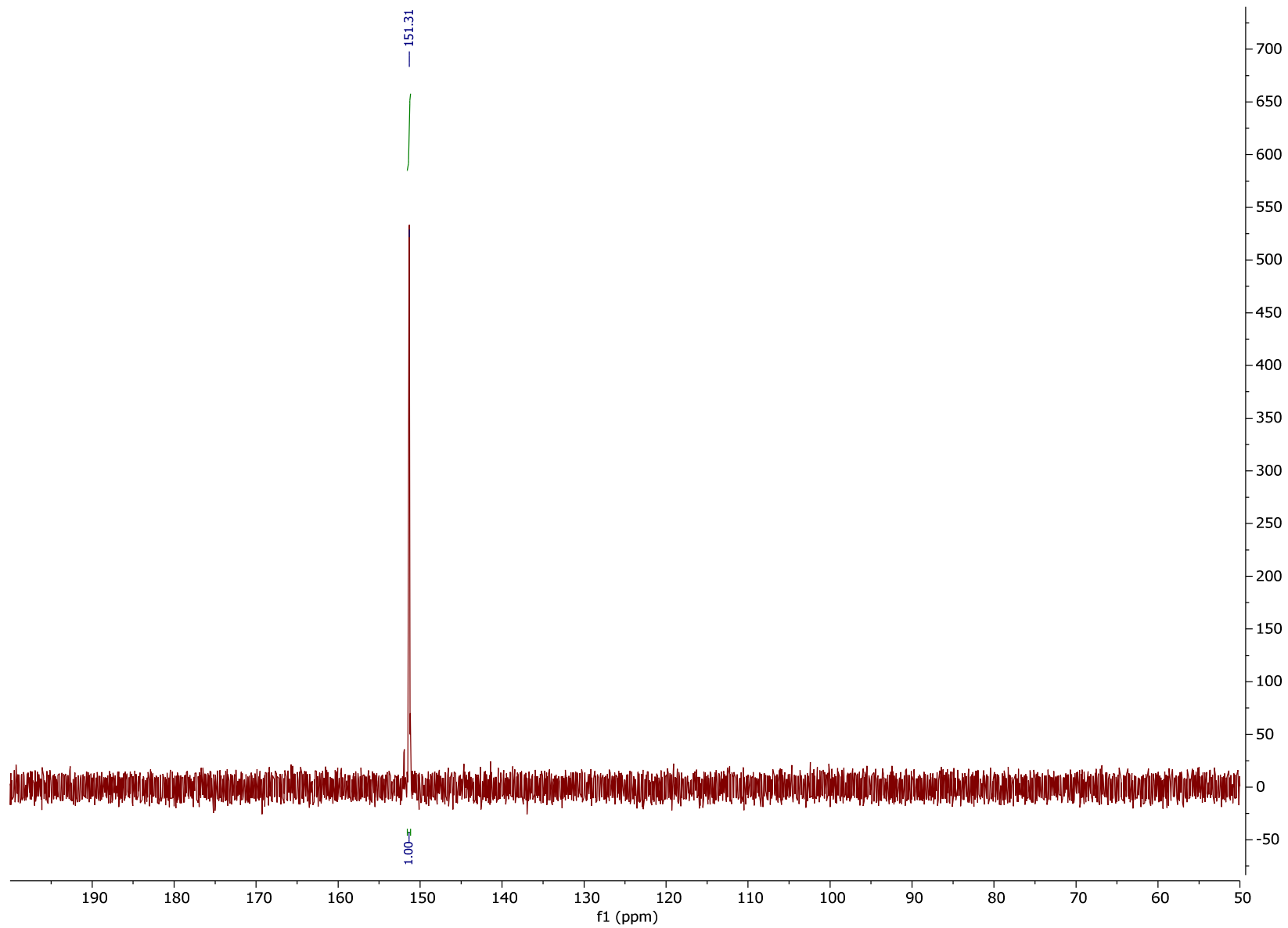
¹H NMR



COSY NMR

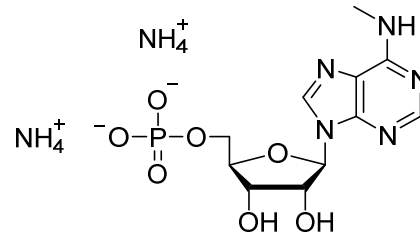


³¹P NMR

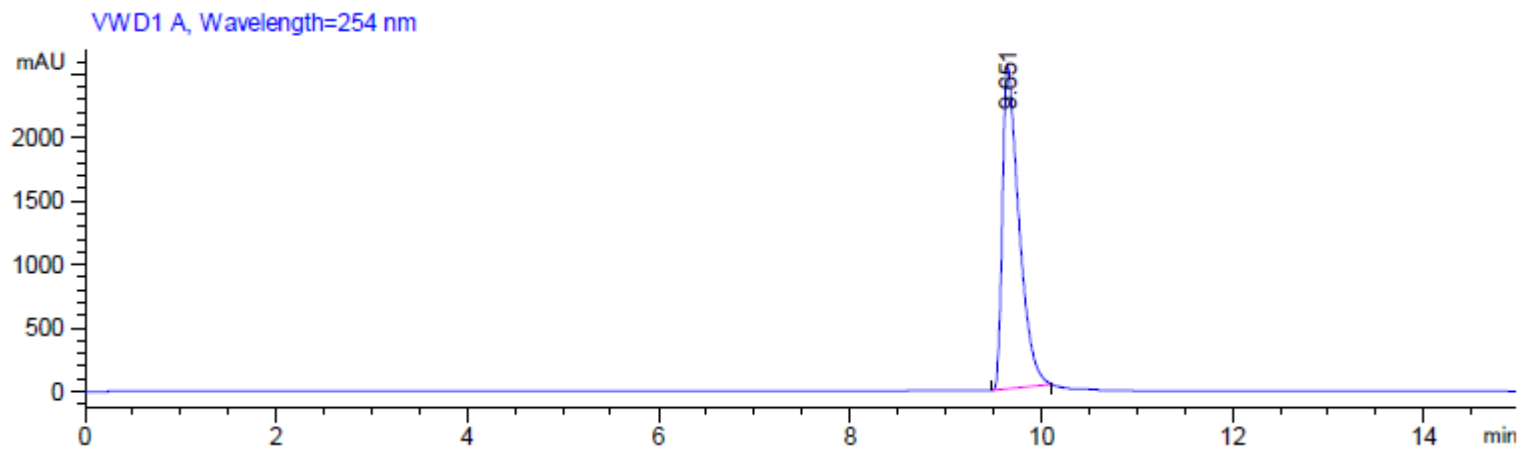


m⁶AMP

Chemical structure

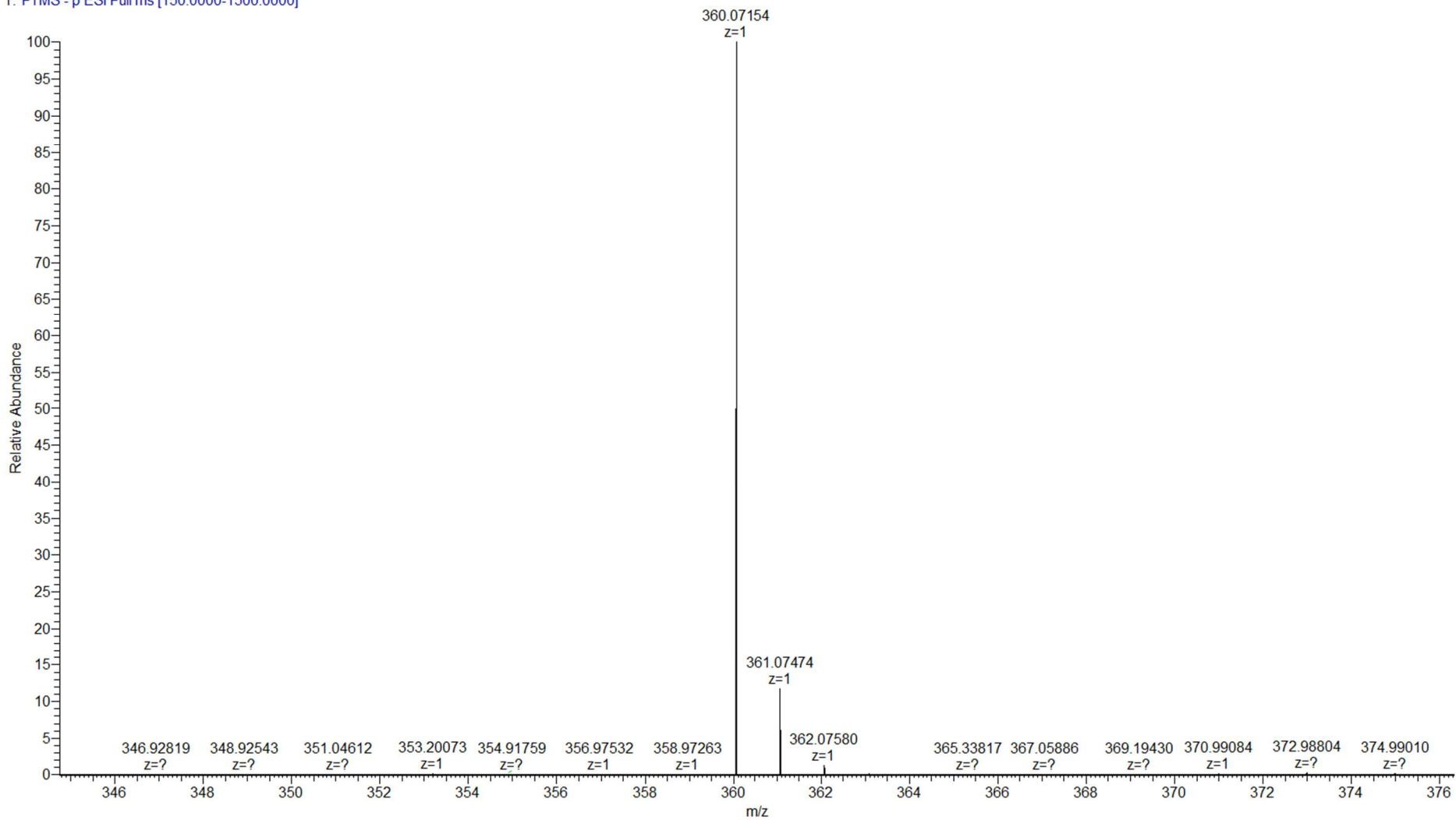


RP HPLC

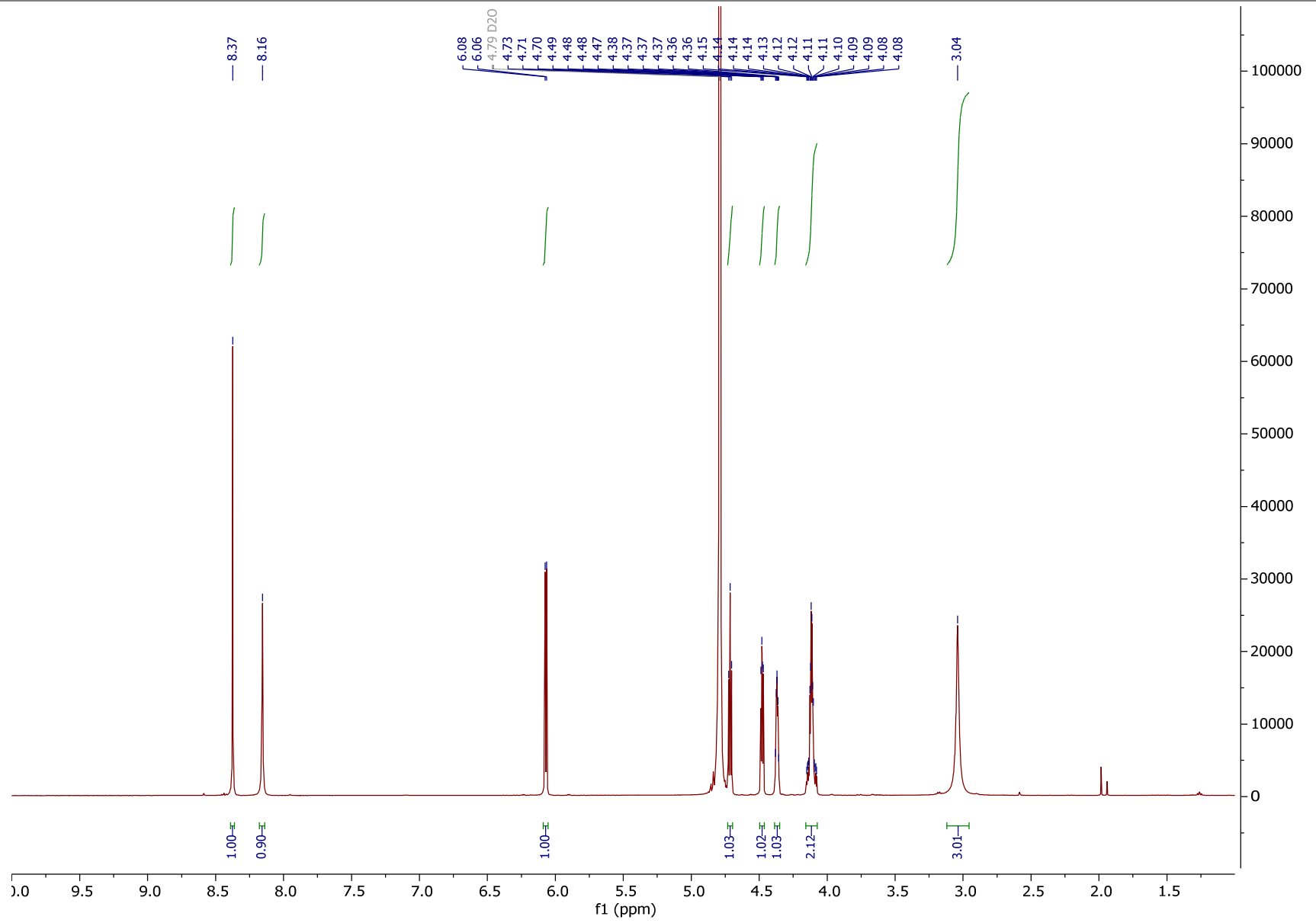


MS (-) ESI
(Calc. [M-H]⁻ C₁₁H₁₅N₅O₇P⁻ 360.07146)

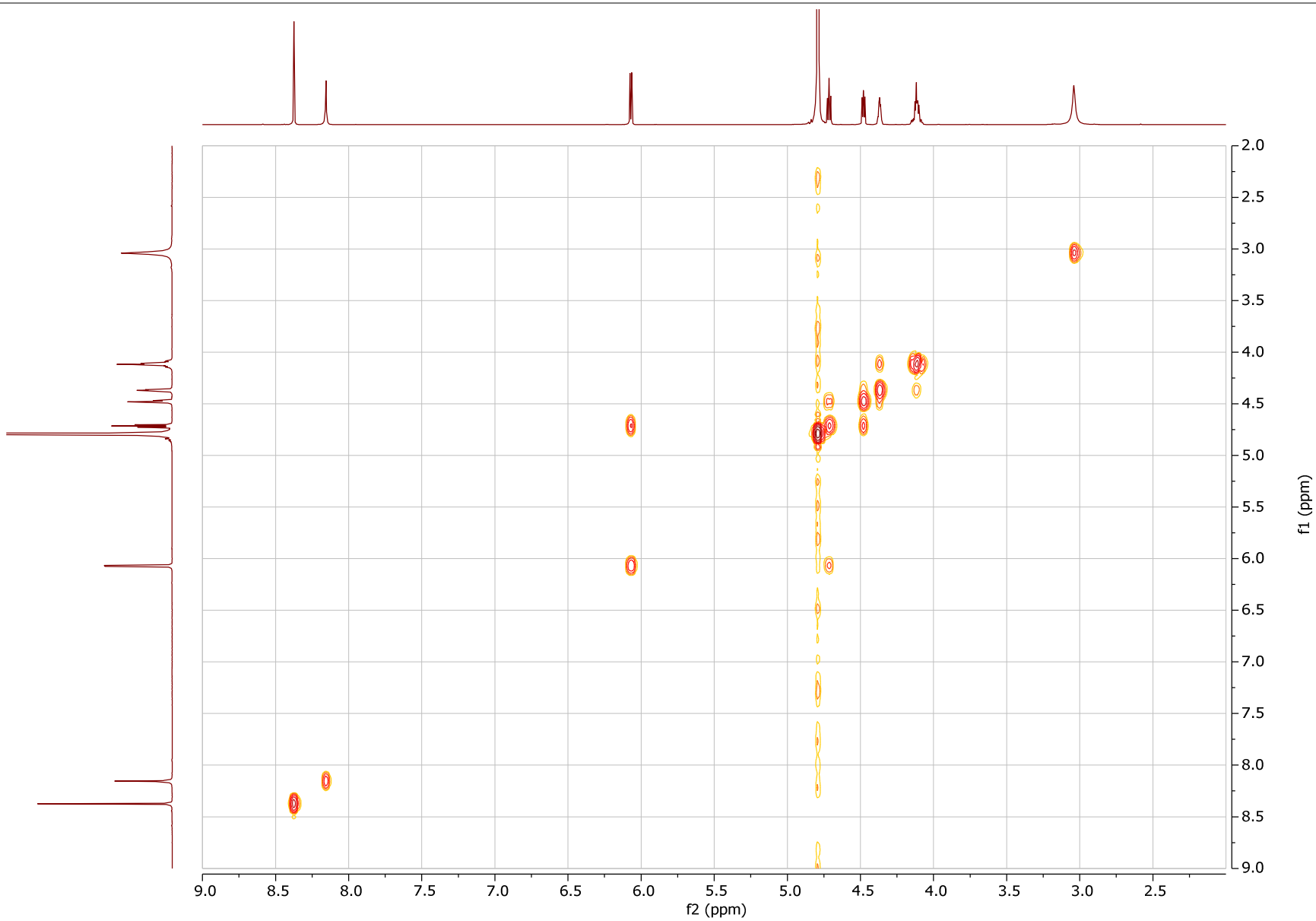
171213_TP_009#11-66 RT: 0.10-0.58 AV: 56 NL: 5.50E7
T: FTMS - p ESI Full ms [150.0000-1500.0000]



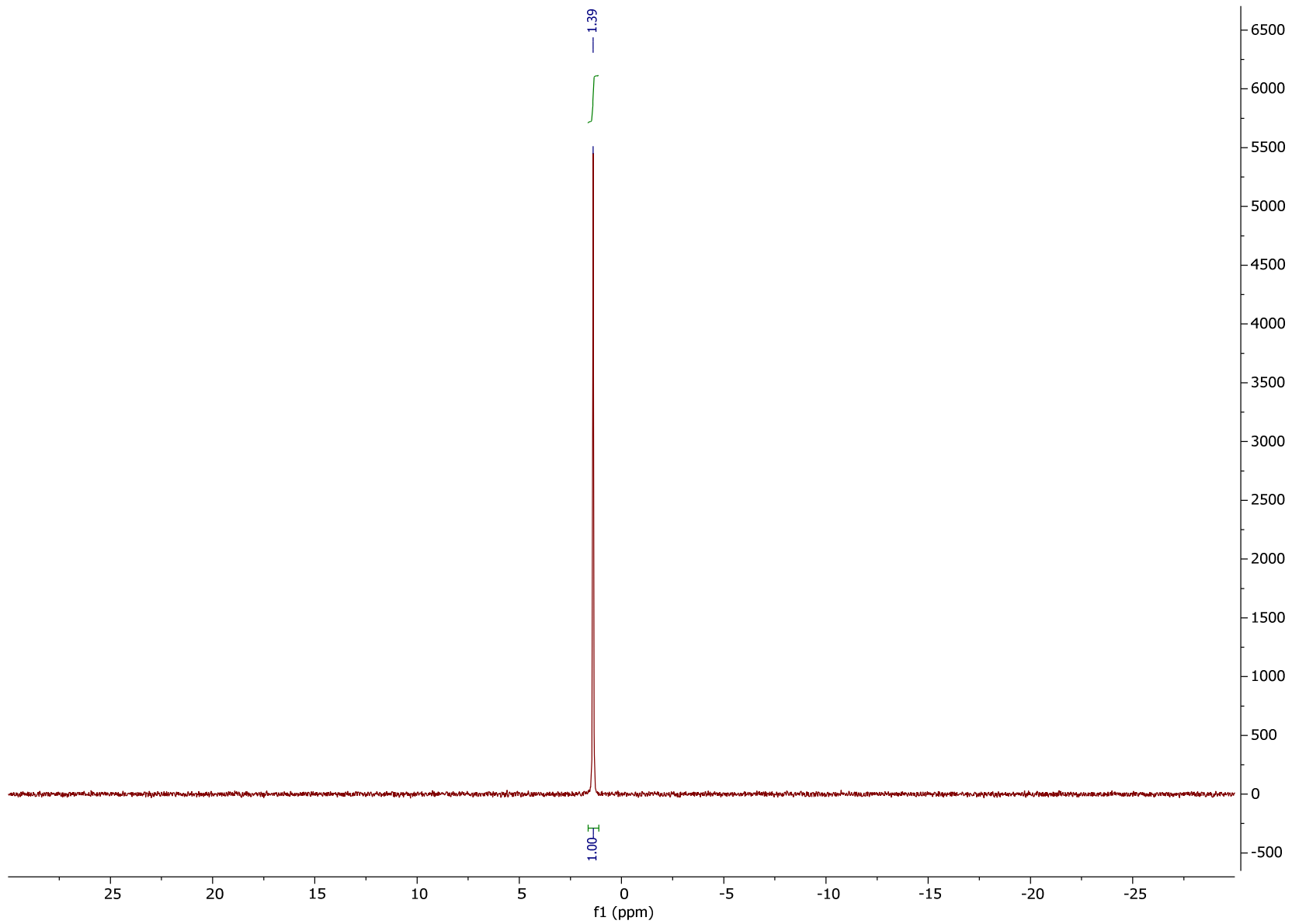
¹H NMR



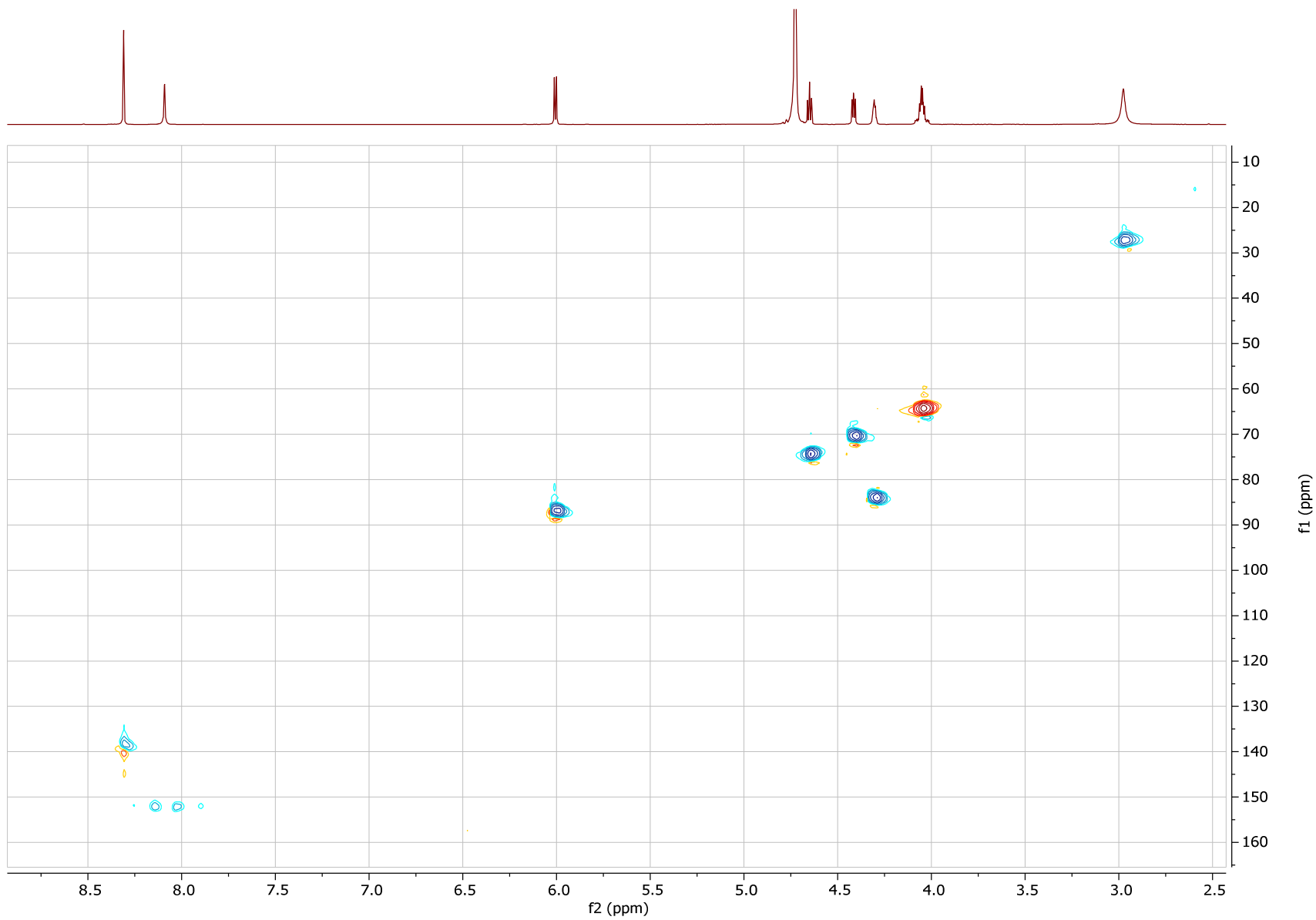
COSY NMR



³¹P NMR

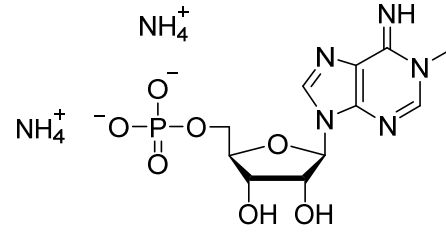


^1H - ^{13}C HSQC

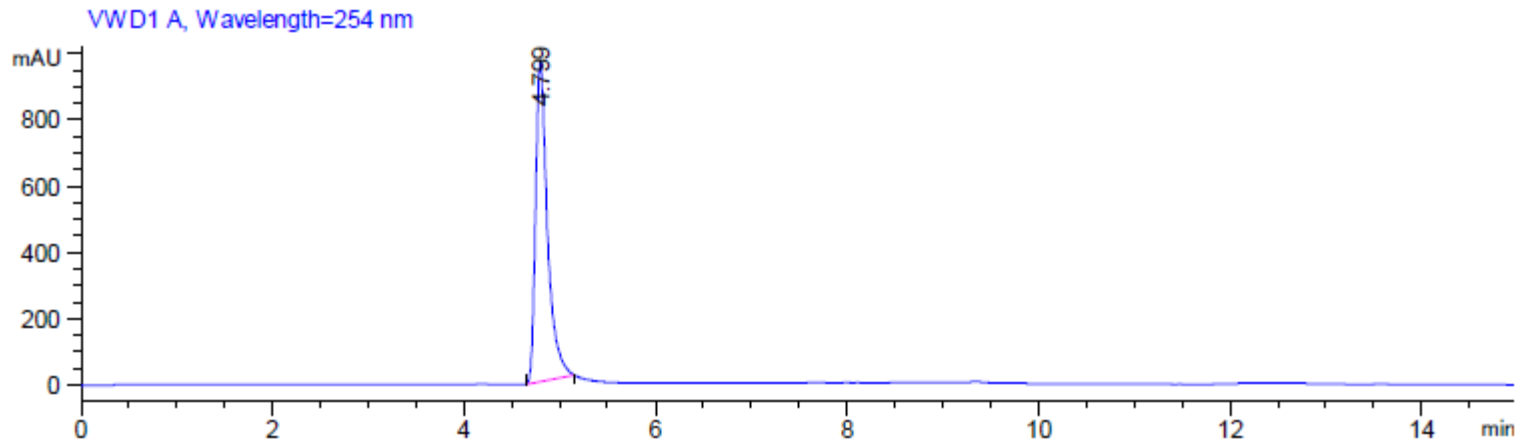


m¹AMP

Chemical structure

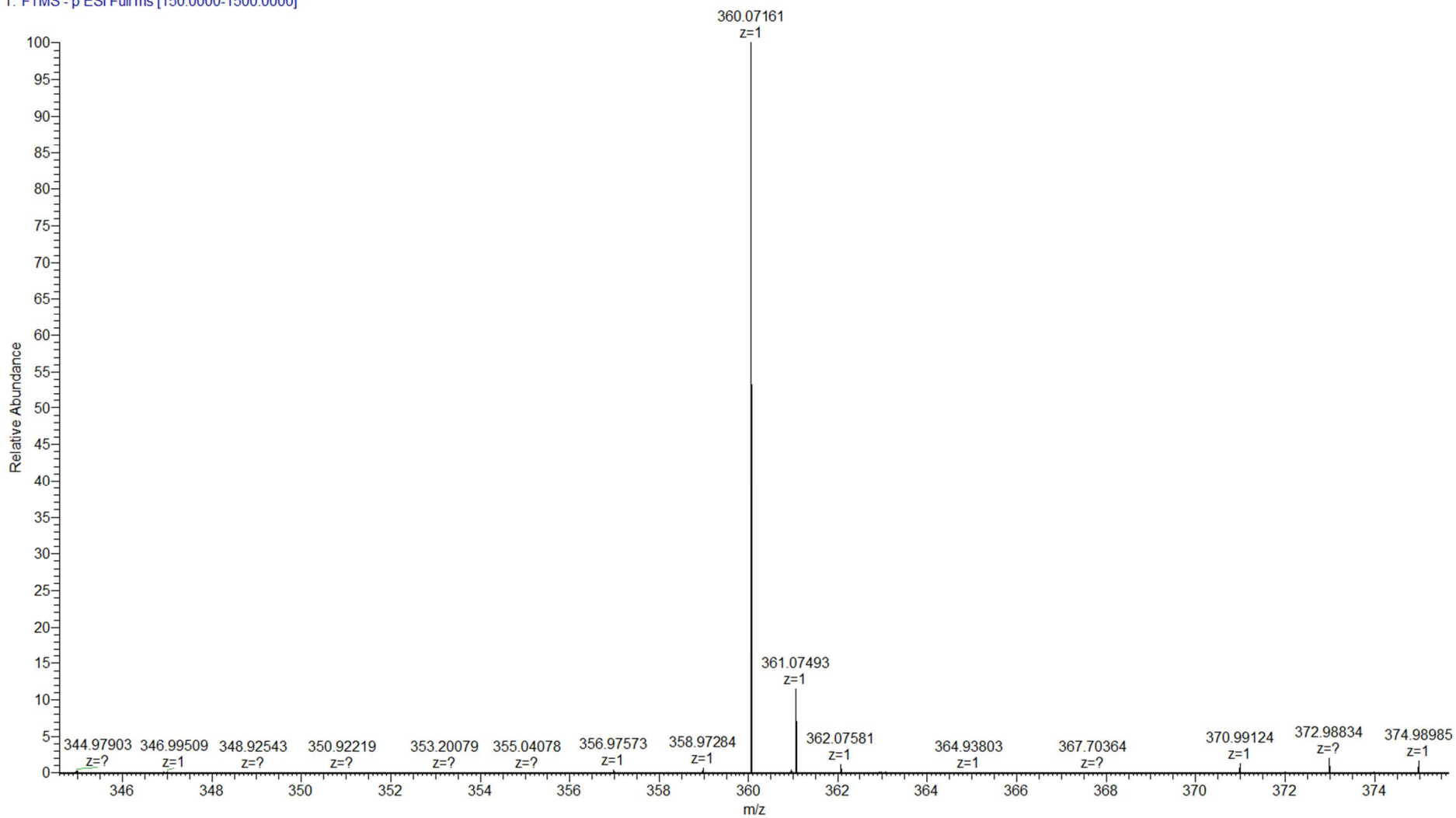


RP HPLC

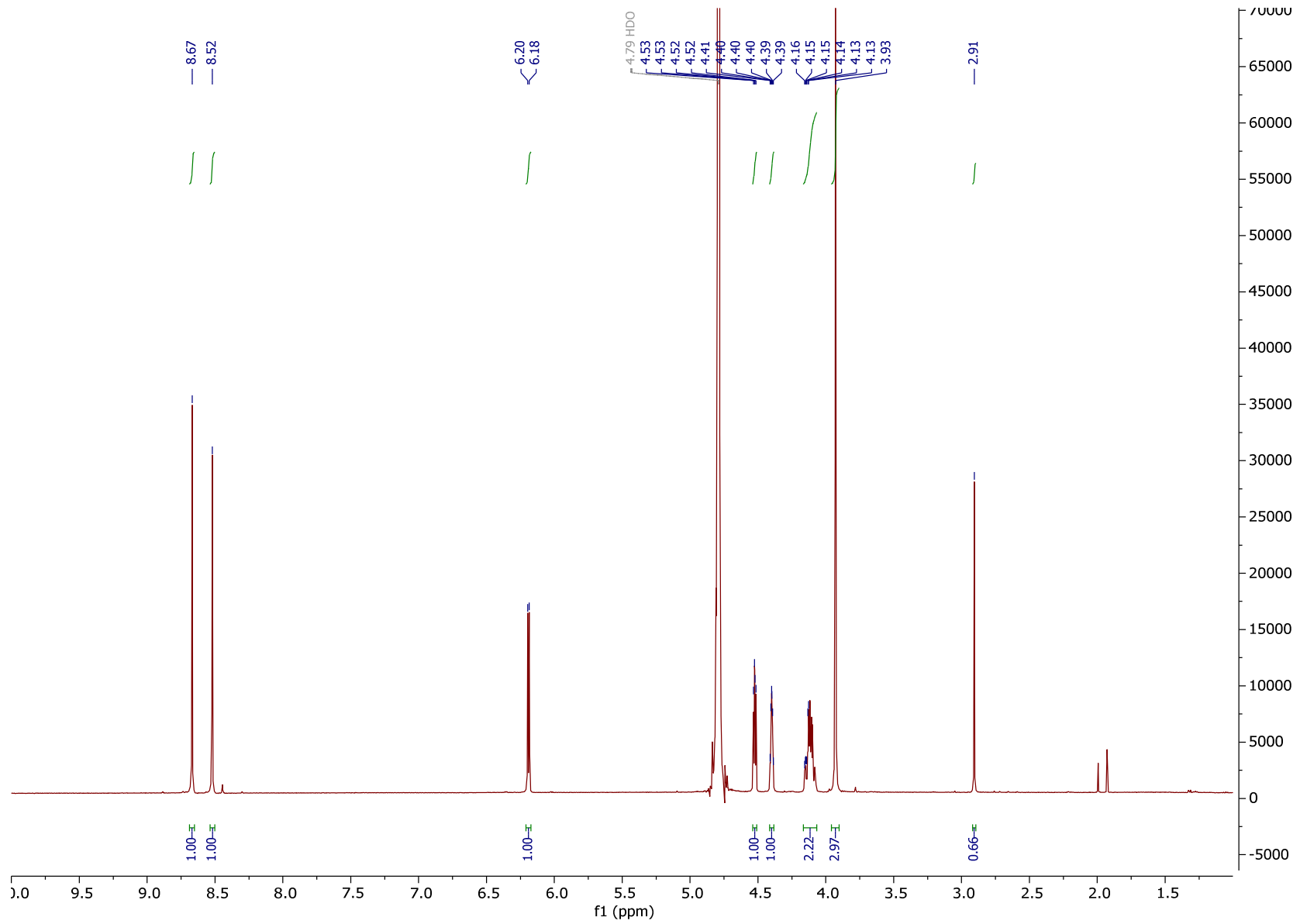


MS (-) ESI
(Calc. [M-H]⁻ C₁₁H₁₅N₅O₇P⁻ 360.07146)

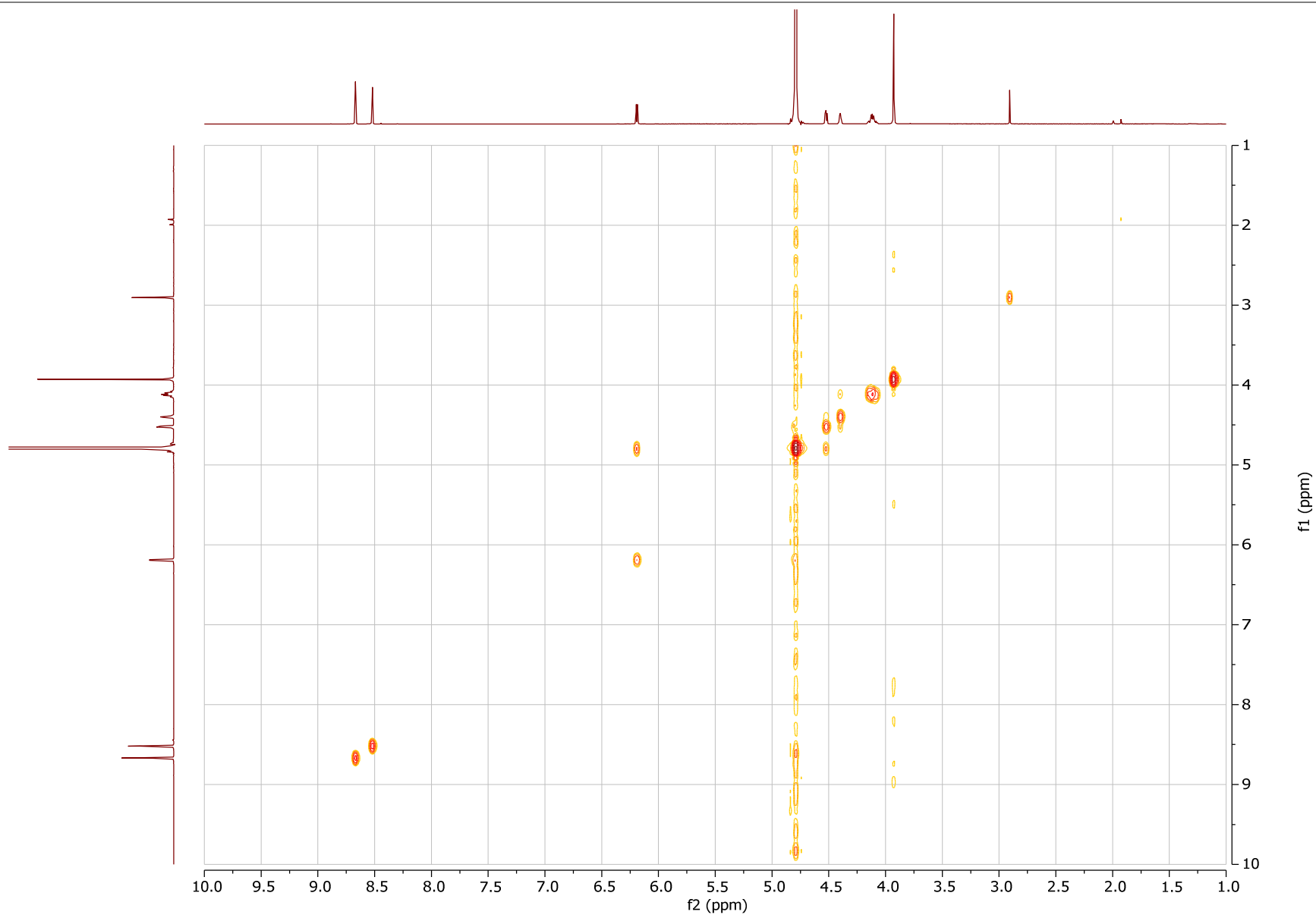
171213_TP_008#58-118 RT: 0.51-1.03 AV: 61 NL: 5.57E6
T: FTMS - p ESI Full ms [150.0000-1500.0000]



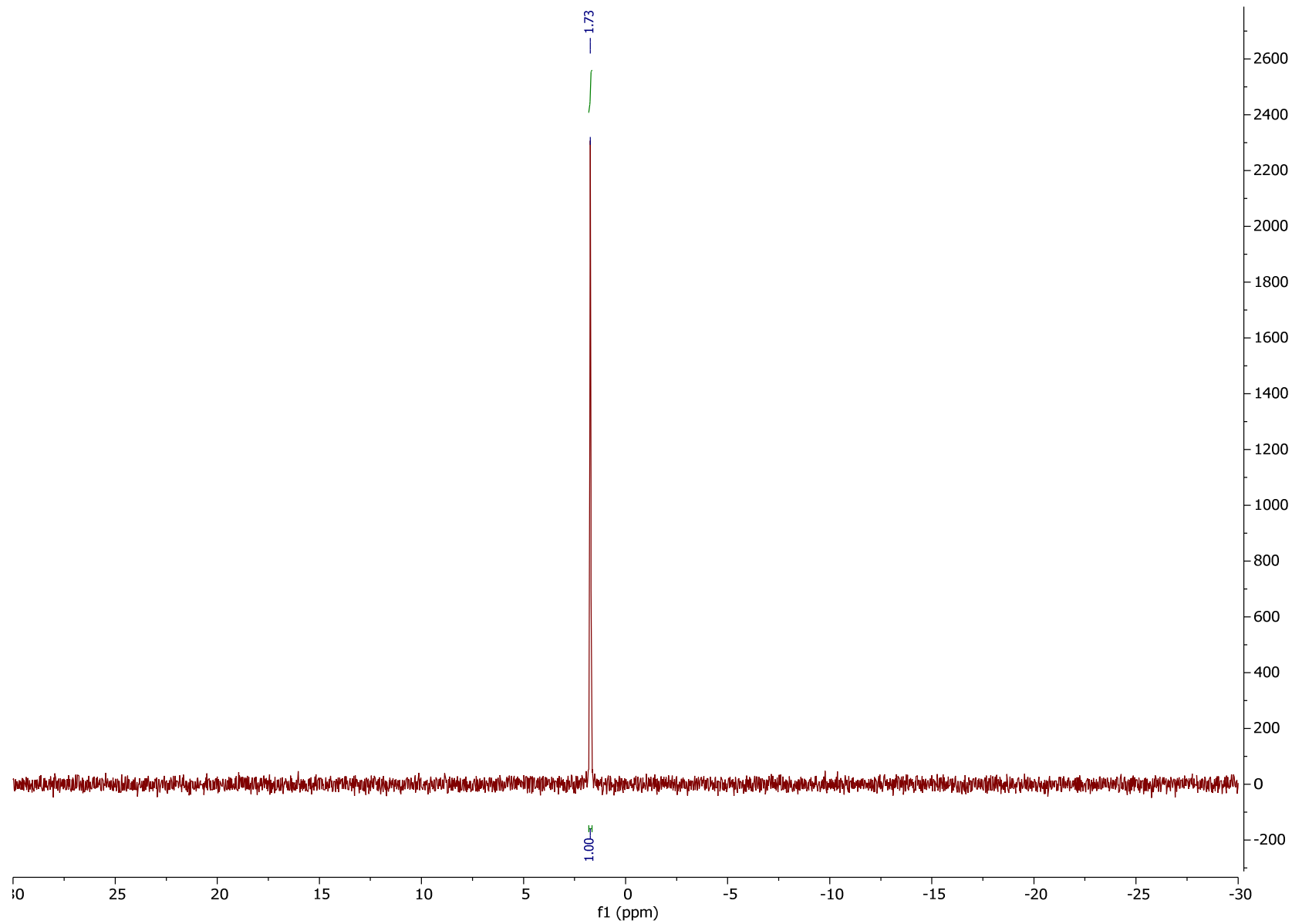
¹H NMR



COSY NMR

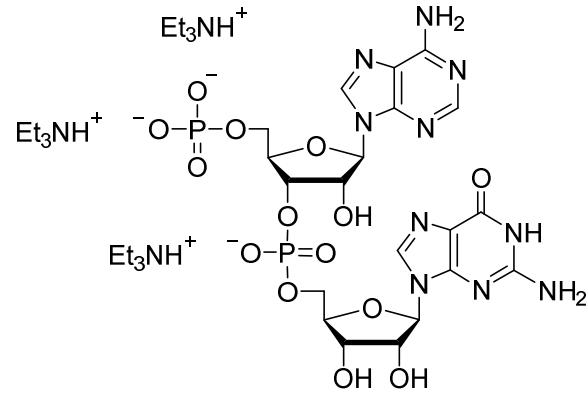


³¹P NMR

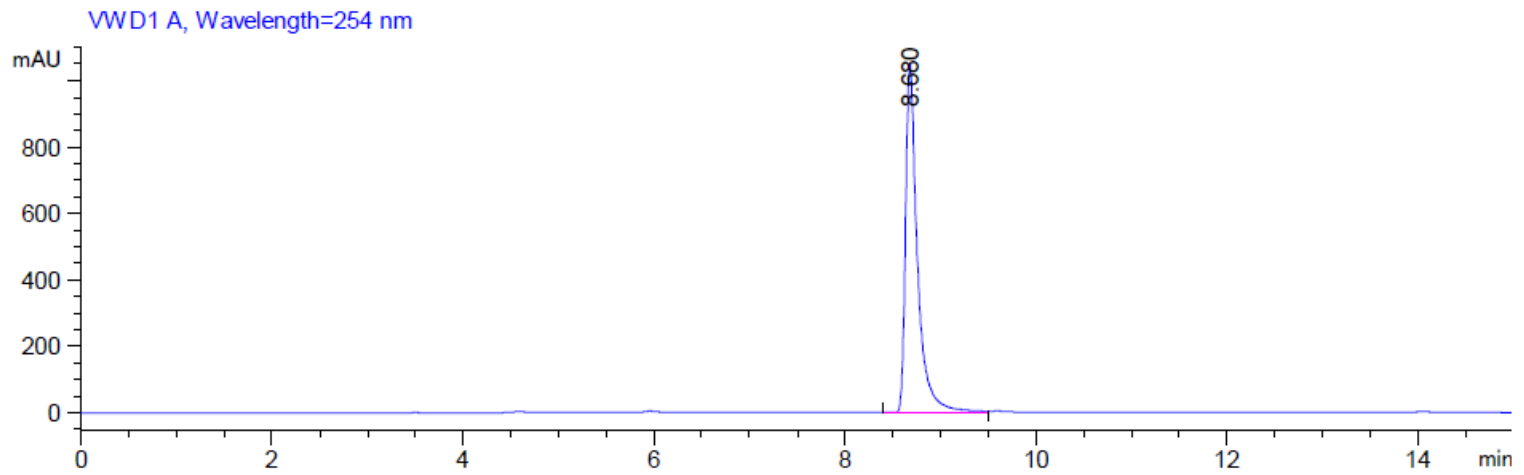


pApG

Chemical structure

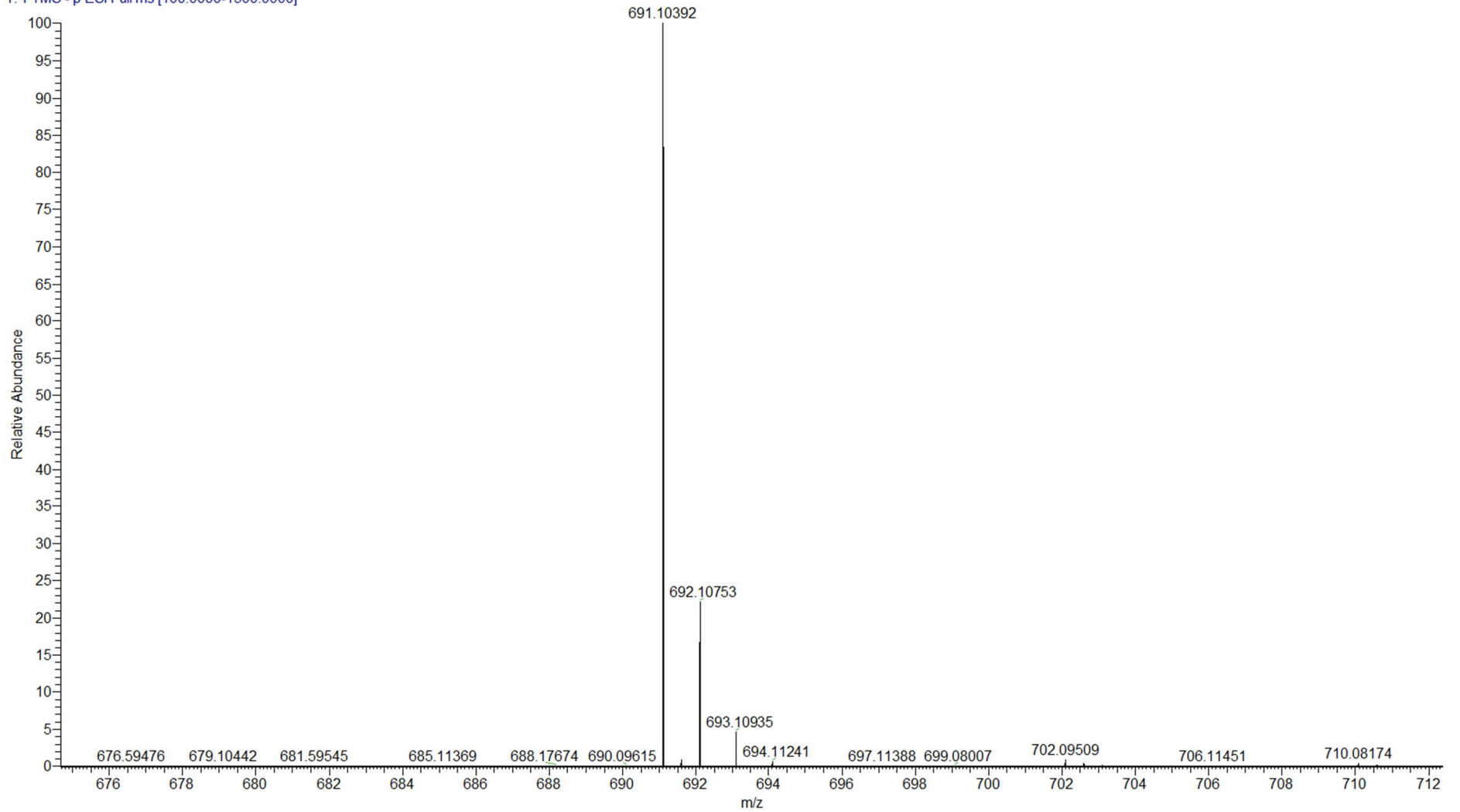


RP HPLC



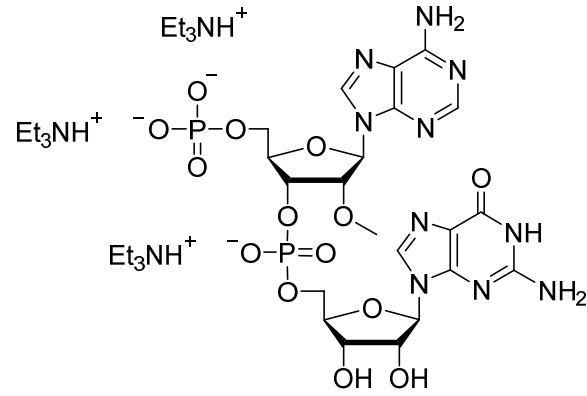
MS (-) ESI
(Calc. [M-H]⁻ C₂₀H₂₅N₁₀O₁₄P₂⁻ 691.10324)

171213_TP_003#85-213 RT: 0.74-1.86 AV: 129 NL: 1.90E7
T: FTMS - p ESI Full ms [100.0000-1500.0000]

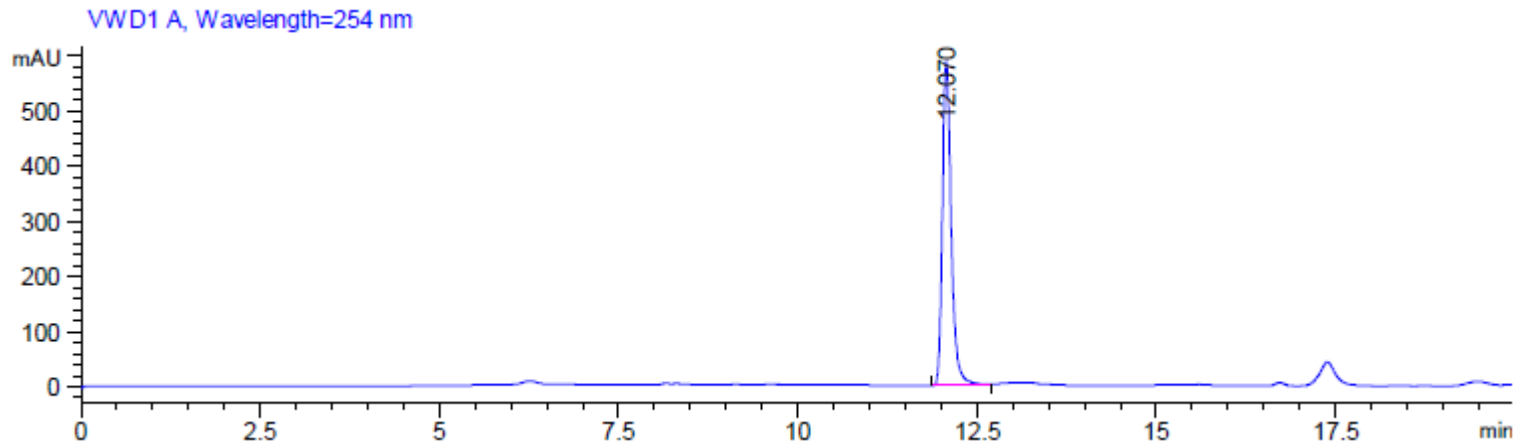


pA_mpG

Chemical structure

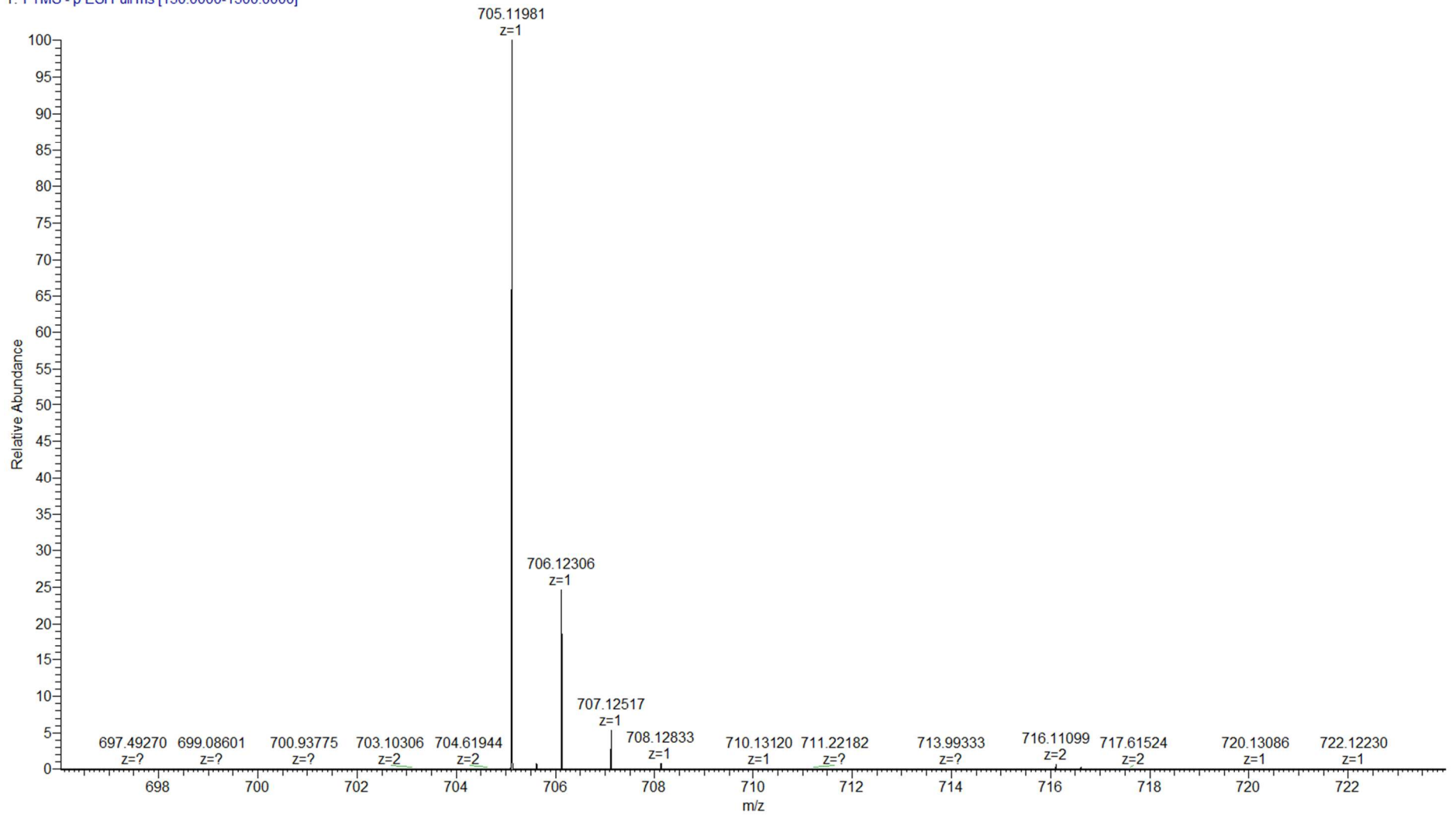


RP HPLC



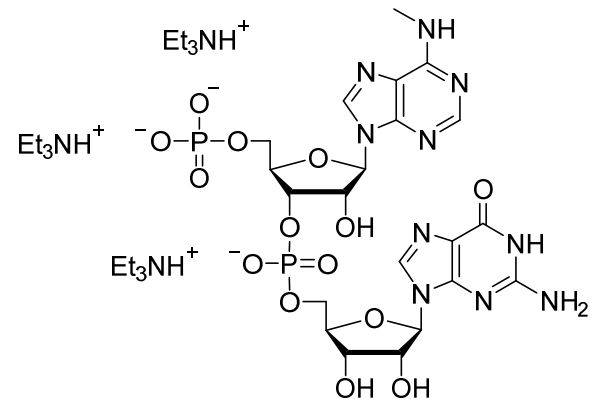
MS (-) ESI
(Calc. [M-H]⁻ C₂₁H₂₇N₁₀O₁₄P₂⁻ 705.11889)

171213_TP_015#15-85 RT: 0.13-0.74 AV: 71 NL: 7.34E6
T: FTMS - p ESI Full ms [150.0000-1500.0000]

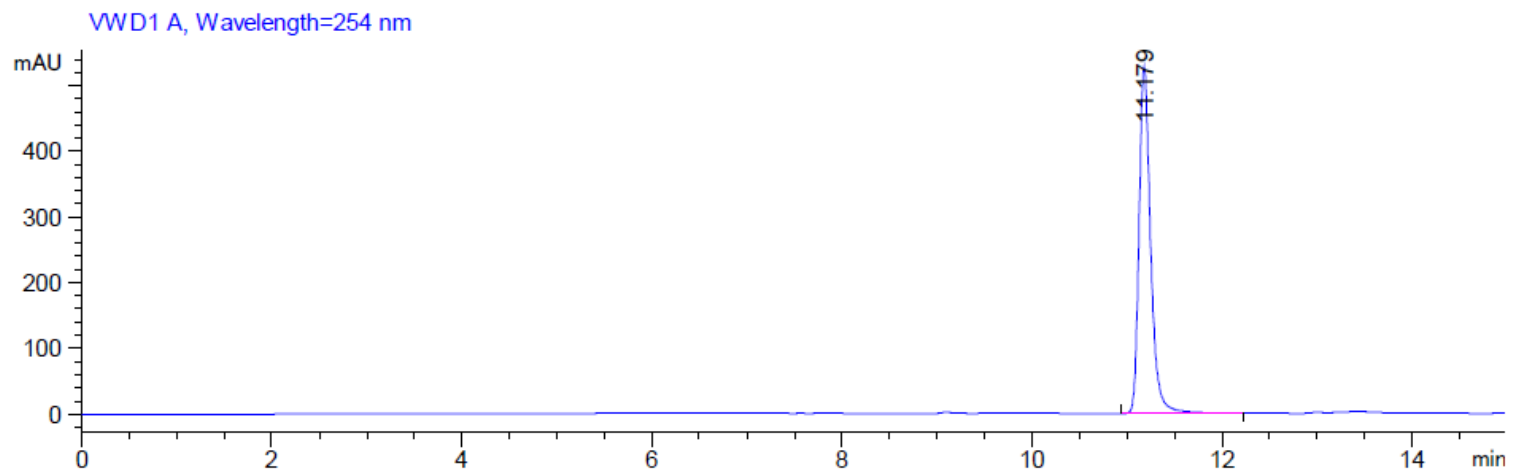


p^{m6}ApG

Chemical structure

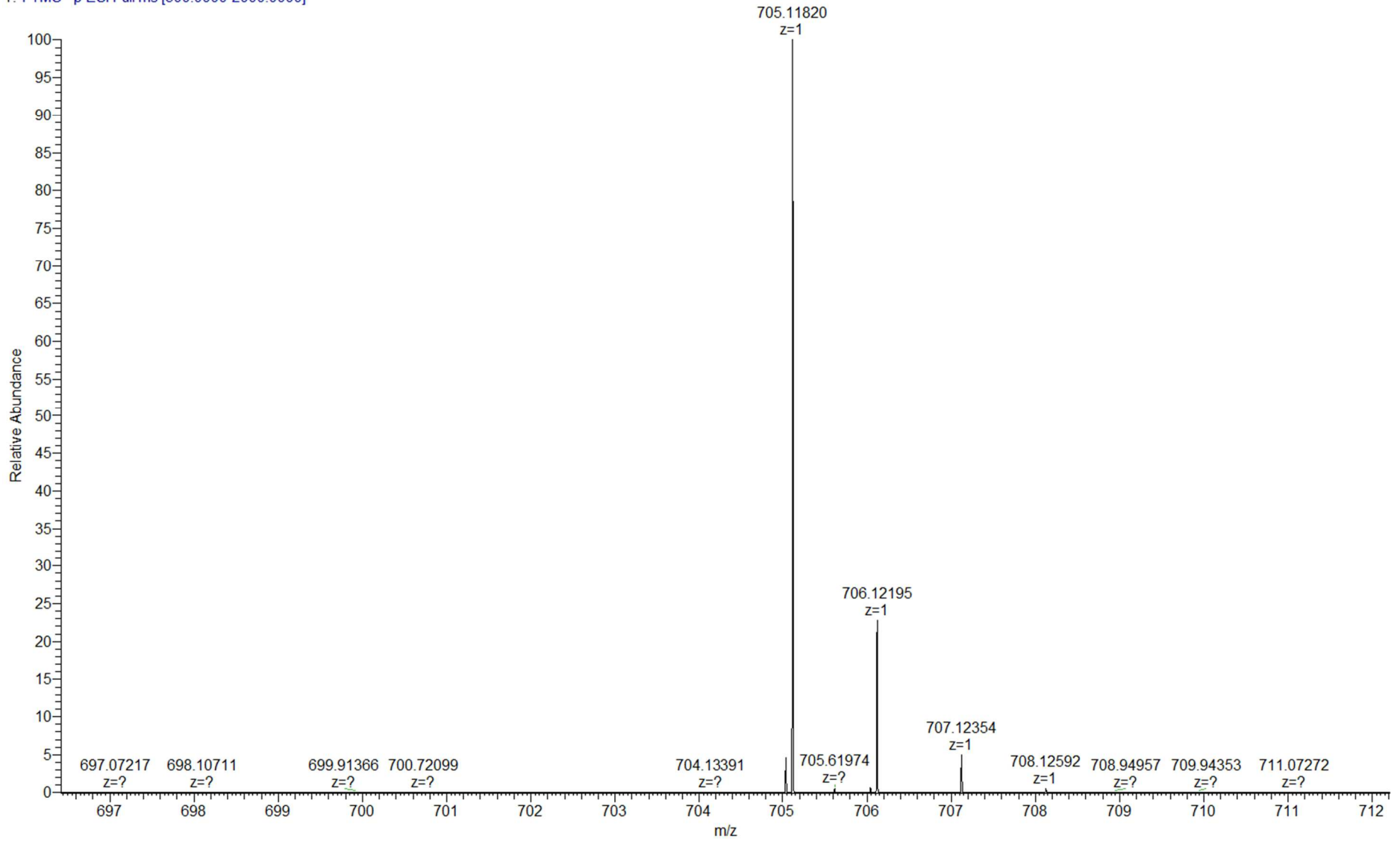


RP HPLC



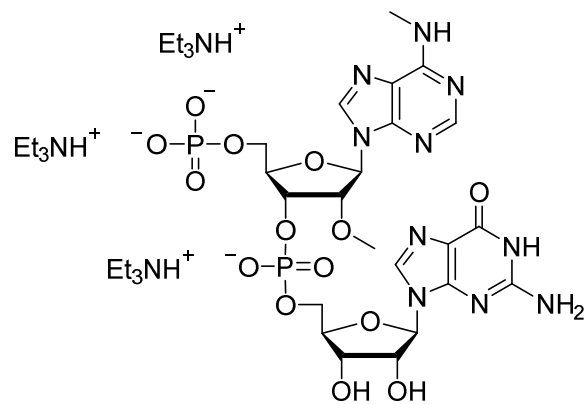
170711_TP001 #1-49 RT: 0.02-0.63 AV: 49 NL: 8.34E4
T: FTMS - p ESI Full ms [300.0000-2000.0000]

MS (-) ESI
(Calc. [M-H]⁻ C₂₁H₂₇N₁₀O₁₄P₂⁻ 705.11889)

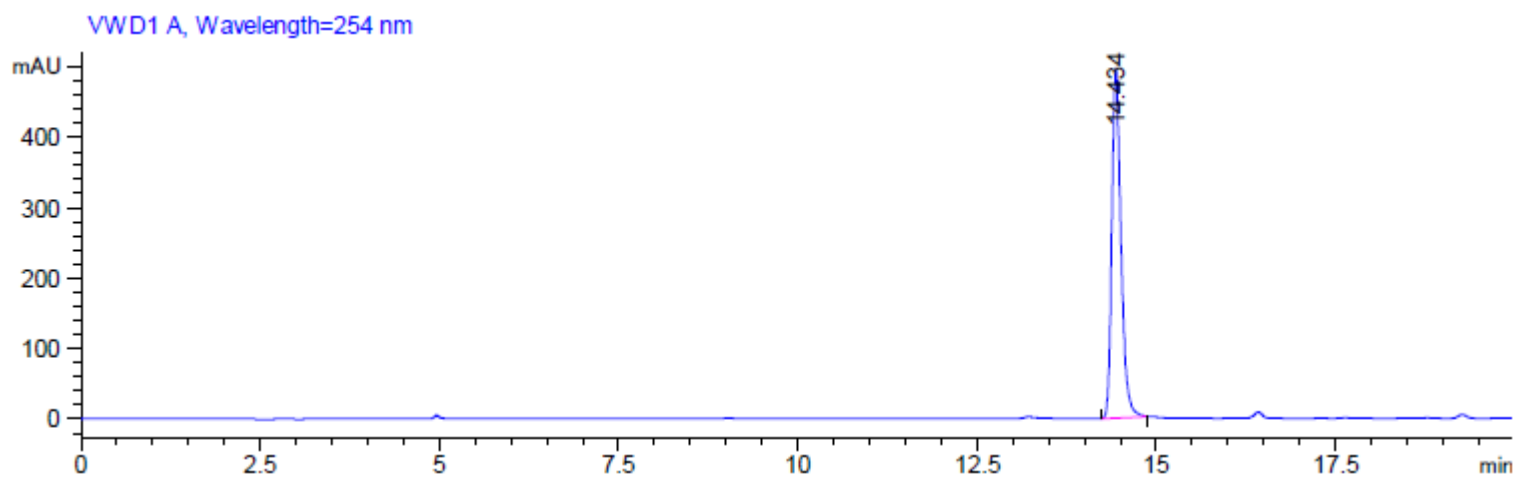


p^{m6}AmpG

Chemical structure

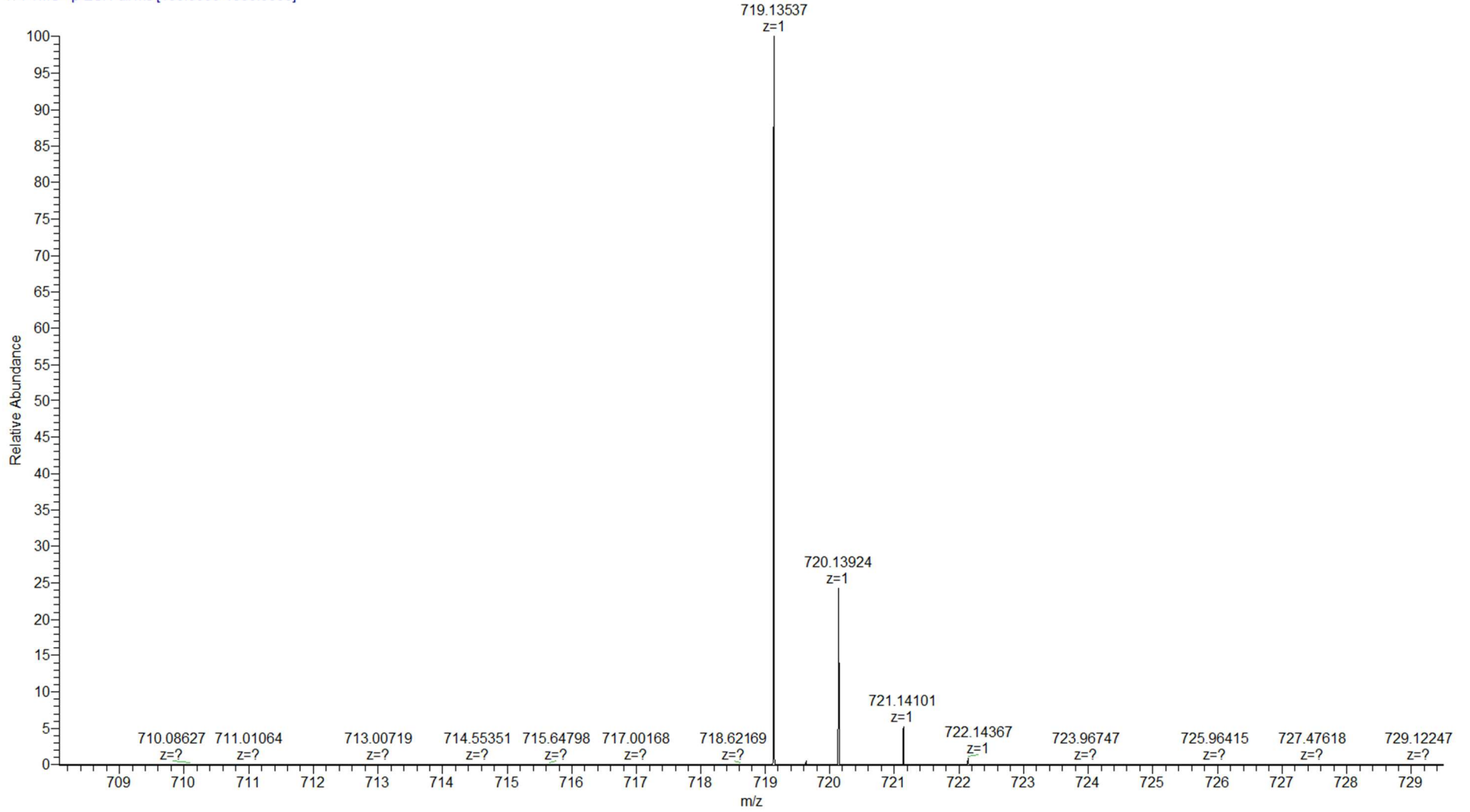


RP HPLC

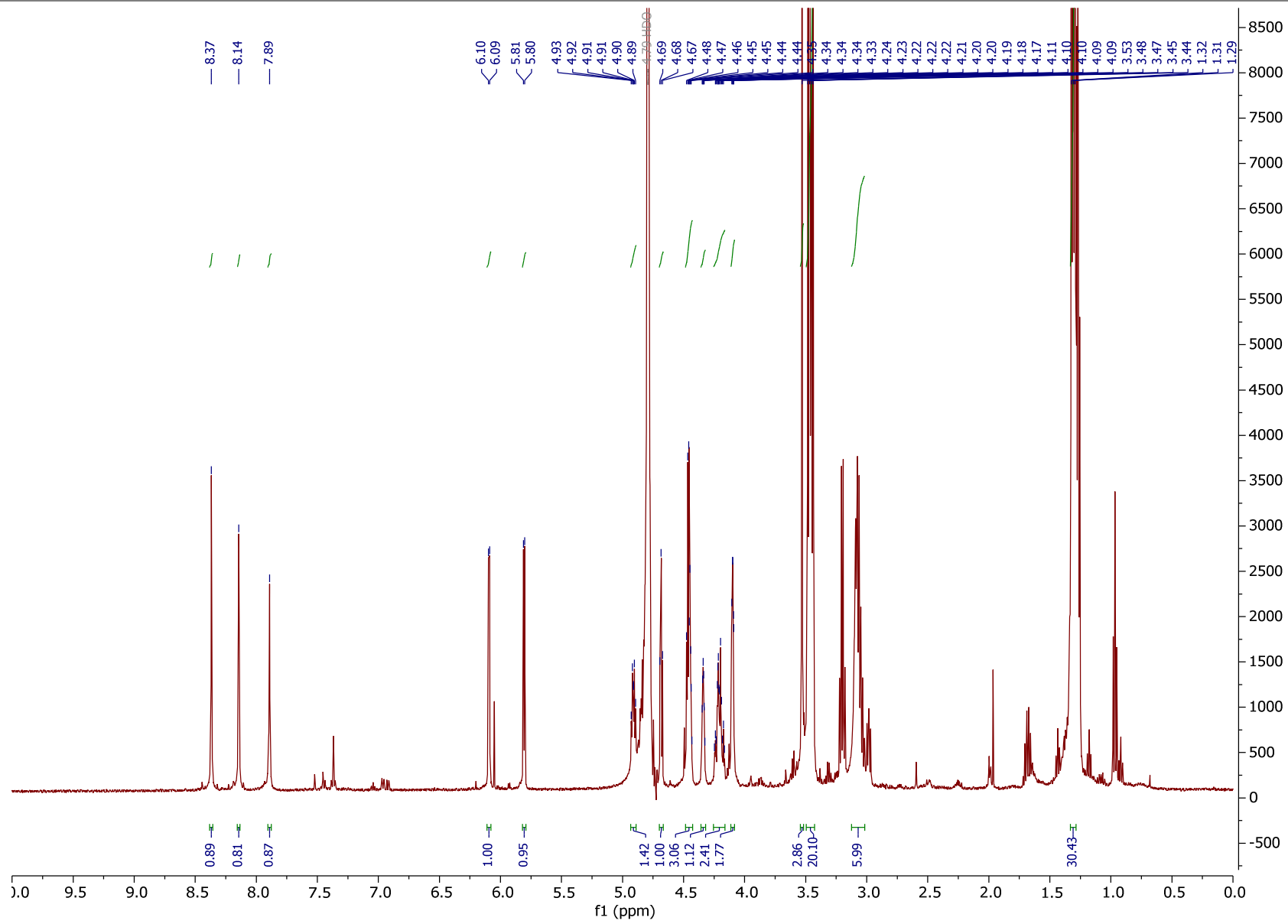


171213_TP_004#29-112 RT: 0.25-0.98 AV: 84 NL: 4.41E6
T: FTMS - p ESI Full ms [100.0000-1500.0000]

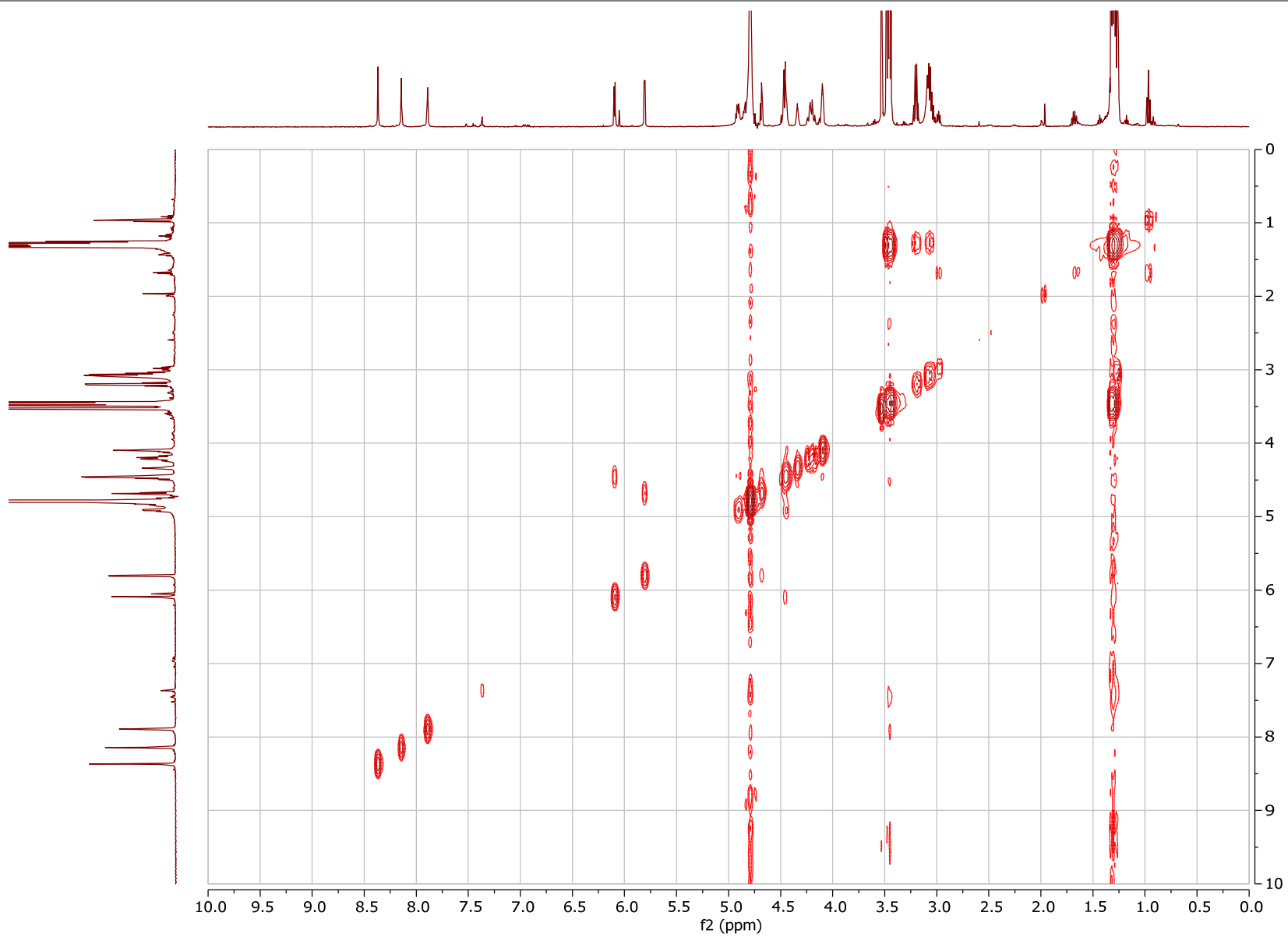
MS (-) ESI
(Calc. [M-H]⁻ C₂₂H₂₈N₁₀O₁₄P₂⁻ 719.13454)



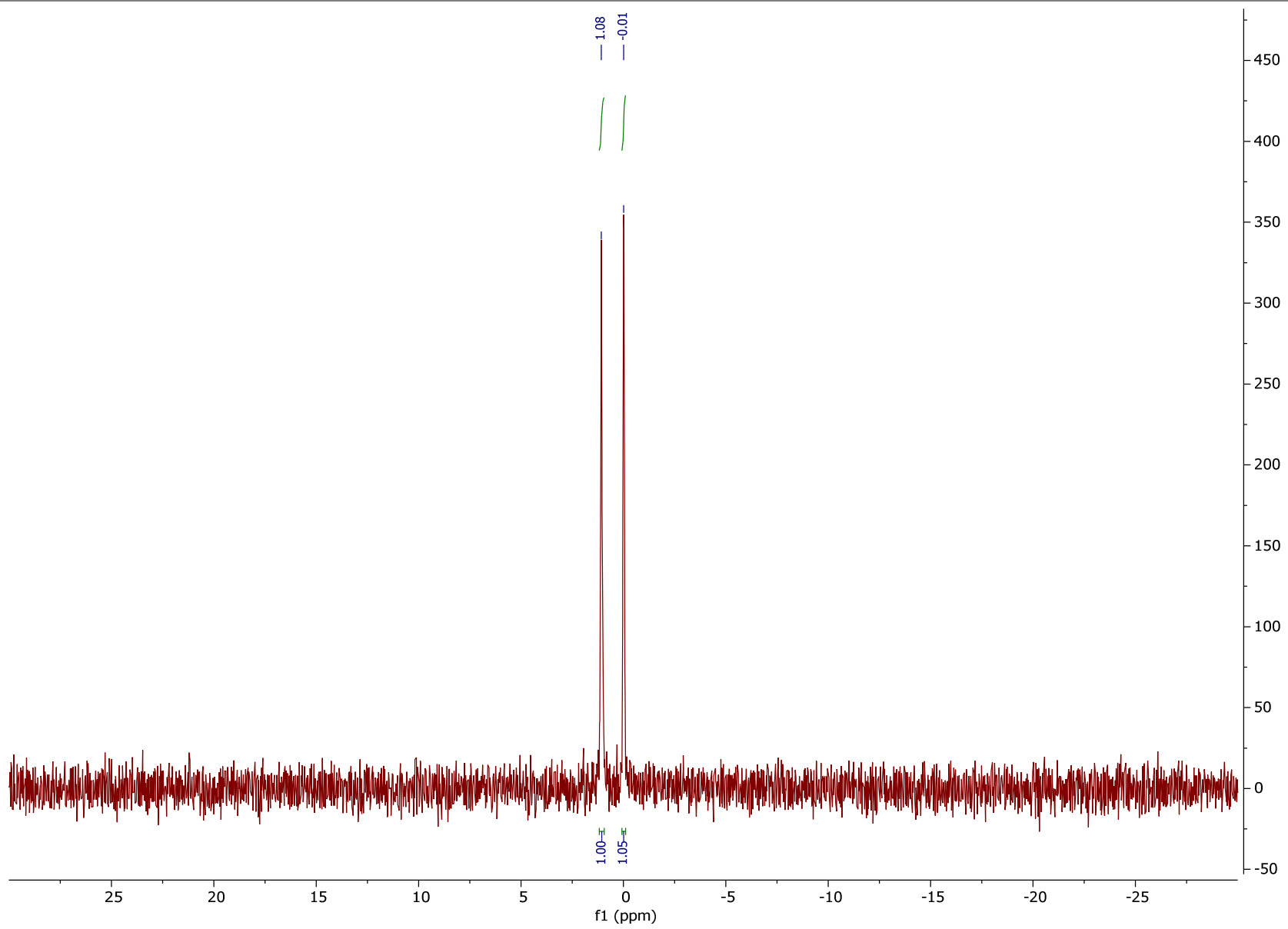
¹H NMR



COSY NMR

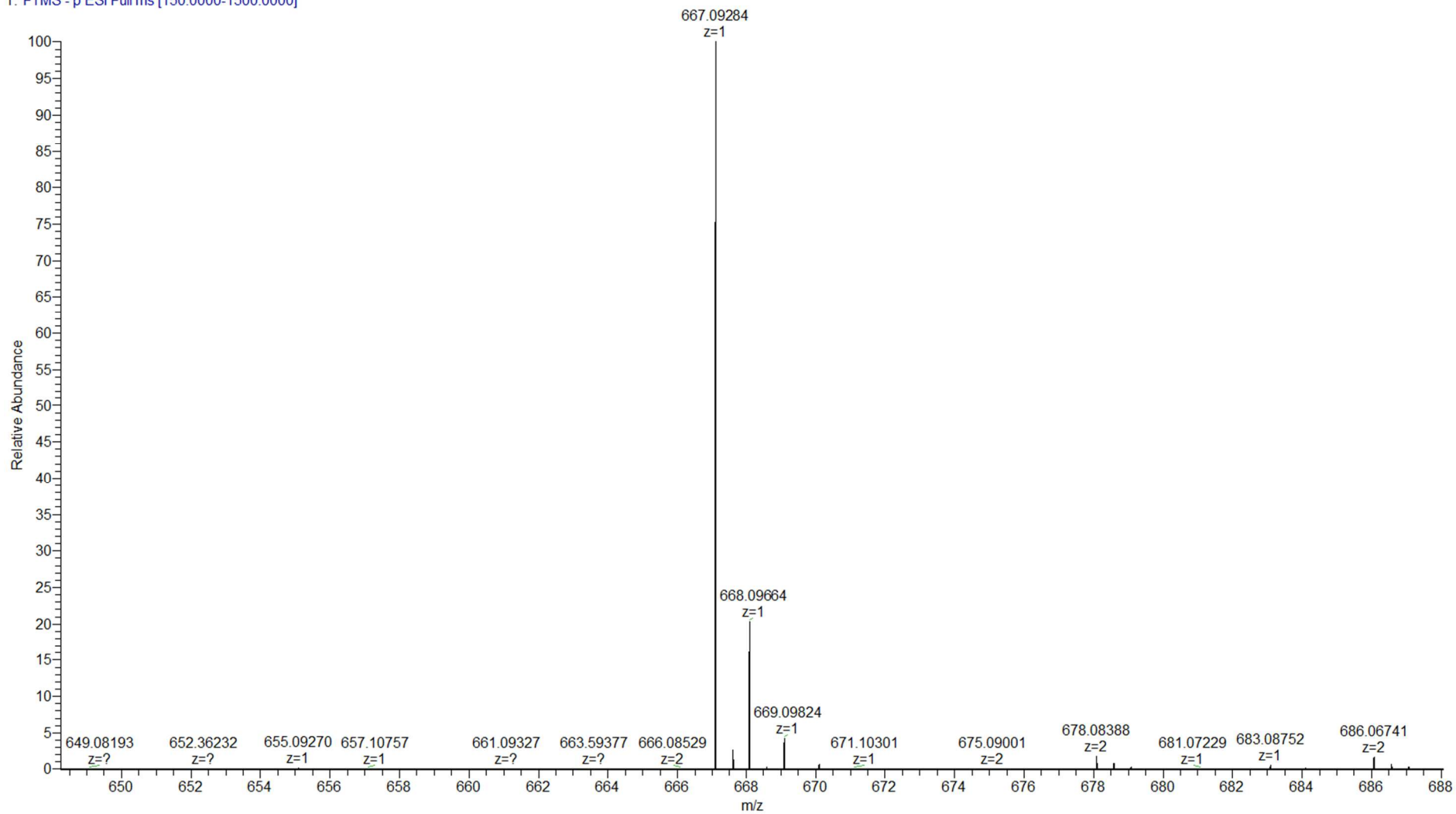


³¹P NMR



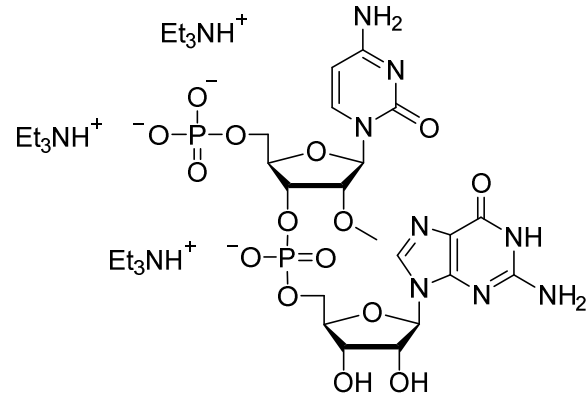
171213_TP_016 #53-162 RT: 0.46-1.41 AV: 110 NL: 3.00E6
T: FTMS - p ESI Full ms [150.0000-1500.0000]

MS (-) ESI
(Calc. [M-H]⁻ C₁₉H₂₅N₈O₁₅P₂⁻ 667.09201)

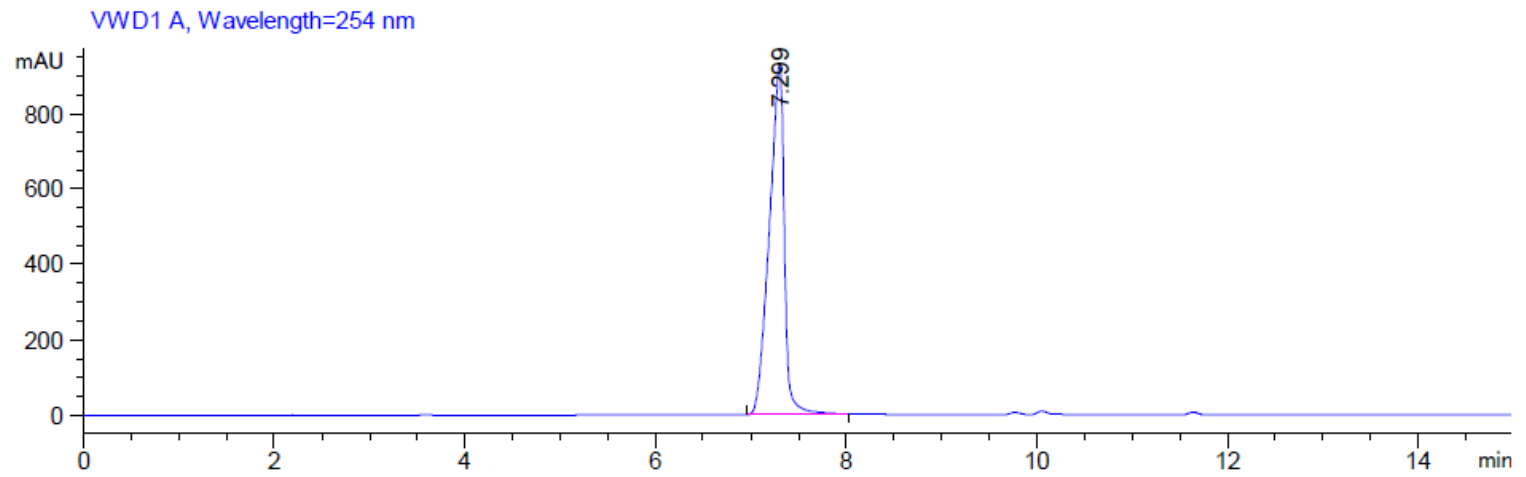


pC_mpG

Chemical structure

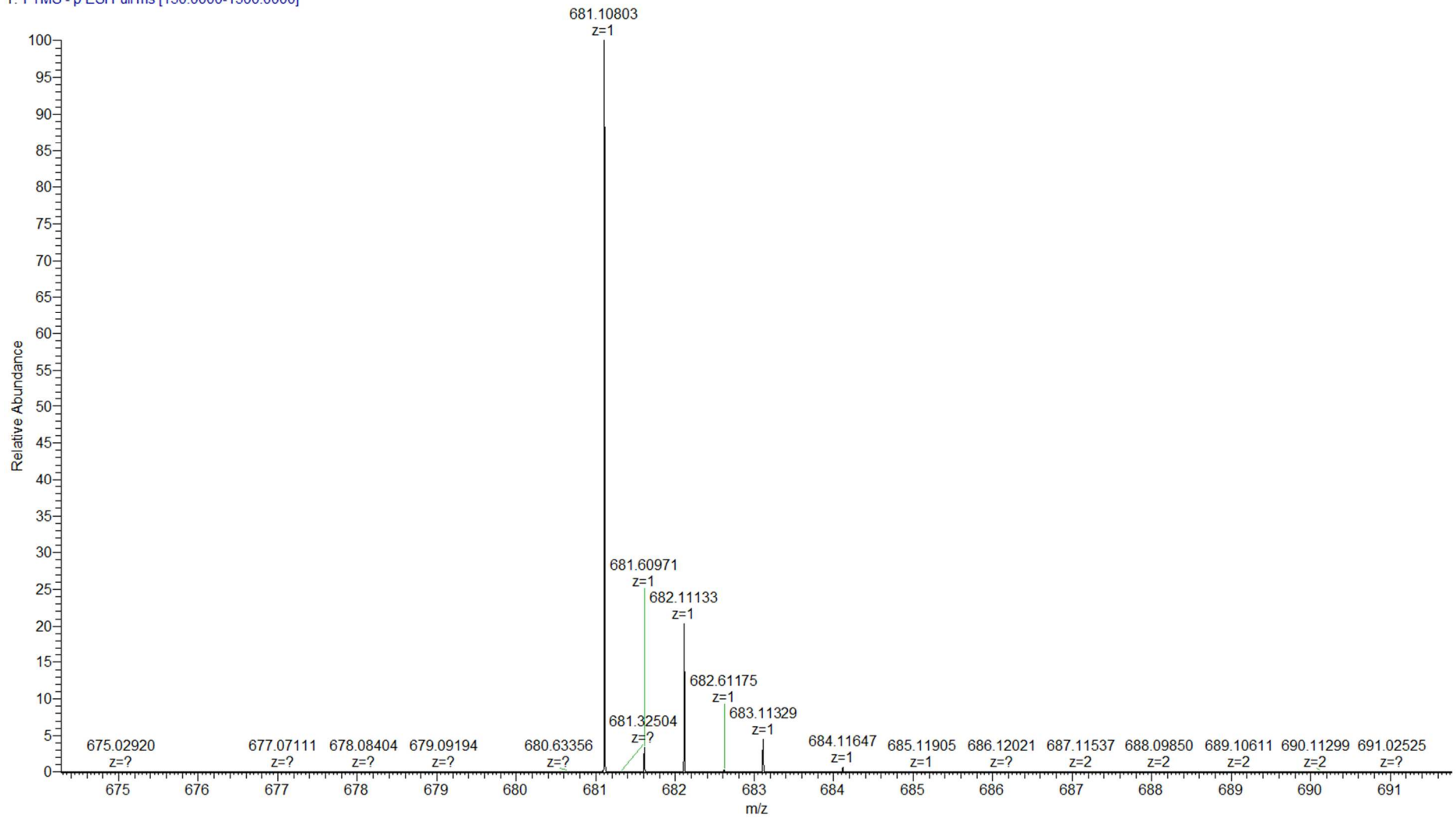


RP HPLC



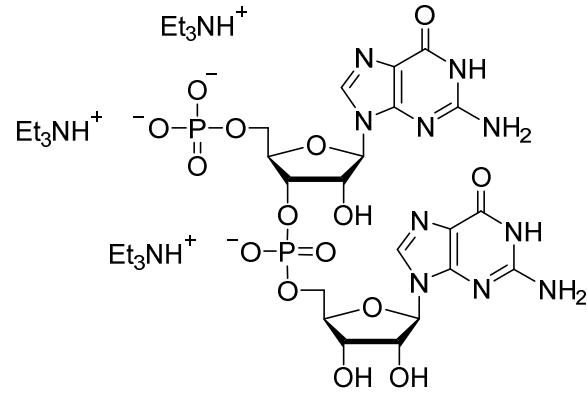
MS (-) ESI
(Calc. [M-H]⁻ C₂₀H₂₇N₈O₁₅P₂⁻ 681.10766)

171213_TP_017 #8-45 RT: 0.07-0.39 AV: 38 NL: 1.13E7
T: FTMS - p ESI Full ms [150.0000-1500.0000]

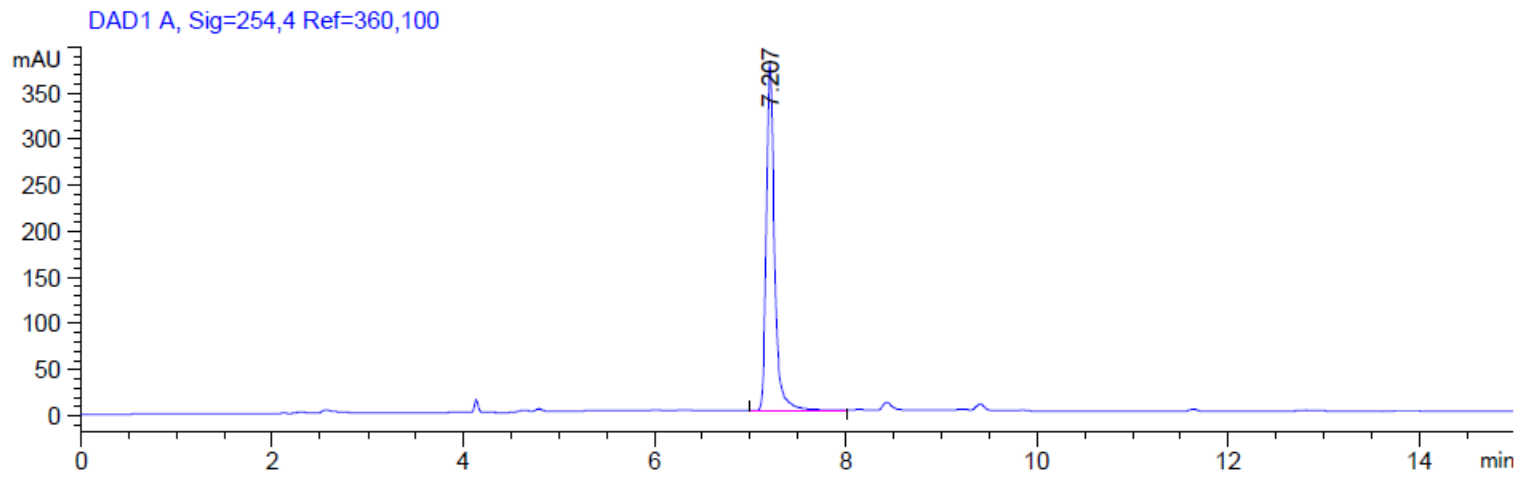


pGpG

Chemical structure

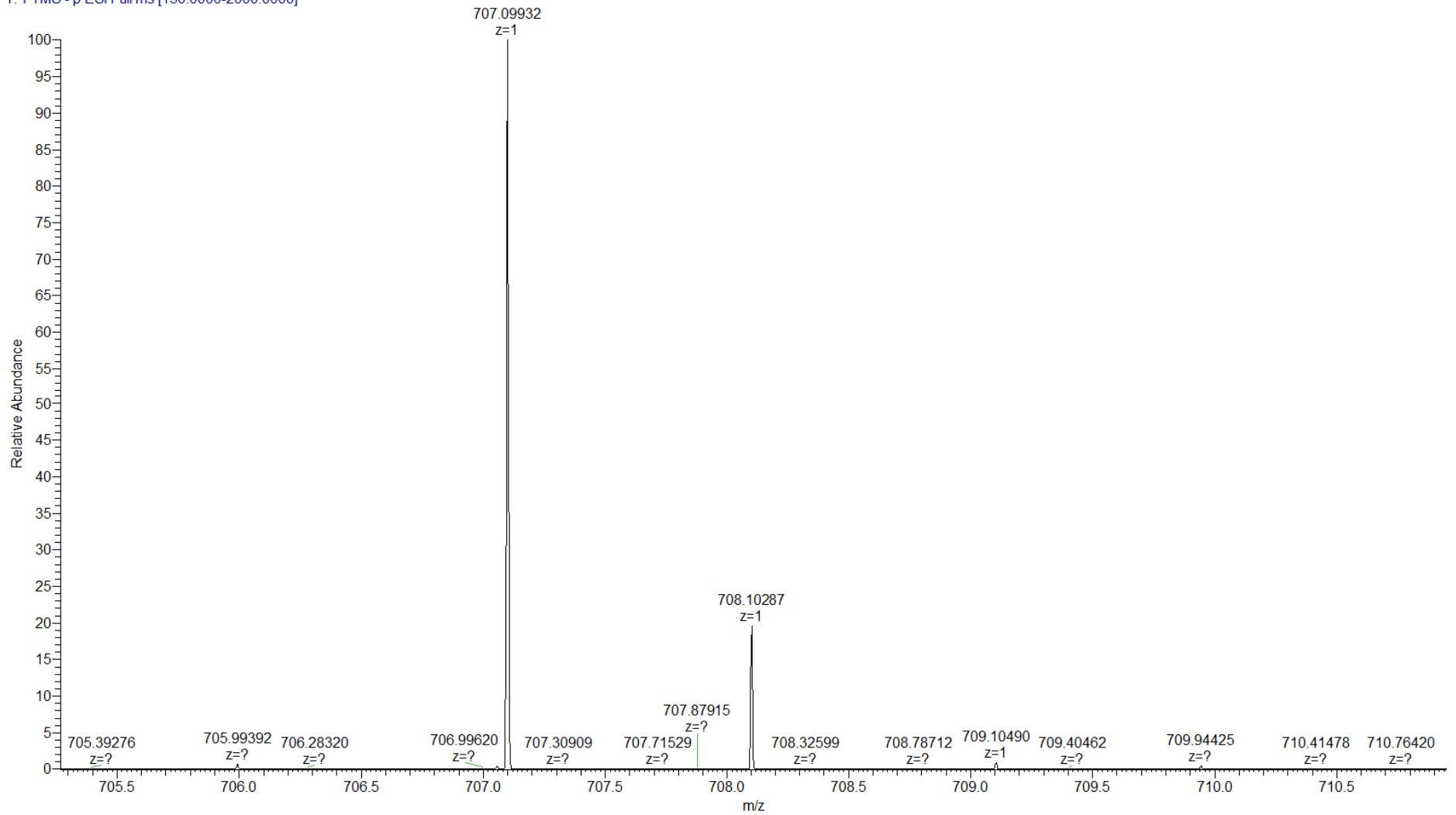


RP HPLC



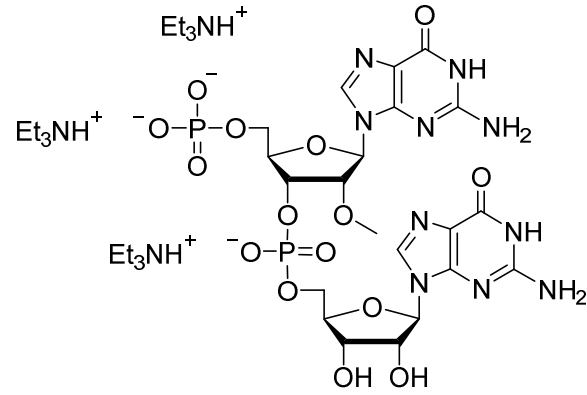
MS (-) ESI
(Calc. [M-H]⁻ C₂₀H₂₅N₁₀O₁₅P₂⁻ 707.09816)

190711_TR_136#416-735 RT: 4.15-7.48 AV: 320 NL: 1.90E4
T: FTMS - p ESI Full ms [150.0000-2000.0000]

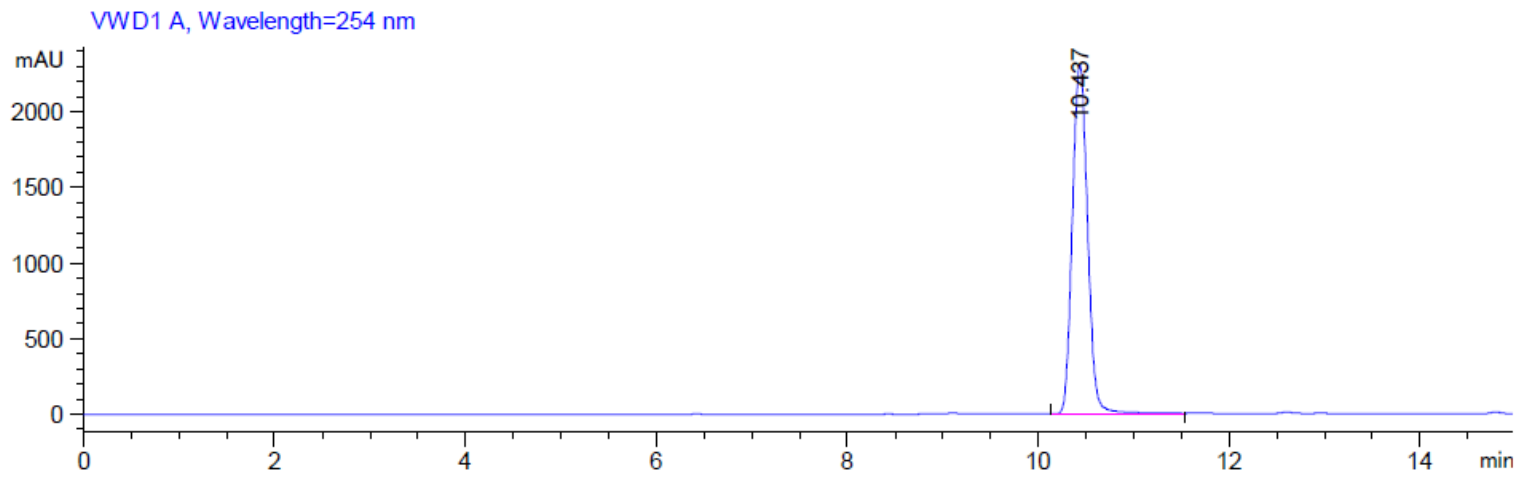


pG_mpG

Chemical structure

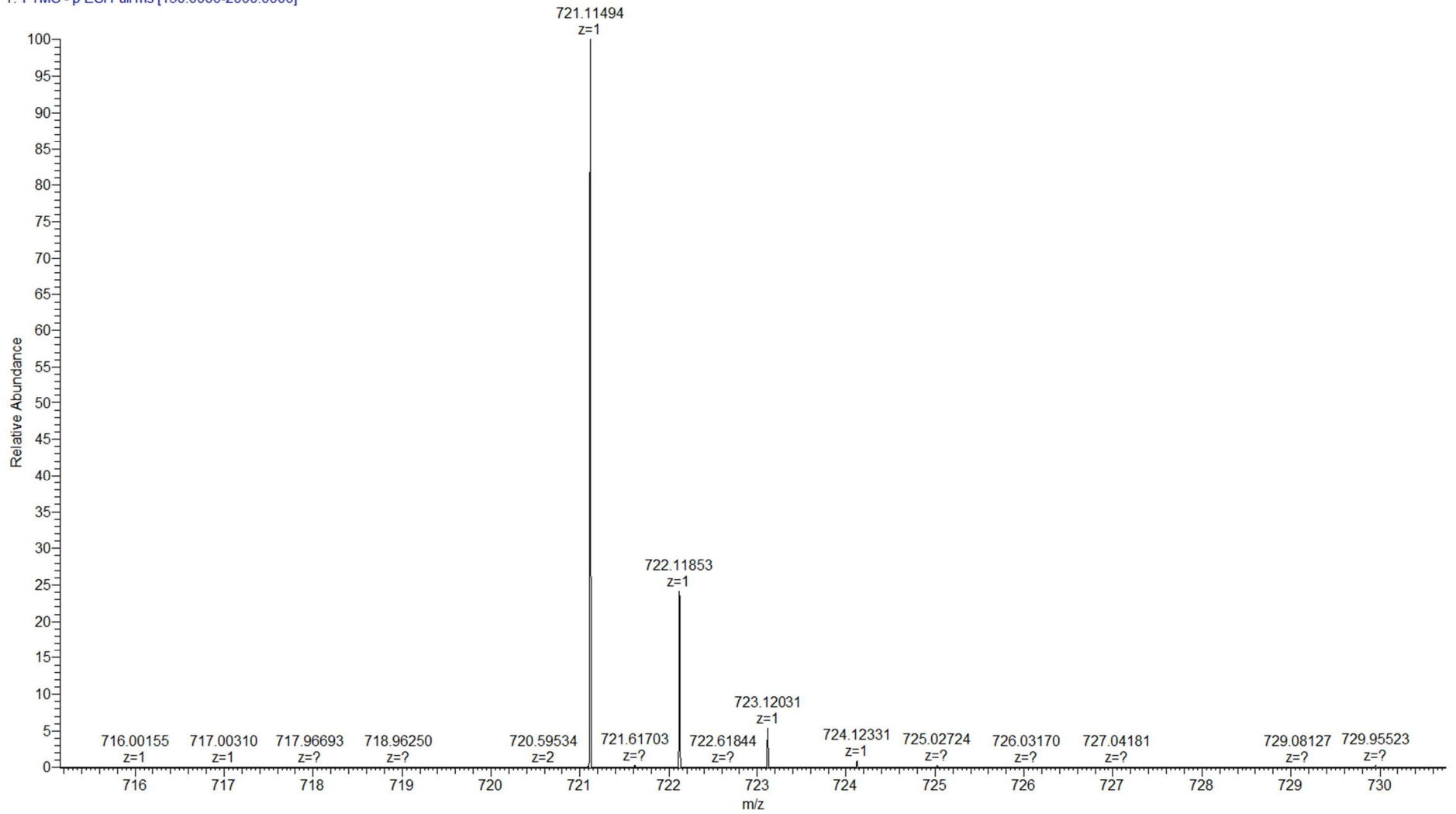


RP HPLC



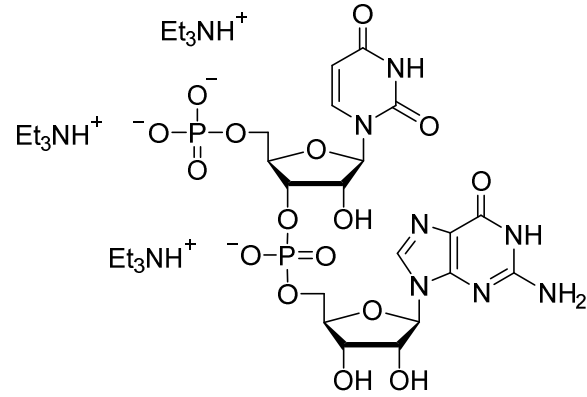
MS (-) ESI
(Calc. [M-H]⁻ C₂₁H₂₇N₁₀O₁₅P₂⁻ 721.11381)

190711_TR_137 #6-39 RT: 0.06-0.40 AV: 34 NL: 1.46E6
T: FTMS - p ESI Full ms [150.0000-2000.0000]

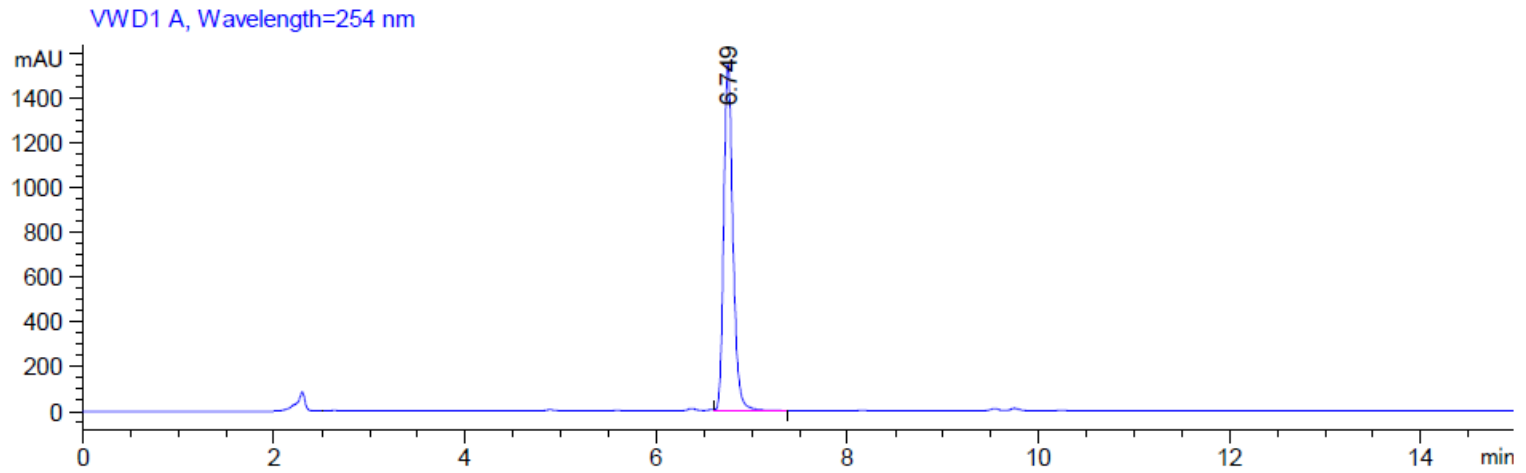


pUpG

Chemical structure

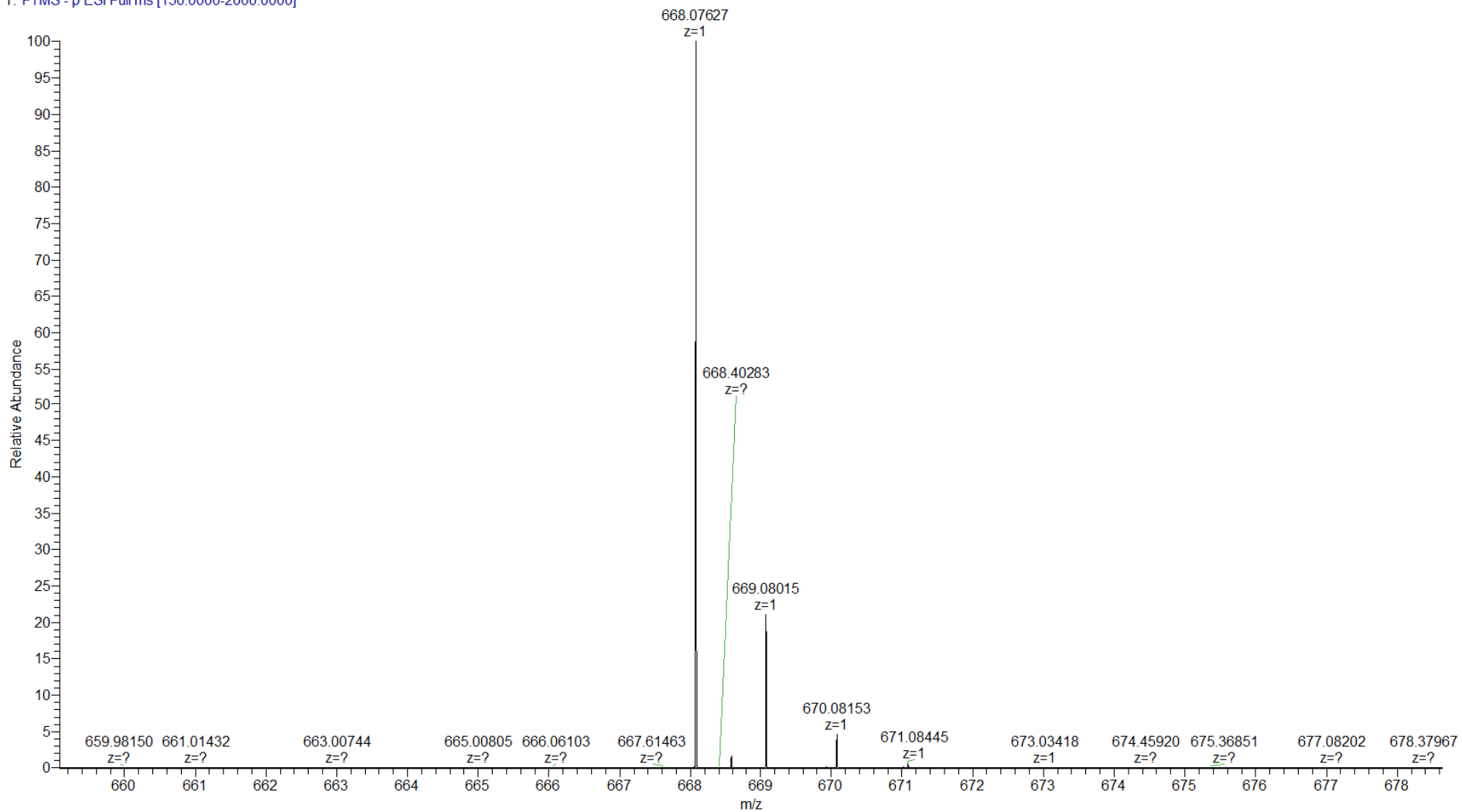


RP HPLC



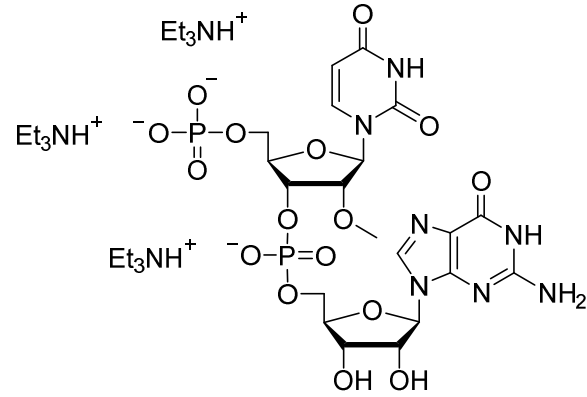
190528_AD_76 #88-160 RT: 0.86-1.56 AV: 73 NL: 5.71E6
T: FTMS - p ESI Full ms [150.0000-2000.0000]

MS (-) ESI
(Calc. [M-H]⁻ C₁₉H₂₄N₇O₁₆P₂⁻ 668.07602)

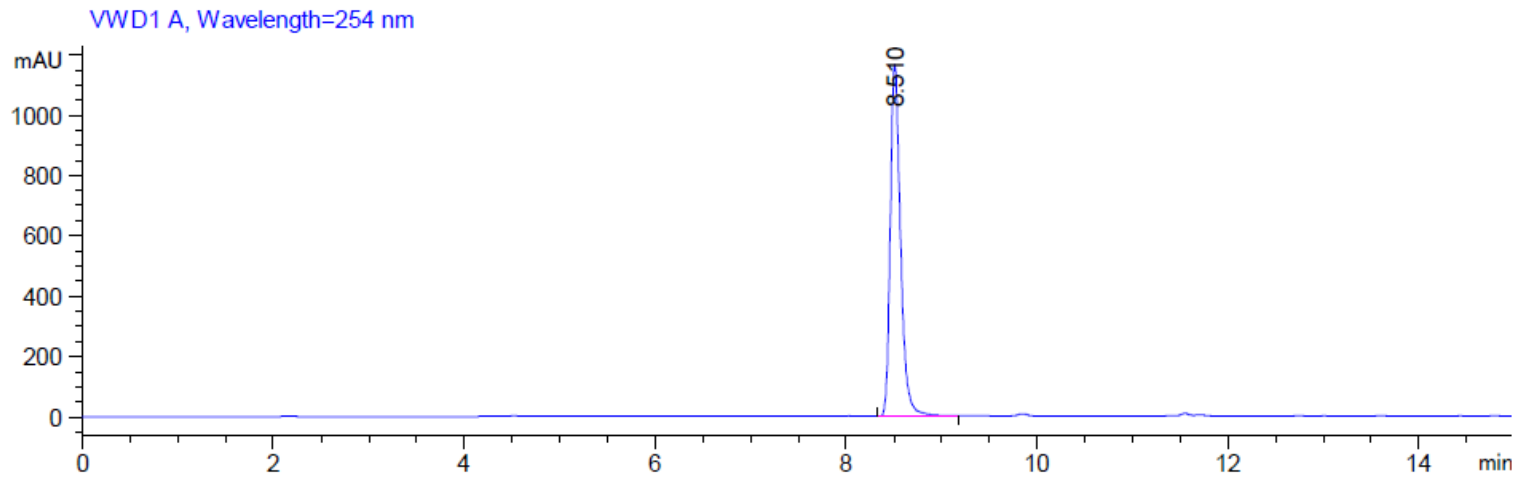


pU_mpG

Chemical structure



RP HPLC



171213_TP_019#44-127 RT: 0.38-1.11 AV: 84 NL: 8.73E6
T: FTMS - p ESI Full ms [150.0000-1500.0000]

MS (-) ESI
(Calc. [M-H]⁻ C₂₀H₂₆N₇O₁₆P₂⁻ 682.09167)

